



2023 CFA® Exam Prep

SchweserNotes™

Quantitative Methods and Economics



LEVEL I BOOK 1

KAPLAN SCHWEISER

Kaplan Schweser's Path to Success

Level I CFA® Exam

CFA®

Welcome

As the head of Advanced Designations at Kaplan Schweser, I am pleased to have the opportunity to help you prepare for the CFA® exam. Kaplan Schweser has decades of experience in delivering the most effective CFA exam prep products in the market and I know you will find them to be invaluable in your studies.

Our products are designed to be an integrated study solution across print and digital media to provide you the best learning experience, whether you are studying with a physical book, online, or on your mobile device.

Our core product, the SchweserNotes™, addresses all of the Topics, Readings, and LOS in the CFA curriculum. Each reading in the SchweserNotes has been broken into smaller, bite-sized modules with Module Quizzes interspersed throughout to help you continually assess your comprehension. After you complete each Topic, take our online Topic Quiz to help you assess your knowledge of the material before you move on to the next section.

All purchasers of the SchweserNotes receive online access to the Kaplan Schweser online platform (our learning management system or LMS) at www.Schweser.com. In the LMS, you will see a dashboard that tracks your overall progress and performance and also includes an Activity Feed, which provides structure and organization to the tasks required to prepare for the CFA exam. You also have access to the SchweserNotes, Module Quizzes, Topic Quizzes, and Mock Exams, as well as the SchweserNotes Videos (if purchased), which contain a short video that complements each module in the SchweserNotes. Look for the icons indicating where video content, Module Quizzes, Topic Quizzes, and Mock Exams are available online. I strongly encourage you to use the dashboard to track your progress and stay motivated.

Again, thank you for trusting Kaplan Schweser with your CFA exam preparation. We're here to help you throughout your journey to become a CFA charterholder.

Regards,

Derek Burkett

Derek Burkett, CFA, FRM, CAIA

Vice President (Advanced Designations)

Contact us for questions about your study package, upgrading your package, purchasing additional study materials, or for additional information:

888.325.5072 (U.S.) | +1 608.779.8327 (Int'l.)

staff@schweser.com | www.schweser.com/cfa

Book 1: Quantitative Methods and Economics

SchweserNotes™ 2023

Level I CFA®



SCHWESERNOTES™ 2023 LEVEL I CFA® BOOK 1: QUANTITATIVE METHODS AND ECONOMICS

©2022 Kaplan, Inc. All rights reserved.

Published in 2022 by Kaplan, Inc.

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

ISBN: 978-1-0788-2585-6

These materials may not be copied without written permission from the author. The unauthorized duplication of these notes is a violation of global copyright laws and the CFA Institute Code of Ethics. Your assistance in pursuing potential violators of this law is greatly appreciated.

Required CFA Institute disclaimer: Kaplan Schweser is a CFA Institute Prep Provider. Only CFA Institute Prep Providers are permitted to make use of CFA Institute copyrighted materials which are the building blocks of the exam. We are also required to update our materials every year and this is validated by CFA Institute.

CFA Institute does not endorse, promote, review or warrant the accuracy or quality of the product and services offered by Kaplan Schweser. CFA Institute®, CFA® and "Chartered Financial Analyst®" are trademarks owned by CFA Institute.

Certain materials contained within this text are the copyrighted property of CFA Institute. The following is the copyright disclosure for these materials: "Copyright, 2022, CFA Institute. Reproduced and republished from 2023 Learning Outcome Statements, Level I, II, and III questions from CFA® Program Materials, CFA Institute Standards of Professional Conduct, and CFA Institute's Global Investment Performance Standards with permission from CFA Institute. All Rights Reserved."

Disclaimer: The SchweserNotes should be used in conjunction with the original readings as set forth by CFA Institute in their 2023 Level I CFA Study Guide. The information contained in these Notes covers topics contained in the readings referenced by CFA Institute and is believed to be accurate. However, their accuracy cannot be guaranteed nor is any warranty conveyed as to your ultimate exam success. The authors of the referenced readings have not endorsed or sponsored these Notes.

CONTENTS

Learning Outcome Statements (LOS)

QUANTITATIVE METHODS

READING 1

The Time Value of Money

Exam Focus

Module 1.1: EAY and Compounding Frequency

Module 1.2: Calculating PV and FV

Module 1.3: Uneven Cash Flows

Module 1.4: Compounding Frequencies

Key Concepts

Answer Key for Module Quizzes

READING 2

Organizing, Visualizing, and Describing Data

Exam Focus

Module 2.1: Organizing Data

Module 2.2: Visualizing Data

Module 2.3: Measures of Central Tendency

Module 2.4: Measures of Location and Dispersion

Module 2.5: Skewness, Kurtosis, and Correlation

Key Concepts

Answer Key for Module Quizzes

READING 3

Probability Concepts

Exam Focus

Module 3.1: Conditional and Joint Probabilities

Module 3.2: Conditional Expectations and Expected Value

Module 3.3: Portfolio Variance, Bayes, and Counting Problems

Key Concepts

Answers to Module Quiz Questions

READING 4

Common Probability Distributions

Exam Focus

Module 4.1: Uniform and Binomial Distributions

Module 4.2: Normal Distributions

Module 4.3: Lognormal, T, Chi-Square, and F Distributions

Key Concepts

Answer Key for Module Quizzes

READING 5

Sampling and Estimation

Exam Focus

Module 5.1: Sampling Methods, Central Limit Theorem, and Standard Error

Module 5.2: Confidence Intervals, Resampling, and Sampling Biases

Key Concepts

Answer Key for Module Quizzes

READING 6

Hypothesis Testing

Exam Focus

Module 6.1: Hypothesis Tests and Types of Errors

Module 6.2: *p*-Values and Tests of Means

Module 6.3: Mean Differences and Difference in Means

Module 6.4: Tests of Variance, Correlation, and Independence

Key Concepts

Answer Key for Module Quizzes

READING 7

Introduction to Linear Regression

Exam Focus

Module 7.1: Linear Regression: Introduction

Module 7.2: Goodness of Fit and Hypothesis Tests

Module 7.3: Predicting Dependent Variables and Functional Forms

Key Concepts

Answer Key for Module Quizzes

Topic Quiz: Quantitative Methods

ECONOMICS

READING 8

Topics in Demand and Supply Analysis

Exam Focus

Module 8.1: Elasticity

Module 8.2: Demand and Supply

Key Concepts

Answer Key for Module Quizzes

READING 9

The Firm and Market Structures

Exam Focus

Module 9.1: Perfect Competition

Module 9.2: Monopolistic Competition

Module 9.3: Oligopoly

Module 9.4: Monopoly and Concentration

Key Concepts

Answer Key for Module Quizzes

READING 10

Aggregate Output, Prices, and Economic Growth

Exam Focus

Module 10.1: GDP, Income, and Expenditures

Module 10.2: Aggregate Demand and Supply

Module 10.3: Macroeconomic Equilibrium and Growth

Key Concepts

Answer Key for Module Quizzes

READING 11

Understanding Business Cycles

Exam Focus

Module 11.1: Business Cycle Phases

Module 11.2: Inflation and Indicators

Key Concepts

Answer Key for Module Quizzes

READING 12

Monetary and Fiscal Policy

Exam Focus

Module 12.1: Money and Inflation

Module 12.2: Monetary Policy

Module 12.3: Fiscal Policy

Key Concepts

Answer Key for Module Quizzes

READING 13

Introduction to Geopolitics

Exam Focus

Module 13.1: Geopolitics and Geopolitical Risk

Key Concepts

Answer Key for Module Quizzes

READING 14

International Trade and Capital Flows

Exam Focus

Module 14.1: International Trade Benefits

Module 14.2: Trade Restrictions

Key Concepts

Answer Key for Module Quizzes

READING 15

Currency Exchange Rates

Exam Focus

Module 15.1: Foreign Exchange Rates

Module 15.2: Forward Exchange Rates

Module 15.3: Managing Exchange Rates

Key Concepts

Answer Key for Module Quizzes

Topic Quiz: Economics

Formulas

Appendices

Appendix A: Areas Under The Normal Curve

Cumulative Z-Table

Appendix B: Student's t-Distribution

Appendix C: F-Table at 5% (Upper Tail)

Appendix D: F-Table at 2.5% (Upper Tail)

Appendix E: Chi-Squared Table

Index

LEARNING OUTCOME STATEMENTS (LOS)

1. The Time Value of Money

The candidate should be able to:

- a. interpret interest rates as required rates of return, discount rates, or opportunity costs.
- b. explain an interest rate as the sum of a real risk-free rate and premiums that compensate investors for bearing distinct types of risk.
- c. calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows.
- d. demonstrate the use of a time line in modeling and solving time value of money problems.
- e. calculate the solution for time value of money problems with different frequencies of compounding.
- f. calculate and interpret the effective annual rate, given the stated annual interest rate and the frequency of compounding.

2. Organizing, Visualizing, and Describing Data

The candidate should be able to:

- a. identify and compare data types.
- b. describe how data are organized for quantitative analysis.
- c. interpret frequency and related distributions.
- d. interpret a contingency table.
- e. describe ways that data may be visualized and evaluate uses of specific visualizations.
- f. describe how to select among visualization types.
- g. calculate and interpret measures of central tendency.
- h. evaluate alternative definitions of mean to address an investment problem.
- i. calculate quantiles and interpret related visualizations.
- j. calculate and interpret measures of dispersion.
- k. calculate and interpret target downside deviation.
- l. interpret skewness.
- m. interpret kurtosis.
- n. interpret correlation between two variables.

3. Probability Concepts

The candidate should be able to:

- a. define a random variable, an outcome, and an event.
- b. identify the two defining properties of probability, including mutually exclusive and exhaustive events, and compare and contrast empirical, subjective, and a priori probabilities.
- c. describe the probability of an event in terms of odds for and against the event.
- d. calculate and interpret conditional probabilities.
- e. demonstrate the application of the multiplication and addition rules for probability.
- f. compare and contrast dependent and independent events.
- g. calculate and interpret an unconditional probability using the total probability rule.
- h. calculate and interpret the expected value, variance, and standard deviation of random variables.
- i. explain the use of conditional expectation in investment applications.
- j. interpret a probability tree and demonstrate its application to investment problems.
- k. calculate and interpret the expected value, variance, standard deviation, covariances, and correlations of portfolio returns.
- l. calculate and interpret the covariances of portfolio returns using the joint probability function.
- m. calculate and interpret an updated probability using Bayes' formula.
- n. identify the most appropriate method to solve a particular counting problem and analyze counting problems using factorial, combination, and permutation concepts.

4. Common Probability Distributions

The candidate should be able to:

- a. define a probability distribution and compare and contrast discrete and continuous random variables and their probability functions.
- b. calculate and interpret probabilities for a random variable given its cumulative distribution function.
- c. describe the properties of a discrete uniform random variable, and calculate and interpret probabilities given the discrete uniform distribution function.

- d. describe the properties of the continuous uniform distribution, and calculate and interpret probabilities given a continuous uniform distribution.
- e. describe the properties of a Bernoulli random variable and a binomial random variable, and calculate and interpret probabilities given the binomial distribution function.
- f. explain the key properties of the normal distribution.
- g. contrast a multivariate distribution and a univariate distribution, and explain the role of correlation in the multivariate normal distribution.
- h. calculate the probability that a normally distributed random variable lies inside a given interval.
- i. explain how to standardize a random variable.
- j. calculate and interpret probabilities using the standard normal distribution.
- k. define shortfall risk, calculate the safety-first ratio, and identify an optimal portfolio using Roy's safety-first criterion.
- l. explain the relationship between normal and lognormal distributions and why the lognormal distribution is used to model asset prices.
- m. calculate and interpret a continuously compounded rate of return, given a specific holding period return.
- n. describe the properties of the Student's t -distribution, and calculate and interpret its degrees of freedom.
- o. describe the properties of the chi-square distribution and the F -distribution, and calculate and interpret their degrees of freedom.
- p. describe Monte Carlo simulation.

5. Sampling and Estimation

The candidate should be able to:

- a. compare and contrast probability samples with non-probability samples and discuss applications of each to an investment problem.
- b. explain sampling error.
- c. compare and contrast simple random, stratified random, cluster, convenience, and judgmental sampling.
- d. explain the central limit theorem and its importance.
- e. calculate and interpret the standard error of the sample mean.
- f. identify and describe desirable properties of an estimator.
- g. contrast a point estimate and a confidence interval estimate of a population parameter.
- h. calculate and interpret a confidence interval for a population mean, given a normal distribution with 1) a known population variance, 2) an unknown population variance, or 3) an unknown population variance and a large sample size.
- i. describe the use of resampling (bootstrap, jackknife) to estimate the sampling distribution of a statistic.
- j. describe the issues regarding selection of the appropriate sample size, data snooping bias, sample selection bias, survivorship bias, look-ahead bias, and time-period bias.

6. Hypothesis Testing

The candidate should be able to:

- a. define a hypothesis, describe the steps of hypothesis testing, and describe and interpret the choice of the null and alternative hypotheses.
- b. compare and contrast one-tailed and two-tailed tests of hypotheses.
- c. explain a test statistic, Type I and Type II errors, a significance level, how significance levels are used in hypothesis testing, and the power of a test.
- d. explain a decision rule and the relation between confidence intervals and hypothesis tests, and determine whether a statistically significant result is also economically meaningful.
- e. explain and interpret the p -value as it relates to hypothesis testing.
- f. describe how to interpret the significance of a test in the context of multiple tests.
- g. identify the appropriate test statistic and interpret the results for a hypothesis test concerning the population mean of both large and small samples when the population is normally or approximately normally distributed and the variance is (1) known or (2) unknown.
- h. identify the appropriate test statistic and interpret the results for a hypothesis test concerning the equality of the population means of two at least approximately normally distributed populations based on independent random samples with equal assumed variances.
- i. identify the appropriate test statistic and interpret the results for a hypothesis test concerning the mean difference of two normally distributed populations.
- j. identify the appropriate test statistic and interpret the results for a hypothesis test concerning (1) the variance of a normally distributed population and (2) the equality of the variances of two normally distributed populations based on two independent random samples.

- k. compare and contrast parametric and nonparametric tests, and describe situations where each is the more appropriate type of test.
- l. explain parametric and nonparametric tests of the hypothesis that the population correlation coefficient equals zero, and determine whether the hypothesis is rejected at a given level of significance.
- m. explain tests of independence based on contingency table data.

7. Introduction to Linear Regression

The candidate should be able to:

- a. describe a simple linear regression model and the roles of the dependent and independent variables in the model.
- b. describe the least squares criterion, how it is used to estimate regression coefficients, and their interpretation.
- c. explain the assumptions underlying the simple linear regression model, and describe how residuals and residual plots indicate if these assumptions may have been violated.
- d. calculate and interpret the coefficient of determination and the F -statistic in a simple linear regression.
- e. describe the use of analysis of variance (ANOVA) in regression analysis, interpret ANOVA results, and calculate and interpret the standard error of estimate in a simple linear regression.
- f. formulate a null and an alternative hypothesis about a population value of a regression coefficient, and determine whether the null hypothesis is rejected at a given level of significance.
- g. calculate and interpret the predicted value for the dependent variable, and a prediction interval for it, given an estimated linear regression model and a value for the independent variable.
- h. describe different functional forms of simple linear regressions.

8. Topics in Demand and Supply Analysis

The candidate should be able to:

- a. calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure.
- b. compare substitution and income effects.
- c. contrast normal goods with inferior goods.
- d. describe the phenomenon of diminishing marginal returns.
- e. determine and interpret breakeven and shutdown points of production.
- f. describe how economies of scale and diseconomies of scale affect costs.

9. The Firm and Market Structures

The candidate should be able to:

- a. describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly.
- b. explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure.
- c. describe a firm's supply function under each market structure.
- d. describe and determine the optimal price and output for firms under each market structure.
- e. describe pricing strategy under each market structure.
- f. explain factors affecting long-run equilibrium under each market structure.
- g. describe the use and limitations of concentration measures in identifying market structure.
- h. identify the type of market structure within which a firm operates.

10. Aggregate Output, Prices, and Economic Growth

The candidate should be able to:

- a. calculate and explain gross domestic product (GDP) using expenditure and income approaches.
- b. compare the sum-of-value-added and value-of-final-output methods of calculating GDP.
- c. compare nominal and real GDP and calculate and interpret the GDP deflator.
- d. compare GDP, national income, personal income, and personal disposable income.
- e. explain the fundamental relationship among saving, investment, the fiscal balance, and the trade balance.
- f. explain how the aggregate demand curve is generated.
- g. explain the aggregate supply curve in the short run and long run.
- h. explain causes of movements along and shifts in aggregate demand and supply curves.
- i. describe how fluctuations in aggregate demand and aggregate supply cause short-run changes in the economy and the business cycle.
- j. distinguish among the following types of macroeconomic equilibria: long-run full employment, short-run recessionary gap, short-run inflationary gap, and short-run stagflation.
- k. explain how a short-run macroeconomic equilibrium may occur at a level above or below full employment.
- l. analyze the effect of combined changes in aggregate supply and demand on the economy.

- m. describe sources, measurement, and sustainability of economic growth.
- n. describe the production function approach to analyzing the sources of economic growth.
- o. define and contrast input growth with growth of total factor productivity as components of economic growth.

11. Understanding Business Cycles

The candidate should be able to:

- a. describe the business cycle and its phases.
- b. describe credit cycles.
- c. describe how resource use, consumer and business activity, housing sector activity, and external trade sector activity vary as an economy moves through the business cycle.
- d. describe theories of the business cycle.
- e. interpret a set of economic indicators, and describe their uses and limitations.
- f. describe types of unemployment, and compare measures of unemployment.
- g. explain inflation, hyperinflation, disinflation, and deflation.
- h. explain the construction of indexes used to measure inflation.
- i. compare inflation measures, including their uses and limitations.
- j. contrast cost-push and demand-pull inflation.

12. Monetary and Fiscal Policy

The candidate should be able to:

- a. compare monetary and fiscal policy.
- b. describe functions and definitions of money.
- c. explain the money creation process.
- d. describe theories of the demand for and supply of money.
- e. describe the Fisher effect.
- f. describe roles and objectives of central banks.
- g. contrast the costs of expected and unexpected inflation.
- h. describe tools used to implement monetary policy.
- i. describe the monetary transmission mechanism.
- j. explain the relationships between monetary policy and economic growth, inflation, interest, and exchange rates.
- k. describe qualities of effective central banks.
- l. contrast the use of inflation, interest rate, and exchange rate targeting by central banks.
- m. determine whether a monetary policy is expansionary or contractionary.
- n. describe limitations of monetary policy.
- o. describe roles and objectives of fiscal policy.
- p. describe the arguments about whether the size of a national debt relative to GDP matters.
- q. describe tools of fiscal policy, including their advantages and disadvantages.
- r. explain the implementation of fiscal policy and difficulties of implementation.
- s. determine whether a fiscal policy is expansionary or contractionary.
- t. explain the interaction of monetary and fiscal policy.

13. Introduction to Geopolitics

The candidate should be able to:

- a. describe geopolitics from a cooperation versus competition perspective.
- b. describe geopolitics and its relationship with globalization.
- c. describe tools of geopolitics and their impact on regions and economies.
- d. describe geopolitical risk and its impact on investments.

14. International Trade and Capital Flows

The candidate should be able to:

- a. compare gross domestic product and gross national product.
- b. describe benefits and costs of international trade.
- c. contrast comparative advantage and absolute advantage.
- d. compare the Ricardian and Heckscher–Ohlin models of trade and the source(s) of comparative advantage in each model.
- e. compare types of trade and capital restrictions and their economic implications.
- f. explain motivations for and advantages of trading blocs, common markets, and economic unions.
- g. describe common objectives of capital restrictions imposed by governments.
- h. describe the balance of payments accounts including their components.
- i. explain how decisions by consumers, firms, and governments affect the balance of payments.

- j. describe functions and objectives of the international organizations that facilitate trade, including the World Bank, the International Monetary Fund, and the World Trade Organization.

15. Currency Exchange Rates

The candidate should be able to:

- a. define an exchange rate and distinguish between nominal and real exchange rates and spot and forward exchange rates.
- b. calculate and interpret the percentage change in a currency relative to another currency.
- c. describe functions of and participants in the foreign exchange market.
- d. calculate and interpret currency cross-rates.
- e. calculate an outright forward quotation from forward quotations expressed on a points basis or in percentage terms.
- f. explain the arbitrage relationship between spot rates, forward rates, and interest rates.
- g. calculate and interpret a forward discount or premium.
- h. calculate and interpret the forward rate consistent with the spot rate and the interest rate in each currency.
- i. describe exchange rate regimes.
- j. explain the effects of exchange rates on countries' international trade and capital flows.

READING 1

THE TIME VALUE OF MONEY

EXAM FOCUS

This reading covers time value of money concepts and applications. Procedures are presented for calculating the future value and present value of a single cash flow, an annuity, and a series of uneven cash flows. The impact of different compounding periods is examined, along with the procedures for solving for other variables in time value of money problems. Your main objective in this chapter is to master time value of money mechanics (i.e., learn how to crunch the numbers). Work all the questions and problems found at the end of this review. Make sure you know how to grind out all the time value of money problems on your calculator. The more rapidly you can do them (correctly), the more time you will have for the more conceptual parts of the exam.

MODULE 1.1: EASY AND COMPOUNDING FREQUENCY



Video covering
this content is
available online.

The concept of **compound interest** or **interest on interest** is deeply embedded in time value of money (TVM) procedures. When an investment is subjected to compound interest, the growth in the value of the investment from period to period reflects not only the interest earned on the original principal amount but also on the interest earned on the previous period's interest earnings—the interest on interest.

TVM applications frequently call for determining the **future value (FV)** of an investment's cash flows as a result of the effects of compound interest. Computing FV involves projecting the cash flows forward, on the basis of an appropriate compound interest rate, to the end of the investment's life. The computation of the **present value (PV)** works in the opposite direction—it brings the cash flows from an investment back to the beginning of the investment's life based on an appropriate compound rate of return.

Being able to measure the PV and/or FV of an investment's cash flows becomes useful when comparing investment alternatives because the value of the investment's cash flows must be measured at some common point in time, typically at the end of the investment horizon (FV) or at the beginning of the investment horizon (PV).

Using a Financial Calculator

It is very important that you be able to use a financial calculator when working TVM problems because the exam is constructed under the assumption that candidates have the ability to do so. There is simply no other way that you will have time to solve TVM problems. *CFA Institute allows only two types of calculators to be used for the exam—the TI BAII Plus® (including the*

BAII Plus Professional) and the HP 12C® (including the HP 12C Platinum). This reading is written primarily with the TI BAII Plus in mind. If you don't already own a calculator, go out and buy a TI BAII Plus! However, if you already own the HP 12C and are comfortable with it, by all means continue to use it.

The TI BAII Plus comes preloaded from the factory with the periods per year function (P/Y) set to 12. This automatically converts the annual interest rate (I/Y) into monthly rates. While appropriate for many loan-type problems, this feature is not suitable for the vast majority of the TVM applications we will be studying. So prior to using our SchweserNotes™, please set your P/Y key to "1" using the following sequence of keystrokes:

[2nd] [P/Y] "1" [ENTER] [2nd] [QUIT]

As long as you do not change the P/Y setting, it will remain set at one period per year until the battery from your calculator is removed (it does not change when you turn the calculator on and off). If you want to check this setting at any time, press [2nd] [P/Y]. The display should read P/Y = 1.0. If it does, press [2nd] [QUIT] to get out of the "programming" mode. If it doesn't, repeat the procedure previously described to set the P/Y key. With P/Y set to equal 1, it is now possible to think of I/Y as the interest rate per compounding period and N as the number of compounding periods under analysis. Thinking of these keys in this way should help you keep things straight as we work through TVM problems.

Before we begin working with financial calculators, you should familiarize yourself with your TI by locating the TVM keys noted below. These are the only keys you need to know to work virtually on all TVM problems.

- N = Number of compounding periods
- I/Y = Interest rate per compounding period
- PV = Present value
- FV = Future value
- PMT = Annuity payments, or constant periodic cash flow
- CPT = Compute



PROFESSOR'S NOTE

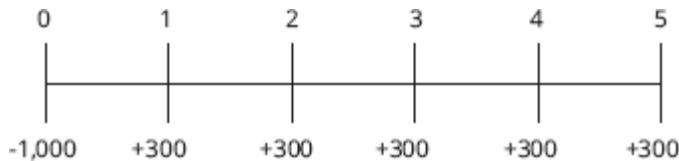
We have provided an online video in the Resource Library on how to use the TI calculator. You can view it by logging in at www.schweser.com.

Time Lines

It is often a good idea to draw a time line before you start to solve a TVM problem. A **time line** is simply a diagram of the cash flows associated with a TVM problem. A cash flow that occurs in the present (today) is put at time zero. Cash outflows (payments) are given a negative sign, and cash inflows (receipts) are given a positive sign. Once the cash flows are assigned to a time line, they may be moved to the beginning of the investment period to calculate the PV through a process called **discounting** or to the end of the period to calculate the FV using a process called **compounding**.

Figure 1.1 illustrates a time line for an investment that costs \$1,000 today (outflow) and will return a stream of cash payments (inflows) of \$300 per year at the end of each of the next five years.

Figure 1.1: Time Line



Please recognize that the cash flows occur at the end of the period depicted on the time line. Furthermore, note that the end of one period is the same as the beginning of the next period. For example, the end of the second year ($t = 2$) is the same as the beginning of the third year, so a cash flow at the beginning of Year 3 appears at time $t = 2$ on the time line. Keeping this convention in mind will help you keep things straight when you are setting up TVM problems.



PROFESSOR'S NOTE

Throughout the problems in this review, rounding differences may occur between the use of different calculators or techniques presented in this document. So don't panic if you are a few cents off in your calculations.

LOS 1.a: Interpret interest rates as required rates of return, discount rates, or opportunity costs.

Interest rates are our measure of the time value of money, although risk differences in financial securities lead to differences in their equilibrium interest rates. Equilibrium interest rates are the **required rate of return** for a particular investment, in the sense that the market rate of return is the return that investors and savers require to get them to willingly lend their funds. Interest rates are also referred to as **discount rates** and, in fact, the terms are often used interchangeably. If an individual can borrow funds at an interest rate of 10%, then that individual should *discount* payments to be made in the future at that rate in order to get their equivalent value in current dollars or other currencies. Finally, we can also view interest rates as the **opportunity cost** of current consumption. If the market rate of interest on 1-year securities is 5%, earning an additional 5% is the opportunity forgone when current consumption is chosen rather than saving (postponing consumption).

LOS 1.b: Explain an interest rate as the sum of a real risk-free rate and premiums that compensate investors for bearing distinct types of risk.

The **real risk-free rate** of interest is a theoretical rate on a single-period loan that has no expectation of inflation in it. When we speak of a real rate of return, we are referring to an investor's increase in purchasing power (after adjusting for inflation). Since expected inflation in future periods is not zero, the rates we observe on U.S. Treasury bills (T-bills), for example, are

risk-free rates but not *real* rates of return. T-bill rates are *nominal risk-free rates* because they contain an *inflation premium*. The approximate relation here is:

$$\text{nominal risk-free rate} = \text{real risk-free rate} + \text{expected inflation rate}$$

Securities may have one or more **types of risk**, and each added risk increases the required rate of return on the security. These types of risk are:

- **Default risk.** The risk that a borrower will not make the promised payments in a timely manner.
- **Liquidity risk.** The risk of receiving less than fair value for an investment if it must be sold for cash quickly.
- **Maturity risk.** As we will cover in detail in the section on debt securities, the prices of longer-term bonds are more volatile than those of shorter-term bonds. Longer maturity bonds have more maturity risk than shorter-term bonds and require a maturity risk premium.

Each of these risk factors is associated with a risk premium that we add to the nominal risk-free rate to adjust for greater default risk, less liquidity, and longer maturity relative to a very liquid, short-term, default risk-free rate such as that on T-bills. We can write:

$$\begin{aligned}\text{nominal rate of interest} &= \text{nominal risk-free rate} \\ &\quad + \text{default risk premium} \\ &\quad + \text{liquidity premium} \\ &\quad + \text{maturity risk premium}\end{aligned}$$



MODULE QUIZ 1.1

1. An interest rate is *best* interpreted as:
 - A. a discount rate or a measure of risk.
 - B. a measure of risk or a required rate of return.
 - C. a required rate of return or the opportunity cost of consumption.
2. An interest rate from which the inflation premium has been subtracted is known as:
 - A. a real interest rate.
 - B. a risk-free interest rate.
 - C. a real risk-free interest rate.

MODULE 1.2: CALCULATING PV AND FV



LOS 1.c: Calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows.

Video covering this content is available online.

Future Value of a Single Sum

Future value is the amount to which a current deposit will grow over time when it is placed in an account paying compound interest. The FV, also called the compound value, is simply an example of compound interest at work.

The formula for the FV of a *single* cash flow is:

$$FV = PV(1 + I/Y)^N$$

where:

PV = amount of money invested today (the present value)

I/Y = rate of return per compounding period

N = total number of compounding periods

In this expression, the investment involves a single cash outflow, PV, which occurs today, at $t = 0$ on the time line. The single sum FV formula will determine the value of an investment at the end of N compounding periods, given that it can earn a fully compounded rate of return, I/Y, over all of the periods.

The factor $(1 + I/Y)^N$ represents the compounding rate on an investment and is frequently referred to as the **future value factor**, or the **future value interest factor**, for a single cash flow at I/Y over N compounding periods. These are the values that appear in interest factor tables, which we will not be using.

EXAMPLE: FV of a single sum

Calculate the FV of a \$200 investment at the end of two years if it earns an annually compounded rate of return of 10%.

Answer:

To solve this problem with your calculator, input the relevant data and compute FV.

$$N = 2; I/Y = 10; PV = -200; CPT \rightarrow FV = \$242$$



PROFESSOR'S NOTE

Note the negative sign on PV. This is not necessary, but it makes the FV come out as a positive number. If you enter PV as a positive number, ignore the negative sign that appears on the FV.

This relatively simple problem could also be solved using the following equation:

$$FV = 200(1 + 0.10)^2 = \$242$$

On the TI calculator, enter 1.10 [x²] 200 [=].

Present Value of a Single Sum

The PV of a single sum is today's value of a cash flow that is to be received at some point in the future. In other words, it is the amount of money that must be invested today, at a given rate of return over a given period of time, in order to end up with a specified FV. As previously mentioned, the process for finding the PV of a cash flow is known as *discounting* (i.e., future cash flows are "discounted" back to the present). The interest rate used in the discounting process is commonly referred to as the **discount rate** but may also be referred to as the **opportunity cost, required rate of return, and the cost of capital**. Whatever you want to call it, it represents the annual compound rate of return that can be earned on an investment.

The relationship between PV and FV can be seen by examining the FV expression stated earlier. Rewriting the FV equation in terms of PV, we get:

$$PV = FV \times \left[\frac{1}{(1 + I/Y)^N} \right] = \frac{FV}{(1 + I/Y)^N}$$

Note that for a single future cash flow, PV is always less than the FV whenever the discount rate is positive.

The quantity $1/(1 + I/Y)^N$ in the PV equation is frequently referred to as the **present value factor**, **present value interest factor**, or **discount factor** for a single cash flow at I/Y over N compounding periods.

EXAMPLE: PV of a single sum

Given a discount rate of 10%, calculate the PV of a \$200 cash flow that will be received in two years.

Answer:

To solve this problem, input the relevant data and compute PV.

$$N = 2; I/Y = 10; FV = 200; CPT \rightarrow PV = -\$165.29 \text{ (ignore the sign)}$$



PROFESSOR'S NOTE

With single sum PV problems, you can either enter FV as a positive number and ignore the negative sign on PV or enter FV as a negative number.

This relatively simple problem could also be solved using the following PV equation:

$$PV = \frac{200}{(1 + 0.10)^2} = \$165.29$$

On the TI, enter $1.10 [y^x] 2 [=] [1/x] [\times] 200 [=]$.

The PV computed here implies that at a rate of 10%, an investor will be indifferent between \$200 in two years and \$165.29 today. Put another way, \$165.29 is the amount that must be invested today at a 10% rate of return in order to generate a cash flow of \$200 at the end of two years.

Annuites

An **annuity** is a stream of *equal cash flows* that occurs at *equal intervals* over a given period. Receiving \$1,000 per year at the end of each of the next eight years is an example of an annuity. There are two types of annuities: **ordinary annuities** and **annuities due**. The *ordinary annuity* is the most common type of annuity. It is characterized by cash flows that occur at the *end* of each compounding period. This is a typical cash flow pattern for many investment and business finance applications. The other type of annuity is called an *annuity due*, where payments or receipts occur at the beginning of each period (i.e., the first payment is today at $t = 0$).

Computing the FV or PV of an annuity with your calculator is no more difficult than it is for a single cash flow. You will know four of the five relevant variables and solve for the fifth (either PV or FV). The difference between single sum and annuity TVM problems is that instead of solving for the PV or FV of a single cash flow, we solve for the PV or FV of a stream of equal

periodic cash flows, where the size of the periodic cash flow is defined by the payment (PMT) variable on your calculator.

EXAMPLE: FV of an ordinary annuity

What is the future value of an ordinary annuity that pays \$200 per year at the end of each of the next three years, given the investment is expected to earn a 10% rate of return?

Answer:

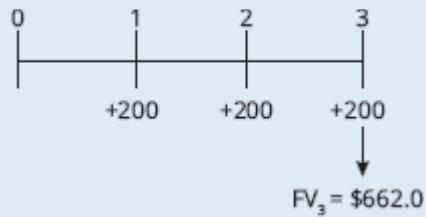
This problem can be solved by entering the relevant data and computing FV.

$$N = 3; I/Y = 10; PMT = -200; CPT \rightarrow FV = \$662$$

Implicit here is that PV = 0; clearing the TVM functions sets both PV and FV to zero.

The time line for the cash flows in this problem is depicted in the following figure:

FV of an Ordinary Annuity



As indicated here, the sum of the compounded values of the individual cash flows in this 3-year ordinary annuity is \$662. Note that the annuity payments themselves amounted to \$600, and the balance is the interest earned at the rate of 10% per year.

To find the PV of an ordinary annuity, we use the future cash flow stream, PMT, that we used with FV annuity problems, but we discount the cash flows back to the present (time $t = 0$) rather than compounding them forward to the terminal date of the annuity.

Here again, the PMT variable is a *single* periodic payment, *not* the total of all the payments (or deposits) in the annuity. The PVA_0 measures the collective PV of a stream of equal cash flows received at the end of each compounding period over a stated number of periods, N, given a specified rate of return, I/Y. The following examples illustrate how to determine the PV of an ordinary annuity using a financial calculator:

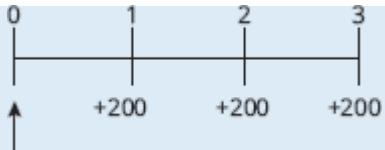
EXAMPLE: PV of an ordinary annuity

What is the PV of an annuity that pays \$200 per year at the end of each of the next three years, given a 10% discount rate?

Answer:

The payments occur at the end of the year, so this annuity is an ordinary annuity. To solve this problem, enter the relevant information and compute PV.

$$N = 3; I/Y = 10; PMT = -200; FV = 0; CPT \rightarrow PV = \$497.37$$



$$PV_0 = \$497.37$$

The \$497.37 computed here represents the amount of money that an investor would need to invest *today* at a 10% rate of return to generate three end-of-year cash flows of \$200 each.

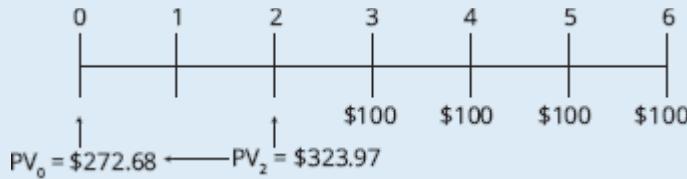
EXAMPLE: PV of an ordinary annuity beginning later than $t = 1$

What is the present value of four \$100 end-of-year payments if the first payment is to be received three years from today and the appropriate rate of return is 9%?

Answer:

The time line for this cash flow stream is shown in the following figure:

PV of an Annuity Beginning at $t = 3$



Step 1: Find the present value of the annuity as of the end of year 2 (PV_2).

Input the relevant data and solve for PV_2 .

$$N = 4; I/Y = 9; PMT = -100; FV = 0; CPT \rightarrow PV = PV_2 = \$323.97$$

Step 2: Find the present value of PV_2 .

Input the relevant data and solve for PV_0 .

$$N = 2; I/Y = 9; PMT = 0; FV = -323.97; CPT \rightarrow PV = PV_0 = \$272.68$$

In this solution, the annuity was treated as an ordinary annuity. The PV was computed one period before the first payment, and we discounted $PV_2 = \$323.97$ over two years. We need to stress this important point. The PV annuity function on your calculator set in "END" mode gives you the value *one period before the annuity begins*. Although the annuity begins at $t = 3$, we discounted the result for only two periods to get the present ($t = 0$) value.

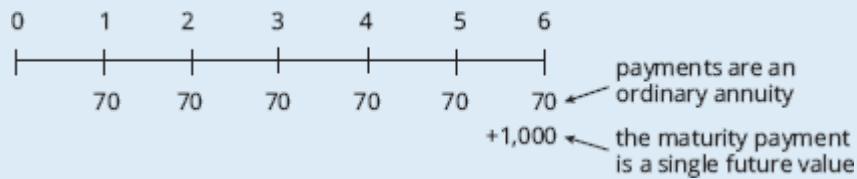
EXAMPLE: PV of a bond's cash flows

A bond will make coupon interest payments of 70 euros (7% of its face value) at the end of each year and will also pay its face value of 1,000 euros at maturity in six years. If the appropriate discount rate is 8%, what is the present value of the bond's promised cash flows?

Answer:

The six annual coupon payments of 70 euros each can be viewed as an ordinary annuity. The maturity value of 1,000 euros is the future value of the bond at the time the last coupon payment is made. On a time line, the promised payment stream is as shown below.

Cash Flows for a 6-Year, 7%, 1,000 Euro Bond



The PV of the bond's cash flows can be broken down into the PV of a 6-payment ordinary annuity, plus the PV of a 1,000 euro lump sum to be received six years from now.

The calculator solution is:

$$N = 6; PMT = 70; I/Y = 8; FV = 1,000; CPT \rightarrow PV = -953.77$$

With a yield to maturity of 8%, the value of the bond is 953.77 euros.

Note that the PMT and FV must have the same sign, since both are cash flows paid to the investor (paid by the bond issuer). The calculated PV will have the opposite sign from PMT and FV.

Future Value of an Annuity Due

Sometimes it is necessary to find the *FV of an annuity due* (FVA_D), an annuity where the annuity payments (or deposits) occur at the beginning of each compounding period. Fortunately, our financial calculators can be used to do this, but with one slight modification—the calculator must be set to the beginning-of-period (BGN) mode. To switch between the BGN and END modes on the TI, press [2nd] [BGN] [2nd] [SET]. When this is done, "BGN" will appear in the upper right corner of the display window. If the display indicates the desired mode, press [2nd] [QUIT]. You will normally want your calculator to be in the ordinary annuity (END) mode, so remember to switch out of BGN mode after working annuity due problems. Note that nothing appears in the upper right corner of the display window when the TI is set to the END mode. It should be mentioned that while annuity due payments are made or received at the beginning of each period, the FV of an annuity due is calculated as of the end of the last period.

Another way to compute the FV of an annuity due is to calculate the FV of an ordinary annuity, and simply multiply the resulting FV by $[1 + \text{periodic compounding rate} (I/Y)]$. Symbolically, this can be expressed as:

$$FVA_D = FVA_0 \times (1 + I/Y)$$

The following examples illustrate how to compute the FV of an annuity due:

EXAMPLE: FV of an annuity due

What is the future value of an annuity that pays \$200 per year at the beginning of each of the next three years, commencing today, if the cash flows can be invested at an annual rate of 10%?

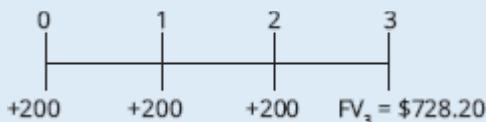
Answer:

Note in the time line in the following figure that the FV is computed as of the end of the last year in the life of the annuity, Year 3, even though the final payment occurs at the beginning of Year 3 (end of Year 2).

To solve this problem, put your calculator in the BGN mode ([2nd] [BGN] [2nd] [SET] [2nd] [QUIT] on the TI or [g] [BEG] on the HP), then input the relevant data and compute FV.

$$N = 3; I/Y = 10; PMT = -200; CPT \rightarrow FV = \$728.20$$

FV of an Annuity Due



Alternatively, we could calculate the FV for an ordinary annuity and multiply it by $(1 + I/Y)$. Leaving your calculator in the END mode, enter the following inputs:

$$N = 3; I/Y = 10; PMT = -200; CPT \rightarrow FVA_0 = \$662.00$$

$$FVA_D = FVA_0 \times (1 + I/Y) = 662 \times 1.10 = \$728.20$$

Present Value of an Annuity Due

While less common than those for ordinary annuities, some problems may require you to find the *PV of an annuity due* (PVA_D). Using a financial calculator, this really shouldn't be much of a problem. With an annuity due, *there is one less discounting period* since the first cash flow occurs at $t = 0$ and thus is already its PV. This implies that, all else equal, the PV of an annuity due will be greater than the PV of an ordinary annuity.

As you will see in the next example, there are two ways to compute the PV of an annuity due. The first is to put the calculator in the BGN mode and then input all the relevant variables (PMT, I/Y, and N) as you normally would. The second, and far easier way, is to treat the cash flow stream as an ordinary annuity over N compounding periods, and simply multiply the resulting PV by $[1 + \text{periodic compounding rate} (I/Y)]$.

Symbolically, this can be stated as:

$$PVA_D = PVA_0 \times (1 + I/Y)$$

The advantage of this second method is that you leave your calculator in the END mode and won't run the risk of forgetting to reset it. Regardless of the procedure used, the computed PV is given as of the beginning of the first period, $t = 0$.

EXAMPLE: PV of an annuity due

Given a discount rate of 10%, what is the present value of an annuity that makes \$200 payments at the beginning of each of the next three years, starting today?

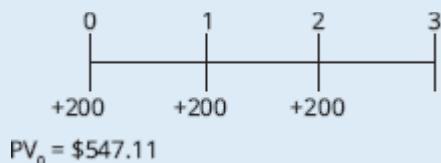
Answer:

First, let's solve this problem using the calculator's BGN mode. Set your calculator to the BGN mode ([2nd] [BGN] [2nd] [SET] [2nd] [QUIT] on the TI or [g] [BEG] on the HP), enter the relevant data, and compute PV.

$$N = 3; I/Y = 10; PMT = -200; CPT \rightarrow PVA_D = \$547.11$$

The time line for this problem is shown in the following figure:

PV of an Annuity Due



Alternatively, this problem can be solved by leaving your calculator in the END mode. First, compute the PV of an ordinary 3-year annuity. Then multiply this PV by (1 + I/Y). To use this approach, enter the relevant inputs and compute PV.

$$N = 3; I/Y = 10; PMT = -200; CPT \rightarrow PVA_0 = \$497.37$$

$$PVA_D = PVA_0 \times (1 + I/Y) = \$497.37 \times 1.10 = \$547.11$$

Present Value of a Perpetuity

A **perpetuity** is a financial instrument that pays a fixed amount of money at set intervals over an *infinite* period of time. In essence, a perpetuity is a perpetual annuity. Most preferred stocks are examples of perpetuities since they promise fixed interest or dividend payments forever. Without going into all the excruciating mathematical details, the discount factor for a perpetuity is just one divided by the appropriate rate of return (i.e., 1/r). Given this, we can compute the PV of a perpetuity.

$$PV_{\text{perpetuity}} = \frac{PMT}{I/Y}$$

The PV of a perpetuity is the fixed periodic cash flow divided by the appropriate periodic rate of return.

As with other TVM applications, it is possible to solve for unknown variables in the PV_{perpetuity} equation. In fact, you can solve for any one of the three relevant variables, given the values for the other two.

EXAMPLE: PV of a perpetuity

Kodon Corporation issues preferred stock that will pay \$4.50 per year in annual dividends beginning next year and plans to follow this dividend policy forever. Given an 8% rate of return, what is the value of Kodon's preferred stock today?

Answer:

Given that the value of the stock is the PV of all future dividends, we have:

$$PV_{\text{perpetuity}} = \frac{4.50}{0.08} = \$56.25$$

Thus, if an investor requires an 8% rate of return, the investor should be willing to pay \$56.25 for each share of Kodon's preferred stock. Note that the PV of a perpetuity is its value one period before its next payment.

EXAMPLE: PV of a deferred perpetuity

Assume the Kodon preferred stock in the preceding examples is scheduled to pay its first dividend in four years, and is non-cumulative (i.e., does not pay any dividends for the first three years). Given an 8% required rate of return, what is the value of Kodon's preferred stock today?

Answer:

As in the previous example, $PV_{\text{perpetuity}} = \frac{4.50}{0.08} = \56.25 , but because the first dividend is paid at $t = 4$, this PV is the value at $t = 3$. To get the value of the preferred stock today, we must discount this value for three periods: $\frac{56.25}{(1.08)^3} = \44.65 .



MODULE QUIZ 1.2

1. The amount an investor will have in 15 years if \$1,000 is invested today at an annual interest rate of 9% will be *closest* to:
 - A. \$1,350.
 - B. \$3,518.
 - C. \$3,642.
2. How much must be invested today, at 8% interest, to accumulate enough to retire a \$10,000 debt due seven years from today?
 - A. \$5,835.
 - B. \$6,123.
 - C. \$8,794.
3. An investor has just won the lottery and will receive \$50,000 per year at the end of each of the next 20 years. At a 10% interest rate, the present value of the winnings is *closest* to:
 - A. \$425,678.
 - B. \$637,241.
 - C. \$2,863,750.
4. An investor is to receive a 15-year, \$8,000 annuity, with the first payment to be received today. At an 11% discount rate, this annuity's worth today is *closest* to:
 - A. \$55,855.
 - B. \$57,527.
 - C. \$63,855.
5. If \$1,000 is invested today and \$1,000 is invested at the beginning of each of the next three years at 12% interest (compounded annually), the amount an investor will have at the end of the fourth year will be *closest* to:
 - A. \$4,779.
 - B. \$5,353.
 - C. \$6,792.
6. Terry Corporation preferred stocks are expected to pay a \$9 annual dividend forever. If the required rate of return on equivalent investments is 11%, a share of Terry preferred should be worth:

- A. \$81.82.
- B. \$99.00.
- C. \$122.22.

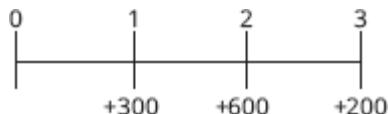
MODULE 1.3: UNEVEN CASH FLOWS



It is not uncommon to have applications in investments and corporate finance where it is necessary to evaluate a cash flow stream that is not equal from period to period. The time line in Figure 1.2 depicts such a cash flow stream.

Video covering this content is available online.

Figure 1.2: Time Line for Uneven Cash Flows



This 3-year cash flow series is not an annuity since the cash flows are different every year. In essence, this series of uneven cash flows is nothing more than a stream of annual single sum cash flows. Thus, to find the PV or FV of this cash flow stream, all we need to do is sum the PVs or FVs of the individual cash flows.

EXAMPLE: Computing the FV of an uneven cash flow series

Using a rate of return of 10%, compute the future value of the 3-year uneven cash flow stream described above at the end of the third year.

Answer:

The FV for the cash flow stream is determined by first computing the FV of each individual cash flow, then summing the FVs of the individual cash flows.

$$FV_1: PV = -300; I/Y = 10; N = 2; CPT \rightarrow FV = FV_1 = 363$$

$$FV_2: PV = -600; I/Y = 10; N = 1; CPT \rightarrow FV = FV_2 = 660$$

$$FV_3: PV = -200; I/Y = 10; N = 0; CPT \rightarrow FV = FV_3 = 200$$

$$\text{FV of cash flow stream} = \sum FV_{\text{individual}} = 1,223$$

EXAMPLE: Computing PV of an uneven cash flow series

Compute the present value of this 3-year uneven cash flow stream described previously using a 10% rate of return.

Answer:

This problem is solved by first computing the PV of each individual cash flow, then summing the PVs of the individual cash flows, which yields the PV of the cash flow stream. Again the signs of the cash flows are preserved.

$$PV_1: FV = 300; I/Y = 10; N = 1; CPT \rightarrow PV = PV_1 = -272.73$$

PV_2 : $FV = 600$; $I/Y = 10$; $N = 2$; $CPT \rightarrow PV = PV_2 = -495.87$

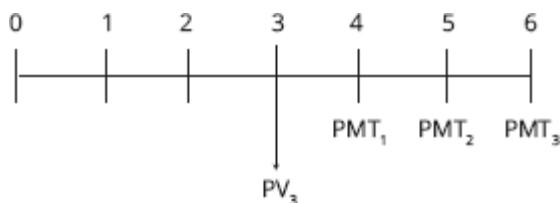
PV_3 : $FV = 200$; $I/Y = 10$; $N = 3$; $CPT \rightarrow PV = PV_3 = -150.26$

PV of cash flow stream = $\sum PV_{\text{individual}} = \918.86

LOS 1.d: Demonstrate the use of a time line in modeling and solving time value of money problems.

In most of the PV problems we have discussed, cash flows were discounted back to the current period. In this case, the PV is said to be indexed to $t = 0$, or the time index is $t = 0$. For example, the PV of a 3-year ordinary annuity that is indexed to $t = 0$ is computed at the beginning of Year 1 ($t = 0$). Contrast this situation with another 3-year ordinary annuity that doesn't start until Year 4 and extends to Year 6. It would not be uncommon to want to know the PV of this annuity at the beginning of Year 4, in which case the time index is $t = 3$. The time line for this annuity is presented in Figure 1.3.

Figure 1.3: Indexing Time Line to Other Than $t = 0$



The following examples will illustrate how to compute I/Y, N, or PMT in annuity problems:

EXAMPLE: Computing an annuity payment needed to achieve a given FV

At an expected rate of return of 7%, how much must be deposited at the end of each year for the next 15 years to accumulate \$3,000?

Answer:

To solve this problem, enter the three relevant known values and compute PMT.

$N = 15$; $I/Y = 7$; $FV = +\$3,000$; $CPT \rightarrow PMT = -\$119.38$ (ignore sign)

EXAMPLE: Computing a loan payment

Suppose you are considering applying for a \$2,000 loan that will be repaid with equal end-of-year payments over the next 13 years. If the annual interest rate for the loan is 6%, how much will your payments be?

Answer:

The size of the end-of-year loan payment can be determined by inputting values for the three known variables and computing PMT.

$$N = 13; I/Y = 6; PV = -2,000; CPT \rightarrow PMT = \$225.92$$

EXAMPLE: Computing the number of periods in an annuity

How many \$100 end-of-year payments are required to accumulate \$920 if the discount rate is 9%?

Answer:

The number of payments necessary can be determined by inputting the relevant data and computing N.

$$I/Y = 9\%; FV = \$920; PMT = -\$100; CPT \rightarrow N = 7 \text{ years}$$

It will take seven annual \$100 payments, compounded at 9% annually, to accrue an investment value of \$920.



PROFESSOR'S NOTE

Remember the sign convention. PMT and FV must have opposite signs or your calculator will issue an error message.

EXAMPLE: Computing the number of years in an ordinary annuity

Suppose you have a \$1,000 ordinary annuity earning an 8% return. How many annual end-of-year \$150 withdrawals can be made?

Answer:

The number of years in the annuity can be determined by entering the three relevant variables and computing N.

$$I/Y = 8\%; PMT = 150; PV = -1,000; CPT \rightarrow N = 9.9 \text{ years}$$

EXAMPLE: Computing the rate of return for an annuity

Suppose you have the opportunity to invest \$100 at the end of each of the next five years in exchange for \$600 at the end of the fifth year. What is the annual rate of return on this investment?

Answer:

The rate of return on this investment can be determined by entering the relevant data and solving for I/Y.

$$N = 5; FV = \$600; PMT = -100; CPT \rightarrow I/Y = 9.13\%$$

EXAMPLE: Computing the discount rate for an annuity

What rate of return will you earn on an ordinary annuity that requires a \$700 deposit today and promises to pay \$100 per year at the end of each of the next 10 years?

Answer:

The discount rate on this annuity is determined by entering the three known values and computing I/Y.

$$N = 10; PV = -700; PMT = 100; CPT \rightarrow I/Y = 7.07\%$$

Funding a Future Obligation

There are many TVM applications where it is necessary to determine the size of the deposit(s) that must be made over a specified period in order to meet a future liability, such as setting up a funding program for future college tuition or a retirement program. In most of these applications, the objective is to determine the size of the payment(s) or deposit(s) necessary to meet a particular monetary goal.

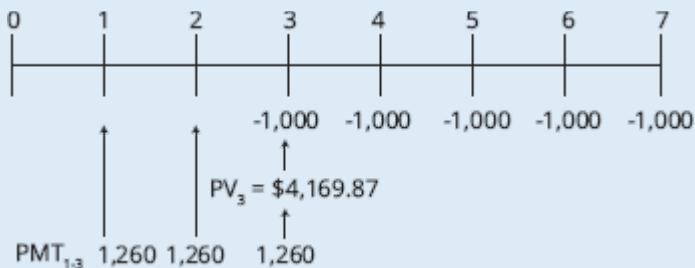
EXAMPLE: Computing the required payment to fund an annuity due

Suppose you must make five annual \$1,000 payments, the first one starting at the beginning of Year 4 (end of Year 3). To accumulate the money to make these payments, you want to make three equal payments into an investment account, the first to be made one year from today. Assuming a 10% rate of return, what is the amount of these three payments?

Answer:

The time line for this annuity problem is shown in the following figure:

Funding an Annuity Due



The first step in this type of problem is to determine the amount of money that must be available at the end of the third year in order to satisfy the payment requirements. This amount is the PV of a 5-year annuity due at the beginning of Year 4 (end of Year 3). To determine this amount, set your calculator to the BGN mode, enter the relevant data, and compute PV.

$$N = 5; I/Y = 10; PMT = -1,000; CPT \rightarrow PV = PV_3 = \$4,169.87$$

PV_3 becomes the FV that you need three years from today from your three equal end-of-year deposits. To determine the amount of the three payments necessary to meet this funding requirement, be sure that your calculator is in the END mode, input the relevant data, and compute PMT.

$$N = 3; I/Y = 10; FV = -4,169.87; CPT \rightarrow PMT = \$1,259.78$$

The second part of this problem is an ordinary annuity. If you changed your calculator to BGN mode and failed to put it back in the END mode, you will get a PMT of \$1,145, which is incorrect.

The Connection Between Present Values, Future Values, and Series of Cash Flows

As we have explained in the discussion of annuities and series of uneven cash flows, the sum of the present values of the cash flows is the present value of the series. The sum of the future values (at some future time = n) of a series of cash flows is the future value of that series of cash flows.

One interpretation of the present value of a series of cash flows is how much would have to be put in the bank today in order to make these future withdrawals and exhaust the account with the final withdrawal. Let's illustrate this with cash flows of \$100 in Year 1, \$200 in Year 2, \$300 in Year 3, and an assumed interest rate of 10%.

Calculate the present value of these three cash flows as:

$$\frac{100}{1.1} + \frac{200}{1.1^2} + \frac{300}{1.1^3} = \$481.59$$

If we put \$481.59 in an account yielding 10%, at the end of the year we would have $481.59 \times 1.1 = \$529.75$. Withdrawing \$100 would leave \$429.75.

Over the second year, the \$429.75 would grow to $429.75 \times 1.1 = \$472.73$. Withdrawing \$200 would leave \$272.73.

Over the third year, \$272.73 would grow to $272.73 \times 1.1 = \$300$, so that the last withdrawal of \$300 would empty the account.

The interpretation of the future value of a series of cash flows is straightforward as well. The FV answers the question, "How much would be in an account when the last of a series of deposits is made?" Using the same three cash flows—\$100, \$200, and \$300—and the same interest rate of 10%, we can calculate the future value of the series as:

$$100(1.1)^2 + 200(1.1) + 300 = \$641$$

This is simply the sum of the $t = 3$ value of each of the cash flows. Note that the $t = 3$ value and the $t = 0$ (present) value of the series are related by the interest rate, $481.59(1.1)^3 = 641$.

The \$100 cash flow (deposit) comes at $t = 1$, so it will earn interest of 10% compounded for two periods (until $t = 3$). The \$200 cash flow (deposit) will earn 10% between $t = 2$ and $t = 3$, and the final cash flow (deposit) of \$300 is made at $t = 3$, so \$300 is the future ($t = 3$) value of that cash flow.

We can also look at the future value in terms of how the account grows over time. At $t = 1$ we deposit \$100, so at $t = 2$ it has grown to \$110 and the \$200 deposit at $t = 2$ makes the account balance \$310. Over the next period, the \$310 grows to $310 \times 1.1 = \$341$ at $t = 3$, and the addition of the final \$300 deposit puts the account balance at \$641. This is, of course, the future value we calculated initially.



PROFESSOR'S NOTE

This last view of the future value of a series of cash flows suggests a quick way to calculate the future value of an uneven cash flow series. The process described previously for the future value of a series of end-of-period payments can be written mathematically as $[(100 \times 1.1) + 200] \times 1.1 + 300 = 641$, and this might be a quick way to do some future value problems. On your TI calculator, you would enter $100 \times 1.1 + 200 = \times 1.1 + 300 =$ to get 641.

Note that questions on the future value of an *annuity due* refer to the amount in the account one period after the last deposit is made. If the three deposits considered here were made at the beginning of each period (at $t = 0, 1, 2$) the amount in the account at the end of three years ($t = 3$) would be 10% higher (i.e., $641 \times 1.1 = \$705.10$).

The **cash flow additivity principle** refers to the fact that present value of any stream of cash flows equals the sum of the present values of the cash flows. There are different applications of this principle in time value of money problems. If we have two series of cash flows, the sum of the present values of the two series is the same as the present values of the two series taken together, adding cash flows that will be paid at the same point in time. We can also divide up a series of cash flows any way we like, and the present value of the “pieces” will equal the present value of the original series.

EXAMPLE: Additivity principle

A security will make the following payments at the end of the next four years: \$100, \$100, \$400, and \$100. Calculate the present value of these cash flows using the concept of the present value of an annuity when the appropriate discount rate is 10%.

Answer:

We can divide the cash flows so that we have:

$t = 1$	$t = 2$	$t = 3$	$t = 4$	
100	100	100	100	cash flow series #1
0	0	300	0	cash flow series #2
\$100	\$100	\$400	\$100	

The additivity principle tells us that to get the present value of the original series, we can just add the present values of series #1 (a 4-period annuity) and series #2 (a single payment three periods from now).

For the annuity: $N = 4$; $PMT = 100$; $FV = 0$; $I/Y = 10$; $CPT \rightarrow PV = -\$316.99$

For the single payment: $N = 3$; $PMT = 0$; $FV = 300$; $I/Y = 10$; $CPT \rightarrow PV = -\$225.39$

The sum of these two values is $316.99 + 225.39 = \$542.38$.

The sum of these two (present) values is identical (except for rounding) to the sum of the present values of the payments of the original series:

$$\frac{100}{1.1} + \frac{100}{1.1^2} + \frac{400}{1.1^3} + \frac{100}{1.1^4} = \$542.38$$



MODULE QUIZ 1.3

1. An analyst estimates that XYZ's earnings will grow from \$3.00 a share to \$4.50 per share over the next eight years. The rate of growth in XYZ's earnings is *closest* to:
 - A. 4.9%.
 - B. 5.2%.
 - C. 6.7%.
2. If \$5,000 is invested in a fund offering a rate of return of 12% per year, approximately how many years will it take for the investment to reach \$10,000?
 - A. 4 years.
 - B. 5 years.
 - C. 6 years.
3. An investment is expected to produce the cash flows of \$500, \$200, and \$800 at the end of the next three years. If the required rate of return is 12%, the present value of this investment is *closest* to:
 - A. \$835.
 - B. \$1,175.
 - C. \$1,235.
4. If \$10,000 is invested today in an account that earns interest at a rate of 9.5%, what is the value of the equal withdrawals that can be taken out of the account at the end of each of the next five years if the investor plans to deplete the account at the end of the time period?
 - A. \$2,453.
 - B. \$2,604.
 - C. \$2,750.
5. Given an 11% rate of return, the amount that must be put into an investment account at the end of each of the next 10 years in order to accumulate \$60,000 to pay for a child's education is *closest* to:
 - A. \$2,500.
 - B. \$3,588.
 - C. \$4,432.
6. An investor will receive an annuity of \$4,000 a year for 10 years. The first payment is to be received five years from today. At a 9% discount rate, this annuity's worth today is *closest* to:
 - A. \$16,684.
 - B. \$18,186.
 - C. \$25,671.

MODULE 1.4: COMPOUNDING FREQUENCIES

LOS 1.e: Calculate the solution for time value of money problems with different frequencies of compounding.

While the conceptual foundations of TVM calculations are not affected by the compounding period, more frequent compounding does have an impact on FV and PV computations.

Specifically, since an increase in the frequency of compounding increases the effective rate of interest, it also *increases* the FV of a given cash flow and *decreases* the PV of a given cash flow.

EXAMPLE: The effect of compounding frequency on FV and PV

Compute the FV one year from now of \$1,000 today and the PV of \$1,000 to be received one year from now using a stated annual interest rate of 6% with a range of compounding periods.

Answer:

Compounding Frequency Effect

Compounding Frequency	Interest Rate per Period	Effective Annual Rate	Future Value	Present Value
Annual ($m = 1$)	6.000%	6.00%	\$1,060.00	\$943.396
Semiannual ($m = 2$)	3.000	6.090	1,060.90	942.596
Quarterly ($m = 4$)	1.500	6.136	1,061.36	942.184
Monthly ($m = 12$)	0.500	6.168	1,061.68	941.905
Daily ($m = 365$)	0.016438	6.183	1,061.83	941.769

There are two ways to use your financial calculator to compute PVs and FVs under different compounding frequencies:

1. Adjust the number of periods per year (P/Y) mode on your calculator to correspond to the compounding frequency (e.g., for quarterly, P/Y = 4). WE DO NOT RECOMMEND THIS APPROACH!
2. Keep the calculator in the annual compounding mode (P/Y = 1) and enter I/Y as the interest rate per compounding period, and N as the number of compounding periods in the investment horizon. Letting m equal the number of compounding periods per year, the basic formulas for the calculator input data are determined as follows:

$$I/Y = \text{the annual interest rate}/m$$

$$N = \text{the number of years} \times m$$

The computations for the FV and PV amounts in the previous example are:

PV_A : $FV = -1,000; I/Y = 6/1 = 6; N = 1 \times 1 = 1;$
 CPT → PV = $PV_A = 943.396$
 PV_S : $FV = -1,000; I/Y = 6/2 = 3; N = 1 \times 2 = 2;$
 CPT → PV = $PV_S = 942.596$
 PV_Q : $FV = -1,000; I/Y = 6/4 = 1.5; N = 1 \times 4 = 4;$
 CPT → PV = $PV_Q = 942.184$
 PV_M : $FV = -1,000; I/Y = 6/12 = 0.5; N = 1 \times 12 = 12;$
 CPT → PV = $PV_M = 941.905$
 PV_D : $FV = -1,000; I/Y = 6/365 = 0.016438; N = 1 \times 365 = 365;$
 CPT → PV = $PV_D = 941.769$
 FV_A : $PV = -1,000; I/Y = 6/1 = 6; N = 1 \times 1 = 1;$
 CPT → FV = $FV_A = 1,060.00$
 FV_S : $PV = -1,000; I/Y = 6/2 = 3; N = 1 \times 2 = 2;$
 CPT → FV = $FV_S = 1,060.90$
 FV_Q : $PV = -1,000; I/Y = 6/4 = 1.5; N = 1 \times 4 = 4;$
 CPT → FV = $FV_Q = 1,061.36$
 FV_M : $PV = -1,000; I/Y = 6/12 = 0.5; N = 1 \times 12 = 12;$
 CPT → FV = $FV_M = 1,061.68$
 FV_D : $PV = -1,000; I/Y = 6/365 = 0.016438; N = 1 \times 365 = 365;$
 CPT → FV = $FV_D = 1,061.83$

EXAMPLE: FV of a single sum using quarterly compounding

Compute the FV of \$2,000 today, five years from today using an interest rate of 12%, compounded quarterly.

Answer:

To solve this problem, enter the relevant data and compute FV:

$$N = 5 \times 4 = 20; I/Y = 12/4 = 3; PV = -\$2,000; \text{CPT} \rightarrow FV = \$3,612.22$$

EXAMPLE: Growth with quarterly compounding

John plans to invest \$2,500 in an account that will earn 8% per year with quarterly compounding. How much will be in the account at the end of two years?

Answer:

There are eight quarterly compounding periods in two years, and the effective quarterly rate is $8/4 = 2\%$. The account will grow to $2,500(1.02)^8 = \$2,929.15$. Alternatively, since the EAR is $1.02^4 - 1 = 0.082432$, we can grow the \$2,500 at 8.2432% for two years to get $2,500(1.082432)^2 = \$2,929.15$, which is the same result.

EXAMPLE: Present value with monthly compounding

Alice would like to have \$5,000 saved in an account at the end of three years. If the return on the account is 9% per year with monthly compounding, how much must Alice deposit today in order to reach her savings goal in three years?

Answer:

The effective monthly rate is $9/12 = 0.75\%$, and we can calculate the present value of \$5,000 three years (36 months) from now as $5,000/(1.0075)^{36} = \$3,820.74$. Alternatively, since the EAR is $1.0075^{12} - 1 = 0.093807$, we can calculate the present value by discounting 5,000 at the EAR for three years. $5,000/1.093807^3 = \$3,820.74$, which is the same result.

LOS 1.f: Calculate and interpret the effective annual rate, given the stated annual interest rate and the frequency of compounding.

Financial institutions usually quote rates as stated annual interest rates, along with a compounding frequency, as opposed to quoting rates as periodic rates—the rate of interest earned over a single compounding period. For example, a bank will quote a savings rate as 8%, compounded quarterly, rather than 2% per quarter. The rate of interest that investors actually realize as a result of compounding is known as the **effective annual rate (EAR)** or **effective annual yield (EAY)**. EAR represents the annual rate of return actually being earned *after adjustments have been made for different compounding periods*.

EAR may be determined as follows:

$$\text{EAR} = (1 + \text{periodic rate})^m - 1$$

where:

periodic rate = stated annual rate/m

m = the number of compounding periods per year

Obviously, the EAR for a stated rate of 8% *compounded annually* is not the same as the EAR for 8% *compounded semiannually*, or *quarterly*. Indeed, whenever compound interest is being used, the stated rate and the actual (effective) rate of interest are equal only when interest is compounded annually. Otherwise, the greater the compounding frequency, the greater the EAR will be in comparison to the stated rate.

The computation of EAR is necessary when comparing investments that have different compounding periods. It allows for an apples-to-apples rate comparison.

EXAMPLE: Computing EAR

Compute EAR if the stated annual rate is 12%, compounded quarterly.

Answer:

Here $m = 4$, so the periodic rate is $\frac{12}{4} = 3\%$.

$$\text{Thus, } \text{EAR} = (1 + 0.03)^4 - 1 = 1.1255 - 1 = 0.1255 = 12.55\%.$$

This solution uses the $[y^x]$ key on your financial calculator. The exact keystrokes on the TI for the above computation are 1.03 $[y^x]$ 4 [=]. On the HP, the strokes are 1.03 [ENTER] 4 $[y^x]$.

EXAMPLE: Computing EARs for a range of compounding frequencies

Using a stated rate of 6%, compute EARs for semiannual, quarterly, monthly, and daily compounding.

Answer:

EAR with:

$$\text{semiannual compounding} = (1 + 0.03)^2 - 1 = 1.06090 - 1 = 0.06090 = 6.090\%$$

$$\text{quarterly compounding} = (1 + 0.015)^4 - 1 = 1.06136 - 1 = 0.06136 = 6.136\%$$

$$\text{monthly compounding} = (1 + 0.005)^{12} - 1 = 1.06168 - 1 = 0.06168 = 6.168\%$$

$$\text{daily compounding} = (1 + 0.00016438)^{365} - 1 = 1.06183 - 1 = 0.06183 = 6.183\%$$

Notice here that the EAR increases as the compounding frequency increases.



PROFESSOR'S NOTE

The limit of shorter and shorter compounding periods is called continuous compounding, which we will address in a later reading.



MODULE QUIZ 1.4

1. What is the effective annual rate for a credit card that charges 18% compounded monthly?
 - A. 15.38%.
 - B. 18.81%.
 - C. 19.56%.
2. Given daily compounding, the growth of \$5,000 invested for one year at 12% interest will be *closest* to:
 - A. \$5,600.
 - B. \$5,628.
 - C. \$5,637.
3. An investor is looking at a \$150,000 home. If 20% must be put down and the balance is financed at a stated annual rate of 9% over the next 30 years, what is the monthly mortgage payment?
 - A. \$799.33.
 - B. \$895.21.
 - C. \$965.55.

KEY CONCEPTS

LOS 1.a

An interest rate can be interpreted as the rate of return required in equilibrium for a particular investment, the discount rate for calculating the present value of future cash flows, or as the opportunity cost of consuming now, rather than saving and investing.

LOS 1.b

The real risk-free rate is a theoretical rate on a single-period loan when there is no expectation of inflation. Nominal risk-free rate = real risk-free rate + expected inflation rate.

Securities may have several risks, and each increases the required rate of return. These include default risk, liquidity risk, and maturity risk.

The required rate of return on a security = real risk-free rate + expected inflation + default risk premium + liquidity premium + maturity risk premium.

LOS 1.c

Future value: $FV = PV(1 + I/Y)^N$

Present value: $PV = FV/(1 + I/Y)^N$

An annuity is a series of equal cash flows that occurs at evenly spaced intervals over time. Ordinary annuity cash flows occur at the end of each time period. Annuity due cash flows occur at the beginning of each time period.

Perpetuities are annuities with infinite lives (perpetual annuities):

$$PV_{\text{perpetuity}} = \frac{PMT}{I/Y}$$

The present (future) value of any series of cash flows is equal to the sum of the present (future) values of the individual cash flows.

LOS 1.d

Constructing a time line showing future cash flows will help in solving many types of TVM problems. Cash flows occur at the end of the period depicted on the time line. The end of one period is the same as the beginning of the next period. For example, a cash flow at the beginning of Year 3 appears at time $t = 2$ on the time line.

LOS 1.e

For non-annual time value of money problems, divide the stated annual interest rate by the number of compounding periods per year, m , and multiply the number of years by the number of compounding periods per year.

LOS 1.f

The effective annual rate when there are m compounding periods =

$$\left(1 + \frac{\text{stated annual rate}}{m}\right)^m - 1. \text{ Each dollar invested will grow to}$$

$$\left(1 + \frac{\text{stated annual rate}}{m}\right)^m \text{ in one year.}$$

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 1.1

1. **C** Interest rates can be interpreted as required rates of return, discount rates, or opportunity costs of current consumption. A risk premium can be, but is not always, a component of an interest rate. (LOS 1.a, 1.b)
2. **A** Real interest rates are those that have been adjusted for inflation. (LOS 1.b)

Module Quiz 1.2

1. **C** $N = 15; I/Y = 9; PV = -1,000; PMT = 0; CPT \rightarrow FV = \$3,642.48$ (LOS 1.c)
2. **A** $N = 7; I/Y = 8; FV = -10,000; PMT = 0; CPT \rightarrow PV = \$5,834.90$ (LOS 1.c)
3. **A** $N = 20; I/Y = 10; PMT = -50,000; FV = 0; CPT \rightarrow PV = \$425,678.19$ (LOS 1.c)
4. **C** This is an annuity due. Switch to BGN mode: $N = 15; PMT = -8,000; I/Y = 11; FV = 0; CPT \rightarrow PV = 63,854.92$. Switch back to END mode. (LOS 1.c)
5. **B** The key to this problem is to recognize that it is a 4-year annuity due, so switch to BGN mode: $N = 4; PMT = -1,000; PV = 0; I/Y = 12; CPT \rightarrow FV = 5,352.84$. Switch back to END mode. (LOS 1.c)
6. **A** $9/0.11 = \$81.82$ (LOS 1.c)

Module Quiz 1.3

1. **B** $N = 8; PV = -3; FV = 4.50; PMT = 0; CPT \rightarrow I/Y = 5.1989$ (LOS 1.d)

2. **C** $PV = -5,000; I/Y = 12; FV = 10,000; PMT = 0; CPT \rightarrow N = 6.12$.

Note to HP 12C users: One known problem with the HP 12C is that it does not have the capability to round. In this particular question, you will come up with 7, although the correct answer is 6.1163. CFA Institute is aware of this problem, and hopefully you will not be faced with a situation on exam day where the incorrect solution from the HP is one of the answer choices. (LOS 1.d)

3. **B** Add up the present values of each single cash flow.

$$PV_1 = N = 1; FV = -500; I/Y = 12; CPT \rightarrow PV = 446.43$$

$$PV_2 = N = 2; FV = -200; I/Y = 12; CPT \rightarrow PV = 159.44$$

$$PV_3 = N = 3; FV = -800; I/Y = 12; CPT \rightarrow PV = 569.42$$

$$\text{Hence, } 446.43 + 159.44 + 569.42 = \$1,175.29. \text{ (LOS 1.d)}$$

4. **B** $PV = -10,000; I/Y = 9.5; N = 5; FV = 0; CPT \rightarrow PMT = \$2,604.36$ (LOS 1.d)

5. **B** $N = 10; I/Y = 11; FV = -60,000; PV = 0; CPT \rightarrow PMT = \$3,588.08$ (LOS 1.d)

6. **B** Two steps: (1) Find the PV of the 10-year annuity: $N = 10; I/Y = 9; PMT = -4,000; FV = 0; CPT \rightarrow PV = 25,670.63$. This is the present value as of the end of Year 4; (2) Discount PV of the annuity back four years: $N = 4; PMT = 0; FV = -25,670.63; I/Y = 9; CPT \rightarrow PV = 18,185.72$. (LOS 1.d)

Module Quiz 1.4

1. **C** $EAR = [(1 + (0.18/12))]^{12} - 1 = 19.56\%$ (LOS 1.f)

2. **C** $N = 1 \times 365 = 365; I/Y = 12/365 = 0.0328767; PMT = 0; PV = -5,000; CPT \rightarrow FV = \$5,637.37$ (LOS 1.e)

3. **C** $N = 30 \times 12 = 360; I/Y = 9/12 = 0.75; PV = -150,000(1 - 0.2) = -120,000; FV = 0; CPT \rightarrow PMT = \965.55 (LOS 1.e)

READING 2

ORGANIZING, VISUALIZING, AND DESCRIBING DATA

EXAM FOCUS

Candidates must learn how to interpret the various types of illustrations used to describe data. The various measures of central tendency, dispersion, and risk are used throughout the CFA curriculum and are essential knowledge for candidates. The concepts of skewness, kurtosis, and correlation are also used extensively.

MODULE 2.1: ORGANIZING DATA



LOS 2.a: Identify and compare data types.

Video covering
this content is
available online.

The term *data* encompasses information in any form. For our use as analysts, we may classify data types from three different perspectives:

- Numerical versus categorical.
- Time series versus cross-sectional.
- Structured versus unstructured.

Numerical and Categorical Data

Numerical data, or **quantitative data**, are values that can be counted or measured. Numerical data may be discrete or continuous. **Discrete data** are countable, such as the months, days, or hours in a year. **Continuous data** can take any fractional value (e.g., the annual percentage return on an investment).



PROFESSOR'S NOTE

In our reading on Common Probability Distributions, we will use this concept to distinguish between discrete and continuous random variables.

Categorical data, or **qualitative data**, consist of labels that can be used to classify a set of data into groups. Categorical data may be nominal or ordinal.

Nominal data are labels that cannot be placed in order logically. For example, fixed-income mutual funds may be classified as corporate bond funds, municipal bond funds, international bond funds, and so on. Even if we assign numbers to the categories (such as the number 1 to a

corporate bond fund, the number 2 to a municipal bond fund, and so on), the numbers are arbitrary.

By contrast, **ordinal data** can be ranked in a logical order. Every item is assigned to one of multiple categories based on a specific characteristic, then these categories are ordered with respect to that characteristic. For example, the ranking of 1,000 small-cap growth stocks by performance may be done by assigning the number 1 to the 100 best-performing stocks, the number 2 to the next 100 best-performing stocks, and so on through the number 10 for the 100 worst-performing stocks. Based on this type of measurement, we can say a stock ranked 3 performed better than a stock ranked 4. However, we cannot conclude that the difference between a 3 and a 4 is the same as the difference between a 4 and a 5.

The key distinction between numerical data and categorical data is that we can perform mathematical operations only on numerical data.

Time Series and Cross-Sectional Data

A **time series** is a set of observations taken periodically, most often at equal intervals over time. Daily closing prices of a stock over the past year and quarterly earnings per share of a company over a five-year period are examples of time series data.

Cross-sectional data refers to a set of comparable observations all taken at one specific point in time. Today's closing prices of the 30 stocks in the Dow Jones Industrial Average and fourth-quarter earnings per share for 10 health care companies are examples of cross-sectional data.

Time series and cross-sectional data may be combined to form **panel data**. Panel data are often presented in tables. Figure 2.1 is an example of panel data for an economic indicator. In this table, each row represents cross-sectional data and each column represents time series data.

Figure 2.1: OECD Composite Leading Indicators, Year-on-Year Growth Rate

	Canada	United States	Japan	France	Germany	Italy	United Kingdom
January 2019	-1.47	-0.90	-0.36	-1.49	-1.34	-1.45	-1.41
February 2019	-1.46	-1.15	-0.45	-1.39	-1.47	-1.51	-1.40
March 2019	-1.43	-1.34	-0.51	-1.27	-1.57	-1.51	-1.36
April 2019	-1.39	-1.48	-0.58	-1.14	-1.67	-1.46	-1.31
May 2019	-1.36	-1.58	-0.67	-1.01	-1.78	-1.40	-1.24
June 2019	-1.32	-1.66	-0.75	-0.85	-1.90	-1.33	-1.12
July 2019	-1.27	-1.71	-0.83	-0.65	-2.02	-1.24	-0.96
August 2019	-1.18	-1.70	-0.91	-0.43	-2.05	-1.15	-0.75
September 2019	-1.03	-1.58	-0.97	-0.23	-1.99	-1.05	-0.49
October 2019	-0.83	-1.35	-1.01	-0.07	-1.82	-0.94	-0.18
November 2019	-0.57	-1.02	-1.00	0.05	-1.57	-0.83	0.16
December 2019	-0.27	-0.64	-0.92	0.11	-1.27	-0.70	0.48

Source: www.oecd.org

Structured and Unstructured Data

Time series, cross-sectional, and panel data are examples of **structured data**—they are organized in a defined way. *Market data*, such as security prices; *fundamental data*, such as

accounting values; and *analytical data*, such as analysts' earnings forecasts, are typically presented as structured data.

Unstructured data refers to information that is presented in a form with no defined structure. Management's commentary in company financial statements is an example of unstructured data. One way of classifying unstructured data is according to how it is generated. Data may be *generated by individuals*, such as posts on social media; *generated by business processes*, such as deposits, withdrawals, and transfers of cash; or *generated by sensors*, such as satellites or traffic cameras. While unstructured data often contains useful information, it usually must be transformed into structured data for analysis.



PROFESSOR'S NOTE

Technologies such as artificial intelligence can be used to analyze unstructured data. We address some of these in our reading on Fintech in Investment Management, in the Portfolio Management topic area.

LOS 2.b: Describe how data are organized for quantitative analysis.

Data are typically organized into arrays for analysis. A time series is an example of a **one-dimensional array** in that it represents a single variable. A key feature of a time series is that new data can be added without affecting the existing data. Sequentially ordered data are used to identify trends, cycles, and other patterns in the data that can be useful for forecasting.

The panel data in Figure 2.1 are an example of a **two-dimensional array**, or a **data table**. While data tables are not limited to this structure, organizing data sequentially with a cross section of observations for each measurement date is often useful for analysis.

LOS 2.c: Interpret frequency and related distributions.

A **frequency distribution** is a tabular presentation of statistical data that aids the analysis of large data sets. Frequency distributions summarize statistical data by assigning them to specified groups, or intervals.



PROFESSOR'S NOTE

Intervals are also known as *classes*.

The following procedure describes how to construct a frequency distribution:

Step 1: Define the intervals. The first step in building a frequency distribution is to define the intervals to which data measurements (observations) will be assigned. An interval is the set of values that an observation may take on. The range of values for each interval must have a lower and upper limit and be all-inclusive and non-overlapping. Intervals must be *mutually exclusive* so that each observation can be placed in only one interval, and the total set of intervals should cover the total range of values for the entire population. The number of intervals used is an important consideration. If too few intervals are used, the

data may be too broadly summarized and important characteristics may be lost. On the other hand, if too many intervals are used, the data may not be summarized enough.

Step 2: Tally the observations. After the intervals have been defined, the observations must be tallied or assigned to their appropriate interval.

Step 3: Count the observations. Having tallied the data set, the number of observations that are assigned to each interval must be counted. The *absolute frequency*, or simply the *frequency*, is the actual number of observations that fall within a given interval.

EXAMPLE: Constructing a frequency distribution

Use the data in Table A to construct a frequency distribution for the returns on Intelco's common stock.

Table A: Annual Returns for Intelco, Inc., Common Stock

10.4%	22.5%	11.1%	-12.4%
9.8%	17.0%	2.8%	8.4%
34.6%	-28.6%	0.6%	5.0%
-17.6%	5.6%	8.9%	40.4%
-1.0%	-4.2%	-5.2%	21.0%

Answer:

Step 1: Defining the interval. For Intelco's stock, the range of returns is 69.0% (-28.6% to 40.4%). Using a return interval of 1% would result in 69 separate intervals, which in this case is too many. So let's use eight non-overlapping intervals with a width of 10%. The lowest return intervals will be $-30\% \leq R_t < -20\%$, and the intervals will increase to $40\% \leq R_t \leq 50\%$.

Step 2: Tally the observations and count the observations within each interval. The tallies and counts of the observations are presented in Table B.

Table B: Tally and Interval Count for Returns Data

Interval	Tallies	Absolute Frequency
$-30\% \leq R_t < -20\%$	/	1
$-20\% \leq R_t < -10\%$	//	2
$-10\% \leq R_t < 0\%$	///	3
$0\% \leq R_t < 10\%$	////////	7
$10\% \leq R_t < 20\%$	///	3
$20\% \leq R_t < 30\%$	//	2
$30\% \leq R_t < 40\%$	/	1
$40\% \leq R_t \leq 50\%$	/	1
Total		20

Tallying and counting the observations generates a frequency distribution that summarizes the pattern of annual returns on Intelco common stock. Notice that the interval with the greatest (absolute) frequency is the ($0\% \leq R_t < 10\%$) interval, which includes seven return

observations. For any frequency distribution, the interval with the greatest frequency is referred to as the **modal interval**.

The **relative frequency** is another useful way to present data. The relative frequency is calculated by dividing the absolute frequency of each return interval by the total number of observations. Simply stated, relative frequency is the percentage of total observations falling within each interval. Continuing with our example, the relative frequencies are presented in Figure 2.2.

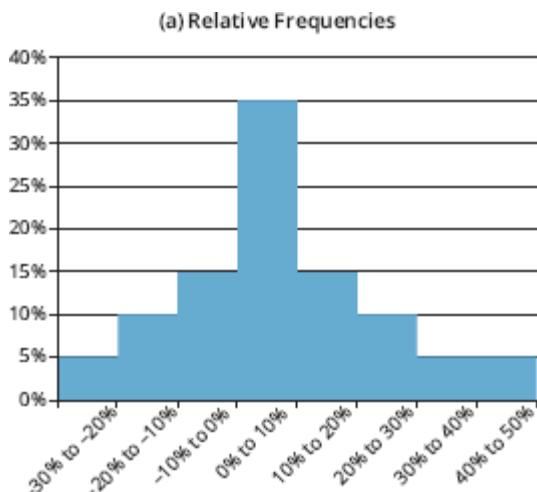
Figure 2.2: Absolute and Relative Frequencies of Intelco Returns

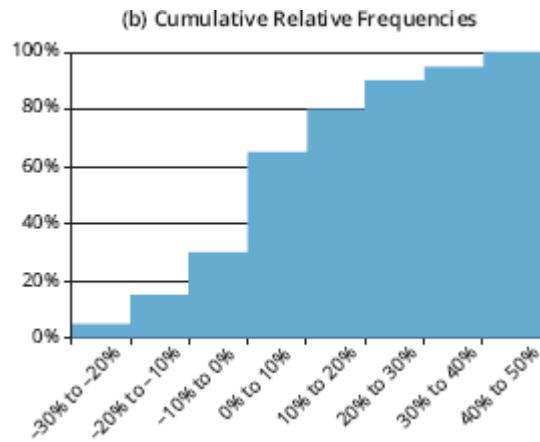
Interval	Absolute Frequency	Relative Frequency
$-30\% \leq R_t < -20\%$	1	$1/20 = 0.05$, or 5%
$-20\% \leq R_t < -10\%$	2	$2/20 = 0.10$, or 10%
$-10\% \leq R_t < 0\%$	3	$3/20 = 0.15$, or 15%
$0\% \leq R_t < 10\%$	7	$7/20 = 0.35$, or 35%
$10\% \leq R_t < 20\%$	3	$3/20 = 0.15$, or 15%
$20\% \leq R_t < 30\%$	2	$2/20 = 0.10$, or 10%
$30\% \leq R_t < 40\%$	1	$1/20 = 0.05$, or 5%
$40\% \leq R_t \leq 50\%$	1	$1/20 = 0.05$, or 5%
Total	20	100%

It is also possible to compute the **cumulative absolute frequency** and **cumulative relative frequency** by summing the absolute or relative frequencies starting at the lowest interval and progressing through the highest. The relative and cumulative relative frequencies for the Intelco stock returns example are presented in Figure 2.3.

The cumulative absolute frequency or cumulative relative frequency for any given interval is the sum of the absolute or relative frequencies up to and including the given interval. For example, the cumulative absolute frequency for $R_t < 10\%$ is $13 = 1 + 2 + 3 + 7$ and the cumulative relative frequency over this range is $5\% + 10\% + 15\% + 35\% = 65\%$.

Figure 2.3: Relative and Cumulative Frequencies of Intelco Returns





LOS 2.d: Interpret a contingency table.

A **contingency table** is a two-dimensional array with which we can analyze two variables at the same time. The rows represent attributes of one of the variables and the columns represent attributes of the other variable. These attributes can be defined using nominal or ordinal data, but there must be a finite number of them.

The data in each cell show the frequency with which we observe two attributes simultaneously. These are known as **joint frequencies** and they can be absolute or relative frequencies. The total of frequencies for a row or a column is termed the **marginal frequency** for that attribute.

For example, Figure 2.4 displays the number of traffic accidents in one year on weekdays at four intersections of a highway. From this table, we can see that accidents occur most frequently on Mondays and Fridays (marginal frequencies of 19 and 18) and that the Front Street intersection has the most accidents (marginal frequency of 25). The Front Street intersection on Mondays experiences the greatest number of accidents (joint frequency of 7).

Figure 2.4: Accidents by Intersection and Day of Week

Intersection	Monday	Tuesday	Wednesday	Thursday	Friday	Total
Palace Street	5	2	1	2	4	14
National Drive	3	2	3	1	3	12
Front Street	7	5	4	3	6	25
Jay Street	4	3	2	3	5	17
Total	19	12	10	9	18	68

One kind of contingency table is a 2-by-2 array called a **confusion matrix**. For each of two possible outcomes, a confusion matrix displays the number of occurrences predicted and the number actually observed.

Figure 2.5 illustrates a confusion matrix for a model that predicted the outcomes of 1,000 events. We can read from this table that the model predicted the outcome would occur 662 times, and it actually occurred 681 times. On 28 occasions, the model predicted the event would occur but it did not, while on 47 occasions, the model predicted the event would not occur when it actually did occur.

Figure 2.5: Confusion Matrix

	Actual Yes	Actual No	
Predicted Yes	634	28	662
Predicted No	<u>47</u>	<u>291</u>	<u>338</u>
	681	319	1,000

Another use of a contingency table is to use the values to determine whether two variables (characteristics), such as firm size and risk, are independent based on a chi-square test statistic.



PROFESSOR'S NOTE

In our reading on Hypothesis Testing, we address statistical tests, including the chi-square test for independence.



MODULE QUIZ 2.1

1. To perform meaningful mathematical analysis, an analyst must use data that are:
 - A. discrete.
 - B. numerical.
 - C. continuous.
2. Which of the following types of data would *most likely* be organized as a two-dimensional array?
 - A. Panel.
 - B. Time series.
 - C. Cross sectional.
3. The intervals in a frequency distribution should always be:
 - A. truncated.
 - B. open-ended.
 - C. non-overlapping.
4. Consider the following contingency table from a political opinion poll:

	Supports Johnson	Supports Williams	Total
Supports Smith	42%	14%	56%
Supports Jones	10%	34%	44%
Total	52%	48%	100%

In this table, the value 34% represents:

- A. a joint frequency.
- B. a marginal frequency.
- C. an absolute frequency.

MODULE 2.2: VISUALIZING DATA



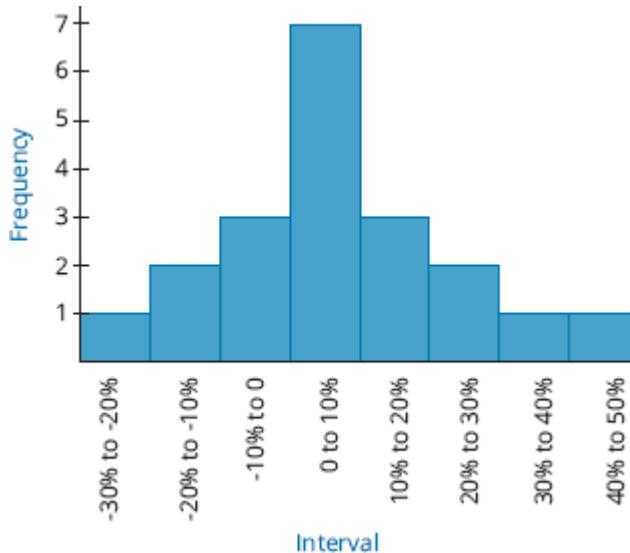
LOS 2.e: Describe ways that data may be visualized and evaluate uses of specific visualizations.

Video covering this content is available online.

A **histogram** is the graphical presentation of the absolute frequency distribution. A histogram is simply a bar chart of continuous data that has been classified into a frequency distribution. The attractive feature of a histogram is that it allows us to quickly see where most of the observations are concentrated.

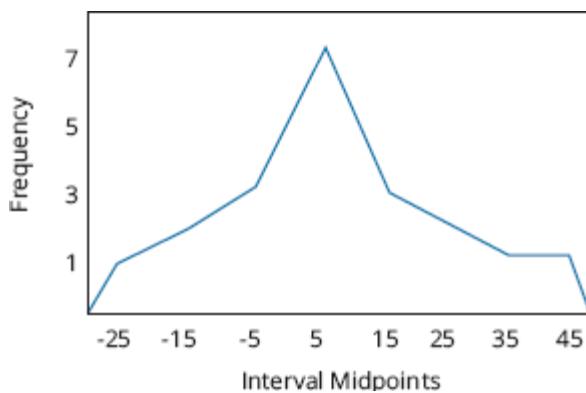
To construct a histogram, the intervals are shown on the horizontal axis and the absolute frequencies are shown on the vertical axis. The histogram for the Intelco returns data from the example presented earlier is provided in Figure 2.6.

Figure 2.6: Histogram of Intelco Stock Return Data



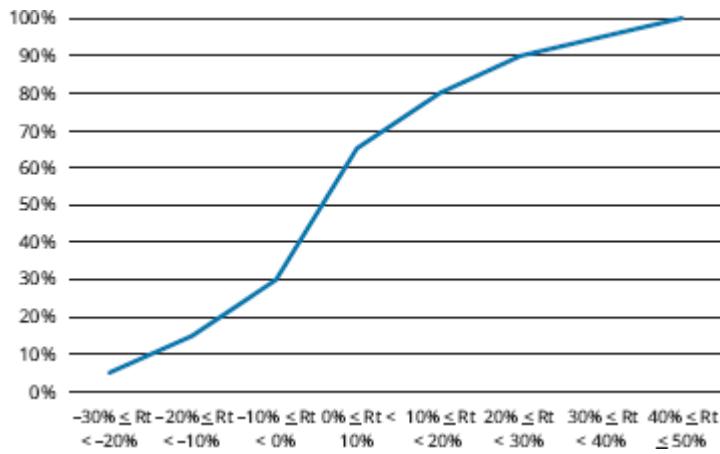
To construct a **frequency polygon**, successive frequencies at the midpoints of the intervals are joined with line segments. A frequency polygon for the Intelco returns data presented previously is illustrated in Figure 2.7.

Figure 2.7: Frequency Polygon of Intelco Stock Return Data



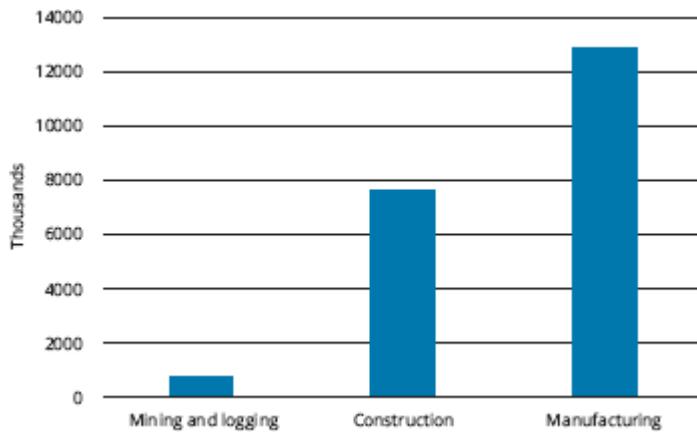
A **cumulative frequency distribution chart** displays either the cumulative absolute frequency or the cumulative relative frequency. Earlier, we showed the cumulative relative frequencies for Intelco as columns. They can also be displayed in a line chart, as in Figure 2.8.

Figure 2.8: Cumulative Relative Frequency Distribution



The histogram shown earlier is an example of a **bar chart**. In general, bar charts are used to illustrate relative sizes, degrees, or magnitudes. The bars can be displayed vertically or horizontally. Figure 2.9 shows a bar chart of employment in goods-producing industry groups in the United States. From this chart, we can see that the construction industries employ about 10 times as many people as the mining and logging industries and that manufacturing payrolls are a bit less than twice as large as construction payrolls.

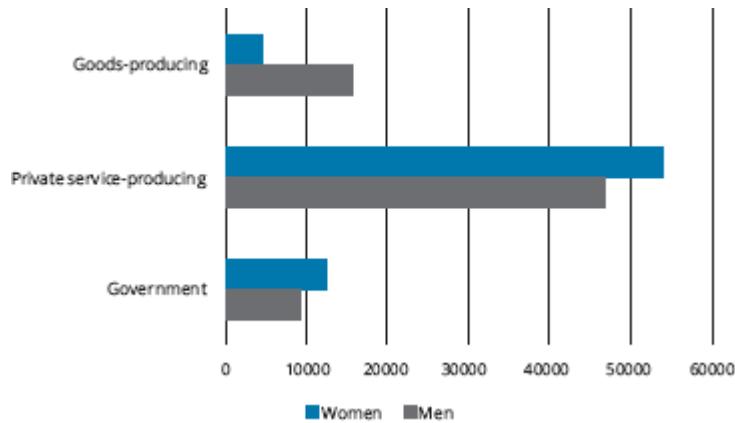
Figure 2.9: Employees on Payrolls in Goods-Producing Industries, January 2020



Source: Bureau of Labor Statistics, stats.bls.gov

A **grouped bar chart** or **clustered bar chart** can illustrate two categories at once, much like a data table. Figure 2.10 displays the number of men and women employed in three segments of the U.S. economy. Here we can see that more men than women are employed in the goods-producing industries, but more women than men are employed in the service-producing industries and government.

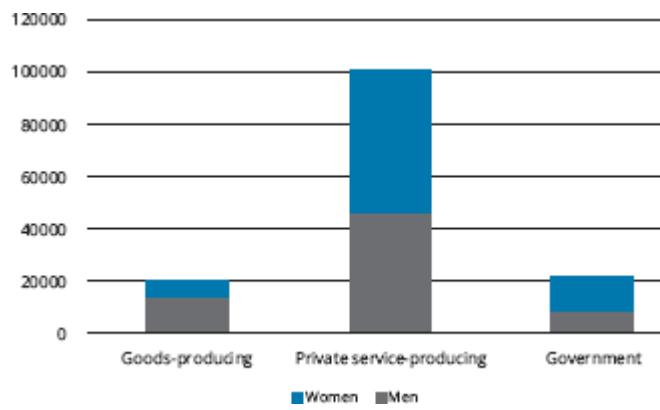
Figure 2.10: Grouped Bar Chart of Employment by Sector, December 2020



Source: Bureau of Labor Statistics, stats.bls.gov, with government payrolls estimated as the difference between total service-producing and private service-producing.

Another way to present two categories at once is with a **stacked bar chart**, as shown in Figure 2.11. In a stacked bar chart, the height of each bar represents the cumulative frequency for a category (such as goods-producing industries) and the colors within each bar represent joint frequencies (such as women employed in government). From this stacked bar chart, we can see the size of the private service-producing sector relative to the other two sectors.

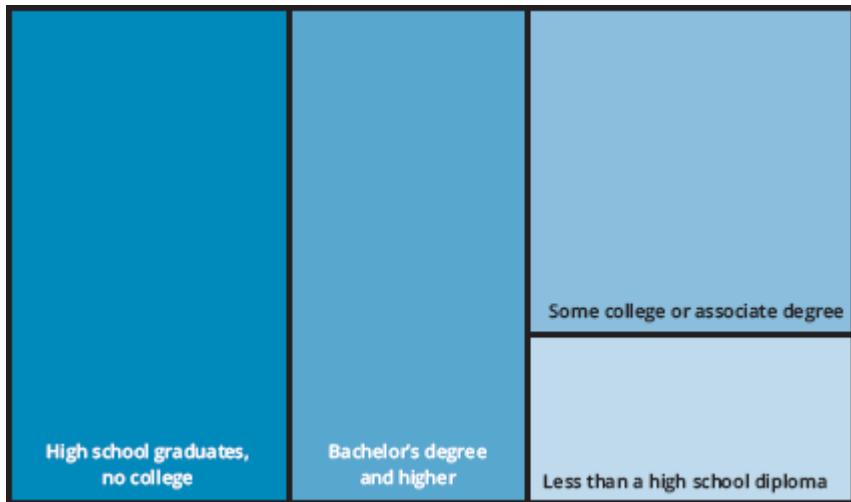
Figure 2.11: Stacked Bar Chart of Employment by Sector, December 2020



Source: Bureau of Labor Statistics, stats.bls.gov, with government payrolls estimated as the difference between total service-producing and private service-producing.

A **tree map** is another method for visualizing the relative sizes of categories. Figure 2.12 is a tree map of labor force categories for the United States by level of education. The filled areas of a tree map may each be divided into subcategories that are displayed in different colors or shades.

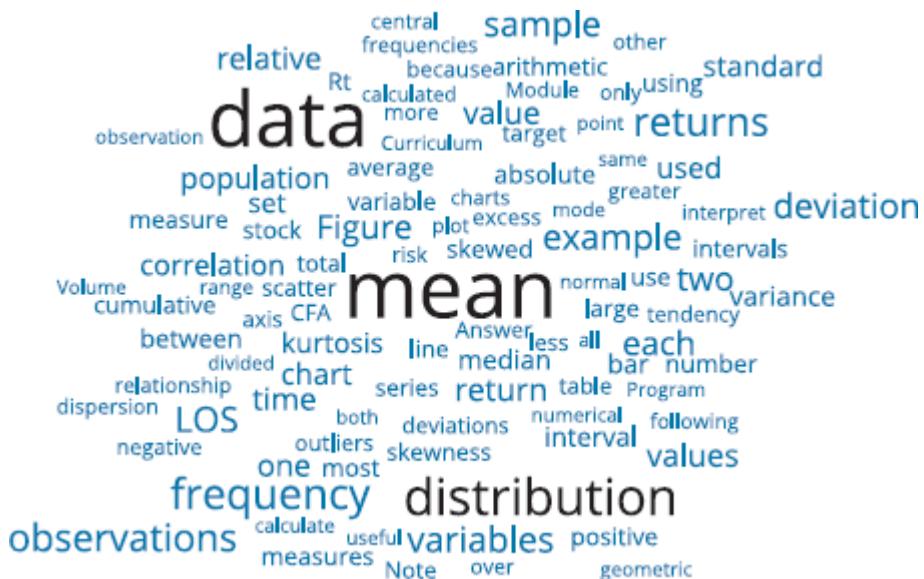
Figure 2.12: Tree Map of Labor Force by Educational Attainment, January 2020



Source: Bureau of Labor Statistics, stats.bls.gov

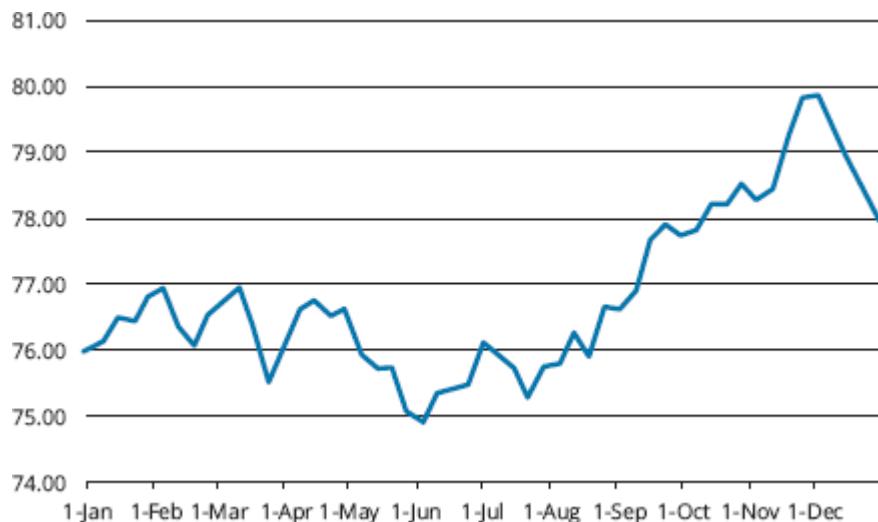
When analyzing text, a useful visualization technique is a **word cloud**. A word cloud is generated by counting the uses of specific words in text data. It displays frequently occurring words, in type sizes that are scaled to the frequency of their use. Figure 2.13 is an example of a word cloud generated from this reading. From this word cloud, we can easily see two of the major concepts this reading addresses: types of data and definitions of the mean.

Figure 2.13: Word Cloud



We have already seen some examples of **line charts**. Line charts are particularly useful for illustrating time series data, such as securities prices. Figure 2.14 is a line chart of weekly closing prices for a hypothetical stock.

Figure 2.14: Line Chart



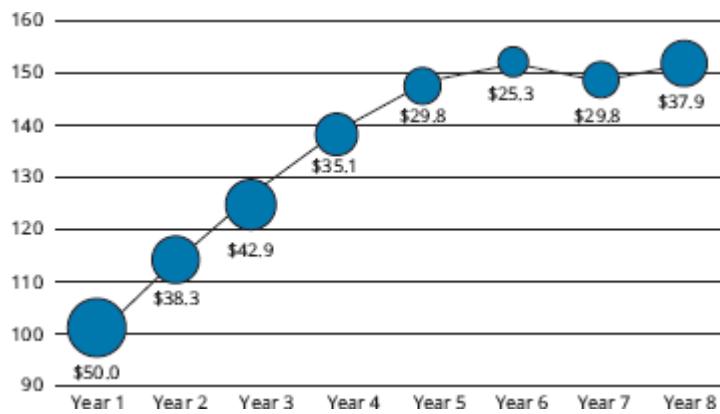
Multiple time series can be displayed on a line chart if their scales are comparable. It is also possible to display two time series on a line chart if their scales are different, by using left and right vertical axes as shown in Figure 2.15. This is one way of showing changes in two variables over time relative to each other.

Figure 2.15: Dual-Scale Line Chart



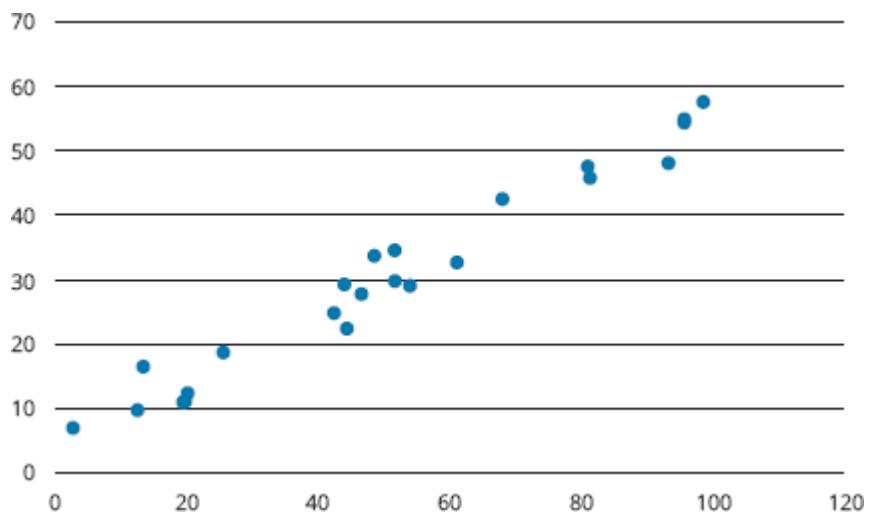
Another technique for adding a dimension to a line chart is to create a **bubble line chart**. For example, in Figure 2.16, a line chart shows total revenues for a company, and at each point, the different-sized bubbles represent revenues per salesperson. From this chart, we can see that revenues per salesperson were declining in years 1 to 6 while total revenues were increasing, but revenues per salesperson started increasing again in years 7 and 8, even though revenues were flat. This suggests the company was adding salespeople during the growth period but has been reducing its sales force recently.

Figure 2.16: Bubble Line Chart



A **scatter plot** is a way of displaying how two variables tend to change in relation to each other. The vertical axis represents one value of a variable and the horizontal axis represents the value of a second variable. Each point in the scatter plot shows the values of both variables at a point in time. Figure 2.17 is a scatter plot for two variables that have a fairly strong positive linear relationship.

Figure 2.17: Scatter Plot



While the relationship in the previous figure appears to be linear, scatter plots can also be useful for identifying nonlinear relationships that are not apparent when using a measure of the strength of a linear relationship, such as the correlation coefficient.

To analyze three variables at the same time, an analyst can create a **scatter plot matrix** that consists of three scatter plots of these variables, each presenting two of the three variables.

A **heat map** uses color and shade to display data frequency. Figure 2.18 is a heat map that uses data from the contingency table example we used previously to examine traffic accidents at several highway intersections. The darker shades indicate more accidents.

Figure 2.18: Heat Map

<u>Intersection</u>	<u>Monday</u>	<u>Tuesday</u>	<u>Wednesday</u>	<u>Thursday</u>	<u>Friday</u>
Palace Street	5	2	1	2	4
National Drive	3	2	3	1	3
Front Street	7	5	4	3	6
Jay Street	4	3	2	3	5

LOS 2.f: Describe how to select among visualization types.

Given the variety of charts to choose from, it can be useful to have a framework for choosing which to use in a specific circumstance. In general, we want to use the simplest chart that will clearly communicate the information to be presented.

We may need a chart to illustrate a relationship between two or more variables, compare two or more variables, or show the distribution of a single variable. Typically the most effective chart types for these purposes are as follows:

- **Relationships.** Scatter plots, scatter plot matrices, and heat maps.
- **Comparisons.** Bar charts, tree maps, and heat maps for comparisons among categories; line charts, dual-scale line charts, and bubble line charts for comparisons over time.
- **Distributions.** Histograms, frequency polygons, and cumulative distribution charts for numerical data; bar charts, tree maps, and heat maps for categorical data; and word clouds for text data.

When creating any chart, an analyst must take care to avoid misrepresentations. Selecting a chart type that is effective for visualizing the underlying data is a good first step. Beyond that, we must avoid potentially misleading practices, such as showing only a time period that supports our analysis while leaving out periods that illustrate the opposite, or choosing the scale of the axes so as to obscure meaningful variations or exaggerate non-meaningful variations in the data.



PROFESSOR'S NOTE

As we will see in our review of Ethical and Professional Standards, presenting selective data to mislead investors is a violation of Standard I(C) Misrepresentation.



MODULE QUIZ 2.2

1. The vertical axis of a histogram shows:
 - A. the frequency with which observations occur.
 - B. the range of observations within each interval.
 - C. the intervals into which the observations are arranged.
2. In which type of bar chart does the height or length of a bar represent the cumulative frequency for its category?
 - A. Stacked bar chart.
 - B. Grouped bar chart.
 - C. Clustered bar chart.

3. An analyst who wants to illustrate the relationships among three variables should *most appropriately* construct:
- a bubble line chart.
 - a scatter plot matrix.
 - a frequency polygon.

MODULE 2.3: MEASURES OF CENTRAL TENDENCY



Video covering this content is available online.

LOS 2.g: Calculate and interpret measures of central tendency.

Measures of central tendency identify the center, or average, of a data set. This central point can then be used to represent the typical, or expected, value in the data set.

To compute the **population mean**, all the observed values in the population are summed (ΣX) and divided by the number of observations in the population, N . Note that the population mean is unique in that a given population only has one mean. The population mean is expressed as:

$$\mu = \frac{\sum_{i=1}^N X_i}{N}$$

The **sample mean** is the sum of all the values in a sample of a population, ΣX , divided by the number of observations in the sample, n . It is used to make *inferences* about the population mean. The sample mean is expressed as:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Note the use of n , the sample size, versus N , the population size.

EXAMPLE: Population mean and sample mean

You have calculated the stock returns for AXZ Corporation over the last five years as 25%, 34%, 19%, 54%, and 17%. Given this information, estimate the mean of the distribution of returns.

Answer:

The sample mean can be used as an estimate of the mean of the distribution:

$$\bar{X} = \text{sample mean} = \frac{25 + 34 + 19 + 54 + 17}{5} = 29.8\%$$

The population mean and sample mean are both examples of **arithmetic means**. The arithmetic mean is the sum of the observation values divided by the number of observations. It is the most widely used measure of central tendency and has the following properties:

- All interval and ratio data sets have an arithmetic mean.
- All data values are considered and included in the arithmetic mean computation.
- A data set has only one arithmetic mean (i.e., the arithmetic mean is unique).

- The sum of the deviations of each observation in the data set from the mean is always zero.

The arithmetic mean is the only measure of central tendency for which the sum of the deviations from the mean is zero. Mathematically, this property can be expressed as follows:

$$\text{sum of mean deviations} = \sum_{i=1}^n (X_i - \bar{X}) = 0$$

Unusually large or small values, **outliers** can have a disproportionate influence on the arithmetic mean. The mean of 1, 2, 3, and 50 is 14 and is not a good indication of what the individual data values really are. On the positive side, the arithmetic mean uses all the information available about the observations. The arithmetic mean of a sample from a population is the best estimate of both the true mean of the sample and of the value of a single future observation.

In some cases, a researcher may decide that outliers should be excluded from a measure of central tendency. One technique for doing so is to use a **trimmed mean**. A trimmed mean excludes a stated percentage of the most extreme observations. A 1% trimmed mean, for example, would discard the lowest 0.5% and the highest 0.5% of the observations.

Another technique is to use a **winsorized mean**. Instead of discarding the highest and lowest observations, we substitute a value for them. To calculate a 90% winsorized mean, for example, we would determine the 5th and 95th percentile of the observations, substitute the 5th percentile for any values lower than that, substitute the 95th percentile for any values higher than that, and then calculate the mean of the revised data set.



PROFESSOR'S NOTE

Percentiles are explained later in this reading.

The computation of a **weighted mean** (or **weighted average**) recognizes that different observations may have a disproportionate influence on the mean. The weighted mean of a set of numbers is computed with the following equation:

$$\bar{X}_w = \sum_{i=1}^n w_i X_i = (w_1 X_1 + w_2 X_2 + \dots + w_n X_n)$$

where:

X_1, X_2, \dots, X_n = observed values

w_1, w_2, \dots, w_n = corresponding weights associated with each of the observations
such that $\sum w_i = 1$

EXAMPLE: Weighted mean as a portfolio return

A portfolio consists of 50% common stocks, 40% bonds, and 10% cash. If the return on common stocks is 12%, the return on bonds is 7%, and the return on cash is 3%, what is the portfolio return?

Answer:

$$\bar{X}_w = w_{\text{stock}} R_{\text{stock}} + w_{\text{bonds}} R_{\text{bonds}} + w_{\text{cash}} R_{\text{cash}}$$

$$\bar{X}_w = (0.50 \times 0.12) + (0.40 \times 0.07) + (0.10 \times 0.03) = 0.091, \text{ or } 9.1\%$$

The example illustrates an extremely important investments concept: *the return for a portfolio is the weighted average of the returns of the individual assets in the portfolio*. Asset weights are market weights, the market value of each asset relative to the market value of the entire portfolio.

The **median** is the midpoint of a data set when the data is arranged in ascending or descending order. Half the observations lie above the median and half are below. To determine the median, arrange the data from the highest to the lowest value, or lowest to highest value, and find the middle observation.

The median is important because the arithmetic mean can be affected by extremely large or small values (outliers). When this occurs, the median is a better measure of central tendency than the mean because it is not affected by extreme values that may actually be the result of errors in the data.

EXAMPLE: The median using an odd number of observations

What is the median return for five portfolio managers with a 10-year annualized total returns record of 30%, 15%, 25%, 21%, and 23%?

Answer:

First, arrange the returns in descending order.

30%, 25%, 23%, 21%, 15%

Then, select the observation that has an equal number of observations above and below it—the one in the middle. For the given data set, the third observation, 23%, is the median value.

EXAMPLE: The median using an even number of observations

Suppose we add a sixth manager to the previous example with a return of 28%. What is the median return?

Answer:

Arranging the returns in descending order gives us:

30%, 28%, 25%, 23%, 21%, 15%

With an even number of observations, there is no single middle value. The median value in this case is the arithmetic mean of the two middle observations, 25% and 23%. Thus, the median return for the six managers is $24.0\% = 0.5(25 + 23)$.

Consider that while we calculated the mean of 1, 2, 3, and 50 as 14, the median is 2.5. If the data were 1, 2, 3, and 4 instead, the arithmetic mean and median would both be 2.5.

The **mode** is the value that occurs most frequently in a data set. A data set may have more than one mode or even no mode. When a distribution has one value that appears most frequently, it is said to be **unimodal**. When a set of data has two or three values that occur most frequently, it is said to be **bimodal** or **trimodal**, respectively.

EXAMPLE: The mode

What is the mode of the following data set?

Data set: [30%, 28%, 25%, 23%, 28%, 15%, 5%]

Answer:

The mode is 28% because it is the value appearing most frequently.

The **geometric mean** is often used when calculating investment returns over multiple periods or when measuring compound growth rates. The general formula for the geometric mean, G , is as follows:

$$G = \sqrt[n]{X_1 \times X_2 \times \dots \times X_n} = (X_1 \times X_2 \times \dots \times X_n)^{1/n}$$

Note that this equation has a solution only if the product under the radical sign is nonnegative.

When calculating the geometric mean for a returns data set, it is necessary to add 1 to each value under the radical and then subtract 1 from the result. The geometric mean return (R_G) can be computed using the following equation:

$$1 + R_G = \sqrt[n]{(1 + R_1) \times (1 + R_2) \times \dots \times (1 + R_n)}$$

where:

R_t = the return for period t

EXAMPLE: Geometric mean return

For the last three years, the returns for Acme Corporation common stock have been -9.34%, 23.45%, and 8.92%. Compute the compound annual rate of return over the three-year period.

Answer:

$$1 + R_G = \sqrt[3]{(1 - 0.0934) \times (1 + 0.2345) \times (1 + 0.0892)}$$

$$1 + R_G = \sqrt[3]{0.9066 \times 1.2345 \times 1.0892} = \sqrt[3]{1.21903}$$

$$= (1.21903)^{1/3} = 1.06825$$

$$R_G = 1.06825 - 1 = 6.825\%$$

Solve this type of problem with your calculator as follows:

- On the TI, enter 1.21903 [y^x] 3 [1/x] [=]
- On the HP, enter 1.21903 [ENTER] 3 [1/x] [y^x]



PROFESSOR'S NOTE

The geometric mean is always less than or equal to the arithmetic mean, and the difference increases as the dispersion of the observations increases. The only time the arithmetic and geometric means are equal is when there is no variability in the observations (i.e., all observations are equal).

A **harmonic mean** is used for certain computations, such as the average cost of shares purchased over time. The harmonic mean is calculated as $\frac{N}{\sum_{i=1}^N \frac{1}{X_i}}$, where there are N values of X_i .

EXAMPLE: Calculating average cost with the harmonic mean

An investor purchases \$1,000 of mutual fund shares each month, and over the last three months, the prices paid per share were \$8, \$9, and \$10. What is the average cost per share?

Answer:

$$\bar{X}_H = \frac{3}{\frac{1}{8} + \frac{1}{9} + \frac{1}{10}} = \$8.926 \text{ per share}$$

To check this result, calculate the total shares purchased as:

$$\frac{1,000}{8} + \frac{1,000}{9} + \frac{1,000}{10} = 336.11 \text{ shares}$$

The average price is $\frac{\$3,000}{336.11} = \8.926 per share .

The previous example illustrates the interpretation of the harmonic mean in its most common application. Note that the average price paid per share (\$8.93) is less than the arithmetic average of the share prices, $\frac{8+9+10}{3} = 9$.

For values that are not all equal, harmonic mean < geometric mean < arithmetic mean. This mathematical fact is the basis for the claimed benefit of purchasing the same dollar amount of mutual fund shares each month or each week. Some refer to this practice as cost averaging.

LOS 2.h: Evaluate alternative definitions of mean to address an investment problem.

Appropriate uses for the various definitions of the mean are as follows:

- **Arithmetic mean.** Estimate the next observation, expected value of a distribution.
- **Geometric mean.** Compound rate of returns over multiple periods.
- **Trimmed mean.** Estimate the mean without the effects of a given percentage of outliers.
- **Winsorized mean.** Decrease the effect of outliers on the mean.
- **Harmonic mean.** Calculate the average share cost from periodic purchases in a fixed dollar amount.



MODULE QUIZ 2.3

1. XYZ Corp. Annual Stock Returns

2015	2016	2017	2018	2019	2020
22%	5%	-7%	11%	2%	11%

What is the arithmetic mean return for XYZ stock?

A. 7.3%.

- B. 8.0%.
- C. 11.0%.

2. XYZ Corp. Annual Stock Returns

2015	2016	2017	2018	2019	2020
22%	5%	-7%	11%	2%	11%

What is the median return for XYZ stock?

- A. 7.3%.
- B. 8.0%.
- C. 11.0%.

3. A data set has 100 observations. Which of the following measures of central tendency will be calculated using a denominator of 100?

- A. The winsorized mean, but not the trimmed mean.
- B. Both the trimmed mean and the winsorized mean.
- C. Neither the trimmed mean nor the winsorized mean.

4. The harmonic mean of 3, 4, and 5 is:

- A. 3.74.
- B. 3.83.
- C. 4.12.

5. XYZ Corp. Annual Stock Returns

2015	2016	2017	2018	2019	2020
22%	5%	-7%	11%	2%	11%

The mean annual return on XYZ stock is *most appropriately* calculated using:

- A. the harmonic mean.
- B. the arithmetic mean.
- C. the geometric mean.

MODULE 2.4: MEASURES OF LOCATION AND DISPERSION



Video covering this content is available online.

LOS 2.i: Calculate quantiles and interpret related visualizations.

Quantile is the general term for a value at or below which a stated proportion of the data in a distribution lies. Examples of quantiles include the following:

- **Quartile.** The distribution is divided into quarters.
- **Quintile.** The distribution is divided into fifths.
- **Decile.** The distribution is divided into tenths.
- **Percentile.** The distribution is divided into hundredths (percents).

Note that any quantile may be expressed as a percentile. For example, the third quartile partitions the distribution at a value such that three-fourths, or 75%, of the observations fall below that value. Thus, the third quartile is the 75th percentile. The difference between the third quartile and the first quartile (25th percentile) is known as the **interquartile range**.

The formula for the position of the observation at a given percentile, y , with n data points sorted in ascending order is:

$$L_y = (n + 1) \frac{y}{100}$$

Quantiles and measures of central tendency are known collectively as **measures of location**.

EXAMPLE: Quartiles

What is the third quartile for the following distribution of returns?

8%, 10%, 12%, 13%, 15%, 17%, 17%, 18%, 19%, 23%

Answer:

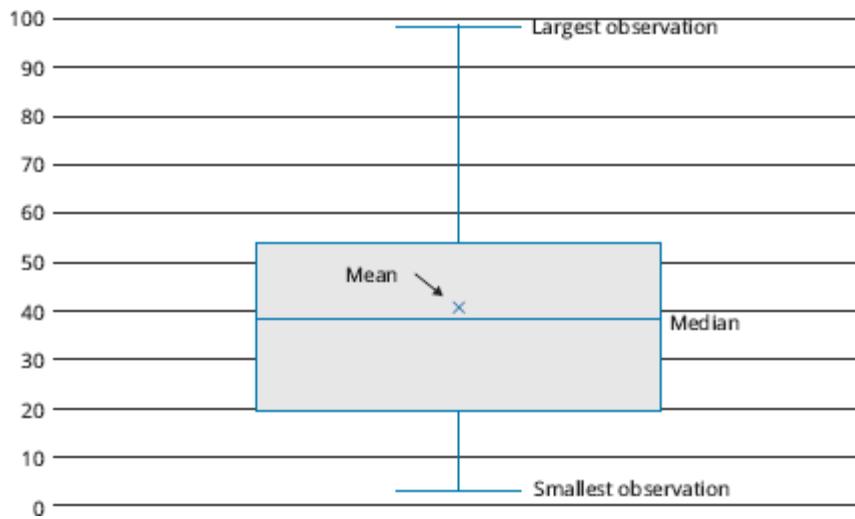
The third quartile is the point below which 75% of the observations lie. Recognizing that there are 10 observations in the data set, the third quartile can be identified as:

$$L_y = (10 + 1) \times \frac{75}{100} = 8.25$$

When the data are arranged in ascending order, the third quartile is a fourth (0.25) of the way from the eighth data point (18%) to the ninth data point (19%), or 18.25%. This means that 75% of all observations lie below 18.25%.

To visualize a data set based on quantiles, we can create a **box and whisker plot**, as shown in Figure 2.19. In a box and whisker plot, the box represents the central portion of the data, such as the interquartile range. The vertical line represents the entire range. In Figure 2.19, we can see that the largest observation is farther away from the center than the smallest observation is. This suggests that the data might include one or more outliers on the high side.

Figure 2.19: Box and Whisker Plot



LOS 2.j: Calculate and interpret measures of dispersion.

Dispersion is defined as the *variability around the central tendency*. The common theme in finance and investments is the tradeoff between reward and variability, where the central tendency is the measure of the reward and dispersion is a measure of risk.

The **range** is a relatively simple measure of variability, but when used with other measures, it provides extremely useful information. The range is the distance between the largest and the smallest value in the data set, or:

$$\text{range} = \text{maximum value} - \text{minimum value}$$

EXAMPLE: The range

What is the range for the 5-year annualized total returns for five investment managers if the managers' individual returns were 30%, 12%, 25%, 20%, and 23%?

Answer:

$$\text{range} = 30 - 12 = 18\%$$

The **mean absolute deviation (MAD)** is the average of the absolute values of the deviations of individual observations from the arithmetic mean:

$$\text{MAD} = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

The computation of the MAD uses the absolute values of each deviation from the mean because the sum of the actual deviations from the arithmetic mean is zero.

EXAMPLE: MAD

What is the MAD of the investment returns for the five managers discussed in the preceding example? How is it interpreted?

Answer:

annualized returns: [30%, 12%, 25%, 20%, 23%]

$$\begin{aligned}\bar{X} &= \frac{[30 + 12 + 25 + 20 + 23]}{5} = 22\% \\ \text{MAD} &= \frac{[|30 - 22| + |12 - 22| + |25 - 22| + |20 - 22| + |23 - 22|]}{5} \\ \text{MAD} &= \frac{[8 + 10 + 3 + 2 + 1]}{5} = 4.8\%\end{aligned}$$

This result can be interpreted to mean that, on average, an individual return will deviate $\pm 4.8\%$ from the mean return of 22%.

The **sample variance**, s^2 , is the measure of dispersion that applies when we are evaluating a sample of n observations from a population. The sample variance is calculated using the following formula:

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

The denominator for s^2 is $n - 1$, one less than the sample size n . Based on the mathematical theory behind statistical procedures, the use of the entire number of sample observations, n , instead of $n - 1$ as the divisor in the computation of s^2 , will systematically *underestimate* the population variance, particularly for small sample sizes. This systematic underestimation causes the sample variance to be a **biased estimator** of the population variance. Using $n - 1$ instead of n in the denominator, however, improves the statistical properties of s^2 as an estimator of the population variance.

EXAMPLE: Sample variance

Assume that the 5-year annualized total returns for the five investment managers used in the preceding examples represent only a sample of the managers at a large investment firm. What is the sample variance of these returns?

Answer:

$$\bar{X} = \frac{[30 + 12 + 25 + 20 + 23]}{5} = 22\%$$

$$s^2 = \frac{[(30 - 22)^2 + (12 - 22)^2 + (25 - 22)^2 + (20 - 22)^2 + (23 - 22)^2]}{5 - 1} = 44.5(\%)^2$$

Thus, the sample variance of $44.5(\%)^2$ can be interpreted to be an unbiased estimator of the population variance. Note that 44.5 “percent squared” is 0.00445 and you will get this value if you put the percentage returns in decimal form [e.g., $(0.30 - 0.22)^2$].

A major problem with using variance is the difficulty of interpreting it. The computed variance, unlike the mean, is in terms of squared units of measurement. How does one interpret squared percentages, squared dollars, or squared yen? This problem is mitigated through the use of the *standard deviation*. The units of standard deviation are the same as the units of the data (e.g., percentage return, dollars, euros). The **sample standard deviation** is the square root of the sample variance. The sample standard deviation, s , is calculated as:

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

EXAMPLE: Sample standard deviation

Compute the sample standard deviation based on the result of the preceding example.

Answer:

Because the sample variance for the preceding example was computed to be $44.5(\%)^2$, the sample standard deviation is:

$$s = [44.5(\%)^2]^{1/2} = 6.67\%, \text{ or } \sqrt{0.00445} = 0.0667$$

The results shown here mean that the sample standard deviation, $s = 6.67\%$, can be interpreted as an unbiased estimator of the population standard deviation, σ .

A direct comparison between two or more measures of dispersion may be difficult. For instance, suppose you are comparing the annual returns distribution for retail stocks with a mean of 8% and an annual returns distribution for a real estate portfolio with a mean of 16%. A direct comparison between the dispersion of the two distributions is not meaningful because of the relatively large difference in their means. To make a meaningful comparison, a relative measure of dispersion must be used. **Relative dispersion** is the amount of variability in a distribution relative to a reference point or benchmark. Relative dispersion is commonly measured with the **coefficient of variation (CV)**, which is computed as:

$$CV = \frac{s_x}{\bar{X}} = \frac{\text{standard deviation of } x}{\text{average value of } x}$$

CV measures the amount of dispersion in a distribution relative to the distribution's mean. It is useful because it enables us to make a direct comparison of dispersion across different sets of data. In an investments setting, the CV is used to measure the risk (variability) per unit of expected return (mean). A lower CV is better.

EXAMPLE: Coefficient of variation

You have just been presented with a report that indicates that the mean monthly return on T-bills is 0.25% with a standard deviation of 0.36%, and the mean monthly return for the S&P 500 is 1.09% with a standard deviation of 7.30%. Your unit manager has asked you to compute the CV for these two investments and to interpret your results.

Answer:

$$CV_{\text{T-bills}} = \frac{0.36}{0.25} = 1.44/$$

$$CV_{\text{S&P 500}} = \frac{7.30}{1.09} = 6.70$$

These results indicate that there is less dispersion (risk) per unit of monthly return for T-bills than for the S&P 500 (1.44 versus 6.70).



PROFESSOR'S NOTE

To remember the formula for CV, remember that the coefficient of variation is a measure of variation, so standard deviation goes in the numerator. CV is variation per unit of return.

LOS 2.k: Calculate and interpret target downside deviation.

When we use variance or standard deviation as risk measures, we calculate risk based on outcomes both above and below the mean. In some situations, it may be more appropriate to consider only outcomes less than the mean (or some other specific value) in calculating a risk measure. In this case, we are measuring **downside risk**.

One measure of downside risk is **target downside deviation**, which is also known as **target semideviation**. Calculating target downside deviation is similar to calculating standard deviation, but in this case, we choose a target value against which to measure each outcome and only include deviations from the target value in our calculation if the outcomes are below that target.

The formula for target downside deviation is stated as:

$$s_{\text{target}} = \sqrt{\frac{\sum_{\text{all } X_i < B}^n (X_i - B)^2}{n - 1}}, \text{ where } B \text{ is the target. Note that the denominator}$$

remains the sample size n minus one, even though we are not using all the observations in the numerator.

EXAMPLE: Target downside deviation

Calculate the target downside deviation based on the data in the preceding examples, for a target return equal to the mean (22%), and for a target return of 24%.

Answer:

Return	Deviation From Mean	Deviation From Target Return
30%	30% - 22% = 8%	30% - 24% = 6%
12%	12% - 22% = -10%	12% - 24% = -12%
25%	25% - 22% = 3%	25% - 24% = 1%
20%	20% - 22% = -2%	20% - 24% = -4%
23%	23% - 22% = 1%	23% - 24% = -1%

$$s_{22\%} = \sqrt{\frac{(-10)^2 + (-2)^2}{5 - 1}} = 5.10\%$$

$$s_{24\%} = \sqrt{\frac{(-12)^2 + (-4)^2 + (-1)^2}{5 - 1}} = 6.34\%$$



MODULE QUIZ 2.4

1. Given the following observations:

2, 4, 5, 6, 7, 9, 10, 11

The 65th percentile is *closest* to:

- A. 5.85.
- B. 6.55.
- C. 8.70.

2. XYZ Corp. Annual Stock Returns

20x1	20x2	20x3	20x4	20x5	20x6
22%	5%	-7%	11%	2%	11%

What is the sample standard deviation?

- A. 9.8%.
- B. 72.4%.
- C. 96.3%.

3. XYZ Corp. Annual Stock Returns

20x1	20x2	20x3	20x4	20x5	20x6
22%	5%	-7%	11%	2%	11%

Assume an investor has a target return of 11% for XYZ stock. What is the stock's target downside deviation?

- A. 9.4%.
- B. 12.1%.
- C. 14.8%.

MODULE 2.5: SKEWNESS, KURTOSIS, AND CORRELATION



Video covering
this content is
available online.

LOS 2.J: Interpret skewness.

A distribution is **symmetrical** if it is shaped identically on both sides of its mean. Distributional symmetry implies that intervals of losses and gains will exhibit the same frequency. For example, a symmetrical distribution with a mean return of zero will have losses in the -6% to -4% interval as frequently as it will have gains in the +4% to +6% interval. The extent to which a returns distribution is symmetrical is important because the degree of symmetry tells analysts if deviations from the mean are more likely to be positive or negative.

Skewness, or skew, refers to the extent to which a distribution is not symmetrical. Nonsymmetrical distributions may be either positively or negatively skewed and result from the occurrence of outliers in the data set. **Outliers** are observations extraordinarily far from the mean, either above or below:

- A *positively skewed* distribution is characterized by outliers greater than the mean (in the upper region, or right tail). A positively skewed distribution is said to be skewed right because of its relatively long upper (right) tail.
- A *negatively skewed* distribution has a disproportionately large amount of outliers less than the mean that fall within its lower (left) tail. A negatively skewed distribution is said to be skewed left because of its long lower tail.

Skewness affects the location of the mean, median, and mode of a distribution:

- For a symmetrical distribution, the mean, median, and mode are equal.
- For a positively skewed, unimodal distribution, the mode is less than the median, which is less than the mean. The mean is affected by outliers; in a positively skewed distribution, there are large, positive outliers, which will tend to pull the mean upward, or more positive. An example of a positively skewed distribution is that of housing prices. Suppose you live in a neighborhood with 100 homes; 99 of them sell for \$100,000 and one sells for \$1,000,000. The median and the mode will be \$100,000, but the mean will be \$109,000. Hence, the mean has been pulled upward (to the right) by the existence of one home (outlier) in the neighborhood.
- For a negatively skewed, unimodal distribution, the mean is less than the median, which is less than the mode. In this case, there are large, negative outliers that tend to pull the mean downward (to the left).

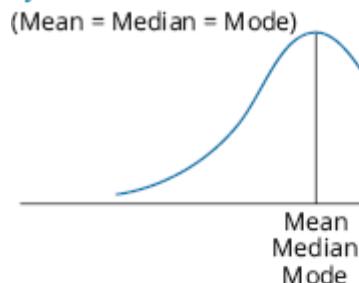


PROFESSOR'S NOTE

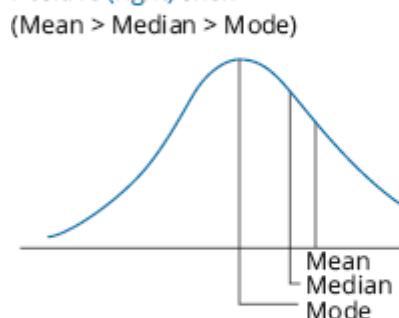
The key to remembering how measures of central tendency are affected by skewed data is to recognize that skew affects the mean more than the median and mode, and the mean is pulled in the direction of the skew. The relative location of the mean, median, and mode for different distribution shapes is shown in Figure 2.20. Note the median is between the other two measures for positively or negatively skewed distributions.

Figure 2.20: Effect of Skewness on Mean, Median, and Mode

Symmetrical

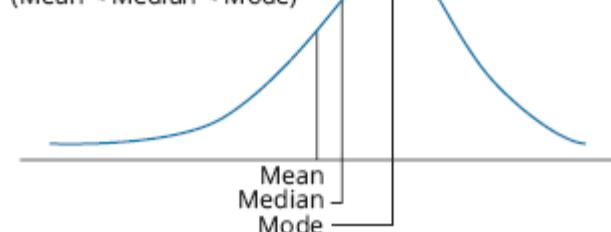


Positive (right) skew



Negative (left) skew

(Mean < Median < Mode)



Sample skewness is equal to the sum of the cubed deviations from the mean divided by the cubed standard deviation and by the number of observations. Sample skewness for large samples is computed as:

$$\text{sample skewness} = \frac{1}{n} \frac{\sum_{i=1}^n (X_i - \bar{X})^3}{s^3}$$

where:

s = sample standard deviation

Note that the denominator is always positive but that the numerator can be positive or negative depending on whether observations above the mean or observations below the mean tend to be farther from the mean on average. When a distribution is right skewed, sample skewness is positive because the deviations above the mean are larger on average. A left-skewed distribution has a negative sample skewness.

Dividing by standard deviation cubed standardizes the statistic and allows **interpretation of the skewness measure**. If relative skewness is equal to zero, the data is not skewed. Positive levels of relative skewness imply a positively skewed distribution, whereas negative values of relative skewness imply a negatively skewed distribution. Values of sample skewness in excess of 0.5 in absolute value are considered significant.



PROFESSOR'S NOTE

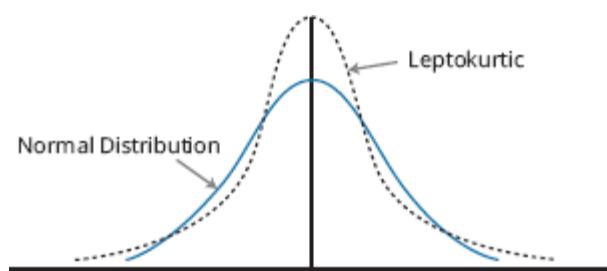
The LOS related to skewness and kurtosis require us to “interpret” these measures, but not to calculate them.

LOS 2.m: Interpret kurtosis.

Kurtosis is a measure of the degree to which a distribution is more or less peaked than a normal distribution. **Leptokurtic** describes a distribution that is more peaked than a normal distribution, whereas **platykurtic** refers to a distribution that is less peaked, or flatter than a normal distribution. A distribution is **mesokurtic** if it has the same kurtosis as a normal distribution.

As indicated in Figure 2.21, a leptokurtic return distribution will have more returns clustered around the mean and more returns with large deviations from the mean (fatter tails). Relative to a normal distribution, a leptokurtic distribution will have a greater percentage of small deviations from the mean and a greater percentage of extremely large deviations from the mean. This means that there is a relatively greater probability of an observed value being either close to the mean or far from the mean. With regard to an investment returns distribution, a greater likelihood of a large deviation from the mean return is often perceived as an increase in risk.

Figure 2.21: Kurtosis



A distribution is said to exhibit **excess kurtosis** if it has either more or less kurtosis than the normal distribution. The computed kurtosis for all normal distributions is three. Statisticians, however, sometimes report excess kurtosis, which is defined as kurtosis minus three. Thus, a normal distribution has excess kurtosis equal to zero, a leptokurtic distribution has excess kurtosis greater than zero, and platykurtic distributions will have excess kurtosis less than zero.

Kurtosis is critical in a risk management setting. Most research about the distribution of securities returns has shown that returns are not normally distributed. Actual securities returns tend to exhibit both skewness and kurtosis. Skewness and kurtosis are critical concepts for risk management because when securities returns are modeled using an assumed normal distribution, the predictions from the models will not take into account the potential for extremely large, negative outcomes. In fact, most risk managers put very little emphasis on the mean and standard deviation of a distribution and focus more on the distribution of returns in the tails of the distribution—that is where the risk is. In general, greater positive kurtosis and more negative skew in returns distributions indicates increased risk.

Sample kurtosis is measured using deviations raised to the *fourth power*:

$$\text{sample kurtosis} = \frac{1}{n} \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{s^4}$$

where:

s = sample standard deviation

To interpret kurtosis, note that it is measured relative to the kurtosis of a normal distribution, which is 3. Positive values of excess kurtosis indicate a distribution that is leptokurtic (more peaked, fat tails), whereas negative values indicate a platykurtic distribution (less peaked, thin tails). We can calculate kurtosis relative to that of a normal distribution as:

$$\text{excess kurtosis} = \text{sample kurtosis} - 3$$

LOS 2.n: Interpret correlation between two variables.

Covariance is a measure of how two variables move together. The calculation of the **sample covariance** is based on the following formula:

$$s_{X,Y} = \frac{\sum_{i=1}^n \{ [X_i - \bar{X}] [Y_i - \bar{Y}] \}}{n - 1}$$

where:

X_i = an observation of variable X

Y_i = an observation of variable Y

\bar{X} = mean of variable X

\bar{Y} = mean of variable Y

n = number of periods

In practice, the covariance is difficult to interpret. The value of covariance depends on the units of the variables. The covariance of daily price changes of two securities priced in yen will be much greater than the covariance they will be if the securities are priced in dollars. Like the variance, the units of covariance are the square of the units used for the data.

Additionally, we cannot interpret the relative strength of the relationship between two variables. Knowing that the covariance of X and Y is 0.8756 tells us only that they tend to move together because the covariance is positive. A standardized measure of the linear relationship between two variables is called the **correlation coefficient**, or simply correlation. The correlation between two variables, X and Y , is calculated as:

$$\rho_{XY} = \frac{s_{XY}}{s_X s_Y}$$
 which implies,

$$s_{XY} = \rho_{XY} s_X s_Y$$

The properties of the correlation of two random variables, X and Y , are summarized here:

- Correlation measures the strength of the linear relationship between two random variables.
- Correlation has no units.
- The correlation ranges from -1 to $+1$. That is, $-1 \leq \rho_{XY} \leq +1$.
- If $\rho^{XY} = 1.0$, the random variables have perfect positive correlation. This means that a movement in one random variable results in a proportional positive movement in the other relative to its mean.
- If $\rho^{XY} = -1.0$, the random variables have perfect negative correlation. This means that a movement in one random variable results in an exact opposite proportional movement in the other relative to its mean.
- If $\rho^{XY} = 0$, there is no linear relationship between the variables, indicating that prediction of Y cannot be made on the basis of X using linear methods.

EXAMPLE: Correlation

The variance of returns on stock A is 0.0028, the variance of returns on stock B is 0.0124, and their covariance of returns is 0.0058. Calculate and interpret the correlation of the returns for stocks A and B.

Answer:

First, it is necessary to convert the variances to standard deviations:

$$s_A = (0.0028)^{1/2} = 0.0529$$

$$s_B = (0.0124)^{1/2} = 0.1114$$

Now, the correlation between the returns of stock A and stock B can be computed as follows:

$$\rho_{AB} = \frac{0.0058}{(0.0529)(0.1114)} = 0.9842$$

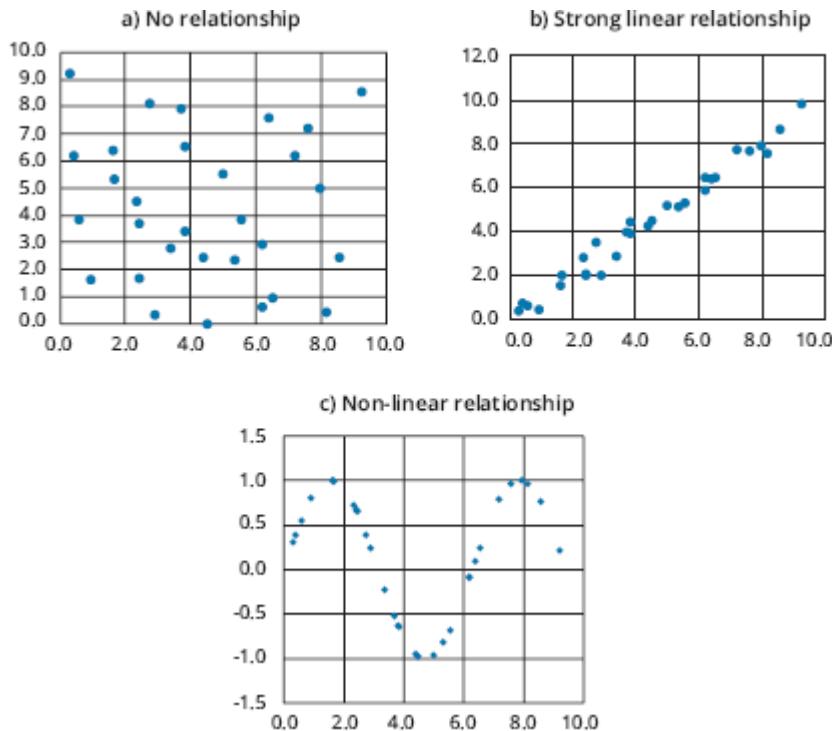
The fact that this value is close to $+1$ indicates that the linear relationship is not only positive but very strong.

Scatter plots are a method for displaying the relationship between two variables. With one variable on the vertical axis and the other on the horizontal axis, their paired observations can each be plotted as a single point. For example, in panel a of Figure 2.22, the point farthest to the upper right shows that when one of the variables (on the horizontal axis) equaled 9.2, the other variable (on the vertical axis) equaled 8.5.

The scatter plot in panel a is typical of two variables that have no clear relationship. Panel b shows two variables that have a strong linear relationship—that is, a high correlation coefficient.

A key advantage of creating scatter plots is that they can reveal nonlinear relationships, which are not described by the correlation coefficient. Panel c illustrates such a relationship. Although the correlation coefficient for these two variables is close to zero, their scatter plot shows clearly that they are related in a predictable way.

Figure 2.22: Scatter plots



Care should be taken when drawing conclusions based on correlation. Causation is not implied just from significant correlation. Even if it were, which variable is causing change in the other is not revealed by correlation. It is more prudent to say that two variables exhibit positive (or negative) association, suggesting that the nature of any causal relationship is to be separately investigated or based on theory that can be subject to additional tests.

One question that can be investigated is the role of outliers (extreme values) in the correlation of two variables. If removing the outliers significantly reduces the calculated correlation, further inquiry is necessary into whether the outliers provide information or are caused by noise (randomness) in the data used.

Spurious correlation refers to correlation that is either the result of chance or present due to changes in both variables over time that is caused by their association with a third variable. For example, we can find instances where two variables that are both related to the inflation rate exhibit significant correlation but for which causation in either direction is not present.

In his book *Spurious Correlation*¹, Tyler Vigen presents the following examples. The correlation between the age of each year's Miss America and the number of films Nicholas Cage appeared in that year is 87%. This seems a bit random. The correlation between the U.S. spending on science, space, and technology and suicides by hanging, strangulation, and suffocation over the 1999–2009 period is 99.87%. Impressive correlation, but both variables increased in an approximately linear fashion over the period.



MODULE QUIZ 2.5

1. Which of the following is *most accurate* regarding a distribution of returns that has a mean greater than its median?
 - A. It is positively skewed.
 - B. It is a symmetric distribution.
 - C. It has positive excess kurtosis.
2. A distribution of returns that has a greater percentage of small deviations from the mean and a greater percentage of extremely large deviations from the mean compared with a normal distribution:
 - A. is positively skewed.
 - B. has positive excess kurtosis.
 - C. has negative excess kurtosis.
3. The correlation between two variables is +0.25. The *most appropriate* way to interpret this value is to say:
 - A. a scatter plot of the two variables is likely to show a strong linear relationship.
 - B. when one variable is above its mean, the other variable tends to be above its mean as well.
 - C. a change in one of the variables usually causes the other variable to change in the same direction.

KEY CONCEPTS

LOS 2.a

We may classify data types from three different perspectives: numerical versus categorical, time series versus cross sectional, and structured versus unstructured.

Numerical, or quantitative, data are values that can be counted or measured and may be discrete or continuous. Categorical, or qualitative, data are labels that can be used to classify a set of data into groups and may be nominal or ordinal.

A time series is a set of observations taken at a sequence of points in time. Cross-sectional data are a set of comparable observations taken at one point in time. Time series and cross-sectional data may be combined to form panel data.

Unstructured data refers to information that is presented in forms that are not regularly structured and may be generated by individuals, business processes, or sensors.

LOS 2.b

Data are typically organized into arrays for analysis. A time series is an example of a one-dimensional array. A data table is an example of a two-dimensional array.

LOS 2.c

A frequency distribution groups observations into classes, or intervals. An interval is a range of values.

Relative frequency is the percentage of total observations falling within an interval.

Cumulative relative frequency for an interval is the sum of the relative frequencies for all values less than or equal to that interval's maximum value.

LOS 2.d

A contingency table is a two-dimensional array with which we can analyze two variables at the same time. The rows represent some attributes of one of the variables and the columns represent those attributes for the other variable. The data in each cell show the joint frequency with which we observe a pair of attributes simultaneously. The total of frequencies for a row or a column is the marginal frequency for that attribute.

LOS 2.e

A histogram is a bar chart of data that has been grouped into a frequency distribution.

A frequency polygon plots the midpoint of each interval on the horizontal axis and the absolute frequency for that interval on the vertical axis, and it connects the midpoints with straight lines.

A cumulative frequency distribution chart is a line chart of the cumulative absolute frequency or the cumulative relative frequency.

Bar charts can be used to illustrate relative sizes, degrees, or magnitudes. A grouped or clustered bar chart can illustrate two categories at once. In a stacked bar chart, the height of each bar represents the cumulative frequency for a category, and the colors within each bar represent joint frequencies. A tree map is another method for visualizing the relative sizes of categories.

A word cloud is generated by counting the uses of specific words in a text file. It displays the words that appear most often, in type sizes that are scaled to the frequency of their use.

Line charts are particularly useful for exhibiting time series. Multiple time series can be displayed on a line chart if their scales are comparable. It is also possible to display two time series on a line chart if their scales are different by using left and right vertical axes. A technique for adding a dimension to a line chart is to create a bubble line chart.

A scatter plot is a way of displaying how two variables tend to change together. The vertical axis represents one variable and the horizontal axis represents a second variable. Each point in the scatter plot shows the values of both variables at one specific point in time.

A heat map uses color and shade to display data frequency.

LOS 2.f

Which chart types tend to be most effective depends on what they are intended to visualize:

- **Relationships.** Scatter plots, scatter plot matrices, and heat maps.
- **Comparisons.** Bar charts, tree maps, and heat maps for comparisons among categories; line charts and bubble line charts for comparisons over time.
- **Distributions.** Histograms, frequency polygons, and cumulative distribution charts for numerical data; bar charts, tree maps, and heat maps for categorical data; and word clouds for unstructured data.

LOS 2.g

The arithmetic mean is the average:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Population mean and sample mean are examples of arithmetic means.

The geometric mean is used to find a compound growth rate:

$$G = \sqrt[n]{X_1 \times X_2 \times \dots \times X_n}$$

The weighted mean weights each value according to its influence:

$$\bar{X}_w = \sum_{i=1}^n w_i X_i$$

The harmonic mean can be used to find an average purchase price, such as dollars per share for equal periodic investments:

$$\bar{X}_H = \frac{N}{\sum_{i=1}^N \frac{1}{X_i}}$$

The median is the midpoint of a data set when the data are arranged from largest to smallest.

The mode of a data set is the value that occurs most frequently.

LOS 2.h

Arithmetic mean is used to estimate expected value, value of a single outcome from a distribution.

Geometric mean is used calculate or estimate periodic compound returns over multiple periods.

Harmonic mean is used to calculate the average price paid with equal periodic investments.

A trimmed mean omits outliers and a winsorized mean replaces outliers with given values, reducing the effect of outliers on the mean in both cases.

LOS 2.i

Quantile is the general term for a value at or below which a stated proportion of the data in a distribution lies. Examples of quantiles include the following:

- Quartile. The distribution is divided into quarters.
- Quintile. The distribution is divided into fifths.
- Decile. The distribution is divided into tenths.
- Percentile. The distribution is divided into hundredths (percents).

LOS 2.j

The range is the difference between the largest and smallest values in a data set.

Mean absolute deviation (MAD) is the average of the absolute values of the deviations from the arithmetic mean:

$$MAD = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

Variance is defined as the mean of the squared deviations from the arithmetic mean or from the expected value of a distribution:

- Sample variance $s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$, where \bar{X} = sample mean and n = sample size.

Standard deviation is the positive square root of the variance and is frequently used as a quantitative measure of risk.

The coefficient of variation for sample data, $CV = \frac{s}{\bar{X}}$, is the ratio of the standard deviation of the sample to its mean (expected value of the underlying distribution).

LOS 2.k

Target downside deviation or semideviation is a measure of downside risk. Calculating target downside deviation is similar to calculating standard deviation, but in this case, we choose a target against which to measure each outcome and only include outcomes below that target when calculating the numerator.

The formula for target downside deviation is:

$$s_{target} = \sqrt{\frac{\sum_{all X_i < B}^n (X_i - B)^2}{n - 1}}, \text{ where } B \text{ is the target value.}$$

LOS 2.1

Skewness describes the degree to which a distribution is not symmetric about its mean. A right-skewed distribution has positive skewness. A left-skewed distribution has negative skewness.

For a positively skewed, unimodal distribution, the mean is greater than the median, which is greater than the mode.

For a negatively skewed, unimodal distribution, the mean is less than the median, which is less than the mode.

LOS 2.m

Kurtosis measures the peakedness of a distribution and the probability of extreme outcomes (thickness of tails):

- Excess kurtosis is measured relative to a normal distribution, which has a kurtosis of 3.
- Positive values of excess kurtosis indicate a distribution that is leptokurtic (fat tails, more peaked), so the probability of extreme outcomes is greater than for a normal distribution.
- Negative values of excess kurtosis indicate a platykurtic distribution (thin tails, less peaked).

LOS 2.n

Correlation is a standardized measure of association between two random variables. It ranges in value from -1 to $+1$ and is equal to $\frac{\text{Cov}_{A,B}}{\sigma_A \sigma_B}$.

Scatterplots are useful for revealing nonlinear relationships that are not measured by correlation.

Correlation does not imply that changes in one variable cause changes in the other. Spurious correlation may result by chance or from the relationships of two variables to a third variable.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 2.1

1. **B** We can perform mathematical operations on numerical data but not on categorical data. Numerical data can be discrete or continuous. (LOS 2.a)
2. **A** Panel data combine time series data with cross-sectional data and are typically organized as data tables, which are two-dimensional arrays. (LOS 2.a,b)
3. **C** Intervals within a frequency distribution should always be non-overlapping and closed-ended so that each data value can be placed into only one interval. Interval widths should be defined so that data are adequately summarized without losing valuable characteristics. (LOS 2.c)
4. **A** The value 34% is the joint probability that a voter supports both Jones and Williams. Because it is stated as a percentage, this value is a relative frequency. The totals for each row and column are marginal frequencies. An absolute frequency is a number of occurrences, not a percentage of occurrences. (LOS 2.d)

Module Quiz 2.2

1. **A** In a histogram, the intervals are on the horizontal axis and the frequency is on the vertical axis. (LOS 2.e)
2. **A** In a stacked bar chart, the height or length of a bar represents the cumulative frequency of a category. In a grouped or clustered bar chart, each category is displayed with bars side by side that together represent the cumulative frequency. (LOS 2.e)
3. **B** With a scatter plot matrix, an analyst can visualize the relationships among three variables by organizing scatter plots of the relationships between each pair of variables. Bubble line charts are typically used to visualize two variables over time. Frequency polygons are best used to visualize distributions. (LOS 2.f)

Module Quiz 2.3

1. **A** $[22\% + 5\% + -7\% + 11\% + 2\% + 11\%] / 6 = 7.3\%$ (LOS 2.g)
2. **B** To find the median, rank the returns in order and take the middle value: -7%, 2%, 5%, 11%, 11%, 22%. In this case, because there is an even number of observations, the median is the average of the two middle values, or $(5\% + 11\%) / 2 = 8.0\%$. (LOS 2.g)
3. **A** The winsorized mean substitutes a value for some of the largest and smallest observations. The trimmed mean removes some of the largest and smallest observations. (LOS 2.g)
4. **B** $\bar{X}_H = \frac{\frac{3}{1} + \frac{1}{4} + \frac{1}{5}}{3 + 4 + 5} = 3.83$
(LOS 2.g)
5. **C** Because returns are compounded, the geometric mean is appropriate.

$$[(1.22)(1.05)(0.93)(1.11)(1.02)(1.11)]^{1/6} - 1 = 6.96\%$$

(LOS 2.h)

Module Quiz 2.4

1. **C** With eight observations, the location of the 65th percentile is:

$$(8 + 1) \times 65/100 = 5.85 \text{ observations}$$

The fifth observation is 7 and the sixth observation is 9, so the value at 5.85 observations is 7 + 0.85(9 - 7) = 8.7. (LOS 2.i)

2. **A** The sample standard deviation is the square root of the sample variance:

$$s = \sqrt{\frac{(22 - 7.3)^2 + (5 - 7.3)^2 + (-7 - 7.3)^2 + (11 - 7.3)^2 + (2 - 7.3)^2 + (11 - 7.3)^2}{6 - 1}}$$

$$= \sqrt{96.3} = 9.8\%$$

(LOS 2.j)

3. A Deviations from the target return:

$$22\% - 11\% = 11\%$$

$$5\% - 11\% = -6\%$$

$$-7\% - 11\% = -18\%$$

$$11\% - 11\% = 0\%$$

$$2\% - 11\% = -9\%$$

$$11\% - 11\% = 0\%$$

$$\text{Target downside deviation} = \sqrt{\frac{(-6)^2 + (-18)^2 + (-9)^2}{6 - 1}} = \sqrt{88.2} = 9.39\%$$

(LOS 2.k)

Module Quiz 2.5

- 1. A** A distribution with a mean greater than its median is positively skewed, or skewed to the right. The skew pulls the mean. Kurtosis deals with the overall shape of a distribution, not its skewness. (LOS 2.l)
- 2. B** A distribution that has a greater percentage of small deviations from the mean and a greater percentage of extremely large deviations from the mean will be leptokurtic and will exhibit excess kurtosis (positive). The distribution will be more peaked and have fatter tails than a normal distribution. (LOS 2.m)
- 3. B** Correlation of +0.25 indicates a positive linear relationship between the variables—one tends to be above its mean when the other is above its mean. The value 0.25 indicates that the linear relationship is not particularly strong. Correlation does not imply causation. (LOS 2.n)

¹ Tyler Vigen, "Spurious Correlations," www.tylervigen.com

READING 3

PROBABILITY CONCEPTS

EXAM FOCUS

This reading covers important terms and concepts associated with probability theory. We describe random variables, events, outcomes, conditional probability, and joint probability, and introduce probability rules such as the addition rule and multiplication rule. Finance practitioners use these rules frequently. We also discuss expected value, standard deviation, covariance, and correlation for individual asset and portfolio returns. A well-prepared candidate will be able to calculate and interpret these widely used measures. This review also discusses counting rules, which lay the foundation for the binomial probability distribution that is covered in the next reading.

MODULE 3.1: CONDITIONAL AND JOINT PROBABILITIES



Video covering
this content is
available online.

LOS 3.a: Define a random variable, an outcome, and an event.

LOS 3.b: Identify the two defining properties of probability, including mutually exclusive and exhaustive events, and compare and contrast empirical, subjective, and a priori probabilities.

- A **random variable** is an uncertain quantity/number.
- An **outcome** is an observed value of a random variable.
- An **event** is a single outcome or a set of outcomes.
- **Mutually exclusive events** are events that cannot both happen at the same time.
- **Exhaustive events** are those that include all possible outcomes.

Consider rolling a 6-sided die one time. The number that comes up is a *random variable*. If you roll a 4, that is an *outcome*. Rolling a 4 is also an *event*, as is rolling an even number. The possible outcomes from 1 to 6 are *mutually exclusive* (you cannot get a 3 and a 5 on the same roll) and *exhaustive* (you cannot roll a 7 or a 0).

There are **two defining properties of probability**:

- The probability of occurrence of any event (E_i) is between 0 and 1 (i.e., $0 \leq P(E_i) \leq 1$).
- If a set of events, E_1, E_2, \dots, E_n , is mutually exclusive and exhaustive, the probabilities of those events sum to 1 (i.e., $\sum P(E_i) = 1$).

The first of the defining properties introduces the term $P(E_i)$, which is shorthand for the “probability of event i .” If $P(E_i) = 0$, the event will never happen. If $P(E_i) = 1$, the event is certain to occur, and the outcome is not random. The probability of rolling any one of the numbers 1–6 with a fair die is $1/6 = 0.1667 = 16.7\%$. The set of events—rolling a number equal to 1, 2, 3, 4, 5, or 6—is an exhaustive set of outcomes (events). The six possible outcomes are mutually exclusive events, if a 2 is rolled, none of the other values can be the result of that roll. The probability of this set of events thus is 100% (equal to 1).

An **empirical probability** is established by analyzing past data (outcomes). An **a priori probability** is determined using a formal reasoning and inspection process (not data).

Inspecting a coin and reasoning that the probability of each side coming up when the coin is flipped is an example of an a priori probability.

A **subjective probability** is the least formal method of developing probabilities and involves the use of personal judgment. An analyst may know many things about a firm’s performance and have expectations about the overall market that are all used to arrive at a subjective probability, such as “I believe there is a 70% probability that Acme Foods will outperform the market this year.” Empirical and a priori probabilities, by contrast, are **objective probabilities**.

LOS 3.c: Describe the probability of an event in terms of odds for and against the event.

Stating the **odds** that an event will or will not occur is an alternative way of expressing probabilities. Consider an event that has a probability of occurrence of 0.125, which is one-eighth. The *odds* that the event will occur are $\frac{0.125}{(1 - 0.125)} = \frac{\frac{1}{8}}{\frac{7}{8}} = \frac{1}{7}$, which we state as “the odds for the event occurring are one-to-seven.” The *odds against* the event occurring are the reciprocal of $1/7$, which is seven-to-one.

We can also get the probability of an event from the odds by reversing these calculations. If we know that the odds for an event are one-to-six, we can compute the probability of occurrence as $\frac{1}{1+6} = \frac{1}{7} = 0.1429 = 14.29\%$. Alternatively, the probability that the event will not occur is $\frac{6}{1+6} = \frac{6}{7} = 0.8571 = 85.71\%$.

LOS 3.d: Calculate and interpret conditional probabilities.

Unconditional probability (a.k.a. *marginal probability*) refers to the probability of an event regardless of the past or future occurrence of other events. If we are concerned with the probability of an economic recession, regardless of the occurrence of changes in interest rates or inflation, we are concerned with the unconditional probability of a recession.

A **conditional probability** is one where the occurrence of one event affects the probability of the occurrence of another event. In symbols we write “the probability of A occurring, given that B has occurred as $\text{Prob}(A|B)$ or $P(A|B)$. For example, we might be concerned with the probability of a recession *given* that the monetary authority has increased interest rates. This is a conditional probability. The key word to watch for here is “given.” Using probability notation, “the probability of A *given* the occurrence of B” is expressed as $P(A | B)$, where the vertical bar

(|) indicates “given,” or “conditional upon.” For our interest rate example above, the probability of a recession *given* an increase in interest rates is expressed as $P(\text{recession} | \text{increase in interest rates})$. A conditional probability of an occurrence is also called its **likelihood**.

Consider a numerical example. If the Fed increases the USD policy rate, there is a 70% probability that a recession will follow. If the Fed does not increase the USD policy rate, there is a 20% probability that a recession will follow. These are both conditional expectations, $P(\text{recession} | \text{rate increase}) = 70\%$, and $P(\text{recession} | \text{no rate increase}) = 20\%$.

LOS 3.e: Demonstrate the application of the multiplication and addition rules for probability.

The **addition rule of probability** is used to determine the probability that at least one of two events will occur:

$$P(A \text{ or } B) = P(A) + P(B) - P(AB)$$

where B_1, B_2, \dots, B_N is a mutually exclusive and exhaustive set of outcomes.

The **joint probability** of two events is the probability that they will both occur. We can calculate this from the conditional probability that A will occur given B occurs (a conditional probability) and the probability that B will occur (the unconditional probability of B). This calculation is sometimes referred to as the **multiplication rule of probability**. Using the notation for conditional and unconditional probabilities, we can express this rule as:

$$P(AB) = P(A | B) \times P(B)$$

This expression is read as follows: “The joint probability of A and B, $P(AB)$, is equal to the conditional probability of A *given* B, $P(A | B)$, times the unconditional probability of B, $P(B)$.”

This relationship can be rearranged to define the conditional probability of A given B as follows:

$$P(A | B) = \frac{P(AB)}{P(B)}$$

EXAMPLE: Multiplication rule of probability

Consider the following information:

- $P(I) = 0.4$, the probability of the monetary authority increasing interest rates (I) is 40%.
- $P(R | I) = 0.7$, the probability of a recession (R) given an increase in interest rates is 70%.

What is $P(RI)$, the joint probability of a recession *and* an increase in interest rates?

Answer:

Applying the multiplication rule, we get the following result:

$$P(RI) = P(R | I) \times P(I)$$

$$P(RI) = 0.7 \times 0.4$$

$$P(RI) = 0.28$$

Don't let the cumbersome notation obscure the simple logic of this result. If an interest rate increase will occur 40% of the time and lead to a recession 70% of the time when it occurs, the joint probability of an interest rate increase and a resulting recession is $(0.4)(0.7) = (0.28) = 28\%$.

Calculating the Probability That at Least One of Two Events Will Occur

The *addition rule* for probabilities is used to determine the probability that at least one of two events will occur. For example, given two events, A and B, the addition rule can be used to determine the probability that either A or B will occur. If the events are *not* mutually exclusive, double counting must be avoided by subtracting the joint probability that *both* A and B will occur from the sum of the unconditional probabilities. This is reflected in the following general expression for the addition rule:

$$P(A \text{ or } B) = P(A) + P(B) - P(AB)$$

For mutually exclusive events, where the joint probability, $P(AB)$, is zero, the probability that either A or B will occur is simply the sum of the unconditional probabilities for each event, $P(A \text{ or } B) = P(A) + P(B)$.

Figure 3.1: Venn Diagram for Events That Are Not Mutually Exclusive

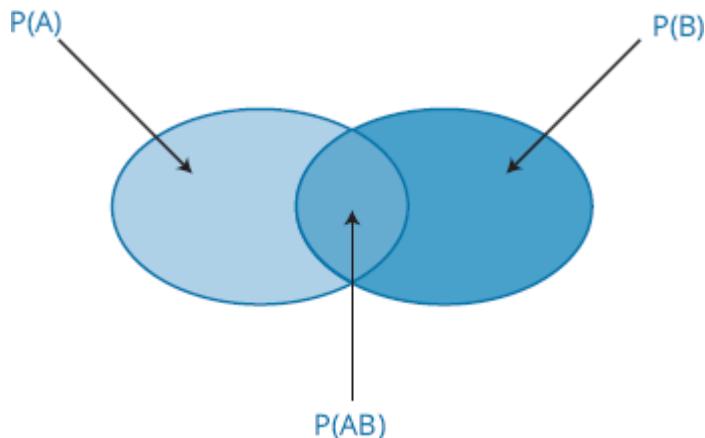


Figure 3.1 illustrates the addition rule with a Venn Diagram and highlights why the joint probability must be subtracted from the sum of the unconditional probabilities. Note that if the events are *mutually exclusive*, the sets do not intersect, $P(AB) = 0$, and the probability that one of the two events will occur is simply $P(A) + P(B)$.

EXAMPLE: Addition rule of probability

Using the information in our previous interest rate and recession example and the fact that the unconditional probability of a recession, $P(R)$, is 34%, determine the probability that either interest rates will increase *or* a recession will occur.

Answer:

Given that $P(R) = 0.34$, $P(I) = 0.40$, and $P(RI) = 0.28$, we can compute $P(R \text{ or } I)$ as follows:

$$P(R \text{ or } I) = P(R) + P(I) - P(RI)$$

$$P(R \text{ or } I) = 0.34 + 0.40 - 0.28$$

$$P(R \text{ or } I) = 0.46$$

Calculating a Joint Probability of Any Number of Independent Events

On the roll of two dice, the joint probability of getting two 4s is calculated as:

$$P(4 \text{ on first die and } 4 \text{ on second die}) = P(4 \text{ on first die}) \times P(4 \text{ on second die}) = 1/6 \times 1/6 = 1/36 = 0.0278$$

On the flip of two coins, the probability of getting two heads is:

$$P(\text{heads on first coin and heads on second coin}) = 1/2 \times 1/2 = 1/4 = 0.25$$

Hint: When dealing with *independent events*, the word *and* indicates multiplication, and the word *or* indicates addition. In probability notation:

$$P(A \text{ or } B) = P(A) + P(B) - P(AB), \text{ and } P(A \text{ and } B) = P(A) \times P(B)$$

The multiplication rule we used to calculate the joint probability of two independent events may be applied to any number of independent events, as the following example illustrates.

EXAMPLE: Joint probability for more than two independent events

What is the probability of rolling three 4s in one simultaneous toss of three dice?

Answer:

Since the probability of rolling a 4 for each die is $1/6$, the probability of rolling three 4s is:

$$P(\text{three 4s on the roll of three dice}) = 1/6 \times 1/6 \times 1/6 = 1/216 = 0.00463$$



MODULE QUIZ 3.1

1. An event that includes all of the possible outcomes is said to be:
 - A. random.
 - B. exclusive.
 - C. exhaustive.
2. Which of the following values *cannot* be the probability of an event?
 - A. 0.00.
 - B. 1.00.
 - C. 1.25.
3. The probability that the DJIA will increase tomorrow is $2/3$. The probability of an increase in the DJIA stated as odds is:
 - A. two-to-one.
 - B. one-to-three.
 - C. two-to-three.
4. The multiplication rule of probability determines the joint probability of two events as the product of:

- A. two conditional probabilities.
 - B. two unconditional probabilities.
 - C. a conditional probability and an unconditional probability.
5. If events A and B are mutually exclusive, then:
- A. $P(A | B) = P(A)$.
 - B. $P(AB) = P(A) \times P(B)$.
 - C. $P(A \text{ or } B) = P(A) + P(B)$.
6. Two mutually exclusive events:
- A. will both occur.
 - B. cannot both occur.
 - C. may both occur.
7. At a charity ball, 800 names are put into a hat. Four of the names are identical. On a random draw, what is the probability that one of these four names will be drawn?
- A. 0.004.
 - B. 0.005.
 - C. 0.010.

MODULE 3.2: CONDITIONAL EXPECTATIONS AND EXPECTED VALUE



Video covering
this content is
available online.

LOS 3.f: Compare and contrast dependent and independent events.

Independent events refer to events for which the occurrence of one has no influence on the occurrence of the others. The definition of independent events can be expressed in terms of conditional probabilities. Events A and B are independent if and only if:

$$P(A | B) = P(A), \text{ or equivalently, } P(B | A) = P(B)$$

If this condition is not satisfied, the events are dependent events (i.e., the occurrence of one is dependent on the occurrence of the other).

In our interest rate and recession example, recall that events I and R are not independent; the occurrence of I affects the probability of the occurrence of R. In this example, the independence conditions for I and R are violated because:

$$P(R) = 0.34, \text{ but } P(R | I) = 0.7; \text{ the probability of a recession is greater when there is an increase in interest rates.}$$

The best examples of independent events are found with the a priori probabilities of dice tosses or coin flips. A die has “no memory.” Therefore, the event of rolling a 4 on the second toss is independent of rolling a 4 on the first toss. This idea may be expressed as:

$$P(4 \text{ on second toss} | 4 \text{ on first toss}) = P(4 \text{ on second toss}) = 1/6 \text{ or } 0.167$$

Because the two events are independent, the conditional probability of a 4 on the second toss is the same as its unconditional probability.

The idea of independent events also applies to flips of a coin: $P(\text{heads on second toss} | \text{heads on first toss}) = P(\text{heads on second toss}) = 1/2 \text{ or } 0.50$

LOS 3.g: Calculate and interpret an unconditional probability using the total probability rule.

The **total probability rule** is used to determine the unconditional probability of an event, given conditional probabilities:

$$P(A) = P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots + P(A | B_N)P(B_N)$$

In general, the unconditional probability of event R, $P(R) = P(R | S_1) \times P(S_1) + P(R | S_2) \times P(S_2) + \dots + P(R | S_N) \times P(S_N)$, where the set of events $\{S_1, S_2, \dots, S_N\}$ is mutually exclusive and exhaustive.

EXAMPLE: An investment application of unconditional probability

Building upon our ongoing example about interest rates and economic recession, we can assume that a recession can only occur with either of the two events—interest rates increase (I) or interest rates do not increase (I^C)—since these events are mutually exclusive and exhaustive. I^C is read “the complement of I,” which means “not I.” Therefore, the probability of I^C is $1 - P(I)$. It is logical, therefore, that the sum of the two joint probabilities must be the unconditional probability of a recession. This can be expressed as follows:

$$P(R) = P(RI) + P(RI^C)$$

Applying the multiplication rule, we may restate this expression as:

$$P(R) = P(R | I) \times P(I) + P(R | I^C) \times P(I^C)$$

Assume that $P(R | I) = 0.70$, $P(R | I^C)$, the probability of recession if interest rates do not rise, is 10% and that $P(I) = 0.40$ so that $P(I^C) = 0.60$. The unconditional probability of a recession can be calculated as follows:

$$\begin{aligned} P(R) &= P(R | I) \times P(I) + P(R | I^C) \times P(I^C) \\ &= (0.70)(0.40) + (0.10)(0.60) \\ &= 0.28 + 0.06 = 0.34 \end{aligned}$$

LOS 3.h: Calculate and interpret the expected value, variance, and standard deviation of random variables.

The **expected value** of a random variable is the weighted average of the possible outcomes for the variable. The mathematical representation for the expected value of random variable X , that can take on any of the values from x_1 to x_n , is:

$$E(X) = \sum P(x_i)x_i = P(x_1)x_1 + P(x_2)x_2 + \dots + P(x_n)x_n$$

EXAMPLE: Expected earnings per share

The probability distribution of EPS for Ron's Stores is given in the figure below. Calculate the expected earnings per share.

EPS Probability Distribution

Probability	Earnings Per Share
10%	£1.80
20%	£1.60
40%	£1.20
30%	£1.00
100%	

Answer:

The expected EPS is simply a weighted average of each possible EPS, where the weights are the probabilities of each possible outcome.

$$E[\text{EPS}] = 0.10(1.80) + 0.20(1.60) + 0.40(1.20) + 0.30(1.00) = £1.28$$

Variance and *standard deviation* measure the dispersion of a random variable around its expected value, sometimes referred to as the **volatility** of a random variable. Variance can be calculated as the probability-weighted sum of the squared deviations from the mean (or expected value). The standard deviation is the positive square root of the variance. The following example illustrates the calculations for a probability model of possible returns.

EXAMPLE: Expected Value, Variance, and Standard Deviation from a probability model

Using the probabilities given in the table below, calculate the expected return on Stock A, the variance of returns on Stock A, and the standard deviation of returns on Stock A.

Event	Probability	R _A	Probability × R _A	R _A - E(R _A)	[R _A - E(R _A)] ²	Probability × [R _A - E(R _A)] ²
Boom	30%	20%	0.06	0.07	0.0049	0.00147
Normal	50%	12%	0.06	-0.01	0.0001	0.00005
Slow	20%	5%	0.01	-0.08	0.0064	0.00128
$E(R_A) =$				$\text{Var}(R_A) =$		
0.13				0.00280		

Answer:

$$E(R_A) = (0.30 \times 0.20) + (0.50 \times 0.12) + (0.20 \times 0.05) = 0.13 = 13\%$$

The expected return for Stock A is the probability-weighted sum of the returns under the three different economic scenarios.

In column 5, we have calculated the differences between the returns under each economic scenario and the expected return of 13%.

In column 6, we squared all the differences from column 5, and in the final column we have multiplied the probabilities of each economic scenario times the squared deviation of returns from the expected returns, and their sum, 0.00280, is the variance of R_A.

The standard deviation of $R_A = \sqrt{0.0028} = 0.0529$

Note that in a previous reading we estimated the standard deviation of a distribution from sample data, rather than from a probability model of returns. For the sample standard deviation, we divided the sum of the squared deviations from the mean by $n - 1$, where n was the size of the sample. Here we have no “ n ,” but we use the probability weights instead, as they describe the entire distribution of outcomes.

LOS 3.i: Explain the use of conditional expectation in investment applications.

Expected values or expected returns can be calculated using conditional probabilities. As the name implies, **conditional expected values** are contingent on the outcome of some other event. An analyst would use a conditional expected value to revise his expectations when new information arrives.

Consider the effect a tariff on steel imports might have on the returns of a domestic steel producer's stock. The stock's conditional expected return, given that the government imposes the tariff, will be higher than the conditional expected return if the tariff is not imposed.

Using the total probability rule, we can estimate the (unconditional) expected return on the stock as the sum of the expected return given no tariff, times the probability a tariff will not be enacted, and the expected return given a tariff, times the probability a tariff will be enacted.

LOS 3.j: Interpret a probability tree and demonstrate its application to investment problems.

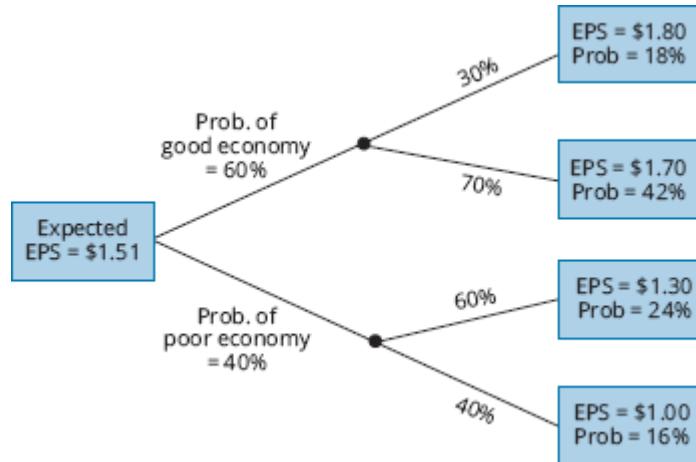
You might well wonder where the returns and probabilities used in calculating expected values come from. A general framework called a **probability tree** is used to show the probabilities of various outcomes. In Figure 3.2, we have shown estimates of EPS for four different events: (1) a good economy and relatively good results at the company, (2) a good economy and relatively poor results at the company, (3) a poor economy and relatively good results at the company, and (4) a poor economy and relatively poor results at the company. Using the rules of probability, we can calculate the probabilities of each of the four EPS outcomes shown in the boxes on the right-hand side of the “tree.”

The expected EPS of \$1.51 is simply calculated as:

$$0.18 \times 1.80 + 0.42 \times 1.70 + 0.24 \times 1.30 + 0.16 \times 1.00 = \$1.51$$

Note that the probabilities of the four possible outcomes sum to 1.

Figure 3.2: A Probability Tree



MODULE QUIZ 3.2

1. Two events are said to be independent if the occurrence of one event:
 - A. means that the second event cannot occur.
 - B. means that the second event is certain to occur.
 - C. does not affect the probability of the occurrence of the other event.
2. An analyst estimates that a share price has an 80% probability of increasing if economic growth exceeds 3%, a 40% probability of increasing if economic growth is between zero and 3%, and a 10% probability of increasing if economic growth is negative. If economic growth has a 25% probability of exceeding 3% and a 25% probability of being negative, what is the probability that the share price increases?
 - A. 22.5%.
 - B. 42.5%.
 - C. 62.5%.
3. $P(A|B) = 40\%$ and $P(B) = 30\%$ and $P(A) = 40\%$. It is *most likely* that:
 - A. A and B are dependent.
 - B. A and B are independent.
 - C. A and B are mutually exclusive.

MODULE 3.3: PORTFOLIO VARIANCE, BAYES, AND COUNTING PROBLEMS



Video covering this content is available online.

LOS 3.k: Calculate and interpret the expected value, variance, standard deviation, covariances, and correlations of portfolio returns.

Portfolio expected return. The expected return of a portfolio composed of n assets with weights, w_i , and expected returns, R_i , can be determined using the following formula:

$$E(R_p) = \sum_{i=1}^n w_i E(R_i) = w_1 E(R_1) + w_2 E(R_2) + \dots + w_n E(R_n)$$

The expected return and variance for a portfolio of assets can be determined using the properties of the individual assets in the portfolio. To do this, it is necessary to establish the portfolio weight for each asset. As indicated in the formula, the weight, w , of portfolio asset i is simply the market value currently invested in the asset divided by the current market value of the entire portfolio.

$$w_i = \frac{\text{market value of investment in asset } i}{\text{market value of the portfolio}}$$

In many finance situations, we are interested in how two random variables move in relation to each other. For investment applications, one of the most frequently analyzed pairs of random variables is the returns of two assets. Investors and managers frequently ask questions such as "what is the relationship between the return for Stock A and Stock B?" or "what is the relationship between the performance of the S&P 500 and that of the automotive industry?"

Covariance is a measure of how two assets move together. It is the expected value of the product of the deviations of the two random variables from their respective expected values. A common symbol for the covariance between random variables X and Y is $\text{Cov}(X,Y)$. Since we will be mostly concerned with the covariance of asset returns, the following formula has been written in terms of the covariance of the return of asset i , R_i , and the return of asset j , R_j :

$$\text{Cov}(R_i, R_j) = E\{[R_i - E(R_i)][R_j - E(R_j)]\}$$

The following are *properties of covariance*:

- The covariance of a random variable with itself is its variance of R_A ; that is, $\text{Cov}(R_A, R_A) = \text{Var}(R_A)$.
- Covariance may range from negative infinity to positive infinity.
- A positive covariance indicates that when one random variable is above its mean, the other random variable tends to be above its mean as well.
- A negative covariance indicates that when one random variable is above its mean, the other random variable tends to be below its mean.

The sample covariance for a sample of returns data can be calculated as:

$$s_{X,Y} = \frac{\sum_{i=1}^n \{ [R_{1,i} - \bar{R}_1] [R_{2,i} - \bar{R}_2] \}}{n - 1}$$

where:

$R_{1,i}$ = an observation of returns on asset 1

$R_{2,i}$ = an observation of returns on asset 2

\bar{R}_1 = mean return of asset 1

\bar{R}_2 = mean of asset 2

n = number of observations in the sample

A **covariance matrix** shows the covariances between returns on a group of assets.

Figure 3.3: Covariance matrix for three assets

Asset	A	B	C
A	$\text{Cov}(R_A, R_A)$	$\text{Cov}(R_A, R_B)$	$\text{Cov}(R_A, R_C)$
B	$\text{Cov}(R_B, R_A)$	$\text{Cov}(R_B, R_B)$	$\text{Cov}(R_B, R_C)$
C	$\text{Cov}(R_C, R_A)$	$\text{Cov}(R_C, R_B)$	$\text{Cov}(R_C, R_C)$

Note that the diagonal terms are the variances of each asset's returns, i.e., $\text{Cov}(R_A, R_A) = \text{Var}(R_A)$.

The covariance between the returns on two assets does not depend on order, i.e., $\text{Cov}(R_A, R_B) = \text{Cov}(R_B, R_A)$, so in this covariance matrix only 3 of the (off-diagonal) covariance terms are unique. In general for n assets, there are n variance terms (on the diagonal) and $n(n - 1)/2$ unique covariance terms.

Portfolio variance. To calculate the variance of portfolio returns, we use the asset weights, returns variances, and returns covariances.

$$\text{Var}(R_p) = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \text{Cov}(R_i, R_j)$$

The variance of a portfolio composed of two risky assets A and B can be expressed as:

$$\begin{aligned} \text{Var}(R_p) &= w_A w_A \text{Cov}(R_A, R_A) + w_A w_B \text{Cov}(R_A, R_B) + w_B w_A \text{Cov}(R_B, R_A) \\ &\quad + w_B w_B \text{Cov}(R_B, R_B) \end{aligned}$$

which we can write more simply as:

$$\text{Var}(R_p) = w_A^2 \text{Var}(R_A) + w_B^2 \text{Var}(R_B) + 2w_A w_B \text{Cov}(R_A, R_B), \text{ or}$$

$$\sigma_p^2 = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + 2w_A w_B \text{Cov}_{AB}.$$

For a 3-asset portfolio, the portfolio variance is:

$$\sigma_p^2 = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + w_C^2 \sigma_C^2 + 2w_A w_B \text{Cov}_{AB} + 2w_A w_C \text{Cov}_{AC} + 2w_B w_C \text{Cov}_{BC}$$

Consider a portfolio with three assets: an index of domestic stocks (60%), an index of domestic bonds (30%), and an index of international equities (10%). A covariance matrix of the three assets is shown below.

Figure 3.4: Covariance matrix for the three assets

Asset	Domestic Stocks	Domestic Bonds	International Equities
Domestic Stocks	400	44	180
Domestic Bonds	44	70	35
International Equities	180	35	450

Portfolio returns variance =

$$(0.6^2)400 + (0.3^2)70 + (0.1^2)450 + 2(0.6)(0.3)44 + 2(0.6)(0.1)180 + 2(0.3)(0.1)35 = 194.34$$

Portfolio returns standard deviation $\sqrt{194.34} = 13.94\%$.

Note that the units of variance and covariance are $\%^2$, i.e., 0.001. When we put these values in as whole numbers (in $\%^2$), the portfolio variance is in $\%^2$, and the standard deviation is in whole percentages. We could also put variance and covariance in as decimals and get both the portfolio returns variance and standard deviation as decimals.

From the formula for portfolio returns variance, we can see that the lower the covariance terms, the lower the portfolio variance (and standard deviation). This is true for positive values of covariance, as well as negative values.

Correlation

Recall that the correlation coefficient for two variables is $\rho_{AB} = \frac{\text{Cov}_{AB}}{\sigma_A \sigma_B}$, so that ($\text{Cov}_{AB} = \rho_{AB} \times \sigma_A \sigma_B$). This can be substituted for Cov_{AB} in our formula for portfolio returns variance. With this substitution we can use a correlation matrix to calculate portfolio returns variance, rather than using covariances.

Figure 3.5: Correlation matrix for the three assets

Asset	Domestic Stocks	Domestic Bonds	International Equities
Domestic Stocks	1.000	0.263	0.424
Domestic Bonds	0.263	1.000	0.197
International Equities	0.424	0.197	1.000

Note that the correlations of asset returns with themselves (the diagonal terms) are all 1.

LOS 3.I: Calculate and interpret the covariances of portfolio returns using the joint probability function.

EXAMPLE: Covariance of returns from a joint probability function

Assume that the economy can be in three possible states (S) next year: boom, normal, or slow economic growth. An expert source has calculated that $P(\text{boom}) = 0.30$, $P(\text{normal}) = 0.50$, and $P(\text{slow}) = 0.20$. The returns for Asset A, R_A , and Asset B, R_B , under each of the economic states are provided in the probability model as follows. What is the covariance of the returns for Asset A and Asset B?

Joint Probability Function

	$R_B = 30\%$	$R_B = 10\%$	$R_B = 0\%$
$R_A = 20\%$	0.30	0	0
$R_A = 12\%$	0	0.50	0
$R_A = 5\%$	0	0	0.20

The table gives us the joint probability of returns on Assets A and B, e.g., the probability that the return on Asset A is 20% and the return on Asset B is 30% is 30% and the probability that the return on Asset A is 12% and the return on Asset B is 10% is 50%.

Answer:

First, we must calculate the expected returns for each of the assets.

$$E(R_A) = (0.3)(0.20) + (0.5)(0.12) + (0.2)(0.05) = 0.13$$

$$E(R_B) = (0.3)(0.30) + (0.5)(0.10) + (0.2)(0.00) = 0.14$$

The covariance can now be computed using the procedure described in the following table.

Covariance Calculation

Probability	R _A	R _B	Probability × [R _A − E(R _A)] × [R _B − E(R _B)]
0.3	0.20	0.30	(0.3)(0.2 − 0.13)(0.3 − 0.14) = 0.00336
0.5	0.12	0.10	(0.5)(0.12 − 0.13)(0.1 − 0.14) = 0.00020
0.2	0.05	0.00	(0.2)(0.05 − 0.13)(0 − 0.14) = 0.00224

Covariance of returns for Asset A and Asset B is

$$0.00336 + 0.00020 + 0.00224 = 0.005784$$

LOS 3.m: Calculate and interpret an updated probability using Bayes' formula.

Bayes' formula is used to update a given set of prior probabilities for a given event in response to the arrival of new information. The rule for updating prior probability of an event is:

$$\text{updated probability} = \frac{\text{probability of new information for a given event}}{\text{unconditional probability of new information}} \times \text{prior probability of event}$$

We can derive Bayes' formula using the multiplication rule and noting that $P(AB) = P(BA)$.

$$P(B|A) \times P(A) = P(BA), \text{ and } P(A|B) \times P(B) = P(AB).$$

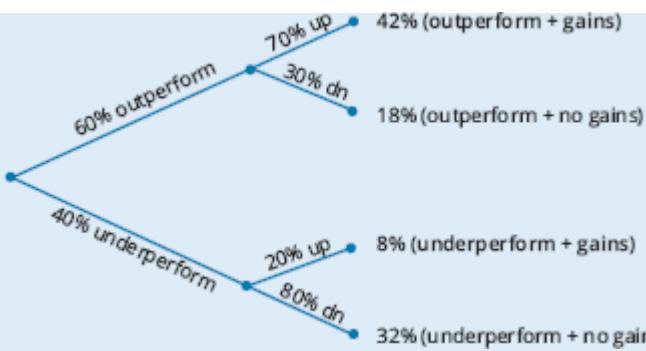
Because $P(BA) = P(AB)$, we can write $P(B|A) P(A) = P(A|B) P(B)$ and $\frac{P(B|A)P(A)}{P(B)}$ which equals $\frac{P(BA)}{P(B)}$ the joint probability of A and B, divided by the unconditional probability of B.

The following example illustrates the use of Bayes' formula. Note that A is outperform and A^C is underperform, $P(BA)$ is (outperform + gains), $P(A^C B)$ is (underperform + gains), and the unconditional probability $P(B)$ is $P(AB) + P(A^C B)$, by the total probability rule.

EXAMPLE: Bayes' formula

There is a 60% probability the economy will outperform, and if it does, there is a 70% probability a stock will go up and a 30% probability the stock will go down. There is a 40% probability the economy will underperform, and if it does, there is a 20% probability the stock in question will increase in value (have gains) and an 80% probability it will not. Given that the stock increased in value, calculate the probability that the economy outperformed.

Answer:



In the figure above, we have multiplied the probabilities to calculate the probabilities of each of the four outcome pairs. Note that these sum to 1. Given that the stock has gains, what is our updated probability of an outperforming economy? We sum the probability of stock gains in both states (outperform and underperform) to get $42\% + 8\% = 50\%$. Given that the stock has gains, the probability that the economy has outperformed is $\frac{42\%}{50\%} = 84\%$.

LOS 3.n: Identify the most appropriate method to solve a particular counting problem and analyze counting problems using factorial, combination, and permutation concepts.

Labeling refers to the situation where there are n items that can each receive one of k different labels. The number of items that receives label 1 is n_1 and the number that receive label 2 is n_2 , and so on, such that $n_1 + n_2 + n_3 + \dots + n_k = n$. The total number of ways that the labels can be assigned is:

$$\frac{n!}{(n_1!) \times (n_2!) \times \dots \times (n_k!)}$$

where:

the symbol “!” stands for **factorial**. For example, $4! = 4 \times 3 \times 2 \times 1 = 24$, and $2! = 2 \times 1 = 2$.

The general expression for n factorial is:

$$n! = n \times (n - 1) \times (n - 2) \times (n - 3) \times \dots \times 1, \text{ where by definition, } 0! = 1$$

Calculator help: On the TI, factorial is [2nd] [x!] (above the multiplication sign). To compute $4!$ on the TI, enter [4][2nd][x!] = 24.

EXAMPLE: Labeling

Consider a portfolio consisting of eight stocks. Your goal is to designate four of the stocks as “long-term holds,” three of the stocks as “short-term holds,” and one stock as “sell.” How many ways can these eight stocks be labeled?

Answer:

There are $8! = 40,320$ total possible sequences that can be followed to assign the three labels to the eight stocks. However, the order that each stock is assigned a label does not matter. For example, it does not matter which of the stocks labeled “long-term” is the first to be labeled. Thus, there are $4!$ ways to assign the long-term label. Continuing this reasoning to the other categories, there are $4! \times 3! \times 1!$ equivalent sequences for assigning the labels. To eliminate the

counting of these redundant sequences, the total number of possible sequences ($8!$) must be divided by the number of redundant sequences ($4! \times 3! \times 1!$).

Thus, the number of *different* ways to label the eight stocks is:

$$\frac{8!}{4! \times 3! \times 1!} = \frac{40,320}{24 \times 6 \times 1} = 280$$

If there are n labels ($k = n$), we have $\frac{n!}{1!} = n!$. The number of ways to assign n different labels to n items is simply $n!$.

A special case of labeling arises when the number of labels equals 2 ($k = 2$). That is, the n items can only be in one of two groups, and $n_1 + n_2 = n$. In this case, we can let $r = n_1$ and $n_2 = n - r$. Since there are only two categories, we usually talk about choosing r items. Then $(n - r)$ items are not chosen. The general formula for labeling when $k = 2$ is called the **combination formula** (or *binomial formula*) and is expressed as:

$${}_nC_r = \frac{n!}{(n - r)!r!}$$

where ${}_nC_r$ is the number of possible ways (combinations) of selecting r items from a set of n items when the order of selection is not important. This is also written $\binom{n}{r}$ and read “ n choose r ”.

Another useful formula is the **permutation formula**. A permutation is a specific ordering of a group of objects. The question of how many different groups of size r in specific order can be chosen from n objects is answered by the permutation formula. The number of permutations of r objects from n objects is:

$${}_nP_r = \frac{n!}{(n - r)!}$$

We will give an example using this formula shortly.



PROFESSOR'S NOTE

The combination formula ${}_nC_r$ and the permutation formula ${}_nP_r$ are both available on the TI calculator. To calculate the number of different groups of three stocks from a list of eight stocks (i.e., ${}_8C_3$), the sequence is 8 [2nd] $[{}_nC_r]$ 3 [=], which yields 56. If we want to know the number of differently ordered groups of three that can be selected from a list of eight, we enter 8 [2nd] $[{}_nP_r]$ 3 [=] to get 336, which is the number of permutations, $\frac{8!}{(8 - 3)!}$. This function is not available on the HP calculator. Remember, current policy permits you to bring both calculators to the exam, if you choose.

EXAMPLE: Number of choices in any order

How many ways can three stocks be sold from an 8-stock portfolio?

Answer:

Here we use the combination formula (or function on your calculator) as order does not matter. Note that this is equivalent to using the labeling formula with two labels, sold and not sold. Thus, the answer is:

$$\frac{8!}{5! \times 3!} = 56$$

In the preceding two examples, ordering did not matter. The order of selection could, however, be important. For example, suppose we want to liquidate only one stock position per week over the next three weeks. Once we choose three particular stocks to sell, the order in which they are sold must be determined. In this case, the concept of permutation comes into play.

The permutation formula implies that there are $r!$ *more* ways to choose r items if the order of selection is *important* than if order is not important. That is, $nP_r = r! \times nC_r$.

EXAMPLE: Permutation

How many ways are there to sell three stocks out of eight if the order of the sales is important?

Answer:

$$nP_r = {}_8P_3 = \frac{8!}{(8-3)!} = \frac{8!}{5!} = 336$$

This is $3!$ times the 56 possible combinations computed in the preceding example for selecting the three stocks when the order was not important.

Five guidelines may be used to determine which counting method to employ when dealing with counting problems:

- The *multiplication rule of counting* is used when there are *two or more groups*. The key is that only *one* item may be selected from each group. If there are k steps required to complete a task and each step can be done in n ways, the number of different ways to complete the task is $n_1 \times n_2 \times \dots \times n_k$.
- *Factorial* is used by itself when there are *no groups*—we are only arranging a given set of n items. Given n items, there are $n!$ ways of arranging them.
- The *labeling formula* applies to *three or more subgroups* of predetermined size. Each element of the entire group must be assigned a place, or label, in one of the three or more subgroups.
- The *combination formula* applies to *only two groups* of predetermined size. Look for the word “choose” or “combination.”
- The *permutation formula* applies to *only two groups* of predetermined size. Look for a specific reference to “order” being important.



MODULE QUIZ 3.3

1. Given the conditional probabilities in the table below and the unconditional probabilities $P(Y = 1) = 0.3$ and $P(Y = 2) = 0.7$, what is the expected value of X ?

x_i	$P(x_i Y = 1)$	$P(x_i Y = 2)$
0	0.2	0.1
5	0.4	0.8
10	0.4	0.1

A. 5.0.

- B. 5.3.
C. 5.7.
2. A discrete uniform distribution (each event has an equal probability of occurrence) has the following possible outcomes for X : [1, 2, 3, 4]. The variance of this distribution is *closest* to:
A. 1.00.
B. 1.25.
C. 2.00.
3. The correlation of returns between Stocks A and B is 0.50. The covariance between these two securities is 0.0043, and the standard deviation of the return of Stock B is 26%. The variance of returns for Stock A is:
A. 0.0011.
B. 0.0331.
C. 0.2656.
4. An analyst believes Davies Company has a 40% probability of earning more than \$2 per share. She estimates that the probability that Davies Company's credit rating will be upgraded is 70% if its earnings per share are greater than \$2 and 20% if its earnings per share are \$2 or less. Given the information that Davies Company's credit rating has been upgraded, what is the updated probability that its earnings per share are greater than \$2?
A. 50%.
B. 60%.
C. 70%.
5. Consider a universe of 10 bonds from which an investor will ultimately purchase six bonds for his portfolio. If the order in which he buys these bonds is not important, how many potential 6-bond combinations are there?
A. 7.
B. 210.
C. 5,040.
6. There are 10 sprinters in the finals of a race. How many different ways can the gold, silver, and bronze medals be awarded?
A. 120.
B. 720.
C. 1,440.

KEY CONCEPTS

LOS 3.a

A random variable is an uncertain value determined by chance.

An outcome is the realization of a random variable.

An event is a set of one or more outcomes. Two events that cannot both occur are termed "mutually exclusive," and a set of events that includes all possible outcomes is an "exhaustive" set of events.

LOS 3.b

The two properties of probability are as follows:

- The sum of the probabilities of all possible mutually exclusive events is 1.
- The probability of any event cannot be greater than 1 or less than 0.

A priori probability measures predetermined probabilities based on well-defined inputs; empirical probability measures probability from observations or experiments; and subjective

probability is an informed guess.

LOS 3.c

Probabilities can be stated as odds that an event will or will not occur. If the probability of an event is A out of B trials (A/B), the “odds for” are A to $(B - A)$ and the “odds against” are $(B - A)$ to A.

LOS 3.d

Unconditional probability (marginal probability) is the probability of an event occurring.

Conditional probability, $P(A | B)$, is the probability of an event A occurring given that event B has occurred.

LOS 3.e

The multiplication rule of probability is used to determine the joint probability of two events:

$$P(AB) = P(A | B) \times P(B)$$

The addition rule of probability is used to determine the probability that at least one of two events will occur:

$$P(A \text{ or } B) = P(A) + P(B) - P(AB)$$

The total probability rule is used to determine the unconditional probability of an event, given conditional probabilities:

$$P(A) = P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots + P(A | B_N)P(B_N)$$

where B_1, B_2, \dots, B_N is a mutually exclusive and exhaustive set of outcomes.

LOS 3.f

The probability of an independent event is unaffected by the occurrence of other events, but the probability of a dependent event is changed by the occurrence of another event. Events A and B are independent if and only if:

$$P(A | B) = P(A), \text{ or equivalently, } P(B | A) = P(B)$$

LOS 3.g

Using the total probability rule, the unconditional probability of A is the probability-weighted sum of the conditional probabilities:

$$P(A) = \sum_{i=1}^n [P_i(B_i)] \times P(A | B_i)$$

where B_i is a set of mutually exclusive and exhaustive events.

LOS 3.h

The expected value of a random variable is the weighted average of its possible outcomes:

$$E(X) = \sum P(x_i)x_i = P(x_1)x_1 + P(x_2)x_2 + \dots + P(x_n)x_n$$

Variance can be calculated as the probability-weighted sum of the squared deviations from the mean or expected value. The standard deviation is the positive square root of the variance.

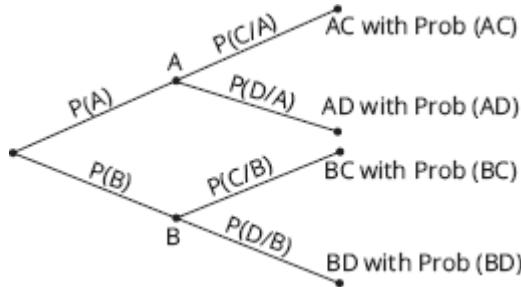
LOS 3.i

Conditional expected values depend on the outcome of some other event.

Forecasts of expected values for a stock's return, earnings, and dividends can be refined, using conditional expected values, when new information arrives that affects the expected outcome.

LOS 3.j

A probability tree shows the probabilities of two events and the conditional probabilities of two subsequent events.



LOS 3.k

The expected value of a random variable, $E(X)$, equals $\sum_{i=1}^n P_i(x_i)X_i$.

The variance of a random variable, $\text{Var}(X)$, equals

$$\sum_{i=1}^n P(X_i)[X_i - E(X)]^2 = \sigma_X^2$$

Standard deviation: $\sigma_X = \sqrt{\sigma_X^2}$.

The expected returns and variance of a 2-asset portfolio are given by:

$$E(R_p) = w_1 E(R_1) + w_2 E(R_2)$$

$$\begin{aligned} \text{Var}(R_p) &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \text{Cov}_{1,2} \\ &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_1 \sigma_2 \rho_{1,2} \end{aligned}$$

LOS 3.l

Given the joint probabilities for X_i and Y_j , i.e., $P(X_i Y_j)$, the covariance is calculated as:

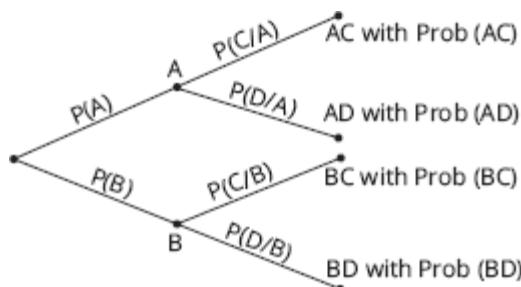
$$\sum_{i=1}^n P(X_i Y_i) [X_i - E(X)] [Y_i - E(Y)]$$

LOS 3.m

Bayes' formula for updating probabilities based on the occurrence of an event O is:

$$P(I|O) = \frac{P(O|I)}{P(O)} \times P(I)$$

Equivalently, based on the tree diagram below, $P(A|C) = \frac{P(AC)}{P(AC) + P(BC)}$



LOS 3.n

The number of ways to order n objects is n factorial, $n! = n \times (n - 1) \times (n - 2) \times \dots \times 1$.

There are $\frac{N!}{n_1! \times n_2! \times \dots \times n_k!}$ ways to assign k different labels to n items, where n_i is the number of items with the label i .

The number of ways to choose a subset of size r from a set of size n when order doesn't matter is $\frac{n!}{(n - r)!r!}$ combinations; when order matters, there are $\frac{n!}{(n - r)!}$ permutations.

ANSWERS TO MODULE QUIZ QUESTIONS

Module Quiz 3.1

1. **C** An event is said to be exhaustive if it includes all possible outcomes. (LOS 3.a)
2. **C** Probabilities may range from 0 (meaning no chance of occurrence) through 1 (which means a sure thing). (LOS 3.b)
3. **A** Odds for E = $P(E) / [1 - P(E)] = \frac{2/3}{1/3} = 2/1$ = two-to-one (LOS 3.c)
4. **C** By the multiplication rule of probability, the joint probability of two events, $P(AB)$, is the product of a conditional probability, $P(A | B)$, and an unconditional probability, $P(B)$. (LOS 3.d, LOS 3.e)
5. **C** There is no intersection of events when events are mutually exclusive. $P(A | B) = P(A) \times P(B)$ is only true for independent events. Note that since A and B are mutually exclusive (cannot both happen), $P(A | B)$ and $P(AB)$ must both be equal to zero. (LOS 3.a, LOS 3.d)
6. **B** One or the other may occur, but not both. (LOS 3.a)
7. **B** $P(\text{name 1 or name 2 or name 3 or name 4}) = 1/800 + 1/800 + 1/800 + 1/800 = 4/800 = 0.005$. (LOS 3.e)

Module Quiz 3.2

1. **C** Two events are said to be independent if the occurrence of one event does not affect the probability of the occurrence of the other event. (LOS 3.f)
2. **B** The three outcomes given for economic growth are mutually exclusive and exhaustive. The probability that economic growth is positive but less than 3% is 100% - 25% - 25% = 50%. Using the total probability rule, the probability that the share price increases is $(80\%)(25\%) + (40\%)(50\%) + (10\%)(25\%) = 42.5\%$. (LOS 3.g)
3. **B** From the values given, $P(A|B) = P(A)$, so A and B are independent. $P(A|B) \times P(B) = P(AB) = 12\%$, so A and B are not mutually exclusive (if they were $P(AB)$ would equal 0). (LOS 3.g)

Module Quiz 3.3

1. **B** $E(X | Y = 1) = (0.2)(0) + (0.4)(5) + (0.4)(10) = 6$

$$E(X | Y = 2) = (0.1)(0) + (0.8)(5) + (0.1)(10) = 5$$

$$E(X) = (0.3)(6) + (0.7)(5) = 5.30$$

(LOS 3.k)

2. **B** Expected value = $(1/4)(1 + 2 + 3 + 4) = 2.5$

$$\text{Variance} = (1/4)[(1 - 2.5)^2 + (2 - 2.5)^2 + (3 - 2.5)^2 + (4 - 2.5)^2] = 1.25$$

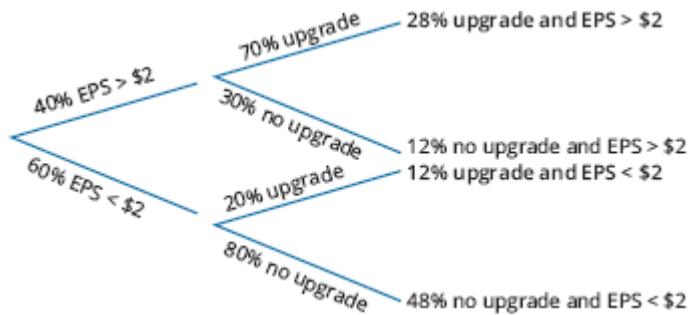
Note that since each observation is equally likely, each has 25% (1/4) chance of occurrence. (LOS 3.k)

3. A



(LOS 3.k)

4. C This is an application of Bayes' formula. As the tree diagram below shows, the updated probability that earnings per share are greater than \$2 is $\frac{28\%}{28\% + 12\%} = 70\%$



(LOS 3.m)

5. B

$${}_n C_r = \frac{n!}{(n-r)!r!} = {}_{10} C_6 = \frac{10!}{(10-6)!6!} = \frac{10!}{4!6!} = 210. \text{ (LOS 3.n)}$$

6. B Since the order of the top-three finishers matters, we need to use the permutation formula.

$${}_{10} P_3 = \frac{10!}{(10-3)!} = 720$$

(LOS 3.n)

READING 4

COMMON PROBABILITY DISTRIBUTIONS

EXAM FOCUS

This reading contains a lot of key material. Learn the difference between discrete and continuous probability distributions. The binomial and normal distributions are the most important here. You must learn the properties of both distributions and memorize the formula for the probability of a particular value when given a binomial probability distribution. Learn what shortfall risk is and how to calculate and use Roy's safety-first criterion. Know how to standardize a normally distributed random variable, use a z-table, determine probabilities using a cumulative distribution function, and construct confidence intervals. Learn the critical values for the often-used confidence intervals. You will use these skills repeatedly in the readings that follow. Additionally, understand the basic features of the lognormal distribution and Monte Carlo simulation. Candidates should know how to get continuously compounded rates of return from holding period returns. The chi-square and F-distribution are introduced here and we will use them in our reading on hypothesis testing.

MODULE 4.1: UNIFORM AND BINOMIAL DISTRIBUTIONS



Video covering this content is available online.

LOS 4.a: Define a probability distribution and compare and contrast discrete and continuous random variables and their probability functions.

A **probability distribution** describes the probabilities of all the possible outcomes for a random variable. The probabilities of all possible outcomes must sum to 1. A simple probability distribution is that for the roll of one fair die, there are six possible outcomes and each one has a probability of 1/6, so they sum to 1. The probability distribution of all the possible returns on the S&P 500 Index for the next year is a more complex version of the same idea.

A **discrete random variable** is one for which the number of possible outcomes can be counted, and for each possible outcome, there is a measurable and positive probability. An example of a discrete random variable is the number of days it will rain in a given month, because there is a countable number of possible outcomes, ranging from zero to the number of days in the month.

A **probability function**, denoted $p(x)$, specifies the probability that a random variable is equal to a specific value. More formally, $p(x)$ is the probability that random variable X takes on the value x , or $p(x) = P(X = x)$.

The two key properties of a probability function are:

- $0 \leq p(x) \leq 1$.
- $\sum p(x) = 1$, the sum of the probabilities for *all* possible outcomes, x , for a random variable, X , equals 1.

EXAMPLE: Evaluating a probability function

Consider the following function: $X = \{1, 2, 3, 4\}$, $p(x) = \frac{x}{10}$, else $p(x) = 0$

Determine whether this function satisfies the conditions for a probability function.

Answer:

Note that all of the probabilities are between 0 and 1, and the sum of all probabilities equals 1:

$$\sum p(x) = \frac{1}{10} + \frac{2}{10} + \frac{3}{10} + \frac{4}{10} = 0.1 + 0.2 + 0.3 + 0.4 = 1$$

Both conditions for a probability function are satisfied.

A **continuous random variable** is one for which the number of possible outcomes is infinite, even if lower and upper bounds exist. The actual amount of daily rainfall between zero and 100 inches is an example of a continuous random variable because the actual amount of rainfall can take on an infinite number of values. Daily rainfall can be measured in inches, half inches, quarter inches, thousandths of inches, or even smaller increments. Thus, the number of possible daily rainfall amounts between zero and 100 inches is essentially infinite.

The assignment of probabilities to the possible outcomes for discrete and continuous random variables provides us with discrete probability distributions and continuous probability distributions. The difference between these types of distributions is most apparent for the following properties:

- For a *discrete distribution*, $p(x) = 0$ when x cannot occur, or $p(x) > 0$ if it can. Recall that $p(x)$ is read: “the probability that random variable $X = x$.” For example, the probability of it raining on 33 days in June is zero because this cannot occur, but the probability of it raining 25 days in June has some positive value.
- For a *continuous distribution*, $p(x) = 0$ even though x can occur. We can only consider $P(x_1 \leq X \leq x_2)$ where x_1 and x_2 are actual numbers. For example, the probability of receiving 2 inches of rain in June is zero because 2 inches is a single point in an infinite range of possible values. On the other hand, the probability of the amount of rain being between 1.99999999 and 2.00000001 inches has some positive value. In the case of continuous distributions, $P(x_1 \leq X \leq x_2) = P(x_1 < X < x_2)$ because $p(x_1) = p(x_2) = 0$.

In finance, some discrete distributions are treated as though they are continuous because the number of possible outcomes is very large. For example, the increase or decrease in the price of a stock traded on an American exchange is recorded in dollars and cents. Yet, the probability of a change of exactly \$1.33 or \$1.34 or any other specific change is almost zero. It is customary, therefore, to speak in terms of the probability of a range of possible price change, say between \$1.00 and \$2.00. In other words $p(\text{price change} = 1.33)$ is essentially zero, but $p(1 < \text{price change} < 2)$ is greater than zero.

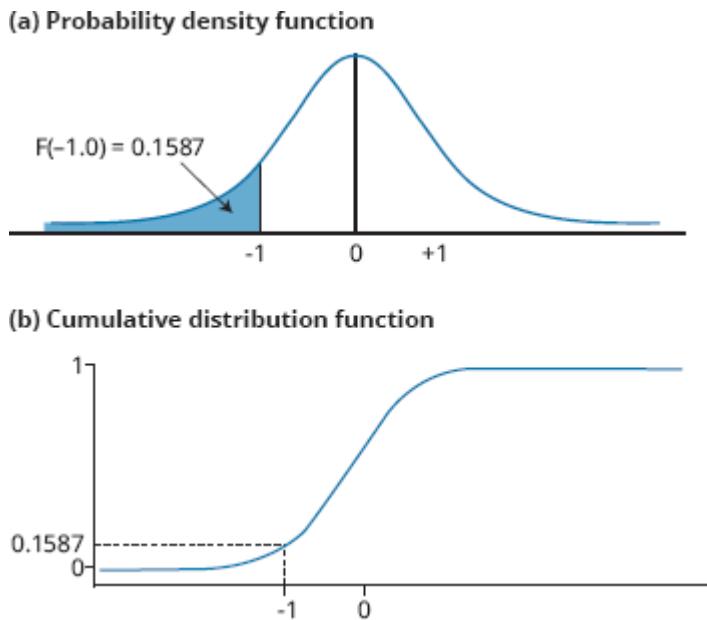
LOS 4.b: Calculate and interpret probabilities for a random variable given its cumulative distribution function.

A **cumulative distribution function (cdf)**, or simply *distribution function*, defines the probability that a random variable, X , takes on a value equal to or less than a specific value, x . It represents the sum, or *cumulative value*, of the probabilities for the outcomes up to and including a specified outcome. The cumulative distribution function for a random variable, X , may be expressed as $F(x) = P(X \leq x)$.

Consider the probability function defined earlier for $X = \{1, 2, 3, 4\}$, $p(x) = x / 10$. For this distribution, $F(3) = 0.6 = 0.1 + 0.2 + 0.3$, and $F(4) = 1 = 0.1 + 0.2 + 0.3 + 0.4$. This means that $F(3)$ is the cumulative probability that outcomes 1, 2, or 3 occur, and $F(4)$ is the cumulative probability that one of the possible outcomes occurs.

Figure 4.1 shows an example of a cumulative distribution function (for a standard normal distribution, described later in this topic). There is a 15.87% probability of a value less than -1 . This is the total area to the left of -1 in the pdf in Panel (a), and the y-axis value of the cdf for a value of -1 in Panel (b).

Figure 4.1: Standard Normal Probability Density and Cumulative Distribution Functions



EXAMPLE: Cumulative distribution function

Return on equity for a firm is defined as a continuous distribution over the range from -20% to $+30\%$ and has a cumulative distribution function of $F(x) = (x + 20) / 50$. Calculate the probability that ROE will be between 0% and 15% .

Answer:

To determine the probability that ROE will be between 0% and 15% , we can first calculate the probability that ROE will be less than or equal to 15% , or $F(15)$, and then subtract the probability that ROE will be less than zero, or $F(0)$.

$$P(0 \leq x \leq 15) = F(15) - F(0)$$

$$F(15) = (15 + 20) / 50 = 0.70$$

$$F(0) = (0 + 20) / 50 = 0.40$$

$$F(15) - F(0) = 0.70 - 0.40 = 0.30 = 30\%$$

LOS 4.c: Describe the properties of a discrete uniform random variable, and calculate and interpret probabilities given the discrete uniform distribution function.

A **discrete uniform random variable** is one for which the probabilities for all possible outcomes for a discrete random variable are equal. For example, consider the *discrete uniform probability distribution* defined as $X = \{1, 2, 3, 4, 5\}$, $p(x) = 0.2$. Here, the probability for each outcome is equal to 0.2 [i.e., $p(1) = p(2) = p(3) = p(4) = p(5) = 0.2$]. Also, the cumulative distribution function for the n th outcome, $F(x_n) = np(x)$, and the probability for a range of outcomes is $p(x)k$, where k is the number of possible outcomes in the range.

EXAMPLE: Discrete uniform distribution

Determine $p(6)$, $F(6)$, and $P(2 \leq X \leq 8)$ for the discrete uniform distribution function defined as:

$$X = \{2, 4, 6, 8, 10\}, p(x) = 0.2$$

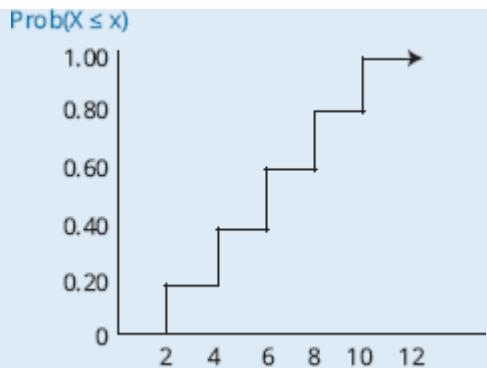
Answer:

$p(6) = 0.2$, since $p(x) = 0.2$ for all x . $F(6) = P(X \leq 6) = np(x) = 3(0.2) = 0.6$. Note that $n = 3$ since 6 is the third outcome in the range of possible outcomes. $P(2 \leq X \leq 8) = 4(0.2) = 0.8$. Note that $k = 4$, since there are four outcomes in the range $2 \leq X \leq 8$. The following figures illustrate the concepts of a probability function and cumulative distribution function for this distribution.

Probability and Cumulative Distribution Functions

$X = x$	Probability of x Prob ($X = x$)	Cumulative Distribution Function Prob ($X < x$)
2	0.20	0.20
4	0.20	0.40
6	0.20	0.60
8	0.20	0.80

Cumulative Distribution Function for $X \sim \text{Uniform} \{2, 4, 6, 8, 10\}$



LOS 4.d: Describe the properties of the continuous uniform distribution, and calculate and interpret probabilities given a continuous uniform distribution.

The **continuous uniform distribution** is defined over a range that spans between some lower limit, a , and some upper limit, b , which serve as the parameters of the distribution. Outcomes can only occur between a and b , and since we are dealing with a continuous distribution, even if $a < x < b$, $P(X = x) = 0$. Formally, the properties of a continuous uniform distribution may be described as follows:

- For all $a \leq x_1 < x_2 \leq b$ (i.e., for all x_1 and x_2 between the boundaries a and b).
- $P(X < a \text{ or } X > b) = 0$ (i.e., the probability of X outside the boundaries is zero).
- $P(x_1 \leq X \leq x_2) = (x_2 - x_1)/(b - a)$. This defines the probability of outcomes between x_1 and x_2 .

Don't miss how simple this is just because the notation is so mathematical. For a continuous uniform distribution, the probability of outcomes in a range that is one-half the whole range is 50%. The probability of outcomes in a range that is one-quarter as large as the whole possible range is 25%.

EXAMPLE: Continuous uniform distribution

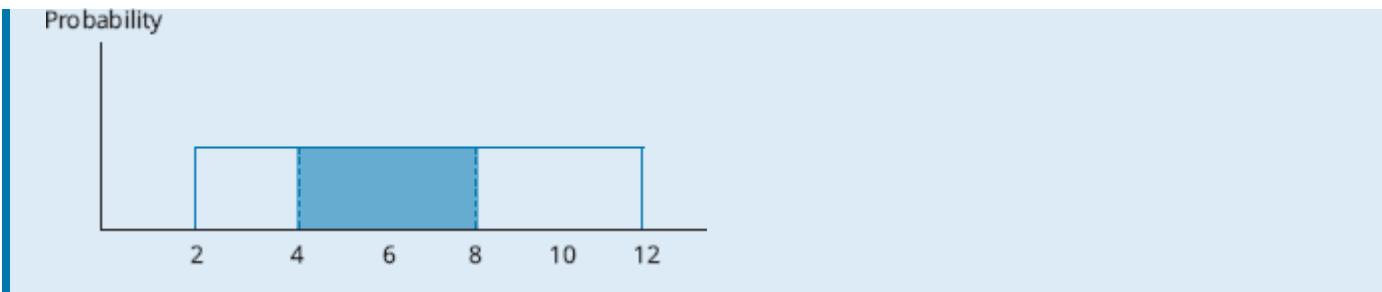
X is uniformly distributed between 2 and 12. **Calculate** the probability that X will be between 4 and 8.

Answer:

$$\frac{8 - 4}{12 - 2} = \frac{4}{10} = 40\%$$

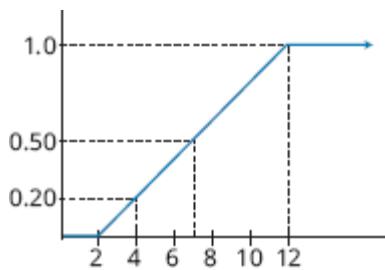
The figure below illustrates this continuous uniform distribution. Note that the area bounded by 4 and 8 is 40% of the total probability between 2 and 12 (which is 100%).

Continuous Uniform Distribution



Because outcomes are equal over equal-size possible intervals, the cdf is linear over the variable's range. The cdf for the distribution in the example, $\text{Prob}(X < x)$, is shown in Figure 4.2.

Figure 4.2: CDF for a Continuous Uniform Variable



LOS 4.e: Describe the properties of a Bernoulli random variable and a binomial random variable, and calculate and interpret probabilities given the binomial distribution function.

A **binomial random variable** may be defined as the number of “successes” in a given number of trials, whereby the outcome can be either “success” or “failure.” The probability of success, p , is constant for each trial, and the trials are independent. A binomial random variable for which the number of trials is 1 is called a **Bernoulli random variable**. Think of a trial as a mini-experiment (or “Bernoulli trial”). The final outcome is the number of successes in a series of n trials. Under these conditions, the binomial probability function defines the probability of x successes in n trials. It can be expressed using the following formula:

$$p(x) = P(X = x) = (\text{number of ways to choose } x \text{ from } n)p^x(1 - p)^{n-x}$$

where:

(number of ways to choose x from n) =

$$\frac{n!}{(n - x)!x!} \text{ which may also be denoted as } \binom{n}{x} \text{ or stated as "n choose } x\text{"}$$

p = the probability of “success” on each trial [don’t confuse it with $p(x)$]

So the probability of exactly x successes in n trials is:

$$p(x) = \frac{n!}{(n - x)!x!} p^x(1 - p)^{n-x}$$

EXAMPLE: Binomial probability

Assuming a binomial distribution, compute the probability of drawing three black beans from a bowl of black and white beans if the probability of selecting a black bean in any given attempt is 0.6. You will draw five beans from the bowl.

Answer:

$$P(X = 3) = p(3) = \frac{5!}{2!3!}(0.6)^3(0.4)^2 = (120 / 12)(0.216)(0.160) = 0.3456$$

Some intuition about these results may help you remember the calculations. Consider that a (very large) bowl of black and white beans has 60% black beans and that each time you select a bean, you replace it in the bowl before drawing again. We want to know the probability of selecting exactly three black beans in five draws, as in the previous example.

One way this might happen is BBBWW. Since the draws are independent, the probability of this is easy to calculate. The probability of drawing a black bean is 60%, and the probability of drawing a white bean is $1 - 60\% = 40\%$. Therefore, the probability of selecting BBBWW, in order, is $0.6 \times 0.6 \times 0.6 \times 0.4 \times 0.4 = 3.456\%$. This is the $p^3(1 - p)^2$ from the formula and p is 60%, the probability of selecting a black bean on any single draw from the bowl. BBBWW is not, however, the only way to choose exactly three black beans in five trials. Another possibility is BBWWB, and a third is BWWBB. Each of these will have exactly the same probability of occurring as our initial outcome, BBBWW. That's why we need to answer the question of how many ways (different orders) there are for us to choose

three black beans in five draws. Using the formula, there are $\frac{5!}{3!(5 - 3)!} = 10$ ways; $10 \times 3.456\% = 34.56\%$, the answer we computed above.

Expected Value and Variance of a Binomial Random Variable

For a given series of n trials, the expected number of successes, or $E(X)$, is given by the following formula:

$$\text{expected value of } X = E(X) = np$$

The intuition is straightforward; if we perform n trials and the probability of success on each trial is p , we expect np successes.

The variance of a binomial random variable is given by:

$$\text{variance of } X = np(1 - p)$$

EXAMPLE: Expected value of a binomial random variable

Based on empirical data, the probability that the Dow Jones Industrial Average (DJIA) will increase on any given day has been determined to equal 0.67. Assuming that the only other outcome is that it decreases, we can state $p(UP) = 0.67$ and $p(DOWN) = 0.33$. Further, assume that movements in the DJIA are independent (i.e., an increase in one day is independent of what happened on another day).

Using the information provided, compute the expected value of the number of up days in a 5-day period.

Answer:

Using binomial terminology, we define success as UP, so $p = 0.67$. Note that the definition of success is critical to any binomial problem.

$$E(X | n = 5, p = 0.67) = (5)(0.67) = 3.35$$

Recall that the “|” symbol means *given*. Hence, the preceding statement is read as: the expected value of X given that $n = 5$, and the probability of success = 67% is 3.35.

We should note that since the binomial distribution is a discrete distribution, the result $X = 3.35$ is not possible. However, if we were to record the results of many 5-day periods, the average number of up days (successes) would converge to 3.35.



MODULE QUIZ 4.1

1. Which of the following is *least likely* an example of a discrete random variable?
 - A. The number of stocks a person owns.
 - B. The time spent by a portfolio manager with a client.
 - C. The number of days it rains in a month in Iowa City.
2. For a continuous random variable X , the probability of any single value of X is:
 - A. one.
 - B. zero.
 - C. determined by the cdf.
3. Which of the following is *least likely* a probability distribution?
 - A. $X = [1, 2, 3, 4]$; $\text{Prob}[X_i] = \frac{x_i^2}{30}$.
 - B. $X = [5, 10]$; $\text{Prob}[X_i] = \frac{8 - x_i}{5}$.
 - C. $X = [5, 10]$; $\text{Prob}[X_i] = \frac{x_i - 3}{9}$.

Use the following table to answer Questions 4 through 7.

Probability distribution of a discrete random variable X								
X	0	1	2	3	4	5	6	7
$P(X)$	0.04	0.11	0.18	0.24	0.14	0.17	0.09	0.03

4. The cdf of 5, or $F(5)$ is:
 - A. 0.17.
 - B. 0.71.
 - C. 0.88.
5. The probability that X is *greater than* 3 is:
 - A. 0.24.
 - B. 0.43.
 - C. 0.67.
6. What is $P(2 \leq X \leq 5)$?
 - A. 0.17.
 - B. 0.38.
 - C. 0.73.
7. The expected value of the random variable X is:
 - A. 3.35.
 - B. 3.70.
 - C. 5.47.
8. A continuous uniform distribution has the parameters $a = 4$ and $b = 10$. The $F(20)$ is:
 - A. 0.25.

- B. 0.50.
C. 1.00.
9. Which of the following is *least likely* a condition of a binomial experiment?
A. There are only two trials.
B. The trials are independent.
C. If p is the probability of success, and q is the probability of failure, then $p + q = 1$.
10. Which of the following statements *least accurately* describes the binomial distribution?
A. It is a discrete distribution.
B. The probability of an outcome of zero is zero.
C. The combination formula is used in computing probabilities.
11. A recent study indicated that 60% of all businesses have a fax machine. From the binomial probability distribution table, the probability that exactly four businesses will have a fax machine in a random selection of six businesses is:
A. 0.138.
B. 0.276.
C. 0.311.
12. Ten percent of all college graduates hired stay with the same company for more than five years. In a random sample of six recently hired college graduates, the probability that exactly two will stay with the same company for more than five years is *closest* to:
A. 0.098.
B. 0.114.
C. 0.185.
13. Assume that 40% of candidates who sit for the CFA[®] examination pass it the first time. Of a random sample of 15 candidates who are sitting for the exam for the first time, what is the expected number of candidates that will pass?
A. 0.375.
B. 4.000.
C. 6.000.

MODULE 4.2: NORMAL DISTRIBUTIONS



LOS 4.f: Explain the key properties of the normal distribution.

Video covering
this content is
available online.

The normal distribution is important for many reasons. Many of the random variables that are relevant to finance and other professional disciplines follow a normal distribution. In the area of investment and portfolio management, the normal distribution plays a central role in portfolio theory.

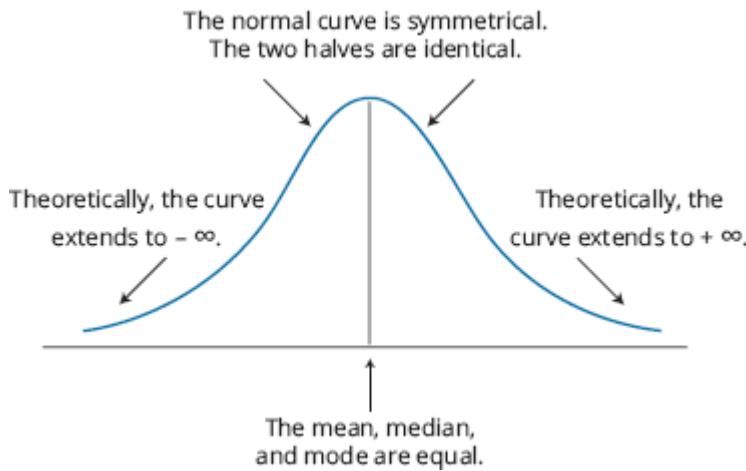
The **normal distribution** has the following key properties:

- It is completely described by its mean, μ , and variance, σ^2 , stated as $X \sim N(\mu, \sigma^2)$. In words, this says that “ X is normally distributed with mean μ and variance σ^2 .”
- Skewness = 0, meaning that the normal distribution is symmetric about its mean, so that $P(X \leq \mu) = P(\mu \leq X) = 0.5$, and mean = median = mode.
- Kurtosis = 3; this is a measure of how flat the distribution is. Recall that excess kurtosis is measured relative to 3, the kurtosis of the normal distribution.
- A linear combination of normally distributed random variables is also normally distributed.

- The probabilities of outcomes farther above and below the mean get smaller and smaller but do not go to zero (the tails get very thin but extend infinitely).

Many of these properties are evident from examining the graph of a normal distribution's probability density function as illustrated in Figure 4.3.

Figure 4.3: Normal Distribution Probability Density Function



LOS 4.g: Contrast a multivariate distribution and a univariate distribution, and explain the role of correlation in the multivariate normal distribution.

Up to this point, our discussion has been strictly focused on **univariate distributions** (i.e., the distribution of a single random variable). In practice, however, the relationships between two or more random variables are often relevant. For instance, investors and investment managers are frequently interested in the interrelationship among the returns of one or more assets. In fact, as you will see in your study of asset pricing models and modern portfolio theory, the return on a given stock and the return on the S&P 500 or some other market index will have special significance. Regardless of the specific variables, the simultaneous analysis of two or more random variables requires an understanding of multivariate distributions.

A **multivariate distribution** specifies the probabilities associated with a group of random variables and is meaningful only when the behavior of each random variable in the group is in some way dependent on the behavior of the others. Both discrete and continuous random variables can have multivariate distributions. Multivariate distributions between two discrete random variables are described using joint probability tables. For continuous random variables, a multivariate *normal* distribution may be used to describe them if all the individual variables follow a normal distribution. As previously mentioned, one of the characteristics of a normal distribution is that a linear combination of normally distributed random variables is normally distributed as well. For example, if the return of each stock in a portfolio is normally distributed, the return on the portfolio will also be normally distributed.

The Role of Correlation in the Multivariate Normal Distribution

Similar to a univariate normal distribution, a multivariate normal distribution can be described by the mean and variance of the individual random variables. Additionally, it is necessary to

specify the correlation between the individual pairs of variables when describing a multivariate distribution. Correlation is the feature that distinguishes a multivariate distribution from a univariate normal distribution. Correlation indicates the strength of the linear relationship between a pair of random variables.

Using asset returns as our random variables, the multivariate normal distribution for the returns on n assets can be completely defined by the following three sets of parameters:

- n means of the n series of returns ($\mu_1, \mu_2, \dots, \mu_n$).
- n variances of the n series of returns ($\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$).
- $0.5n(n - 1)$ pair-wise correlations.

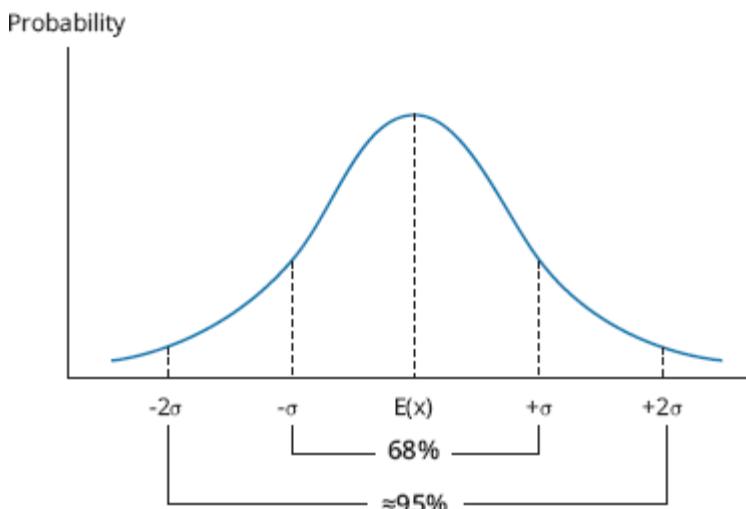
For example, if there are two assets, $n = 2$, then the multivariate returns distribution can be described with two means, two variances, and one correlation [$0.5(2)(2 - 1) = 1$]. If there are four assets, $n = 4$, the multivariate distribution can be described with four means, four variances, and six correlations [$0.5(4)(4 - 1) = 6$]. When building a portfolio of assets, all other things being equal, it is desirable to combine assets having low returns correlation because this will result in a portfolio with a lower variance than one composed of assets with higher correlations.

LOS 4.h: Calculate the probability that a normally distributed random variable lies inside a given interval.

A **confidence interval** is a range of values around the expected outcome within which we expect the actual outcome to be some specified percentage of the time. A 95% confidence interval is a range that we expect the random variable to be in 95% of the time. For a normal distribution, this interval is based on the expected value (sometimes called a point estimate) of the random variable and on its variability, which we measure with standard deviation.

Confidence intervals for a normal distribution are illustrated in Figure 4.4. For any normally distributed random variable, 68% of the outcomes are within one standard deviation of the expected value (mean), and approximately 95% of the outcomes are within two standard deviations of the expected value.

Figure 4.4: Confidence Intervals for a Normal Distribution



In practice, we will not know the actual values for the mean and standard deviation of the distribution, but will have estimated them as \bar{X} and s . The three confidence intervals of most interest are given by the following:

- The 90% confidence interval for \bar{X} is $\bar{X} - 1.65s$ to $\bar{X} + 1.65s$.
- The 95% confidence interval for \bar{X} is $\bar{X} - 1.96s$ to $\bar{X} + 1.96s$.
- The 99% confidence interval for \bar{X} is $\bar{X} - 2.58s$ to $\bar{X} + 2.58s$.

EXAMPLE: Confidence intervals

The average return of a mutual fund is 10.5% per year and the standard deviation of annual returns is 18%. If returns are approximately normal, what is the 95% confidence interval for the mutual fund return next year?

Answer:

Here μ and σ are 10.5% and 18%, respectively. Thus, the 95% confidence interval for the return, R , is:

$$10.5 \pm 1.96(18) = -24.78\% \text{ to } 45.78\%$$

Symbolically, this result can be expressed as:

$$P(-24.78 < R < 45.78) = 0.95 \text{ or } 95\%$$

The interpretation is that the annual return is expected to be within this interval 95% of the time, or 95 out of 100 years.

LOS 4.i: Explain how to standardize a random variable.

The **standard normal distribution** is a normal distribution that has been standardized so that it has a mean of zero and a standard deviation of 1 [i.e., $N \sim (0,1)$]. To standardize an observation from a given normal distribution, the *z-value* of the observation must be calculated. The z-value represents the number of standard deviations a given observation is from the population mean. *Standardization* is the process of converting an observed value for a random variable to its z-value. The following formula is used to standardize a random variable:

$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}} = \frac{x - \mu}{\sigma}$$



PROFESSOR'S NOTE

The term *z-value* will be used for a standardized observation in this document. The terms *z-score* and *z-statistic* are also commonly used.

EXAMPLE: Standardizing a random variable (calculating z-values)

Assume that the annual earnings per share (EPS) for a population of firms are normally distributed with a mean of \$6 and a standard deviation of \$2.

What are the z-values for EPS of \$2 and \$8?

Answer:

If $\text{EPS} = x = \$8$, then $z = (x - \mu) / \sigma = (\$8 - \$6) / \$2 = +1$

If $\text{EPS} = x = \$2$, then $z = (x - \mu) / \sigma = (\$2 - \$6) / \$2 = -2$

Here, $z = +1$ indicates that an EPS of \$8 is one standard deviation above the mean, and $z = -2$ means that an EPS of \$2 is two standard deviations below the mean.

LOS 4.j: Calculate and interpret probabilities using the standard normal distribution.

Now we will show how to use standardized values (z-values) and a table of probabilities for Z to determine probabilities. A portion of a table of the cumulative distribution function for a standard normal distribution is shown in Figure 4.5. We will refer to this table as the z-table, as it contains values generated using the cumulative distribution function for a standard normal distribution, denoted by $F(Z)$. Thus, the values in the z-table are the probabilities of observing a z-value that is less than a given value, z [i.e., $P(Z < z)$]. The numbers in the first column are z-values that have only one decimal place. The columns to the right supply probabilities for z-values with two decimal places.

Note that the z-table in Figure 4.5 only provides probabilities for positive z-values. This is not a problem because we know from the symmetry of the standard normal distribution that $F(-Z) = 1 - F(Z)$. The tables in the back of many texts actually provide probabilities for negative z-values, but we will work with only the positive portion of the table because this may be all you get on the exam. In Figure 4.5, we can find the probability that a standard normal random variable will be less than 1.66, for example. The table value is 95.15%. The probability that the random variable will be less than -1.66 is simply $1 - 0.9515 = 0.0485 = 4.85\%$, which is also the probability that the variable will be greater than +1.66.

Figure 4.5: Cumulative Probabilities for a Standard Normal Distribution

Cdf Values for the Standard Normal Distribution: The z-Table

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.5	.6915	Please note that several of the rows have been deleted to save space.*								
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

*A complete cumulative standard normal table is included in Appendix A.

PROFESSOR'S NOTE

 When you use the standard normal probabilities, you have formulated the problem in terms of standard deviations from the mean. Consider a security with returns that are approximately normal, an expected return of 10%, and standard deviation of returns of 12%. The probability of returns greater than 30% is calculated based on the number of standard deviations that 30% is above the expected return of 10%. 30% is 20% above the expected return of 10%, which is $20 / 12 = 1.67$ standard deviations above the mean. We look up the probability of returns less than 1.67 standard deviations above the mean (0.9525 or 95.25% from Figure 4.5) and calculate the probability of returns more than 1.67 standard deviations above the mean as $1 - 0.9525 = 4.75\%$.

EXAMPLE: Using the z-table (1)

Considering again EPS distributed with $\mu = \$6$ and $\sigma = \$2$, what is the probability that EPS will be \$9.70 or more?

Answer:

Here we want to know $P(\text{EPS} > \$9.70)$, which is the area under the curve to the right of the z-value corresponding to $\text{EPS} = \$9.70$ (see the following figure).

The z-value for $\text{EPS} = \$9.70$ is:

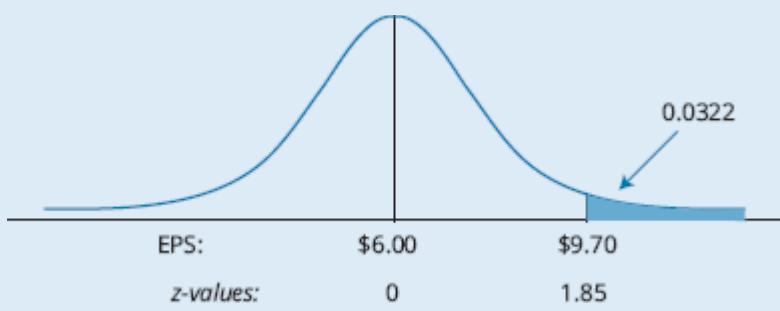
$$z = \frac{x - \mu}{\sigma} = \frac{9.70 - 6}{2} = 1.85$$

That is, \$9.70 is 1.85 standard deviations above the mean EPS value of \$6.

From the z-table we have $F(1.85) = 0.9678$, but this is $P(\text{EPS} \leq 9.70)$. We want $P(\text{EPS} > 9.70)$, which is $1 - P(\text{EPS} \leq 9.70)$.

$$P(\text{EPS} > 9.70) = 1 - 0.9678 = 0.0322, \text{ or } 3.2\%$$

P(EPS > \$9.70)



EXAMPLE: Using the z-table (2)

Using the distribution of EPS with $\mu = \$6$ and $\sigma = \$2$ again, what percentage of the observed EPS values are likely to be less than \$4.10?

Answer:

As shown graphically in the following figure, we want to know $P(\text{EPS} < \$4.10)$. This requires a 2-step approach like the one taken in the preceding example.

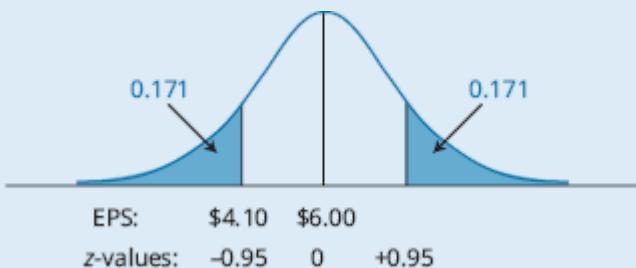
First, the corresponding z-value must be determined as follows:

$$z = \frac{\$4.10 - \$6}{\$2} = -0.95,$$

so \$4.10 is 0.95 standard deviations below the mean of \$6.00.

Now, from the z-table for negative values in the back of this book, we find that $F(-0.95) = 0.1711$, or 17.11%.

Finding a Left-Tail Probability



The z-table gives us the probability that the outcome will be more than 0.95 standard deviations below the mean.

LOS 4.k: Define shortfall risk, calculate the safety-first ratio, and identify an optimal portfolio using Roy's safety-first criterion.

Shortfall risk is the probability that a portfolio value or return will fall below a particular (target) value or return over a given time period.

Roy's safety-first criterion states that the optimal portfolio minimizes the probability that the return of the portfolio falls below some minimum acceptable level. This minimum acceptable level is called the **threshold level**. Symbolically, Roy's safety-first criterion can be stated as:

$$\text{minimize } P(R_p < R_L)$$

where:

R_p = portfolio return

R_L = threshold level return

If portfolio returns are normally distributed, then Roy's safety-first criterion can be stated as:

$$\text{maximize the SFRatio, where SFRatio} = \frac{E(R_p) - R_L}{\sigma_p}$$

The reasoning behind the safety-first criterion is illustrated in Figure 4.6. Assume an investor is choosing between two portfolios: Portfolio A with expected return of 12% and standard deviation of returns of 18%, and Portfolio B with expected return of 10% and standard deviation of returns of 12%. The investor has stated that he wants to minimize the probability of losing money (negative returns). Assuming that returns are normally distributed, the portfolio with the larger SFR using 0% as the threshold return (R_L) will be the one with the lower probability of negative returns.

Figure 4.6: The Safety-First Criterion and Shortfall Risk

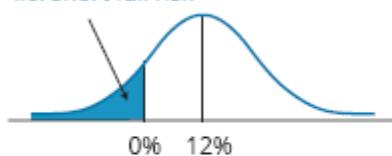
A. Normally Distributed Returns

Portfolio A: $E(R) = 12\%$ $\sigma_A = 18\%$

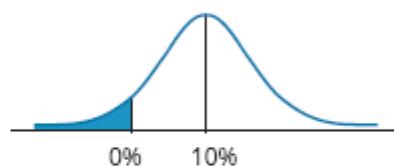
Portfolio B: $E(R) = 10\%$ $\sigma_B = 12\%$

Probability of returns < 0%

- i.e. short fall risk

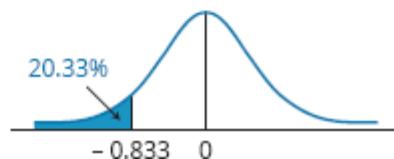
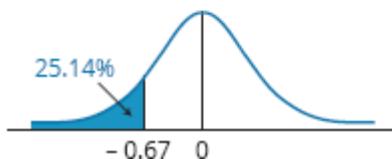


$$SFR_A = \frac{12 - 0}{18} = 0.667$$



$$SFR_B = \frac{10 - 0}{12} = 0.833$$

B. Standard Normal



Panel B of Figure 4.6 relates the SFRatio to the standard normal distribution. Note that the SFR is the number of standard deviations *below* the mean. Thus, the portfolio with the larger SFR has the lower probability of returns below the threshold return, which is a return of 0% in our example. Using a z-table for negative values, we can find the probabilities in the left-hand tails as indicated. These probabilities (25% for Portfolio A and 20% for Portfolio B) are also the shortfall risk for a target return of 0%, that is, the probability of negative returns. Portfolio B has the higher SFR which means it has the lower probability of negative returns.

In summary, when choosing among portfolios with normally distributed returns using Roy's safety-first criterion, there are two steps:

$$\text{Step 1: Calculate the SFRatio} = \frac{E(R_p) - R_L}{\sigma_p}$$

Step 2: Choose the portfolio that has the largest SFRatio.

EXAMPLE: Roy's safety-first criterion

For the next year, the managers of a \$120 million college endowment plan have set a minimum acceptable end-of-year portfolio value of \$123.6 million. Three portfolios are being considered which have the expected returns and standard deviation shown in the first two rows of the following table. Determine which of these portfolios is the most desirable using Roy's safety-first criterion and the probability that the portfolio value will fall short of the target amount.

Answer:

The threshold return is $R_L = (123.6 - 120) / 120 = 0.030 = 3\%$. The SFRs are shown in the table below. As indicated, the best choice is Portfolio A because it has the largest SFR.

Roy's Safety-First Ratios

Portfolio	Portfolio A	Portfolio B	Portfolio C
$E(R_p)$	9%	11%	6.6%
σ_p	12%	20%	8.2%
SFRatio	$0.5 = \frac{9 - 3}{12}$	$0.4 = \frac{11 - 3}{20}$	$0.44 = \frac{6.6 - 3}{8.2}$

The probability of an ending value for Portfolio A less than \$123.6 million (a return less than 3%) is simply $F(-0.5)$, which we can find on the z-table for negative values. The probability is $0.3085 = 30.85\%$.



MODULE QUIZ 4.2

1. A key property of a normal distribution is that it:
 - has zero skewness.
 - is asymmetrical.
 - has zero kurtosis.
2. Which of the following parameters is necessary to describe a multivariate normal distribution?
 - Beta.
 - Correlation.
 - Degrees of freedom.

Use the following table to answer Question 3.

z	0.00	0.01	0.02	0.03	0.04
1.0	0.8413	0.8438	0.8461	0.8485	0.8508
1.1	0.8643	0.8665	0.8686	0.8708	0.8729
1.2	0.8849	0.8869	0.8888	0.8907	0.8925

3. A study of hedge fund investors found that their annual household incomes are normally distributed with a mean of \$175,000 and a standard deviation of \$25,000. The percentage of hedge fund investors that have incomes greater than \$150,000 is *closest* to:
 - 34.13%.
 - 68.26%.
 - 84.13%.
4. For the standard normal distribution, the z -value gives the distance between the mean and a point in terms of:
 - the variance.
 - the standard deviation.
 - the center of the curve.
5. For a standard normal distribution, $F(0)$ is:
 - 0.0.
 - 0.1.
 - 0.5.

Use the following table to answer Questions 6 and 7.

Portfolio	Portfolio A	Portfolio B	Portfolio C
$E(R_p)$	5%	11%	18%
σ_p	8%	21%	40%

6. Given a threshold level of return of 4%, use Roy's safety-first criterion to choose the optimal portfolio.
- Portfolio A.
 - Portfolio B.
 - Portfolio C.
7. Given a threshold level of return of 0%, use Roy's safety-first criterion to choose the optimal portfolio.
- Portfolio A.
 - Portfolio B.
 - Portfolio C.

MODULE 4.3: LOGNORMAL, T, CHI-SQUARE, AND F DISTRIBUTIONS



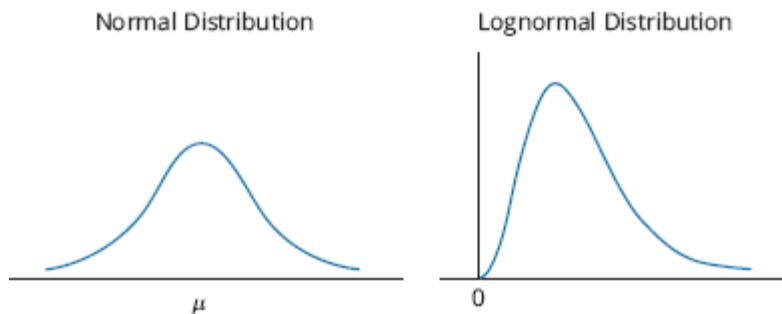
Video covering
this content is
available online.

LOS 4.1: Explain the relationship between normal and lognormal distributions and why the lognormal distribution is used to model asset prices.

The **lognormal distribution** is generated by the function e^x , where x is normally distributed. Since the natural logarithm, \ln , of e^x is x , the logarithms of lognormally distributed random variables are normally distributed, thus the name.

Figure 4.7 illustrates the differences between a normal distribution and a lognormal distribution.

Figure 4.7: Normal vs. Lognormal Distributions



In Figure 4.7, we can see that:

- The lognormal distribution is skewed to the right.
- The lognormal distribution is bounded from below by zero so that it is useful for modeling asset prices, which never take negative values.

If we used a normal distribution of returns to model asset prices over time, we would admit the possibility of returns less than -100%, which would admit the possibility of asset prices less than zero. Using a lognormal distribution to model *price relatives* avoids this problem. A price relative is just the end-of-period price of the asset divided by the beginning price (S_1/S_0) and is equal to (1 + the holding period return). To get the end-of-period asset price, we can simply multiply the price relative times the beginning-of-period asset price. Since a lognormal distribution takes a minimum value of zero, end-of-period asset prices cannot be less than zero. A price relative of zero corresponds to a holding period return of -100% (i.e., the asset price has gone to zero). Recall that we used price relatives as the up-move and down-move (multiplier) terms in constructing a binomial tree for stock price changes over a number of periods.

LOS 4.m: Calculate and interpret a continuously compounded rate of return, given a specific holding period return.

Discretely compounded returns are just the compound returns we are familiar with, given some discrete compounding period, such as semiannual or quarterly. Recall that the more frequent the compounding, the greater the effective annual return. For a stated rate of 10%, semiannual compounding results in an effective yield of $\left(1 + \frac{0.10}{2}\right)^2 - 1 = 10.25\%$ and monthly compounding results in an effective yield of $\left(1 + \frac{0.10}{12}\right)^{12} - 1 = 10.47\%$. Daily or even hourly compounding will produce still larger effective yields. The limit of this exercise, as the compounding periods get shorter and shorter, is called **continuous compounding**. The effective annual rate, based on continuous compounding for a stated annual rate of R_{cc} , can be calculated from the formula:

$$\text{effective annual rate} = e^{R_{cc}} - 1$$

Based on a stated rate of 10%, the effective rate with continuous compounding is $e^{0.10} - 1 = 10.5171\%$. Please verify this by entering 0.1 in your calculator and finding the e^x function.

Since the natural log, \ln , of e^x is x , we can get the continuously compounded rate from an effective annual rate by using the \ln calculator function. Using our previous example, $\ln(1 + 10.517\%) = \ln 1.105171 = 10\%$. Verify this by entering 1.105171 in your calculator and then entering the \ln key.

We can use this method to find the continuously compounded rate that will generate a particular holding period return. If we are given a holding period return of 12.5% for the year, the equivalent continuously compounded rate is $\ln 1.125 = 11.778\%$. Since the calculation is based on 1 plus the holding period return, we can also do the calculation directly from the *price relative*. The price relative is just the end-of-period value divided by the beginning-of-period value. The continuously compounded rate of return is:

$$\ln\left(\frac{S_1}{S_0}\right) = \ln(1 + HPR) = R_{cc}$$

EXAMPLE: Calculating continuously compounded returns

A stock was purchased for \$100 and sold one year later for \$120. Calculate the investor's annual rate of return on a continuously compounded basis.

Answer:

$$\ln\left(\frac{120}{100}\right) = 18.232\%$$

If we had been given the return (20%) instead, the calculation is:

$$\ln(1 + 0.20) = 18.232\%$$

One property of continuously compounded rates of return is that they are additive for multiple periods. Note that the (effective) holding period return over two years is calculated by doubling the continuously compounded annual rate. If $R_{cc} = 10\%$, the (effective) holding period return over two years is $e^{(0.10)^2} - 1 = 22.14\%$. In general, the holding period return after T years, when the annual continuously compounded rate is R_{cc} , is given by:

$$HPR_T = e^{R_{cc} \times T} - 1$$

Given investment results over a 2-year period, we can calculate the 2-year continuously compounded return and divide by 2 to get the annual rate. Consider an investment that appreciated from \$1,000 to \$1,221.40 over a 2-year period. The 2-year continuously compounded rate is $\ln(1,221.40 / 1,000) = 20\%$, and the annual continuously compounded rate (R_{cc}) is $20\% / 2 = 10\%$.

LOS 4.n: Describe the properties of the Student's *t*-distribution, and calculate and interpret its degrees of freedom.

Student's *t*-distribution, or simply the *t*-distribution, is a bell-shaped probability distribution that is symmetrical about its mean. It is the appropriate distribution to use when constructing confidence intervals based on *small samples* ($n < 30$) from populations with *unknown variance* and a normal, or approximately normal, distribution. It may also be appropriate to use the *t*-distribution when the population variance is unknown and the sample size is large enough that the central limit theorem will assure that the sampling distribution is approximately normal.

Student's *t*-distribution has the following properties:

- It is symmetrical.
- It is defined by a single parameter, the **degrees of freedom (df)**, where the degrees of freedom are equal to the number of sample observations minus 1, $n - 1$, for sample means.
- It has more probability in the tails ("fatter tails") than the normal distribution.
- As the degrees of freedom (the sample size) gets larger, the shape of the *t*-distribution more closely approaches a standard normal distribution.

When *compared to the normal distribution*, the *t*-distribution is flatter with more area under the tails (i.e., it has fatter tails). As the degrees of freedom for the *t*-distribution increase, however, its shape approaches that of the normal distribution.

The degrees of freedom for tests based on sample means are $n - 1$ because, given the mean, only $n - 1$ observations can be unique.

The t -distribution is a symmetrical distribution that is centered about zero. The shape of the t -distribution is dependent on the number of degrees of freedom, and degrees of freedom are based on the number of sample observations. The t -distribution is flatter and has thicker tails than the standard normal distribution. As the number of observations increases (i.e., the degrees of freedom increase), the t -distribution becomes more spiked and its tails become thinner. As the number of degrees of freedom increases without bound, the t -distribution converges to the standard normal distribution (z-distribution). The thickness of the tails relative to those of the z-distribution is important in hypothesis testing because thicker tails mean more observations away from the center of the distribution (more outliers). Hence, hypothesis testing using the t -distribution makes it more difficult to reject the null relative to hypothesis testing using the z-distribution.

The table in Figure 4.8 contains one-tailed critical values for the t -distribution at the 0.05 and 0.025 levels of significance with various degrees of freedom (df). Note that, unlike the z-table, the t -values are contained within the table, and the probabilities are located at the column headings.

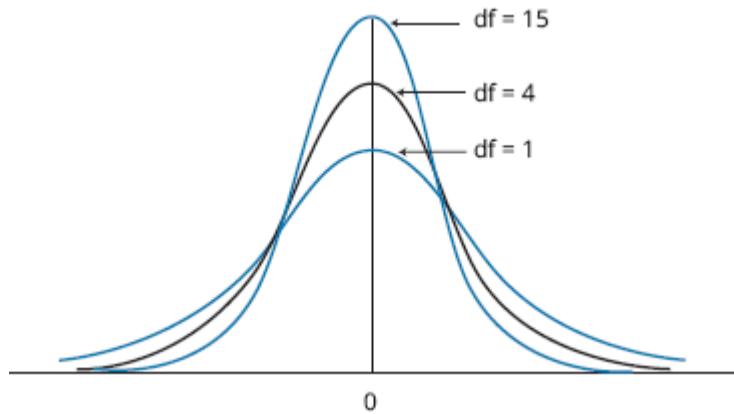
Note that the significance level for a two-tailed test is two times the one-tail probabilities. As degrees of freedom increase, the values under the column for one-tail probabilities of 2.5% approach 1.96, our critical value for a two-tailed test for a normally distributed variable at the 5% significance level. We can see that the critical values increase as the degrees of freedom decrease. Confidence intervals for a t -statistic are wider than those for a z-statistic. Because the tails of the t -distribution are fatter than those for a normally distributed test statistic, we must increase the width of the confidence interval to leave a given percentage of the outcomes in each tail.

Figure 4.8: Table of Critical t -Values

One-Tailed Probabilities, p		
df	$p = 0.05$	$p = 0.025$
5	2.015	2.571
10	1.812	2.228
15	1.753	2.131
20	1.725	2.086
25	1.708	2.060
30	1.697	2.042
40	1.684	2.021
50	1.676	2.009
60	1.671	2.000
70	1.667	1.994
80	1.664	1.990
90	1.662	1.987
100	1.660	1.984
120	1.658	1.980
∞	1.645	1.960

Figure 4.9 illustrates the different shapes of the *t*-distribution associated with different degrees of freedom. The tendency is for the *t*-distribution to look more and more like the normal distribution as the degrees of freedom increase. Practically speaking, the greater the degrees of freedom, the greater the percentage of observations near the center of the distribution and the lower the percentage of observations in the tails, which are thinner as degrees of freedom increase. This means that confidence intervals for a random variable that follows a *t*-distribution must be wider (narrower) when the degrees of freedom are less (more) for a given significance level.

Figure 4.9: *t*-Distributions for Different Degrees of Freedom (df)



LOS 4.0: Describe the properties of the chi-square distribution and the *F*-distribution, and calculate and interpret their degrees of freedom.

Like the *t*-distribution, a **chi-square distribution** (χ^2) is a family of distributions, each based on degrees of freedom. The chi-square distribution is the distribution of the sum of the squared values of n random variables, and k , the degrees of freedom, is equal to $n - 1$.

Because it is the sum of squared values, the chi-square distribution is bounded from below by zero. It is typically asymmetric, but its symmetry increases with the degrees of freedom. As degrees of freedom get larger, the chi-square distribution approaches the normal distribution in shape. The chi-square distribution is often used in tests of the value of the variance of a normally distributed population.

The ***F*-distribution** is the distribution of the quotient of two (appropriately scaled)

$$\text{independent chi-square variables with degrees of freedom } m \text{ and } n: F = \frac{\chi^2/m}{\chi^2/n}$$

where the numerator is a χ^2 variable with m degrees of freedom and the denominator is a χ^2 variable with n degrees of freedom. A common use of the *F*-distribution is to determine the probability that the variances of two independent normal distributions are equal.

The table of values for the *F*-distribution is constructed with the degrees of freedom for the numerator in the top margin, the degrees of freedom for the denominator in the side margin, and

the F -distribution values at the intersections of the degrees of freedom. Each F -distribution table is given for a specific level of significance.

The F -distribution, because it is the ratio of two chi-square values, cannot take on negative values. Therefore, like the chi-square distribution, it is bounded from below by zero. The F -distribution is also asymmetric. As the numerator and denominator degrees of freedom increase, the F -distribution becomes more symmetric and its shape becomes more like the bell curve of a normal distribution.



PROFESSOR'S NOTE

Tables for chi-square and F distributions appear in the Appendix to this book. We will use these distributions in our reading on Hypothesis Testing.

LOS 4.p: Describe Monte Carlo simulation.

Monte Carlo simulation is a technique based on the repeated generation of one or more risk factors that affect security values, in order to generate a distribution of security values. For each of the risk factors, the analyst must specify the parameters of the probability distribution that the risk factor is assumed to follow. A computer is then used to generate random values for each risk factor based on its assumed probability distributions. Each set of randomly generated risk factors is used with a pricing model to value the security. This procedure is repeated many times (100s, 1,000s, or 10,000s), and the distribution of simulated asset values is used to draw inferences about the expected (mean) value of the security and possibly the variance of security values about the mean as well.

As an example, consider the valuation of stock options that can only be exercised on a particular date. The main risk factor is the value of the stock itself, but interest rates could affect the valuation as well. The simulation procedure would be to:

1. Specify the probability distributions of stock prices and of the relevant interest rate, as well as the parameters (mean, variance, possibly skewness) of the distributions.
2. Randomly generate values for both stock prices and interest rates.
3. Value the options for each pair of risk factor values.
4. After many iterations, calculate the mean option value and use that as your estimate of the option's value.

Monte Carlo simulation is used to:

- Value complex securities.
- Simulate the profits/losses from a trading strategy.
- Calculate estimates of value at risk (VaR) to determine the riskiness of a portfolio of assets and liabilities.
- Simulate pension fund assets and liabilities over time to examine the variability of the difference between the two.
- Value portfolios of assets that have nonnormal returns distributions.

The limitations of Monte Carlo simulation are that it is fairly complex and will provide answers that are no better than the assumptions about the distributions of the risk factors and the pricing/valuation model that is used. Also, simulation is not an analytic method, but a statistical one, and cannot provide the insights that analytic methods can.



MODULE QUIZ 4.3

1. For a lognormal distribution:
 - A. the mean equals the median.
 - B. the probability of a negative outcome is zero.
 - C. the probability of a positive outcome is 50%.
2. If a stock's initial price is \$20 and its year-end price is \$23, then its continuously compounded annual (stated) rate of return is:
 - A. 13.64%.
 - B. 13.98%.
 - C. 15.00%.
3. A stock doubled in value last year. Its continuously compounded return over the period was *closest* to:
 - A. 18.2%.
 - B. 69.3%.
 - C. 100.0%.
4. Which of the following is *least likely* a property of Student's *t*-distribution?
 - A. As the degrees of freedom get larger, the variance approaches zero.
 - B. It is defined by a single parameter, the degrees of freedom, which is equal to $n - 1$.
 - C. It has more probability in the tails and less at the peak than a standard normal distribution.
5. Which of the following statements about the *F*-distribution and chi-square distribution is *least accurate*? Both distributions:
 - A. are typically asymmetrical.
 - B. are bounded from below by zero.
 - C. have means that are less than their standard deviations.

KEY CONCEPTS

LOS 4.a

A probability distribution lists all the possible outcomes of an experiment, along with their associated probabilities.

A discrete random variable has positive probabilities associated with a finite number of outcomes.

A continuous random variable has positive probabilities associated with a range of outcome values—the probability of any single value is zero.

LOS 4.b

Given the cumulative distribution function for a random variable, the probability that an outcome will be less than or equal to a specific value is represented by the area under the probability distribution to the left of that value.

LOS 4.c

A discrete uniform distribution is one where there are n discrete, equally likely outcomes.

For a discrete uniform distribution with n possible outcomes, the probability for each outcome equals $1/n$.

LOS 4.d

A continuous uniform distribution is one where the probability of X occurring in a possible range is the length of the range relative to the total of all possible values. Letting a and b be the lower and upper limit of the uniform distribution, respectively, then for:

$$a \leq x_1 \leq x_2 \leq b, P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b - a}$$

LOS 4.e

The binomial distribution is a probability distribution for a binomial (discrete) random variable that has two possible outcomes.

For a binomial distribution, if the probability of success is p , the probability of x successes in n trials is:

$$p(x) = P(X = x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} = {}_n C_x \times p^x (1-p)^{n-x}$$

LOS 4.f

The normal probability distribution and normal curve have the following characteristics:

- The normal curve is symmetrical and bell-shaped with a single peak at the exact center of the distribution.
- Mean = median = mode, and all are in the exact center of the distribution.
- The normal distribution can be completely defined by its mean and standard deviation because the skew is always zero and kurtosis is always 3.

LOS 4.g

Multivariate distributions describe the probabilities for more than one random variable, whereas a univariate distribution is for a single random variable.

The correlation(s) of a multivariate distribution describes the relation between the outcomes of its variables relative to their expected values.

LOS 4.h

A confidence interval is a range within which we have a given level of confidence of finding a point estimate (e.g., the 90% confidence interval for X is  to 

Confidence intervals for any normally distributed random variable are:

- 90%: $\mu \pm 1.65$ standard deviations.
- 95%: $\mu \pm 1.96$ standard deviations.
- 99%: $\mu \pm 2.58$ standard deviations.

The probability that a normally distributed random variable X will be within A standard deviations of its mean, μ , [i.e., $P(\mu - A\sigma \leq X \leq \mu + A\sigma)$], may be calculated as $F(A) - F(-A)$, where $F(A)$ is the cumulative standard normal probability of A , or as $1 - 2[F(-A)]$.

LOS 4.i

The standard normal probability distribution has a mean of 0 and a standard deviation of 1.

A normally distributed random variable X can be standardized as $Z = \frac{X - \mu}{\sigma}$ and Z will be normally distributed with mean = 0 and standard deviation 1.

LOS 4.j

The z-table is used to find the probability that X will be less than or equal to a given value.

- $P(X < x) = F(x) = F\left[\frac{x - \mu}{\sigma}\right] = F(z)$, which is found in the standard normal probability table.
- $P(X > x) = 1 - P(X < x) = 1 - F(z)$.

LOS 4.k

Shortfall risk is the probability that a portfolio's value (or return) will fall below a specific value over a given period of time.

The safety-first ratio for portfolio P , based on a target return R_T , is:

$$\text{SFRatio} = \frac{E(R_P) - R_T}{\sigma_P}$$

Greater safety-first ratios are preferred and indicate a smaller shortfall probability. Roy's safety-first criterion states that the optimal portfolio minimizes shortfall risk.

LOS 4.l

If x is normally distributed, e^x follows a lognormal distribution. A lognormal distribution is often used to model asset prices, since a lognormal random variable cannot be negative and can take on any positive value.

LOS 4.m

As we decrease the length of discrete compounding periods (e.g., from quarterly to monthly) the effective annual rate increases. As the length of the compounding period in discrete compounding gets shorter and shorter, the compounding becomes continuous, where the effective annual rate = $e^i - 1$.

For a holding period return (HPR) over any period, the equivalent continuously compounded rate over the period is $\ln(1 + \text{HPR})$.

LOS 4.n

The t -distribution is similar, but not identical, to the normal distribution in shape—it is defined by the degrees of freedom and has fatter tails compared to the normal distribution.

Degrees of freedom for the t -distribution are equal to $n - 1$. Student's t -distribution is closer to the normal distribution when degrees of freedom are greater, and confidence intervals are narrower when degrees of freedom are greater.

LOS 4.o

A chi-square distribution has $n - 1$ degrees of freedom, is asymmetric, is bounded from below by zero, and becomes more symmetric and approaches a bell-curve shape as its degrees of freedom increase.

A ratio of two chi-square random variables with degrees of freedom m and n follows an F -distribution with degrees of freedom m and n . An F -distribution is asymmetric, bounded from

below by zero, and approaches a bell-curve shape as its degrees of freedom increase.

LOS 4.p

Monte Carlo simulation uses randomly generated values for risk factors, based on their assumed distributions, to produce a distribution of possible security values. Its limitations are that it is fairly complex and will provide answers that are no better than the assumptions used.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 4.1

1. **B** Time is usually a continuous random variable; the others are discrete. (LOS 4.a)
2. **B** For a continuous distribution $p(x) = 0$ for all X ; only ranges of value of X have positive probabilities. (LOS 4.a)
3. **B** $\frac{8-5}{5} + \frac{8-10}{5} = \frac{1}{5}$, and $\frac{8-10}{5}$ is negative, so this satisfies neither of the requirements for a probability distribution. The others have $P[X_i]$ between zero and 1 and $\sum P[X_i] = 1$, and thus satisfy both requirements for a probability distribution. (LOS 4.a)
4. **C** $(0.04 + 0.11 + 0.18 + 0.24 + 0.14 + 0.17) = 0.88$ (LOS 4.b)
5. **B** $(0.14 + 0.17 + 0.09 + 0.03) = 0.43$ (LOS 4.b)
6. **C** $(0.18 + 0.24 + 0.14 + 0.17) = 0.73$ (LOS 4.b)
7. **A** $0 + 1(0.11) + 2(0.18) + 3(0.24) + 4(0.14) + 5(0.17) + 6(0.09) + 7(0.03) = 3.35$ (LOS 4.b)
8. **C** $F(x)$ is the cumulative probability, $P(x < 20)$ here. Because all the observations in this distribution are between 4 and 10, the probability of an outcome less than 20 is 100%. (LOS 4.d)
9. **A** There may be any number of independent trials, each with only two possible outcomes. (LOS 4.e)
10. **B** With only two possible outcomes, there must be some positive probability for each. If this were not the case, the variable in question would not be a random variable, and a probability distribution would be meaningless. It does not matter if one of the possible outcomes happens to be zero. (LOS 4.e)
11. **C** Success = having a fax machine. $[6! / 4!(6 - 4)!](0.6)^4(0.4)^{6-4} = 15(0.1296)(0.16) = 0.311$. (LOS 4.e)
12. **A** Success = staying for five years. $[6! / 2!(6 - 2)!](0.10)^2(0.90)^{6-2} = 15(0.01)(0.656) = 0.0984$. (LOS 4.e)
13. **C** Success = passing the exam. Then, $E(\text{success}) = np = 15 \times 0.4 = 6$. (LOS 4.e)

Module Quiz 4.2

1. **A** Normal distributions are symmetrical (i.e., have zero skewness) and their kurtosis is equal to 3. (LOS 4.f)
2. **B** To describe a multivariate normal distribution, we must consider the correlations among the variables, as well as the means and variances of the variables. (LOS 4.g)
3. **C** $1 - F(-1) = F(1) = 0.8413$. There is an 84.13% probability that a randomly chosen income is not more than one standard deviation below the mean. (LOS 4.h)
4. **B** This is true by the formula for z . (LOS 4.i)
5. **C** By the symmetry of the z -distribution and $F(0) = 0.5$. Half the distribution lies on each side of the mean. (LOS 4.j)

6. **C** $SFR = (18 - 4) / 40 = 0.35$ is the largest value. (LOS 4.k)

7. **A** $SFR = (5 - 0) / 8 = 0.625$ is the largest value. (LOS 4.k)

Module Quiz 4.3

1. **B** A lognormally distributed variable is never negative. (LOS 4.l)
2. **B** $\ln(23 / 20) = 0.1398$ (LOS 4.m)
3. **B** $\ln(2) = 0.6931$ (LOS 4.m)
4. **A** As the degrees of freedom get larger, the t -distribution approaches the normal distribution. As the degrees of freedom fall, the peak of the t -distribution flattens and its tails get fatter (more probability in the tails—that's why, all else the same, the critical t increases as the df decreases). (LOS 4.n)
5. **C** There is no consistent relationship between the mean and standard deviation of the chi-square distribution or F -distribution. (LOS 4.o)

READING 5

SAMPLING AND ESTIMATION

EXAM FOCUS

This reading covers sampling and making inferences about population parameters (and other statistics) from sample data. It is essential that you know the central limit theorem, for it allows us to use sampling statistics to construct confidence intervals for point estimates of population means. Make sure you can calculate confidence intervals for population means given sample parameter estimates and a level of significance, and know when it is appropriate to use the z -statistic versus the t -statistic. You should also understand the various procedures for selecting samples, and recognize the sources of bias in selecting sample data.

MODULE 5.1: SAMPLING METHODS, CENTRAL LIMIT THEOREM, AND STANDARD ERROR



Video covering this content is available online.

LOS 5.a: Compare and contrast probability samples with non-probability samples and discuss applications of each to an investment problem.

Probability sampling refers to selecting a sample when we know the probability of each sample member in the overall population. With **random sampling**, each item is assumed to have the same probability of being selected. If we have a population of data and select our sample by using a computer to randomly select a number of observations from the population, each data point has an equal probability of being selected and we call this **simple random sampling**. If we want to estimate the mean profitability for a population of firms, this may be an appropriate method.

Non-probability sampling is based on either low cost and easy access to some data items, or on using the judgment of the researcher in selecting specific data items. Less randomness in selection may lead to greater sampling error.

LOS 5.b: Explain sampling error.

Sampling error is the difference between a sample statistic (such as the mean, variance, or standard deviation of the sample) and its corresponding population parameter (the true mean, variance, or standard deviation of the population). For example, the sampling error for the mean is as follows:

$$\text{sampling error of the mean} = \text{sample mean} - \text{population mean} = x - \mu$$

It is important to recognize that the sample statistic itself is a random variable and therefore has a probability distribution. The **sampling distribution** of the sample statistic is a probability distribution of all possible sample statistics computed from a set of equal-size samples that were randomly drawn from the same population. Think of it as the probability distribution of a statistic from many samples.

For example, suppose a random sample of 100 bonds is selected from a population of a major municipal bond index consisting of 1,000 bonds, and then the mean return of the 100-bond sample is calculated. Repeating this process many times will result in many different estimates of the population mean return (i.e., one for each sample). The distribution of these estimates of the mean is the *sampling distribution of the mean*.

It is important to note that this sampling distribution is distinct from the distribution of the actual prices of the 1,000 bonds in the underlying population and will have different parameters.

LOS 5.c: Compare and contrast simple random, stratified random, cluster, convenience, and judgmental sampling.

Probability Sampling Methods

Simple random sampling is a method of selecting a sample in such a way that each item or person in the population being studied has the same likelihood of being included in the sample. As an example of simple random sampling, assume that you want to draw a sample of five items out of a group of 50 items. This can be accomplished by numbering each of the 50 items, placing them in a hat, and shaking the hat. Next, one number can be drawn randomly from the hat.

Repeating this process (experiment) four more times results in a set of five numbers. The five drawn numbers (items) comprise a simple random sample from the population. In applications like this one, a random-number table or a computer random-number generator is often used to create the sample. Another way to form an approximately random sample is **systematic sampling**, selecting every *n*th member from a population.

Stratified random sampling uses a classification system to separate the population into smaller groups based on one or more distinguishing characteristics. From each subgroup, or stratum, a random sample is taken and the results are pooled. The size of the samples from each stratum is based on the size of the stratum relative to the population.

Stratified sampling is often used in bond indexing because of the difficulty and cost of completely replicating the entire population of bonds. In this case, bonds in a population are categorized (stratified) according to major bond risk factors such as duration, maturity, coupon rate, and the like. Then, samples are drawn from each separate category and combined to form a final sample.

To see how this works, suppose you want to construct a portfolio of 100 bonds that is indexed to a major municipal bond index of 1,000 bonds, using a stratified random sampling approach.

First, the entire population of 1,000 municipal bonds in the index can be classified on the basis of maturity and coupon rate. Then, cells (stratum) can be created for different maturity/coupon combinations, and random samples can be drawn from each of the maturity/coupon cells. To sample from a cell containing 50 bonds with 2- to 4-year maturities and coupon rates less than

5%, we would select five bonds. The number of bonds drawn from a given cell corresponds to the cell's weight relative to the population (index), or $(50 / 1000) \times 100 = 5$ bonds. This process is repeated for all the maturity/coupon cells, and the individual samples are combined to form the portfolio.

By using stratified sampling, we guarantee that we sample five bonds from this cell. If we had used simple random sampling, there would be no guarantee that we would sample any of the bonds in the cell. Or, we may have selected more than five bonds from this cell.

Cluster sampling is also based on subsets of a population, but in this case we are assuming that each subset (cluster) is representative of the overall population with respect to the item we are sampling. For example, we may have data on personal incomes for a state's residents by county. The data for each county is a cluster.

In **one-stage cluster sampling**, a random sample of clusters is selected and all the data in those clusters comprise the sample. In **two-stage cluster sampling**, random samples from each of the selected clusters comprise the sample. Contrast this with stratified random sampling, in which random samples are selected from every subgroup.

To the extent that the subgroups do not have the same distribution as the entire population of the characteristic we are interested in, cluster sampling will have greater sampling error than simple random sampling. Two-stage cluster sampling can be expected to have greater sampling error than one-stage cluster sampling. Lower cost and less time required to assemble the sample are the primary advantages of cluster sampling, and it may be most appropriate for a smaller pilot study.

Non-Probability Sampling Methods

Convenience sampling refers to selecting sample data based on its ease of access, using data that are readily available. Because such a sample is typically not random, sampling error will be greater. It is most appropriate for an initial look at the data prior to adopting a sampling method with less sampling error.

Judgmental sampling refers to samples for which each observation is selected from a larger data set by the researcher, based on her experience and judgment. As an example, a researcher interested in assessing company compliance with accounting standards may have experience suggesting that evidence of noncompliance is typically found in certain ratios derived from the financial statements. The researcher may select only data on these items. Researcher bias (or simply poor judgment) may lead to samples that have excessive sampling error. In the absence of bias or poor judgment, judgmental sampling may produce a more representative sample or allow the researcher to focus on a sample that offers good data on the characteristic or statistic of interest.

An important consideration when sampling is ensuring that the distribution of data of interest is constant for the whole population being sampled. For example, judging a characteristic U.S. banks using data from 2005 to 2015 may not be appropriate. It may well be that regulatory reform of the banking industry after the financial crisis of 2007–2008 resulted in significant changes in banking practices, so that the mean of a statistic precrisis and its mean value across the population of banks post-crisis are quite different. Pooling the data over the whole period

from 2005 to 2015 would not be appropriate if this is the case, and the sample mean calculated from it would not be a good estimate of either precrisis or post-crisis mean values.

LOS 5.d: Explain the central limit theorem and its importance.

The **central limit theorem** states that for simple random samples of size n from a *population* with a mean μ and a finite variance σ^2 , the sampling distribution of the sample mean \bar{x} approaches a normal probability distribution with mean μ and a variance equal to $\frac{\sigma^2}{n}$ as the sample size becomes large.

The central limit theorem is extremely useful because the normal distribution is relatively easy to apply to hypothesis testing and to the construction of confidence intervals. Specific inferences about the population mean can be made from the sample mean, *regardless of the population's distribution*, as long as the sample size is "sufficiently large," which usually means $n \geq 30$.

Important properties of the central limit theorem include the following:

- If the sample size n is sufficiently large ($n \geq 30$), the sampling distribution of the sample means will be approximately normal. Remember what's going on here, random samples of size n are repeatedly being taken from an overall larger population. Each of these random samples has its own mean, which is itself a random variable, and this set of sample means has a distribution that is approximately normal.
 - The mean of the population, μ , and the mean of the distribution of all possible sample means are equal.
 - The variance of the distribution of sample means is $\frac{\sigma^2}{n}$, the population variance divided by the sample size.
-

LOS 5.e: Calculate and interpret the standard error of the sample mean.

The **standard error of the sample mean** is the standard deviation of the distribution of the sample means.

When the standard deviation of the population, σ , is *known*, the standard error of the sample mean is calculated as:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

where:

$\sigma_{\bar{x}}$ = standard error of the sample mean

σ = standard deviation of the population

n = size of the sample

EXAMPLE: Standard error of sample mean (known population variance)

The mean hourly wage for Iowa farm workers is \$13.50 with a *population standard deviation* of \$2.90. Calculate and interpret the standard error of the sample mean for a sample size of 30.

Answer:

Because the population standard deviation, σ , is known, the standard error of the sample mean is expressed as:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{\$2.90}{\sqrt{30}} = \$0.53$$

This means that if we were to take many samples of size 30 from the Iowa farm worker population and prepare a sampling distribution of the sample means, we would get a distribution with an expected mean of \$13.50 and standard error (standard deviation of the sample means) of \$0.53.

Practically speaking, the *population's standard deviation is almost never known*. Instead, the standard error of the sample mean must be estimated by dividing the standard deviation of the sample by \sqrt{n} :

$$s_{\bar{X}} = \frac{s}{\sqrt{n}}$$

EXAMPLE: Standard error of sample mean (unknown population variance)

Suppose a sample contains the past 30 monthly returns for McCreary, Inc. The mean return is 2% and the *sample* standard deviation is 20%. Calculate and interpret the standard error of the sample mean.

Answer:

Since σ is unknown, the standard error of the sample mean is:

$$s_{\bar{X}} = \frac{s}{\sqrt{n}} = \frac{20\%}{\sqrt{30}} = 3.6\%$$

This implies that if we took all possible samples of size 30 from McCreary's monthly returns and prepared a sampling distribution of the sample means, the mean would be 2% with a standard error of 3.6%.

EXAMPLE: Standard error of sample mean (unknown population variance)

Continuing with our example, suppose that instead of a sample size of 30, we take a sample of the past 200 monthly returns for McCreary, Inc. In order to highlight the effect of sample size on the sample standard error, let's assume that the mean return and standard deviation of this larger sample remain at 2% and 20%, respectively. Now, calculate the standard error of the sample mean for the 200-return sample.

Answer:

The standard error of the sample mean is computed as:

$$s_{\bar{X}} = \frac{s}{\sqrt{n}} = \frac{20\%}{\sqrt{200}} = 1.4\%$$

The result of the preceding two examples illustrates an important property of sampling distributions. Notice that the value of the standard error of the sample mean decreased from 3.6% to 1.4% as the sample size increased from 30 to 200. This is because as the sample size increases, the sample mean gets closer, on average, to the true mean of the population. In other words, the distribution of the sample means about the population mean gets smaller and smaller, so the standard error of the sample mean decreases.



PROFESSOR'S NOTE

I get a lot of questions about when to use σ and $\frac{\sigma}{\sqrt{n}}$. Just remember that the standard deviation of the means of multiple samples is less than the standard deviation of single observations. If the standard deviation of monthly stock returns is 2%, the standard error (deviation) of the average monthly return over the next six months is $\frac{2\%}{\sqrt{6}} = 0.82\%$. The average of several observations of a random variable will be less widely dispersed (have lower standard deviation) around the expected value than will a single observation of the random variable.

LOS 5.f: Identify and describe desirable properties of an estimator.

Regardless of whether we are concerned with point estimates or confidence intervals, there are certain statistical properties that make some estimates more desirable than others. These desirable properties of an estimator are **unbiasedness**, **efficiency**, and **consistency**.

- An *unbiased* estimator is one for which the expected value of the estimator is equal to the parameter you are trying to estimate. For example, because the expected value of the sample mean is equal to the population mean [$E(\bar{X}) = \mu$], the sample mean is an unbiased estimator of the population mean.
- An unbiased estimator is also *efficient* if the variance of its sampling distribution is smaller than all the other unbiased estimators of the parameter you are trying to estimate. The sample mean, for example, is an unbiased and efficient estimator of the population mean.
- A *consistent* estimator is one for which the accuracy of the parameter estimate increases as the sample size increases. As the sample size increases, the standard error of the sample mean falls, and the sampling distribution bunches more closely around the population mean. In fact, as the sample size approaches infinity, the standard error approaches zero.



MODULE QUIZ 5.1

1. An important difference between two-stage cluster sampling and stratified random sampling is that compared to stratified random sampling, two-stage cluster sampling:
 - A. uses all members of each sub-group (strata).
 - B. takes random samples all sub-groups (strata).
 - C. will not preserve differences in a characteristic across sub-groups.
2. Sampling error is defined as:
 - A. an error that occurs when a sample of less than 30 elements is drawn.
 - B. an error that occurs during collection, recording, and tabulation of data.

- C. the difference between the value of a sample statistic and the value of the corresponding population parameter.
3. The mean age of all CFA candidates is 28 years. The mean age of a random sample of 100 candidates is found to be 26.5 years. The difference of 1.5 years is called:
- the random error.
 - the sampling error.
 - the population error.
4. A simple random sample is a sample drawn in such a way that each member of the population has:
- some chance of being selected in the sample.
 - an equal chance of being included in the sample.
 - a 1% chance of being included in the sample.
5. To apply the central limit theorem to the sampling distribution of the sample mean, the sample is usually considered to be large if n is *greater* than:
- 20.
 - 25.
 - 30.
6. If n is large and the population standard deviation is unknown, the standard error of the sampling distribution of the sample mean is *equal* to:
- the sample standard deviation divided by the sample size.
 - the population standard deviation multiplied by the sample size.
 - the sample standard deviation divided by the square root of the sample size.
7. The standard error of the sampling distribution of the sample mean for a sample size of n drawn from a population with a mean of μ and a standard deviation of σ is:
- sample standard deviation divided by the sample size.
 - sample standard deviation divided by the square root of the sample size.
 - population standard deviation divided by the square root of the sample size.
8. Assume that a population has a mean of 14 with a standard deviation of 2. If a random sample of 49 observations is drawn from this population, the standard error of the sample mean is *closest* to:
- 0.04.
 - 0.29.
 - 2.00.
9. The population's mean is 30 and the mean of a sample of size 100 is 28.5. The variance of the sample is 25. The standard error of the sample mean is *closest* to:
- 0.05.
 - 0.25.
 - 0.50.
10. Which of the following is *least likely* a desirable property of an estimator?
- Reliability.
 - Efficiency.
 - Consistency.

MODULE 5.2: CONFIDENCE INTERVALS, RESAMPLING, AND SAMPLING BIASES



Video covering
this content is
available online.

LOS 5.g: Contrast a point estimate and a confidence interval estimate of a population parameter.

LOS 5.h: Calculate and interpret a confidence interval for a population mean, given a normal distribution with 1) a known population variance, 2) an unknown population variance, or 3) an unknown population variance and a large sample size.

Point estimates are single (sample) values used to estimate population parameters. The formula used to compute the point estimate is called the estimator. For example, the sample mean, \bar{x} , is an estimator of the population mean μ and is computed using the familiar formula:

$$\bar{x} = \frac{\sum x}{n}$$

The value generated with this calculation for a given sample is called the *point estimate* of the mean.

A **confidence interval** is a range of values in which the population parameter is expected to lie. Confidence interval estimates result in a range of values within which the actual value of a parameter will lie, given the probability of $1 - \alpha$. Here, alpha, α , is called the *level of significance* for the confidence interval, and the probability $1 - \alpha$ is referred to as the *degree of confidence*. For example, we might estimate that the population mean of random variables will range from 15 to 25 with a 95% degree of confidence, or at the 5% level of significance.

Confidence intervals are usually constructed by adding or subtracting an appropriate value from the point estimate. In general, confidence intervals take on the following form:

$$\text{point estimate} \pm (\text{reliability factor} \times \text{standard error})$$

where:

point estimate = value of a sample statistic of the population parameter

reliability factor = number that depends on the sampling distribution of the point estimate and the probability that the point estimate falls in the confidence interval, $(1 - \alpha)$

standard error = standard error of the point estimate

If the population has a *normal distribution with a known variance*, a **confidence interval for the population mean** can be calculated as:

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

where:

\bar{x} = *point estimate* of the population mean (sample mean).

$z_{\alpha/2}$ = *reliability factor*, a standard normal random variable for which the probability in the right-hand tail of the distribution is $\alpha/2$. In other words, this is the *z-score* that leaves $\alpha/2$ of probability in the upper tail.

$\frac{\sigma}{\sqrt{n}}$ = the *standard error* of the sample mean where σ is the known standard deviation of the population, and n is the sample size.

The most commonly used standard normal distribution reliability factors are:

$z_{\alpha/2} = 1.645$ for 90% confidence intervals (the significance level is 10%, 5% in each tail).

$z_{\alpha/2} = 1.960$ for 95% confidence intervals (the significance level is 5%, 2.5% in each tail).

$z_{\alpha/2} = 2.575$ for 99% confidence intervals (the significance level is 1%, 0.5% in each tail).

Do these numbers look familiar? They should! In our review of common probability distributions, we found the probability under the standard normal curve between $z = -1.96$ and

$z = +1.96$ to be 0.95, or 95%. Owing to symmetry, this leaves a probability of 0.025 under each tail of the curve beyond $z = -1.96$ or $z = +1.96$, for a total of 0.05, or 5%—just what we need for a significance level of 0.05, or 5%.

EXAMPLE: Confidence interval

Consider a practice exam that was administered to 36 Level I candidates. Their mean score on this practice exam was 80. Assuming a population standard deviation equal to 15, construct and interpret a 99% confidence interval for the mean score on the practice exam for all candidates. *Note that, in this example, the population standard deviation is known, so we don't have to estimate it.*

Answer:

At a confidence level of 99%, $z_{\alpha/2} = z_{0.005} = 2.58$. So, the 99% confidence interval is calculated as follows:

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 80 \pm 2.58 \frac{15}{\sqrt{36}} = 80 \pm 6.45$$

Thus, the 99% confidence interval ranges from 73.55 to 86.45.

EXAMPLE: Confidence intervals for a population mean and for a single observation

Annual returns on energy stocks are approximately normally distributed with a mean of 9% and standard deviation of 6%. Construct a 90% confidence interval for the annual returns of a randomly selected energy stock and a 90% confidence interval for the mean of the annual returns for a sample of 12 energy stocks.

Answer:

A 90% confidence interval for a single observation is 1.645 *standard deviations* from the sample mean.

$$9\% \pm 1.645(6\%) = -0.87\% \text{ to } 18.87\%$$

A 90% confidence interval for the population mean is 1.645 *standard errors* from the sample mean.

$$9\% \pm 1.645 \frac{6\%}{\sqrt{12}} = 6.15\% \text{ to } 11.85\%$$

Confidence intervals can be interpreted from a probabilistic perspective or a practical perspective. With regard to the outcome of the practice exam example, these two perspectives can be described as follows:

- *Probabilistic interpretation.* After repeatedly taking samples of CFA candidates, administering the practice exam, and constructing confidence intervals for each sample's mean, 99% of the resulting confidence intervals will, in the long run, include the population mean.
- *Practical interpretation.* We are 99% confident that the population mean score is between 73.55 and 86.45 for candidates from this population.

Confidence Intervals for the Population Mean: Normal With Unknown Variance

If the distribution of the *population is normal with unknown variance*, we can use the *t*-distribution to construct a confidence interval:

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

where:

\bar{x} = the point estimate of the population mean

$t_{\alpha/2}$ = the *t*-reliability factor (a.k.a. *t*-statistic or critical *t*-value) corresponding to a *t*-distributed random variable with $n - 1$ degrees of freedom, where n is the sample size. The area under the tail of the *t*-distribution to the right of $t_{\alpha/2}$ is $\alpha/2$.

$\frac{s}{\sqrt{n}}$ = standard error of the sample mean

s = sample standard deviation

Unlike the standard normal distribution, the reliability factors for the *t*-distribution depend on the sample size, so we cannot rely on a commonly used set of reliability factors. Instead, reliability factors for the *t*-distribution have to be looked up in a table of Student's *t*-distribution, like the one at the back of this book.

Owing to the relatively fatter tails of the *t*-distribution, confidence intervals constructed using *t*-reliability factors ($t_{\alpha/2}$) will be more conservative (wider) than those constructed using *z*-reliability factors.

EXAMPLE: Confidence intervals

Let's return to the McCreary, Inc., example. Recall that we took a sample of the past 30 monthly stock returns for McCreary, Inc., and determined that the mean return was 2% and the sample standard deviation was 20%. Since the population variance is unknown, the standard error of the sample was estimated to be:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{20\%}{\sqrt{30}} = 3.6\%$$

Now, let's construct a 95% confidence interval for the mean monthly return.

Answer:

Here, we will use the *t*-reliability factor because the population variance is unknown. Since there are 30 observations, the degrees of freedom are $29 = 30 - 1$. Remember, because this is a two-tailed test at the 95% confidence level, the probability under each tail must be $\alpha/2 = 2.5\%$, for a total of 5%. So, referencing the one-tailed probabilities for Student's *t*-distribution at the back of this book, we find the critical *t*-value (reliability factor) for $\alpha/2 = 0.025$ and $df = 29$ to be $t_{29, 2.5} = 2.045$. Thus, the 95% confidence interval for the population mean is:

$$2\% \pm 2.045 \left(\frac{20\%}{\sqrt{30}} \right) = 2\% \pm 2.045(3.6\%) = 2\% \pm 7.4\%$$

Thus, the 95% confidence has a lower limit of -5.4% and an upper limit of +9.4%.

We can interpret this confidence interval by saying that we are 95% confident that the population mean monthly return for McCreary stock is between -5.4% and +9.4%.



PROFESSOR'S NOTE

You should practice looking up reliability factors (a.k.a. critical t -values or t -statistics) in a t -table. The first step is always to compute the degrees of freedom, which is $n - 1$. The second step is to find the appropriate level of alpha or significance. This depends on whether the test you're concerned with is one-tailed (use α) or two-tailed (use $\alpha/2$). In this review, our tests will always be two-tailed because confidence intervals are designed to compute an upper and lower limit. Thus, we will use $\alpha/2$. To look up $t_{29, 2.5}$, find the 29 df row and match it with the 0.025 column; $t = 2.045$ is the result. We'll do more of this in our study of hypothesis testing.

Confidence Interval for a Population Mean When the Population Variance Is Unknown, Given a Large Sample From Any Type of Distribution

We now know that the z -statistic should be used to construct confidence intervals when the population distribution is normal and the variance is known, and the t -statistic should be used when the distribution is normal but the variance is unknown. But what do we do when the distribution is *nonnormal*?

As it turns out, the size of the sample influences whether or not we can construct the appropriate confidence interval for the sample mean.

- If the *distribution is nonnormal* but the *population variance is known*, the z -statistic can be used as long as the sample size is large ($n \geq 30$). We can do this because the central limit theorem assures us that the distribution of the sample mean is approximately normal when the sample is large.
- If the *distribution is nonnormal* and the *population variance is unknown*, the t -statistic can be used as long as the sample size is large ($n \geq 30$). It is also acceptable to use the z -statistic, although use of the t -statistic is more conservative.

This means that if we are sampling from a nonnormal distribution (which is sometimes the case in finance), *we cannot create a confidence interval if the sample size is less than 30*. So, all else equal, make sure you have a sample of at least 30, and the larger, the better.

Figure 5.1 summarizes this discussion.



PROFESSOR'S NOTE

You should commit the criteria in the following table to memory.

Figure 5.1: Criteria for Selecting the Appropriate Test Statistic

When sampling from a:	Test Statistic	
	Small Sample (n < 30)	Large Sample (n ≥ 30)
Normal distribution with known variance	z-statistic	z-statistic
Normal distribution with unknown variance	t-statistic	t-statistic*
Nonnormal distribution with known variance	not available	z-statistic
Nonnormal distribution with unknown variance	not available	t-statistic*

*The z-statistic is theoretically acceptable here, but use of the t-statistic is more conservative.

All of the preceding analysis depends on the sample we draw from the population being random. If the sample isn't random, the central limit theorem doesn't apply, our estimates won't have the desirable properties, and we can't form unbiased confidence intervals. Surprisingly, creating a *random sample* is not as easy as one might believe. There are a number of potential mistakes in sampling methods that can bias the results. These biases are particularly problematic in financial research, where available historical data are plentiful, but the creation of new sample data by experimentation is restricted.

LOS 5.i: Describe the use of resampling (bootstrap, jackknife) to estimate the sampling distribution of a statistic.

Previously, we used the sample variance to calculate the standard error of our estimate of the mean. The standard error provides better estimates of the distribution of sample means when the sample is unbiased and the distribution of sample means is approximately normal.

Two alternative methods of estimating the standard error of the sample mean involve resampling of the data. The first of these, termed the **jackknife**, calculates multiple sample means, each with one of the observations removed from the sample. The standard deviation of these sample means can then be used as an estimate of the standard error of sample means. The jackknife is a computationally simple tool and can be used when the number of observations available is relatively small. It can remove bias from statistical estimates.

The jackknife (so named because it is a handy and readily available tool) was developed when computational power was not as available and low-cost as it has become. A **bootstrap** method is more computationally demanding but has some advantages. To estimate the standard error of the sample mean, we draw repeated samples of size n from the full data set (replacing the sampled observations each time). We can then directly calculate the standard deviation of these sample means as our estimate of the standard error of the sample mean.

The bootstrap method can improve accuracy compared to using only the data in a single sample, and can be used to construct confidence intervals for a variety of statistics in addition to the mean, such as the median. It can be used to estimate the distributions of complex statistics, including those that do not have an analytic form.

LOS 5.j: Describe the issues regarding selection of the appropriate sample size, data snooping bias, sample selection bias, survivorship bias, look-ahead bias, and time-period bias.

We have seen so far that a larger sample reduces the sampling error and the standard deviation of the sample statistic around its true (population) value. Confidence intervals are narrower when samples are larger and the standard errors of the point estimates of population parameters are less.

There are two limitations on this idea of “larger is better” when it comes to selecting an appropriate sample size. One is that larger samples may contain observations from a different population (distribution). If we include observations that come from a different population (one with a different population parameter), we will not necessarily improve, and may even reduce, the precision of our population parameter estimates. The other consideration is cost. The costs of using a larger sample must be weighed against the value of the increase in precision from the increase in sample size. Both of these factors suggest that the largest possible sample size is not always the most appropriate choice.

Data snooping occurs when analysts repeatedly use the same database to search for patterns or trading rules until one that “works” is discovered. For example, empirical research has provided evidence that value stocks appear to outperform growth stocks. Some researchers argue that this anomaly is actually the product of data snooping. Because the data set of historical stock returns is quite limited, it is difficult to know for sure whether the difference between value and growth stock returns is a true economic phenomenon or simply a chance pattern that was stumbled upon after repeatedly looking for any identifiable pattern in the data.

Data-snooping bias refers to results where the statistical significance of the pattern is overestimated because the results were found through data snooping.

When reading research findings that suggest a profitable trading strategy, make sure you heed the following warning signs of data snooping:

- Evidence that many different variables were tested, most of which are unreported, until significant ones were found.
- The lack of any economic theory that is consistent with the empirical results.

The best way to avoid data snooping is to test a potentially profitable trading rule on a data set different from the one you used to develop the rule (i.e., use out-of-sample data).

Sample selection bias occurs when some data is systematically excluded from the analysis, usually because of the lack of availability. This practice renders the observed sample to be nonrandom, and any conclusions drawn from this sample can't be applied to the population because the observed sample and the portion of the population that was not observed are different.

Survivorship bias is the most common form of sample selection bias. A good example of the existence of survivorship bias in investments is the study of mutual fund performance. Most mutual fund databases, like Morningstar®'s, only include funds currently in existence—the “survivors.” They do not include funds that have ceased to exist due to closure or merger.

This would not be a problem if the characteristics of the surviving funds and the missing funds were the same; then the sample of survivor funds would still be a random sample drawn from the population of mutual funds. As one would expect, however, and as evidence has shown, the funds that are dropped from the sample have lower returns relative to the surviving funds. Thus,

the surviving sample is biased toward the better funds (i.e., it is not random). The analysis of a mutual fund sample with survivorship bias will yield results that overestimate the average mutual fund return because the database only includes the better-performing funds. The solution to survivorship bias is to use a sample of funds that all started at the same time and not drop funds that have been dropped from the sample.

Look-ahead bias occurs when a study tests a relationship using sample data that was not available on the test date. For example, consider the test of a trading rule that is based on the price-to-book ratio at the end of the fiscal year. Stock prices are available for all companies at the same point in time, while end-of-year book values may not be available until 30 to 60 days after the fiscal year ends. In order to account for this bias, a study that uses price-to-book value ratios to test trading strategies might estimate the book value as reported at fiscal year end and the market value two months later.

Time-period bias can result if the time period over which the data is gathered is either too short or too long. If the time period is too short, research results may reflect phenomena specific to that time period, or perhaps even data mining. If the time period is too long, the fundamental economic relationships that underlie the results may have changed.

For example, research findings may indicate that small stocks outperformed large stocks during 1980–1985. This may well be the result of time-period bias—in this case, using too short a time period. It's not clear whether this relationship will continue in the future or if it is just an isolated occurrence.

On the other hand, a study that quantifies the relationship between inflation and unemployment during the period from 1940 to 2000 will also result in time-period bias—because this period is too long, and it covers a fundamental change in the relationship between inflation and unemployment that occurred in the 1980s. In this case, the data should be divided into two subsamples that span the period before and after the change.



MODULE QUIZ 5.2

1. A random sample of 100 computer store customers spent an average of \$75 at the store. Assuming the distribution is normal and the population standard deviation is \$20, the 95% confidence interval for the population mean is *closest* to:
 - A. \$71.08 to \$78.92.
 - B. \$73.89 to \$80.11.
 - C. \$74.56 to \$79.44.
2. Best Computers, Inc., sells computers and computer parts by mail. A sample of 25 recent orders showed the mean time taken to ship these orders was 70 hours with a sample standard deviation of 14 hours. Assuming the population is normally distributed, the 99% confidence interval for the population mean is:
 - A. 70 ± 2.80 hours.
 - B. 70 ± 6.98 hours.
 - C. 70 ± 7.83 hours.
3. What is the *most appropriate* test statistic for constructing confidence intervals for the population mean when the population is normally distributed, but the variance is unknown?
 - A. The z -statistic at α with n degrees of freedom.
 - B. The t -statistic at $\alpha/2$ with n degrees of freedom.
 - C. The t -statistic at $\alpha/2$ with $n - 1$ degrees of freedom.

4. When constructing a confidence interval for the population mean of a nonnormal distribution when the population variance is unknown and the sample size is large ($n > 30$), an analyst may acceptably use:
 - A. either a z -statistic or a t -statistic.
 - B. only a z -statistic at α with n degrees of freedom.
 - C. only a t -statistic at $\alpha/2$ with n degrees of freedom.
5. Jenny Fox evaluates managers who have a cross-sectional population standard deviation of returns of 8%. If returns are independent across managers, how large of a sample does Fox need so the standard error of sample means is 1.265%?
 - A. 7.
 - B. 30.
 - C. 40.
6. Annual returns on small stocks have a population mean of 12% and a population standard deviation of 20%. If the returns are normally distributed, a 90% confidence interval on mean returns over a 5-year period is:
 - A. 5.40% to 18.60%.
 - B. -2.75% to 26.75%.
 - C. -5.52% to 29.52%.
7. Which of the following techniques to improve the accuracy of confidence intervals on a statistic is *most* computationally demanding?
 - A. The jackknife.
 - B. Systematic resampling.
 - C. Bootstrapping.
8. An analyst who uses historical data that was not publicly available at the time period being studied will have a sample with:
 - A. look-ahead bias.
 - B. time-period bias.
 - C. sample selection bias.
9. Which of the following is *most closely* associated with survivorship bias?
 - A. Price-to-book studies.
 - B. Stratified bond sampling studies.
 - C. Mutual fund performance studies.

KEY CONCEPTS

LOS 5.a

Probability sampling refers to sampling methods based on randomly chosen samples and assuming that members of a population are equally likely to be chosen for the samples.

Non-probability sampling refers to choosing sample data that are not random but based on low cost and availability of the sample data, or specifically chosen based on the experience and judgment of the researcher.

LOS 5.b

Sampling error is the difference between a sample statistic and its corresponding population parameter (e.g., the sample mean minus the population mean).

LOS 5.c

Simple random sampling is a method of selecting a sample in such a way that each item or person in the population being studied has the same probability of being included in the sample.

Stratified random sampling involves randomly selecting samples proportionally from subgroups that are formed based on one or more distinguishing characteristics of the data, so that random samples from the subgroups will have the same distribution of these characteristics as the overall population.

Cluster sampling is also based on subgroups (not necessarily based on data characteristics) of a larger data set. In one-stage cluster sampling, the sample is formed from randomly chosen clusters (subsets) of the overall data set. In two-stage cluster sampling, random samples are taken from each of the randomly chosen clusters (subgroups).

Convenience sampling refers to selecting sample data based on its ease of access, using data that are readily available. Judgmental sampling refers to samples for which each observation is selected from a larger data set by the researcher, based on her experience and judgment. Both are examples of non-probability sampling and are non-random.

LOS 5.d

The central limit theorem states that for a population with a mean μ and a finite variance σ^2 , the sampling distribution of the sample mean of all possible samples of size n (for $n \geq 30$) will be approximately normally distributed with a mean equal to μ and a variance equal to σ^2 / n .

LOS 5.e

The standard error of the sample mean is the standard deviation of the distribution of the sample means and is calculated as $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$, where σ , the population standard deviation, is known, and as $s_{\bar{X}} = \frac{s}{\sqrt{n}}$, where s , the sample standard deviation, is used because the population standard deviation is unknown.

LOS 5.f

Desirable statistical properties of an estimator include unbiasedness (sign of estimation error is random), efficiency (lower sampling error than any other unbiased estimator), and consistency (variance of sampling error decreases with larger sample size).

LOS 5.g

Point estimates are single-value estimates of population parameters. An estimator is a formula used to compute a point estimate.

Confidence intervals are ranges of values, within which the actual value of the parameter will lie with a given probability.

$$\text{confidence interval} = \text{point estimate} \pm (\text{reliability factor} \times \text{standard error})$$

The reliability factor is a number that depends on the sampling distribution of the point estimate and the probability that the point estimate falls in the confidence interval.

LOS 5.h

For a normally distributed population, a confidence interval for its mean can be constructed using a z -statistic when variance is known, and a t -statistic when the variance is unknown. The z -statistic is acceptable in the case of a normal population with an unknown variance if the sample size is large (30+).

In general, we have:

- $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ when the variance is known, and
- $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$ when the variance is unknown and the sample standard deviation must be used.

LOS 5.i

Two resampling techniques to improve our estimates of the distribution of sample statistics are the jackknife and bootstrapping. With the jackknife, we calculate n sample means, one with each observation in a sample of size n removed, and base our estimate on the standard error of sample means of size n . It can remove bias from our estimates based on the sample standard deviation without resampling.

With bootstrapping, we use the distribution of sample means (or other statistics) from a large number of samples of size n , drawn from a large data set. Bootstrapping can improve our estimates of the distribution of various sample statistics and provide such estimates when analytical methods will not.

LOS 5.j

Increasing the sample size will generally improve parameter estimates and narrow confidence intervals. The cost of more data must be weighed against these benefits, and adding data that is not generated by the same distribution will not necessarily improve accuracy or narrow confidence intervals.

Potential mistakes in the sampling method can bias results. These biases include data snooping (significant relationships that have occurred by chance), sample selection bias (selection is nonrandom), look-ahead bias (basing the test at a point in time on data not available at that time), survivorship bias (using only surviving mutual funds, hedge funds, etc.), and time-period bias (the relation does not hold over other time periods).

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 5.1

- C** With cluster sampling, the randomly selected subgroups may have different distributions of the relevant characteristic relative to the entire population. Cluster sampling uses only randomly selected subgroups, whereas stratified random sampling samples all subgroups to match the distribution of characteristics across the entire population. (LOS 5.a)
- C** An example might be the difference between a particular sample mean and the average value of the overall population. (LOS 5.b)
- B** The sampling error is the difference between the population parameter and the sample statistic. (LOS 5.b)
- B** In a simple random sample, each element of the population has an equal probability of being selected. Choice C allows for an equal chance, but only if there are 100 elements in the population from which the random sample is drawn. (LOS 5.c)
- C** Sample sizes of 30 or greater are typically considered large. (LOS 5.d)
- C** The formula for the standard error when the population standard deviation is unknown is $s_{\bar{x}} = \frac{s}{\sqrt{n}}$. (LOS 5.e)
- C** The formula for the standard error when the population standard deviation is known is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. (LOS 5.e)

8. **B** $s_{\bar{X}} = \frac{s}{\sqrt{n}}$. Given $s = 2$, $s_{\bar{X}} = \frac{2}{\sqrt{49}} = \frac{2}{7} = 0.2857$. (LOS 5.e)

9. **C** $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$. Given $\sigma^2 = 25$, $\sigma_{\bar{X}} = \frac{5}{\sqrt{100}} = \frac{5}{10} = 0.5$. (LOS 5.e)

10. **A** Efficiency, consistency, and unbiasedness are desirable properties of an estimator. (LOS 5.f)

Module Quiz 5.2

1. **A** Since the population variance is known and $n \geq 30$, the confidence interval is determined as $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$. $z_{\alpha/2} = z_{0.025} = 1.96$. So, the confidence interval is $75 \pm 1.96(20/10) = 75 \pm 3.92 = 71.08$ to 78.92 . (LOS 5.h)

2. **C** Since the population variance is unknown and $n < 30$, the confidence interval is determined as $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$. Look up $t_{\alpha/2}$ and $df = n - 1$ to get critical t -value. $t_{0.01/2}$ and $df = 24$ is 2.797. So, the confidence interval is $70 \pm 2.797(14 / 5) = 70 \pm 7.83$. (LOS 5.h)

3. **C** Use the t -statistic at $\alpha/2$ and $n - 1$ degrees of freedom when the population variance is unknown. While the z -statistic is acceptable when the sample size is large, sample size is not given here, and the t -statistic is always appropriate under these conditions. (LOS 5.h)

4. **A** When the sample size is large, and the central limit theorem can be relied on to assure a sampling distribution that is normal, either the t -statistic or the z -statistic is acceptable for constructing confidence intervals for the population mean. The t -statistic, however, will provide a more conservative range (wider) at a given level of significance. (LOS 5.h)

5. **C** $1.265 = \frac{8}{\sqrt{n}}$, $n = \left(\frac{8}{1.265}\right)^2 \approx 40$. (LOS 5.h)

6. **B** With a known population standard deviation of returns and a normally distributed population, we can use the z -distribution. The sample mean for a sample of five years will have a standard deviation of $\frac{20}{\sqrt{5}} = 8.94\%$. A 90% confidence interval around the mean return of 12% is $12\% \pm 1.65(8.94\%) = -2.75\%$ to 26.75% . (LOS 5.h)

7. **C** Bootstrapping, repeatedly drawing samples of equal size from a large data set, is more computationally demanding than the jackknife. We have not defined "systematic resampling" as a specific technique. (LOS 5.i)

8. **A** The primary example of look-ahead bias is using year-end financial information in conjunction with market pricing data to compute ratios like the price/earnings (P/E). The E in the denominator is typically not available for 30–60 days after the end of the period. Hence, data that was available on the test date (P) is mixed with information that was not available (E). That is, the P is "ahead" of the E. (LOS 5.j)

9. **C** Mutual fund performance studies are most closely associated with survivorship bias because only the better-performing funds remain in the sample over time. (LOS 5.j)

READING 6

HYPOTHESIS TESTING

EXAM FOCUS

This review addresses common hypothesis tests. Included are tests about population means, population variances, differences in means, mean differences, differences in variances, correlation, and independence. Make sure you understand what a *p*-value is and how to interpret one. The various hypothesis tests have test statistics with different distributions. These distributions include standard normal (z-test), Student's *t*-distribution (*t*-test), chi-square, and *F*-distributions. Candidates should understand the basic characteristics of each of these distributions and how to apply them. Memorize the standard hypothesis testing procedure presented in this review. Finally, there are some non-parametric tests that you should understand and be able to interpret.

MODULE 6.1: HYPOTHESIS TESTS AND TYPES OF ERRORS



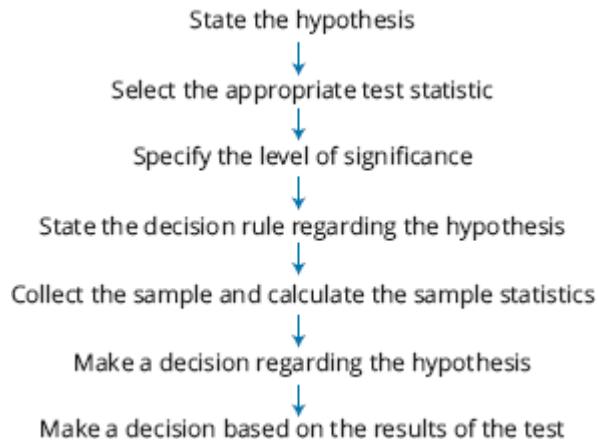
Video covering this content is available online.

LOS 6.a: Define a hypothesis, describe the steps of hypothesis testing, and describe and interpret the choice of the null and alternative hypotheses.

A hypothesis is a statement about the value of a population parameter developed for the purpose of testing a theory or belief. Hypotheses are stated in terms of the population parameter to be tested, like the population mean, μ . For example, a researcher may be interested in the mean daily return on stock options. Hence, the hypothesis may be that the mean daily return on a portfolio of stock options is positive.

Hypothesis testing procedures, based on sample statistics and probability theory, are used to determine whether a hypothesis is a reasonable statement and should not be rejected or if it is an unreasonable statement and should be rejected. The process of hypothesis testing consists of a series of steps shown in Figure 6.1.

Figure 6.1: Hypothesis Testing Procedure*



*Source: Wayne W. Daniel and James C. Terrell, Business Statistics, Basic Concepts and Methodology, Houghton Mifflin, Boston, 1997.

The Null Hypothesis and Alternative Hypothesis

The **null hypothesis**, designated H_0 , is the hypothesis that the researcher wants to reject. It is the hypothesis that is actually tested and is the basis for the selection of the test statistics. The null is generally stated as a simple statement about a population parameter. Typical statements of the null hypothesis for the population mean include $H_0: \mu = \mu_0$, $H_0: \mu \leq \mu_0$, and $H_0: \mu \geq \mu_0$, where μ is the population mean and μ_0 is the hypothesized value of the population mean.



PROFESSOR'S NOTE

The null hypothesis always includes the “equal to” condition.

The **alternative hypothesis**, designated H_a , is what is concluded if there is sufficient evidence to reject the null hypothesis. It is usually the alternative hypothesis that you are really trying to assess. Why? Because you can never really prove anything with statistics, when the null hypothesis is discredited, the implication is that the alternative hypothesis is valid.

LOS 6.b: Compare and contrast one-tailed and two-tailed tests of hypotheses.

The alternative hypothesis can be one-sided or two-sided. A one-sided test is referred to as a **one-tailed test**, and a two-sided test is referred to as a **two-tailed test**. Whether the test is one- or two-sided depends on the proposition being tested. If a researcher wants to test whether the return on stock options is greater than zero, a one-tailed test should be used. However, a two-tailed test should be used if the research question is whether the return on options is simply different from zero. Two-sided tests allow for deviation on both sides of the hypothesized value (zero). In practice, most hypothesis tests are constructed as two-tailed tests.

A **two-tailed test** for the population mean may be structured as:

$$H_0: \mu = \mu_0 \text{ versus } H_a: \mu \neq \mu_0$$

Since the alternative hypothesis allows for values above and below the hypothesized parameter, a two-tailed test uses two **critical values** (or **rejection points**).

The *general decision rule for a two-tailed test* is:

Reject H_0 if:

test statistic > upper critical value or

test statistic < lower critical value

Let's look at the development of the decision rule for a two-tailed test using a z -distributed test statistic (a z -test) at a 5% level of significance, $\alpha = 0.05$.

- At $\alpha = 0.05$, the computed test statistic is compared with the critical z -values of ± 1.96 . The values of ± 1.96 correspond to $\pm z_{\alpha/2} = \pm z_{0.025}$, which is the range of z -values within which 95% of the probability lies. These values are obtained from the cumulative probability table for the standard normal distribution (z -table), which is included at the back of this book.
- If the computed test statistic falls outside the range of critical z -values (i.e., test statistic > 1.96 , or test statistic < -1.96), we reject the null and conclude that the sample statistic is sufficiently different from the hypothesized value.
- If the computed test statistic falls within the range ± 1.96 , we conclude that the sample statistic is not sufficiently different from the hypothesized value ($\mu = \mu_0$ in this case), and we fail to reject the null hypothesis.

The *decision rule* (rejection rule) for a two-tailed z -test at $\alpha = 0.05$ can be stated as:

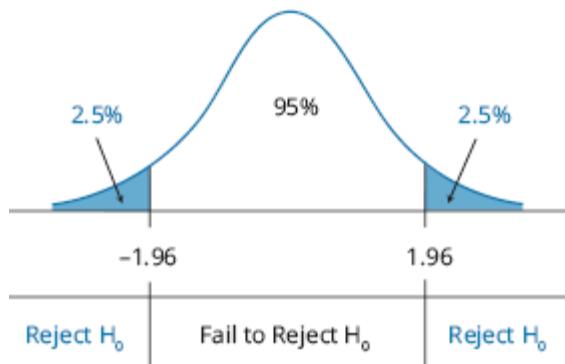
Reject H_0 if:

test statistic < -1.96 or

test statistic > 1.96

Figure 6.2 shows the standard normal distribution for a two-tailed hypothesis test using the z -distribution. Notice that the significance level of 0.05 means that there is $0.05 / 2 = 0.025$ probability (area) under each tail of the distribution beyond ± 1.96 .

Figure 6.2: Two-Tailed Hypothesis Test Using the Standard Normal (z) Distribution



For a **one-tailed hypothesis test** of the population mean, the null and alternative hypotheses are either:

Upper tail: $H_0: \mu \leq \mu_0$ versus $H_a: \mu > \mu_0$, or

Lower tail: $H_0: \mu \geq \mu_0$ versus $H_a: \mu < \mu_0$

The appropriate set of hypotheses depends on whether we believe the population mean, μ , to be greater than (upper tail) or less than (lower tail) the hypothesized value, μ_0 . Using a z -test at the 5% level of significance, the computed test statistic is compared with the critical values of 1.645 for the upper tail tests (i.e., $H_a: \mu > \mu_0$) or -1.645 for lower tail tests (i.e., $H_a: \mu < \mu_0$). These

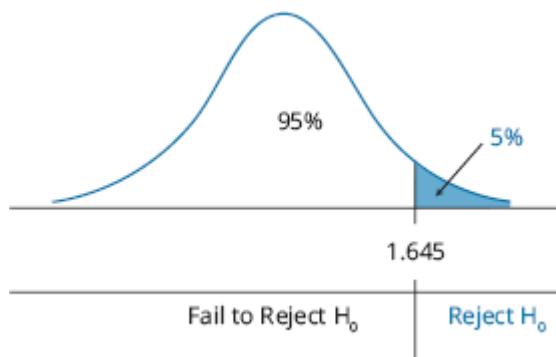
critical values are obtained from a z -table, where $-z_{0.05} = -1.645$ corresponds to a cumulative probability equal to 5%, and the $z_{0.05} = 1.645$ corresponds to a cumulative probability of 95% ($1 - 0.05$).

Let's use the upper tail test structure where $H_0: \mu \leq \mu_0$ and $H_a: \mu > \mu_0$.

- If the calculated test statistic is greater than 1.645, we conclude that the sample statistic is sufficiently greater than the hypothesized value. In other words, we reject the null hypothesis.
- If the calculated test statistic is less than 1.645, we conclude that the sample statistic is not sufficiently different from the hypothesized value, and we fail to reject the null hypothesis.

Figure 6.3 shows the standard normal distribution and the rejection region for a one-tailed test (upper tail) at the 5% level of significance.

Figure 6.3: One-Tailed Hypothesis Test Using the Standard Normal (z) Distribution



The Choice of the Null and Alternative Hypotheses

The most common null hypothesis will be an “equal to” hypothesis. Combined with a “not equal to” alternative, this will require a two-tailed test. The alternative is often the hoped-for hypothesis. The null will include the “equal to” sign and the alternative will include the “not equal to” sign. When the null is that a coefficient is equal to zero, we hope to reject it and show the significance of the relationship.

When the null is less than or equal to, the (mutually exclusive) alternative is framed as greater than, and a one-tail test is appropriate. If we are trying to demonstrate that a return is greater than the risk-free rate, this would be the correct formulation. We will have set up the null and alternative hypothesis so that rejection of the null will lead to acceptance of the alternative, our goal in performing the test. As with a two-tailed test, the null for a one-tailed test will include the “equal to” sign (i.e., either “greater than or equal to” or “less than or equal to”). The alternative will include the opposite sign to the null—either “less than” or “greater than.”

LOS 6.c: Explain a test statistic, Type I and Type II errors, a significance level, how significance levels are used in hypothesis testing, and the power of a test.

Hypothesis testing involves two statistics: the test statistic calculated from the sample data and the *critical value* of the test statistic. The value of the computed test statistic relative to the critical value is a key step in assessing the validity of a hypothesis.

A test statistic is calculated by comparing the point estimate of the population parameter with the hypothesized value of the parameter (i.e., the value specified in the null hypothesis). With reference to our option return example, this means we are concerned with the difference between the mean return of the sample (i.e., $\bar{x} = 0.001$) and the hypothesized mean return (i.e., $\mu_0 = 0$). As indicated in the following expression, the **test statistic** is the difference between the sample statistic and the hypothesized value, scaled by the standard error of the sample statistic.

$$\text{test statistic} = \frac{\text{sample statistic} - \text{hypothesized value}}{\text{standard error of the sample statistic}}$$

The standard error of the sample statistic is the adjusted standard deviation of the sample. When the sample statistic is the sample mean, \bar{x} , the standard error of the sample statistic for sample size n , is calculated as:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

when the population standard deviation, σ , is known, or

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

when the population standard deviation, σ , is not known. In this case, it is estimated using the standard deviation of the sample, s .



PROFESSOR'S NOTE

Don't be confused by the notation here. A lot of the literature you will encounter in your studies simply uses the term $\sigma_{\bar{x}}$ for the standard error of the test statistic, regardless of whether the population standard deviation or sample standard deviation was used in its computation.

As you will soon see, a test statistic is a random variable that may follow one of several distributions, depending on the characteristics of the sample and the population. We will look at four distributions for test statistics: the t -distribution, the z -distribution (standard normal distribution), the chi-square distribution, and the F -distribution. The critical value for the appropriate test statistic—the value against which the computed test statistic is compared—depends on its distribution.

Type I and Type II Errors

Keep in mind that hypothesis testing is used to make inferences about the parameters of a given population on the basis of statistics computed for a sample that is drawn from that population. We must be aware that there is some probability that the sample, in some way, does not represent the population, and any conclusion based on the sample about the population may be made in error.

When drawing inferences from a hypothesis test, there are two types of errors:

- **Type I error:** the rejection of the null hypothesis when it is actually true.
- **Type II error:** the failure to reject the null hypothesis when it is actually false.

The **significance level** is the probability of making a Type I error (rejecting the null when it is true) and is designated by the Greek letter alpha (α). For instance, a significance level of 5% ($\alpha = 0.05$) means there is a 5% chance of rejecting a true null hypothesis. When conducting

hypothesis tests, a significance level must be specified in order to identify the critical values needed to evaluate the test statistic.

The Power of a Test

While the significance level of a test is the probability of rejecting the null hypothesis when it is true, the **power of a test** is the probability of correctly rejecting the null hypothesis when it is false. The power of a test is actually one minus the probability of making a Type II error, or $1 - P(\text{Type II error})$. In other words, the probability of rejecting the null when it is false (power of the test) equals one minus the probability of *not* rejecting the null when it is false (Type II error). When more than one test statistic may be used, the power of the test for the competing test statistics may be useful in deciding which test statistic to use. Ordinarily, we wish to use the test statistic that provides the most powerful test among all possible tests.

Figure 6.4 shows the relationship between the level of significance, the power of a test, and the two types of errors.

Figure 6.4: Type I and Type II Errors in Hypothesis Testing

Decision	True Condition	
	H_0 is true	H_0 is false
Do not reject H_0	Correct decision	Incorrect decision Type II error
Reject H_0	Incorrect decision Type I error Significance level, α , $= P(\text{Type I error})$	Correct decision Power of the test $= 1 - P(\text{Type II error})$

Sample size and the choice of significance level (Type I error probability) will together determine the probability of a Type II error. The relation is not simple, however, and calculating the probability of a Type II error in practice is quite difficult. Decreasing the significance level (probability of a Type I error) from 5% to 1%, for example, will increase the probability of failing to reject a false null (Type II error) and therefore reduce the power of the test.

Conversely, for a given sample size, we can increase the power of a test only with the cost that the probability of rejecting a true null (Type I error) increases. For a given significance level, we can decrease the probability of a Type II error and increase the power of a test, only by increasing the sample size.

LOS 6.d: Explain a decision rule and the relation between confidence intervals and hypothesis tests, and determine whether a statistically significant result is also economically meaningful.

The decision for a hypothesis test is to either reject the null hypothesis or fail to reject the null hypothesis. Note that it is statistically incorrect to say “accept” the null hypothesis; it can only be supported or rejected. The **decision rule** for rejecting or failing to reject the null hypothesis is based on the distribution of the test statistic. For example, if the test statistic follows a normal distribution, the decision rule is based on critical values determined from the standard normal

distribution (z-distribution). Regardless of the appropriate distribution, it must be determined if a one-tailed or two-tailed hypothesis test is appropriate before a decision rule (rejection rule) can be determined.

A decision rule is specific and quantitative. Once we have determined whether a one- or two-tailed test is appropriate, the significance level we require, and the distribution of the test statistic, we can calculate the exact critical value for the test statistic. Then we have a decision rule of the following form: if the test statistic is (greater, less than) the value X, reject the null.

The Relation Between Confidence Intervals and Hypothesis Tests

A confidence interval is a range of values within which the researcher believes the true population parameter may lie.

A confidence interval is determined as:

$$\left\{ \left[\frac{\text{sample statistic}}{\text{value}} - \left(\frac{\text{critical}}{\text{error}} \right) \left(\frac{\text{standard}}{\text{error}} \right) \right] \leq \frac{\text{population parameter}}{\text{sample statistic}} \leq \left[\frac{\text{sample statistic}}{\text{value}} + \left(\frac{\text{critical}}{\text{error}} \right) \left(\frac{\text{standard}}{\text{error}} \right) \right] \right\}$$

The interpretation of a confidence interval is that for a level of confidence of 95%; for example, there is a 95% probability that the true population parameter is contained in the interval.

From the previous expression, we see that a confidence interval and a hypothesis test are linked by the critical value. For example, a 95% confidence interval uses a critical value associated with a given distribution at the 5% level of significance. Similarly, a hypothesis test would compare a test statistic to a critical value at the 5% level of significance. To see this relationship more clearly, the expression for the confidence interval can be manipulated and restated as:

$$-\text{critical value} \leq \text{test statistic} \leq +\text{critical value}$$

This is the range within which we fail to reject the null for a two-tailed hypothesis test at a given level of significance.

EXAMPLE: Confidence intervals and two-tailed hypothesis tests

A researcher has gathered data on the daily returns on a portfolio of call options over a recent 250-day period. The mean daily return has been 0.1%, and the sample standard deviation of daily portfolio returns is 0.25%. The researcher believes that the mean daily portfolio return is not equal to zero.

1. Construct a 95% confidence interval for the population mean daily return over the 250-day sample period.
2. Construct a hypothesis test of the researcher's belief.

Answer:

1. Given a sample size of 250 with a standard deviation of 0.25%, the standard error can be computed as $s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{0.25\%}{\sqrt{250}} = 0.0158\%$.

At the 5% level of significance, the critical z -values for the confidence interval are $z_{0.025} = 1.96$ and $-z_{0.025} = -1.96$. Thus, given a sample mean equal to 0.1%, the 95% confidence interval for the population mean is:

$$0.1 - 1.96(0.0158) \leq \mu \leq 0.1 + 1.96(0.0158), \text{ or}$$

$$0.069\% \leq \mu \leq 0.131\%$$

2. First we need to specify the null and alternative hypotheses. Note that the null hypothesis must contain “equal to.” That is, it can be $=$, \leq , or \geq .

$$H_0: \mu_0 = 0 \text{ versus } H_a: \mu_0 \neq 0$$

Since the null hypothesis is an equality, this is a two-tailed test. At a 5% level of significance, the critical z -values for a two-tailed test are ± 1.96 , so the decision rule can be stated as:

Reject H_0 if test statistic < -1.96 or test statistic $> +1.96$

Using the standard error of the sample mean we calculated above, our test statistic is:

$$\frac{0.001}{\left(\frac{0.0025}{\sqrt{250}}\right)} = \frac{0.001}{0.000158} = 6.33$$

Since $6.33 > 1.96$, we reject the null hypothesis that the mean daily option return is equal to zero.

Notice the similarity of this analysis with our confidence interval. We rejected the hypothesis $\mu = 0$ because the sample mean of 0.1% is more than 1.96 standard errors from zero. Based on the 95% confidence interval, we reject $\mu = 0$ because zero is more than 1.96 standard errors from the sample mean of 0.1%.

Whether a Statistically Significant Result is Also Economically Meaningful

Statistical significance does not necessarily imply **economic significance**. For example, we may have tested a null hypothesis that a strategy of going long all the stocks that satisfy some criteria and shorting all the stocks that do not satisfy the criteria resulted in returns that were less than or equal to zero over a 20-year period. Assume we have rejected the null in favor of the alternative hypothesis that the returns to the strategy are greater than zero (positive). This does not necessarily mean that investing in that strategy will result in economically meaningful positive returns. Several factors must be considered.

One important consideration is transactions costs. Once we consider the costs of buying and selling the securities, we may find that the mean positive returns to the strategy are not enough to generate positive returns. Taxes are another factor that may make a seemingly attractive strategy a poor one in practice. A third reason that statistically significant results may not be economically significant is risk. In the above strategy, we have additional risk from short sales (they may have to be closed out earlier than in the test strategy). Since the statistically significant results were for a period of 20 years, it may be the case that there is significant variation from year to year in the returns from the strategy, even though the mean strategy

return is greater than zero. This variation in returns from period to period is an additional risk to the strategy that is not accounted for in our test of statistical significance.

Any of these factors could make committing funds to a strategy unattractive, even though the statistical evidence of positive returns is highly significant. By the nature of statistical tests, a very large sample size can result in highly (statistically) significant results that are quite small in absolute terms.



MODULE QUIZ 6.1

1. To test whether the mean of a population is greater than 20, the appropriate null hypothesis is that the population mean is:
 - A. less than 20.
 - B. greater than 20.
 - C. less than or equal to 20.
2. Which of the following statements about hypothesis testing is *most accurate*?
 - A Type II error is rejecting the null when it is actually true.
 - The significance level equals one minus the probability of a Type I error.
 - A two-tailed test with a significance level of 5% has z -critical values of ± 1.96 .
3. For a hypothesis test with a probability of a Type II error of 60% and a probability of a Type I error of 5%, which of the following statements is *most accurate*?
 - The power of the test is 40%, and there is a 5% probability that the test statistic will exceed the critical value(s).
 - There is a 95% probability that the test statistic will be between the critical values if this is a two-tailed test.
 - There is a 5% probability that the null hypothesis will be rejected when actually true, and the probability of rejecting the null when it is false is 40%.
4. If the significance level of a test is 0.05 and the probability of a Type II error is 0.15, what is the power of the test?
 - 0.850.
 - 0.950.
 - 0.975.

MODULE 6.2: *P*-VALUES AND TESTS OF MEANS



LOS 6.e: Explain and interpret the *p*-value as it relates to hypothesis testing.

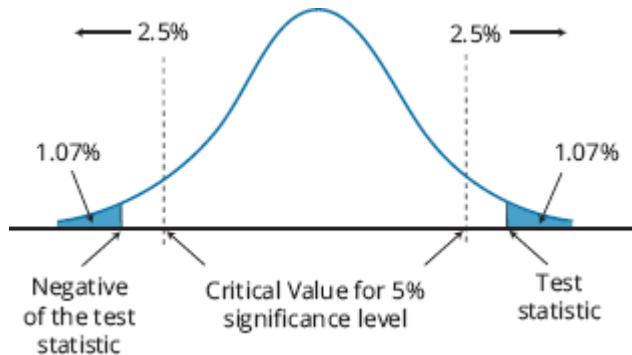
Video covering this content is available online.

The ***p*-value** is the probability of obtaining a test statistic that would lead to a rejection of the null hypothesis, assuming the null hypothesis is true. It is the smallest level of significance for which the null hypothesis can be rejected. For one-tailed tests, the *p*-value is the probability that lies above the computed test statistic for upper tail tests or below the computed test statistic for lower tail tests. For two-tailed tests, the *p*-value is the probability that lies above the positive value of the computed test statistic *plus* the probability that lies below the negative value of the computed test statistic.

Consider a two-tailed hypothesis test about the mean value of a random variable at the 5% significance level where the test statistic is 2.3, greater than the upper critical value of 1.96. If we consult the Z-table, we find the probability of getting a value greater than 2.3 is $(1 - 0.9893)$

$= 1.07\%$. Since it's a two-tailed test, our p -value is $2 \times 1.07 = 2.14\%$, as illustrated in Figure 6.5. At a 3%, 4%, or 5% significance level, we would reject the null hypothesis, but at a 2% or 1% significance level, we would not. Many researchers report p -values without selecting a significance level and allow the reader to judge how strong the evidence for rejection is.

Figure 6.5: Two-Tailed Hypothesis Test With p -Value = 2.14%



LOS 6.f: Describe how to interpret the significance of a test in the context of multiple tests.

Recall that the probability of a Type I error is the probability that a true null hypothesis will be rejected (and is also the significance level of a test). Statisticians refer to these incorrect rejections of the null hypothesis as **false positives**. For a test of the hypothesis that the mean return to an investment strategy is equal to zero, with a significance level of 5%, we will get a false positive 5% of the time, on average. That is, our test statistic will be outside the critical values (in the tails of the distribution) and the p -value of our test statistic will be less than 0.05. If we do a single test, this conclusion is correct.

While we might think that if we get more than 5 false positives with 100 tests, we should reject the null hypothesis, with multiple tests this may be misleading. There is an adjustment to the p -values of our tests, however, that will improve the accuracy of our conclusions.

The procedure is illustrated in the following example. Consider 20 tests at the 10% significance level, for which we get 4 test statistics with p -values less than 10%. This is more than the 10% of 20 that we would expect to get if the null hypothesis is true, and we might believe we should reject the null based on these results. The accepted method to use in this case is to rank the p -values in ascending order and calculate an adjusted significance number for each test with a p -value less than 10%. We then compare these adjusted significance numbers to the reported p -values. We will only count tests as actual rejections if their adjusted significance based on p -value rank is greater than or equal to their reported p -values.

Figure 6.6 illustrates an example of this procedure. We list the p -values for the tests that are less than 10% in ascending order. The adjusted significance levels are calculated with the following formula:

$$\text{Adjusted significance} = \alpha \times \frac{\text{Rank of } p\text{-value}}{\text{Number of tests}}$$

Figure 6.6: Adjusted significance

Rank (Test #)	<i>p</i> -value	Adjusted Significance	<i>p</i> -value \leq adjusted significance
1 (Test 12)	0.004	0.005	Yes
2 (Test 4)	0.010	0.010	Yes
3 (Test 17)	0.053	0.015	No
4 (Test 9)	0.076	0.020	No

From the results reported in Figure 6.6, we see that only two of the tests should actually be counted as rejections. Because only 2 of the 20 tests (tests 12 and 4) qualify as actual rejections based on comparison of their *p*-values with the adjusted significance values for their rank, our rejection rate is 10%. When the null hypothesis is true, two rejections from 20 tests is just what we would expect with a significance level of 10%. In this case, we will not reject the null hypothesis.



PROFESSOR'S NOTE

The LOS here says, “Describe how to interpret the significance of a test ...” It does not indicate that calculations will be required. Perhaps if you just remember that we compare the reported *p*-values (ranked from lowest to highest) to the adjusted significance levels (significance level times rank / number of tests), and then count only those tests for which the *p*-values are less than their adjusted significance levels as rejections, you’ll be able to handle any questions based on this LOS.

LOS 6.g: Identify the appropriate test statistic and interpret the results for a hypothesis test concerning the population mean of both large and small samples when the population is normally or approximately normally distributed and the variance is (1) known or (2) unknown.

When hypothesis testing, the choice between using a critical value based on the *t*-distribution or the *z*-distribution depends on sample size, the distribution of the population, and whether or not the variance of the population is known.

The *t*-Test

The *t*-test is a widely used hypothesis test that employs a test statistic that is distributed according to a *t*-distribution. Following are the rules for when it is appropriate to use the *t*-test for hypothesis tests of the population mean.

Use the t-test if the population variance is unknown and either of the following conditions exist:

- The sample is large ($n \geq 30$).
- The sample is small (less than 30), but the distribution of the population is normal or approximately normal.

If the sample is small and the distribution is nonnormal, we have no reliable statistical test.

The computed value for the test statistic based on the *t*-distribution is referred to as the *t*-statistic. For hypothesis tests of a population mean, a *t*-statistic with $n - 1$ degrees of freedom is computed as:

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

where:
 \bar{x} = sample mean

μ_0 = hypothesized population mean (i.e., the null)

s = standard deviation of the sample

n = sample size



PROFESSOR'S NOTE

This computation is not new. It is the same test statistic computation that we have been performing all along. Note the use of the sample standard deviation, s , in the standard error term in the denominator.

To conduct a t -test, the t -statistic is compared to a critical t -value at the desired level of significance with the appropriate degrees of freedom.

In the real world, the underlying variance of the population is rarely known, so the t -test enjoys widespread application.

The z -Test

The z -test is the appropriate hypothesis test of the population mean when the *population is normally distributed with known variance*. The computed test statistic used with the z -test is referred to as the z -statistic. The z -statistic for a hypothesis test for a population mean is computed as follows:

$$z\text{-statistic} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

where:
 \bar{x} = sample mean

μ_0 = hypothesized population mean

σ = standard deviation of the *population*

n = sample size

To test a hypothesis, the z -statistic is compared to the critical z -value corresponding to the significance of the test. Critical z -values for the most common levels of significance are displayed in Figure 6.7. You should have these memorized by now.

Figure 6.7: Critical z -Values

Level of Significance	Two-Tailed Test	One-Tailed Test
0.10 = 10%	± 1.65	+1.28 or -1.28
0.05 = 5%	± 1.96	+1.65 or -1.65
0.01 = 1%	± 2.58	+2.33 or -2.33

When the *sample size is large* and the *population variance is unknown*, the z -statistic is:

$$z\text{-statistic} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

where:

\bar{x} = sample mean

μ_0 = hypothesized population mean

s = standard deviation of the *sample*

n = sample size

Note the use of the sample standard deviation, s , versus the population standard deviation, σ . Remember, this is acceptable if the sample size is large, although the t -statistic is the more conservative measure when the population variance is unknown.

EXAMPLE: z-test or t-test?

Referring to our previous option portfolio mean return example, determine which test statistic (z or t) should be used.

Answer:

The population variance for our sample of returns is unknown. Hence, the t -distribution is appropriate. With 250 observations, however, the sample is considered to be large, so the z -distribution would also be acceptable. Because our sample is so large, the critical values for the t and z are almost identical. Hence, there is almost no difference in the likelihood of rejecting a true null.

EXAMPLE: The z-test

When your company's gizmo machine is working properly, the mean length of gizmos is 2.5 inches. However, from time to time the machine gets out of alignment and produces gizmos that are either too long or too short. When this happens, production is stopped and the machine is adjusted. To check the machine, the quality control department takes a gizmo sample each day. Today, a random sample of 49 gizmos showed a mean length of 2.49 inches. The population standard deviation is known to be 0.021 inches. Using a 5% significance level, determine if the machine should be shut down and adjusted.

Answer:

Let μ be the mean length of all gizmos made by this machine, and let x be the corresponding mean for the sample.

Let's follow the hypothesis testing procedure presented earlier in Figure 6.1. Again, you should know this process!

Statement of hypothesis. For the information provided, the null and alternative hypotheses are appropriately structured as:

$H_0: \mu = 2.5$ (The machine does not need an adjustment.)

$H_a: \mu \neq 2.5$ (The machine needs an adjustment.)

Note that since this is a two-tailed test, H_a allows for values above and below 2.5.

Select the appropriate test statistic. Since the population variance is known and the sample size is > 30 , the z-statistic is the appropriate test statistic. The z-statistic is computed as:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Specify the level of significance. The level of significance is given at 5%, implying that we are willing to accept a 5% probability of rejecting a true null hypothesis.

State the decision rule regarding the hypothesis. The \neq sign in the alternative hypothesis indicates that the test is two-tailed with two rejection regions, one in each tail of the standard normal distribution curve. Because the total area of both rejection regions combined is 0.05 (the significance level), the area of the rejection region in each tail is 0.025. You should know that the critical z-values for $\pm z_{0.025}$ are ± 1.96 . This means that the null hypothesis should not be rejected if the computed z-statistic lies between -1.96 and $+1.96$ and should be rejected if it lies outside of these critical values. The decision rule can be stated as:

Reject H_0 if $-z_{0.025} > z\text{-statistic} > z_{0.025}$, or equivalently,

Reject H_0 if: $-1.96 > z\text{-statistic} > +1.96$

Collect the sample and calculate the test statistic. The value of x from the sample is 2.49. Since σ is given as 0.021, we calculate the z-statistic using σ as follows:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{2.49 - 2.5}{0.021 / \sqrt{49}} = \frac{-0.01}{0.003} = -3.33$$

Make a decision regarding the hypothesis. The calculated value of the z-statistic is -3.33 . Since this value is less than the critical value, $-z_{0.025} = -1.96$, it falls in the rejection region in the left tail of the z-distribution. Hence, there is sufficient evidence to reject H_0 .

Make a decision based on the results of the test. Based on the sample information and the results of the test, it is concluded that the machine is out of adjustment and should be shut down for repair.

EXAMPLE: One-tailed test

Using the data from the previous example and a 5% significance level, test the hypothesis that the mean length of gizmos is less than 2.5 inches.

Answer:

In this case, we use a one-tailed test with the following structure:

$$H_0: \mu \geq 2.5 \text{ versus } H_a: \mu < 2.5$$

The appropriate decision rule for this one-tailed test at a significance level of 5% is:

Reject H_0 if test statistic < -1.645

The test statistic is computed in the same way, regardless of whether we are using a one-tailed or a two-tailed test. From the previous example, we know that the test statistic for the gizmo

sample is -3.33 . Because $-3.33 < -1.645$, we can reject the null hypothesis and conclude that the mean length is statistically less than 2.5 at a 5% level of significance.



MODULE QUIZ 6.2

1. A researcher has 28 quarterly excess returns to an investment strategy and believes these returns are approximately normally distributed. The mean return on this sample is 1.645% and the standard deviation is 5.29% . For a test with a 5% significance level of the hypothesis that excess returns are less than or equal to zero, the researcher should:
 - A. reject the null hypothesis because the critical value for the test is 1.645 .
 - B. not draw any conclusion because the sample size is less than 30 .
 - C. fail to reject the null because the critical value is greater than 1.645 .
2. An analyst wants to test a hypothesis concerning the population mean of monthly returns for a composite that has existed for 24 months. The analyst may appropriately use:
 - A. a t -test but not a z -test if returns for the composite are normally distributed.
 - B. either a t -test or a z -test if returns for the composite are normally distributed.
 - C. a t -test but not a z -test, regardless of the distribution of returns for the composite.

Use the following segment of Student's t -distribution for Question 3.

Level of Significance for One-Tailed Test				
df	0.100	0.050	0.025	0.01
Level of Significance for Two-Tailed Test				
df	0.20	0.10	0.05	0.02
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602

3. From a sample of 14 observations, an analyst calculates a t -statistic to test a hypothesis that the population mean is equal to zero. If the analyst chooses a 5% significance level, the appropriate critical value is:
 - A. less than 1.80 .
 - B. greater than 2.15 .
 - C. between 1.80 and 2.15 .

MODULE 6.3: MEAN DIFFERENCES AND DIFFERENCE IN MEANS



Video covering this content is available online.

LOS 6.h: Identify the appropriate test statistic and interpret the results for a hypothesis test concerning the equality of the population means of two at least approximately normally distributed populations based on independent random samples with equal assumed variances.

Up to this point, we have been concerned with tests of a single population mean. In practice, we frequently want to know if there is a difference between the means of two populations. The t -

test for differences between means requires that we are reasonably certain that our samples are independent and that they are taken from two populations that are normally distributed.



PROFESSOR'S NOTE

Please note the language of the LOS here. Candidates must “Identify the appropriate test statistic and interpret the results of a hypothesis test....” Certainly you should know that this is a *t*-test, and that we reject the hypothesis of equality when the test statistic is outside the critical *t*-values. Don’t worry about memorizing the following formulas.

A pooled variance is used with the *t*-test for testing the hypothesis that the means of two normally distributed populations are equal, when the variances of the populations are unknown but assumed to be equal.

Assuming independent samples, the *t*-statistic is computed as:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

where:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

s_1^2 = variance of the first sample

s_2^2 = variance of the second sample

n_1 = number of observations in the first sample

n_2 = number of observations in the second sample

Note: The degrees of freedom, *df*, is $(n_1 + n_2 - 2)$.

Since we assume that the variances are equal, we just add the variances of the two sample means in order to calculate the standard error in the denominator.

The intuition here is straightforward. If the sample means are very close together, the numerator of the *t*-statistic (and the *t*-statistic itself) are small, and we do not reject equality. If the sample means are far apart, the numerator of the *t*-statistic (and the *t*-statistic itself) are large, and we reject equality. Perhaps not as easy to remember is the fact that this test is only valid for two populations that are independent and normally distributed.

EXAMPLE: Difference between means – equal variances

Sue Smith is investigating whether the abnormal returns for acquiring firms during merger announcement periods differ for horizontal and vertical mergers. She estimates the abnormal returns for a sample of acquiring firms associated with horizontal mergers and a sample of acquiring firms involved in vertical mergers. Smith finds that abnormal returns from horizontal mergers have a mean of 1.0% and a standard deviation of 1.0%, while abnormal returns from vertical mergers have a mean of 2.5% and a standard deviation of 2.0%.

Smith assumes that the samples are independent, the population means are normally distributed, and the population variances are equal.

Smith calculates the t -statistic as -5.474 and the degrees of freedom as 120 . Using a 5% significance level, should Smith reject or fail to reject the null hypothesis that the abnormal returns to acquiring firms during the announcement period are the same for horizontal and vertical mergers?

Answer:

Since this is a two-tailed test, the structure of the hypotheses takes the following form:

$$H_0: \mu_1 - \mu_2 = 0 \text{ versus } H_a: \mu_1 - \mu_2 \neq 0$$

where:

μ_1 = the mean of the abnormal returns for the horizontal mergers

μ_2 = the mean of the abnormal returns for the vertical mergers

From the following t -table segment, the critical t -value for a 5% level of significance at $\alpha / 2 = p = 0.025$ with $df = 120$, is 1.980 .

Partial t -Table

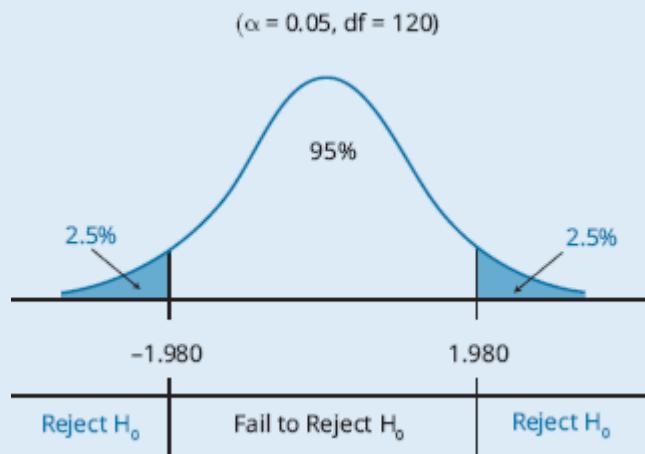
df	One-Tailed Probabilities (p)		
	$p = 0.10$	$p = 0.05$	$p = 0.025$
110	1.289	1.659	1.982
120	1.289	1.658	1.980
200	1.286	1.653	1.972

Thus, the decision rule can be stated as:

Reject H_0 if t -statistic < -1.980 or t -statistic > 1.980

The rejection region for this test is illustrated in the following figure.

Decision Rule for Two-Tailed t -Test



Since the test statistic, -5.474 , falls to the left of the lower critical t -value, Smith can reject the null hypothesis and conclude that mean abnormal returns are different for horizontal and vertical mergers.

LOS 6.i: Identify the appropriate test statistic and interpret the results for a hypothesis test concerning the mean difference of two normally distributed populations.

While the test in the previous section was of the difference between the means of two independent samples, sometimes our samples may be dependent. If the observations in the two samples both depend on some other factor, we can construct a “paired comparisons” test of whether the means of the differences between observations for the two samples are different. Dependence may result from an event that affects both sets of observations for a number of companies or because observations for two firms over time are both influenced by market returns or economic conditions.

For an example of a paired comparisons test, consider a test of whether the returns on two steel firms were equal over a 5-year period. We can't use the difference in means test because we have reason to believe that the samples are not independent. To some extent, both will depend on the returns on the overall market (market risk) and the conditions in the steel industry (industry-specific risk). In this case, our pairs will be the returns on each firm over the same time periods, so we use the differences in monthly returns for the two companies. The paired comparisons test is just a test of whether the average difference between monthly returns is significantly different from zero, based on the standard error of the differences in monthly returns.

Remember, the paired comparisons test also requires that the sample data be normally distributed. Although we frequently just want to test the hypothesis that the mean of the differences in the pairs is zero ($\mu_{dz} = 0$), the general form of the test for any hypothesized mean difference, μ_{dz} , is as follows:

$$H_0: \mu_d = \mu_{dz} \text{ versus } H_a: \mu_d \neq \mu_{dz}$$

where:

μ_d = mean of the population of paired differences

μ_{dz} = hypothesized mean of paired differences, which is commonly zero

For one-tail tests, the hypotheses are structured as either:

$$H_0: \mu_d \leq \mu_{dz} \text{ versus } H_a: \mu_d > \mu_{dz}, \text{ or } H_0: \mu_d \geq \mu_{dz} \text{ versus } H_a: \mu_d < \mu_{dz}$$

For the paired comparisons test, the *t*-statistic with $n - 1$ degrees of freedom is computed as:

$$t = \frac{\bar{d} - \mu_{dz}}{s_{\bar{d}}}$$

where:

$$\bar{d} = \text{sample mean difference} = \frac{1}{n} \sum_{i=1}^n d_i$$

d_i = difference between the i th pair of observations

$$s_{\bar{d}} = \text{standard error of the mean difference} = \frac{s_d}{\sqrt{n}}$$

$$s_d = \text{sample standard deviation} = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$$

n = the number of paired observations

EXAMPLE: Paired comparisons test

Joe Andrews is examining changes in estimated betas for the common stock of companies in the telecommunications industry before and after deregulation. Andrews believes that the betas may decline because of deregulation since companies are no longer subject to the uncertainties of rate regulation or that they may increase because there is more uncertainty regarding competition in the industry. Andrews calculates a t -statistic of 10.26 for this hypothesis test, based on a sample size of 39. Using a 5% significance level, determine whether there is a change in betas.

Answer:

Because the mean difference may be positive or negative, a two-tailed test is in order here. Thus, the hypotheses are structured as:

$$H_0: \mu_d = 0 \text{ versus } H_a: \mu_d \neq 0$$

There are $39 - 1 = 38$ degrees of freedom. Using the t -distribution, the two-tailed critical t -values for a 5% level of significance with $df = 38$ is ± 2.024 . As indicated in the following table, the critical t -value of 2.024 is located at the intersection of the $p = 0.025$ column and the $df = 38$ row. The one-tailed probability of 0.025 is used because we need 2.5% in each tail for 5% significance with a two-tailed test.

Partial t -Table

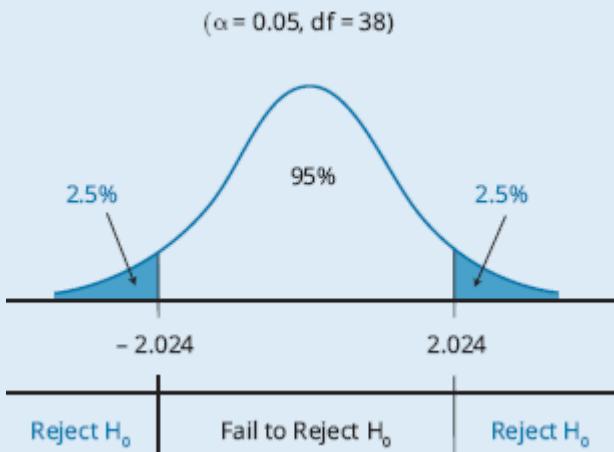
df	One-Tailed Probabilities (p)		
	$p = 0.10$	$p = 0.05$	$p = 0.025$
38	1.304	1.686	2.024
39	1.304	1.685	2.023
40	1.303	1.684	2.021

Thus, the decision rule becomes:

Reject H_0 if t -statistic < -2.024 , or t -statistic > 2.024

This decision rule is illustrated in the following figure.

Decision Rule for a Two-Tailed Paired Comparisons Test



The test statistic, 10.26, is greater than the critical t -value, 2.024—it falls in the rejection region to the right of 2.024 in the previous figure. Thus, we reject the null hypothesis of no

difference, concluding that there *is* a statistically significant difference between mean firm betas before and after deregulation.

Keep in mind that we have been describing two distinct hypothesis tests, one about the significance of the difference between the means of two populations and one about the significance of the mean of the differences between pairs of observations. Here are rules for when these tests may be applied:

- The test of the differences in means is used when there are two *independent samples*.
- A test of the significance of the mean of the differences between paired observations is used when the samples are *not independent*.



PROFESSOR'S NOTE

The LOS here say "Identify the appropriate test statistic and interpret the results ..." I can't believe candidates are expected to memorize these formulas (or that you would be a better analyst if you did). You should instead focus on the fact that both of these tests involve *t*-statistics and depend on the degrees of freedom. Also note that when samples are independent, you can use the difference in means test, and when they are dependent, we must use the paired comparison (mean differences) test. In that case, with a null hypothesis that there is no difference in means, the test statistic is simply the mean of the differences between each pair of observations, divided by the standard error of those differences. This is just a straightforward *t*-test of whether the mean of a sample is zero, which might be considered "fair game" for the exam.



MODULE QUIZ 6.3

1. Which of the following assumptions is *least likely* required for the difference in means test based on two samples?
 - A. The two samples are independent.
 - B. The two populations are normally distributed.
 - C. The two populations have equal variances.
2. William Adams wants to test whether the mean monthly returns over the last five years are the same for two stocks. If he assumes that the returns distributions are normal and have equal variances, the type of test and test statistic are *best* described as:
 - A. paired comparisons test, *t*-statistic.
 - B. paired comparisons test, *F*-statistic.
 - C. difference in means test, *t*-statistic.

MODULE 6.4: TESTS OF VARIANCE, CORRELATION, AND INDEPENDENCE



Video covering this content is available online.

LOS 6.j: Identify the appropriate test statistic and interpret the results for a hypothesis test concerning (1) the variance of a normally distributed population and (2) the equality of the variances of two normally distributed populations based on two independent random samples.

The *chi-square test* is used for hypothesis tests concerning the variance of a normally distributed population. Letting σ^2 represent the true population variance and σ_0^2 represent the hypothesized variance, the hypotheses for a two-tailed test of a single population variance are structured as:

$$H_0: \sigma^2 = \sigma_0^2 \text{ versus } H_a: \sigma^2 \neq \sigma_0^2$$

The hypotheses for one-tailed tests are structured as:

$$H_0: \sigma^2 \leq \sigma_0^2 \text{ versus } H_a: \sigma^2 > \sigma_0^2 \text{ or}$$

$$H_0: \sigma^2 \geq \sigma_0^2 \text{ versus } H_a: \sigma^2 < \sigma_0^2$$

Hypothesis testing of the population variance requires the use of a chi-square distributed test statistic, denoted χ^2 . The **chi-square distribution** is asymmetrical and approaches the normal distribution in shape as the degrees of freedom increase.

To illustrate the chi-square distribution, consider a two-tailed test with a 5% level of significance and 30 degrees of freedom. As displayed in Figure 6.8, the critical chi-square values are 16.791 and 46.979 for the lower and upper bounds, respectively. These values are obtained from a chi-square table, which is used in the same manner as a *t*-table. A portion of a chi-square table is presented in Figure 6.9.

Note that the chi-square values in Figure 6.9 correspond to the probabilities in the right tail of the distribution. As such, the 16.791 in Figure 6.8 is from the column headed 0.975 because 95% + 2.5% of the probability is to the right of it. The 46.979 is from the column headed 0.025 because only 2.5% probability is to the right of it. Similarly, at a 5% level of significance with 10 degrees of freedom, Figure 6.9 shows that the critical chi-square values for a two-tailed test are 3.247 and 20.483.

Figure 6.8: Decision Rule for a Two-Tailed Chi-Square Test

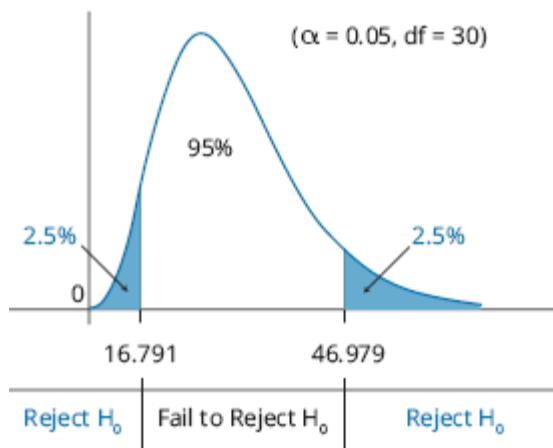


Figure 6.9: Chi-Square Table

Degrees of Freedom	Probability in Right Tail					
	0.975	0.95	0.90	0.1	0.05	0.025
9	2.700	3.325	4.168	14.684	16.919	19.023
10	3.247	3.940	4.865	15.987	18.307	20.483
11	3.816	4.575	5.578	17.275	19.675	21.920
30	16.791	18.493	20.599	40.256	43.773	46.979

The chi-square test statistic, χ^2 , with $n - 1$ degrees of freedom, is computed as:

$$\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma_0^2}$$

where:

n = sample size

s^2 = sample variance

σ_0^2 = hypothesized value for the population variance.

Similar to other hypothesis tests, the chi-square test compares the test statistic, χ^2_{n-1} , to a critical chi-square value at a given level of significance and $n - 1$ degrees of freedom. Note that since the chi-square distribution is bounded below by zero, chi-square values cannot be negative.

EXAMPLE: Chi-square test for a single population variance

Historically, High-Return Equity Fund has advertised that its monthly returns have a standard deviation equal to 4%. This was based on estimates from the 2005–2013 period. High-Return wants to verify whether this claim still adequately describes the standard deviation of the fund's returns. High-Return collected monthly returns for the 24-month period between 2013 and 2015 and measured a standard deviation of monthly returns of 3.8%. High-Return calculates a test statistic of 20.76. Using a 5% significance level, determine if the more recent standard deviation is different from the advertised standard deviation.

Answer:

The null hypothesis is that the standard deviation is equal to 4% and, therefore, the variance of monthly returns for the population is $(0.04)^2 = 0.0016$. Since High-Return simply wants to test whether the standard deviation has changed, up or down, a two-sided test should be used. The hypothesis test structure takes the form:

$$H_0: \sigma_0^2 = 0.0016 \text{ versus } H_a: \sigma^2 \neq 0.0016$$

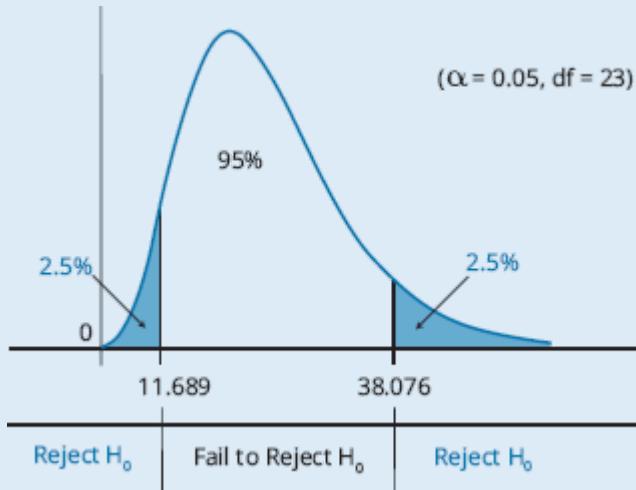
The appropriate test statistic for tests of variance is a chi-square statistic.

With a 24-month sample, there are 23 degrees of freedom. Using the table of chi-square values in Appendix E at the back of this book, for 23 degrees of freedom and probabilities of 0.975 and 0.025, we find two critical values, 11.689 and 38.076. Thus, the decision rule is:

Reject H_0 if $\chi^2 < 11.689$, or $\chi^2 > 38.076$

This decision rule is illustrated in the following figure.

Decision Rule for a Two-Tailed Chi-Square Test of a Single Population Variance



Since the computed test statistic, χ^2 , falls between the two critical values, we cannot reject the null hypothesis that the variance is equal to 0.0016. The recently measured standard deviation is close enough to the advertised standard deviation that we cannot say that it is different from 4%, at a 5% level of significance.

Testing the Equality of the Variances of Two Normally Distributed Populations, Based on Two Independent Random Samples

The hypotheses concerned with the equality of the variances of two populations are tested with an F -distributed test statistic. Hypothesis testing using a test statistic that follows an F -distribution is referred to as the F -test. The F -test is used under the assumption that the populations from which samples are drawn are normally distributed and that the samples are independent.

If we let σ_1^2 and σ_2^2 represent the variances of normal Population 1 and Population 2, respectively, the hypotheses for the two-tailed F -test of differences in the variances can be structured as:

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ versus } H_a: \sigma_1^2 \neq \sigma_2^2$$

and the one-sided test structures can be specified as:

$$H_0: \sigma_1^2 \leq \sigma_2^2 \text{ versus } H_a: \sigma_1^2 > \sigma_2^2, \text{ or } H_0: \sigma_1^2 \geq \sigma_2^2 \text{ versus } H_a: \sigma_1^2 < \sigma_2^2$$

The test statistic for the F -test is the ratio of the sample variances. The F -statistic is computed as:

$$F = \frac{s_1^2}{s_2^2}$$

where:

s_1^2 = variance of the sample of n_1 observations drawn from Population 1

s_2^2 = variance of the sample of n_2 observations drawn from Population 2

Note that $n_1 - 1$ and $n_2 - 1$ are the degrees of freedom used to identify the appropriate critical value from the F -table (provided in the Appendix).

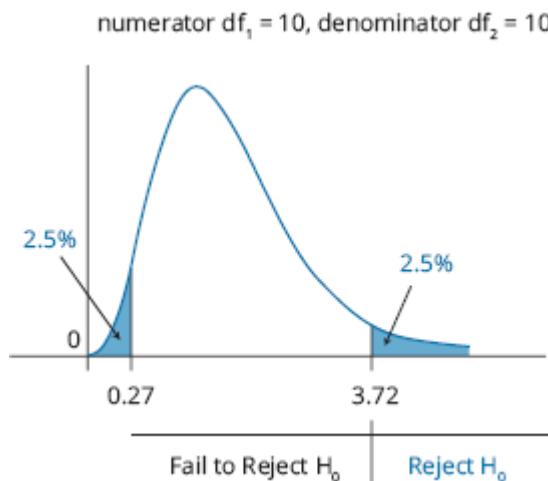


PROFESSOR'S NOTE

Always put the larger variance in the numerator (s_1^2). Following this convention means we only have to consider the critical value for the right-hand tail.

An F -distribution is presented in Figure 6.10. As indicated, the F -distribution is right-skewed and is bounded by zero on the left-hand side. The shape of the F -distribution is determined by *two separate degrees of freedom*, the numerator degrees of freedom, df_1 , and the denominator degrees of freedom, df_2 .

Figure 6.10: F -Distribution



Note that when the sample variances are equal, the value of the test statistic is 1. The upper critical value is always greater than one (the numerator is significantly greater than the denominator), and the lower critical value is always less than one (the numerator is significantly smaller than the denominator). In fact, the lower critical value is the reciprocal of the upper critical value. For this reason, in practice we put the larger sample variance in the numerator and consider only the upper critical value.

EXAMPLE: F -test for equal variances

Annie Cower is examining the earnings for two different industries. Cower suspects that the variance of earnings in the textile industry is different from the variance of earnings in the paper industry. To confirm this suspicion, Cower has looked at a sample of 31 textile manufacturers and a sample of 41 paper companies. She measured the sample standard deviation of earnings across the textile industry to be \$4.30 and that of the paper industry companies to be \$3.80. Cower calculates a test statistic of 1.2805. Using a 5% significance level, determine if the earnings of the textile industry have a different standard deviation than those of the paper industry.

Answer:

In this example, we are concerned with whether the variance of earnings for companies in the textile industry is equal to the variance of earnings for companies in the paper industry. As such, the test hypotheses can be appropriately structured as:

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ versus } H_a: \sigma_1^2 \neq \sigma_2^2$$

For tests of difference between variances, the appropriate test statistic is:

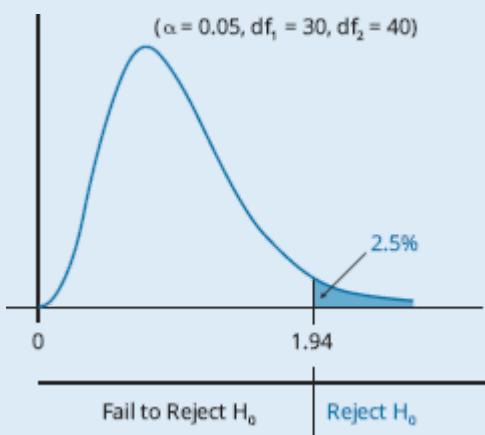
$$F = \frac{s_1^2}{s_2^2}$$

where s_1^2 is the larger sample variance.

Using the sample sizes for the two industries, the critical F -value for our test is found to be 1.94. This value is obtained from the table of the F -distribution for 2.5% in the upper tail, with $df_1 = 30$ and $df_2 = 40$. Thus, if the computed F -statistic is greater than the critical value of 1.94, the null hypothesis is rejected. The decision rule, illustrated in the following figure, can be stated as:

Reject H_0 if $F > 1.94$

Decision Rule for F -Test



Since the calculated F -statistic of 1.2805 is less than the critical F -statistic of 1.94, Cower cannot reject the null hypothesis. Cower should conclude that the earnings variances of the industries are not significantly different from one another at a 5% level of significance.

LOS 6.k: Compare and contrast parametric and nonparametric tests, and describe situations where each is the more appropriate type of test.

Parametric tests rely on assumptions regarding the distribution of the population and are specific to population parameters. For example, the z-test relies upon a mean and a standard deviation to define the normal distribution. The z-test also requires that either the sample is large, relying on the central limit theorem to assure a normal sampling distribution, or that the population is normally distributed.

Nonparametric tests either do not consider a particular population parameter or have few assumptions about the population that is sampled. Nonparametric tests are used when there is concern about quantities other than the parameters of a distribution or when the assumptions of parametric tests can't be supported. They are also used when the data are not suitable for parametric tests (e.g., ranked observations).

Situations where a nonparametric test is called for are the following:

1. The assumptions about the distribution of the random variable that support a parametric test are not met. An example would be a hypothesis test of the mean value for a variable that comes from a distribution that is not normal and is of small size so that neither the *t*-test nor the *z*-test is appropriate.
 2. When data are ranks (an ordinal measurement scale) rather than values.
 3. The hypothesis does not involve the parameters of the distribution, such as testing whether a variable is normally distributed. We can use a nonparametric test, called a runs test, to determine whether data are random. A runs test provides an estimate of the probability that a series of changes (e.g., +, +, -, -, +, -,...) are random.
-

LOS 6.l: Explain parametric and nonparametric tests of the hypothesis that the population correlation coefficient equals zero, and determine whether the hypothesis is rejected at a given level of significance.

Correlation measures the strength of the relationship between two variables. If the correlation between two variables is zero, there is no linear relationship between them. When the sample correlation coefficient for two variables is different from zero, we must address the question of whether the true population correlation coefficient (ρ) is equal to zero. The appropriate test statistic for the hypothesis that the population correlation equals zero, when the two variables are normally distributed, is:

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$
, where r = sample correlation and n = sample size. This test statistic follows a *t*-distribution with $n - 2$ degrees of freedom. It is worth noting that the test statistic increases, not only with the sample correlation coefficient, but also with sample size.

EXAMPLE: Test of the hypothesis that the population correlation coefficient equals zero

A researcher computes the sample correlation coefficient for two normally distributed random variables as 0.35, based on a sample size of 42. Determine whether to reject the hypothesis that the population correlation coefficient is equal to zero at a 5% significance level.

Answer:

Our test statistic is
$$\frac{0.35\sqrt{42-2}}{\sqrt{1-0.35^2}} = 2.363$$
. Using the *t*-table with $42 - 2 = 40$ degrees of freedom for a two-tailed test and a significance level of 5%, we can find the critical value of 2.021. Because our computed test statistic of 2.363 is greater than 2.021, we reject the hypothesis that the population mean is zero and conclude that it is not equal to zero. That is, the two populations are correlated, in this case positively.



PROFESSOR'S NOTE

The correlation coefficient we refer to here is the Pearson correlation coefficient, which is a measure of the linear relationship between two variables. There are other

correlation coefficients that better measure the strength of any non-linear relationship between two variables.

The **Spearman rank correlation test**, a non-parametric test, can be used to test whether two sets of ranks are correlated. Ranks are simply ordered values. If there is a tie (equal values), the ranks are shared, so if 2nd and 3rd rank is the same, the ranks are shared and each gets a rank if $(2 + 3) / 2 = 2.5$.

The Spearman rank correlation, r_s , (when all ranks are integer values) is calculated as:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

where:

r_s = rank correlation

n = sample size

d_i = difference between two ranks

We can test the significance of the Spearman rank correlation calculated with the above formula using the same test statistic we used for estimating the significance of a parametric (Pearson) correlation coefficient:

$$\frac{r_s \sqrt{n - 2}}{\sqrt{1 - r_s^2}}$$

When the sample size is greater than 30, the test statistic follows a t -distribution with $n - 2$ degrees of freedom.

LOS 6.m: Explain tests of independence based on contingency table data.

A contingency or two-way table shows the number of observations from a sample that have a combination of two characteristics. Figure 6.11 is a contingency table where the characteristics are earnings growth (low, medium, or high) and dividend yield (low, medium, or high). We can use the data in the table to test the hypothesis that the two characteristics, earnings growth and dividend yield, are independent of each other.

Figure 6.11: Contingency Table for Categorical Data

Earnings Growth	Dividend Yield			Total
	Low	Medium	High	
Low	28	53	42	123
Medium	42	32	39	113
High	49	25	14	88
Total	119	110	95	324

We index our three categories of earnings growth from low to high with $i = 1, 2$, or 3 , and our three categories of dividend yield from low to high with $j = 1, 2$, or 3 . From the table, we see in cell 1,1 that 28 firms have both low earnings growth and low dividend yield. We see in cell 3,2 that 25 firms have high earnings growth and medium dividends.

For our test, we are going to compare the actual table values to what the values would be if the two characteristics were independent. The test statistic is a chi-square test statistic calculated as:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

where

O_{ij} = the number of observations in cell i, j , row i and column j (i.e., observed frequency)

E_{ij} = the expected number of observations for cell i, j

r = the number of row categories

c = the number of column categories

The degrees of freedom are $[(r - 1)(c - 1)]$, which is 4 in our example for dividend yield and earnings growth.

E_{ij} , the expected number of observations in cell i, j , is: $\frac{\text{total for row } i \times \text{total for column } j}{\text{total for all columns and rows}}$.

The expected number of observations for cell 2,2 is $\frac{110 \times 113}{324} = 38.4$.

In calculating our test statistic, the term for cell 2,2 is then $\frac{(32 - 38.4)^2}{38.4} = 1.067$.

Figure 6.12 shows the expected frequencies for each pair of categories in our earnings growth and dividend yield contingency table.

Figure 6.12: Contingency Table for Expected Frequencies

Earnings Growth	Dividend Yield		
	Low	Medium	High
Low	45.2	41.8	36.1
Medium	41.5	38.4	33.1
High	32.3	29.9	25.8

For our test statistic, we sum, for all nine cells, the squared difference between the expected frequency and observed frequency, divided by the expected frequency. The resulting sum is 27.43.

Our degrees of freedom are $(3 - 1) \times (3 - 1) = 4$. The critical value for a significance level of 5% (from the chi-square table in the Appendix) with 4 degrees of freedom is 9.488. Based on our sample data, we can reject the hypothesis that the earnings growth and dividend yield categories are independent.



MODULE QUIZ 6.4

1. The appropriate test statistic for a test of the equality of variances for two normally distributed random variables, based on two independent random samples, is:
 - A. the t test.
 - B. the F test.
 - C. the χ^2 test.
2. The appropriate test statistic to test the hypothesis that the variance of a normally distributed population is equal to 13 is:

- A. the t -test.
 - B. the F -test.
 - C. the χ^2 test.
3. For a parametric test of whether a correlation coefficient is equal to zero, it is *least likely* that:
- A. degrees of freedom are $n - 1$.
 - B. the test statistic follows a t -distribution.
 - C. the test statistic increases with a greater sample size.
4. The test statistic for a Spearman rank correlation test for a sample size greater than 30 follows:
- A. a t -distribution.
 - B. a normal distribution.
 - C. a chi-square distribution.
5. A contingency table can be used to test:
- A. a null hypothesis that rank correlations are equal to zero.
 - B. whether multiple characteristics of a population are independent.
 - C. the number of p -values from multiple tests that are less than adjusted critical values.

KEY CONCEPTS

LOS 6.a

The hypothesis testing process requires a statement of a null and an alternative hypothesis, the selection of the appropriate test statistic, specification of the significance level, a decision rule, the calculation of a sample statistic, a decision regarding the hypotheses based on the test, and a decision based on the test results.

The null hypothesis is what the researcher wants to reject. The alternative hypothesis is what the researcher wants to support, and it is accepted when the null hypothesis is rejected.

LOS 6.b

A two-tailed test results from a two-sided alternative hypothesis (e.g., $H_a: \mu \neq \mu_0$). A one-tailed test results from a one-sided alternative hypothesis (e.g., $H_a: \mu > \mu_0$, or $H_a: \mu < \mu_0$).

LOS 6.c

The test statistic is the value that a decision about a hypothesis will be based on. For a test about the value of the mean of a distribution:

$$\text{test statistic} = \frac{\text{sample mean} - \text{hypothesized mean}}{\text{standard error of the sample mean}}$$

A Type I error is the rejection of the null hypothesis when it is actually true, while a Type II error is the failure to reject the null hypothesis when it is actually false.

The significance level can be interpreted as the probability that a test statistic will reject the null hypothesis by chance when it is actually true (i.e., the probability of a Type I error). A significance level must be specified to select the critical values for the test.

The power of a test is the probability of rejecting the null when it is false. The power of a test = $1 - P(\text{Type II error})$.

LOS 6.d

Hypothesis testing compares a computed test statistic to a critical value at a stated level of significance, which is the decision rule for the test.

A hypothesis about a population parameter is rejected when the sample statistic lies outside a confidence interval around the hypothesized value for the chosen level of significance.

Statistical significance does not necessarily imply economic significance. Even though a test statistic is significant statistically, the size of the gains to a strategy to exploit a statistically significant result may be absolutely small or simply not great enough to outweigh transaction costs.

LOS 6.e

The p -value for a hypothesis test is the smallest significance level for which the hypothesis would be rejected. For example, a p -value of 7% means the hypothesis can be rejected at the 10% significance level but cannot be rejected at the 5% significance level.

LOS 6.f

When multiple tests are performed on different samples from a population, the p -values of each test are ranked, from lowest to highest, and compared to the adjusted critical values for each rank. When the proportion of the total number of ranked tests for which reported p -values are less than their adjusted critical values is greater than the significance level, the null hypothesis is rejected.

LOS 6.g

With unknown population variance, the t -statistic is used for tests about the mean of a normally distributed population: $t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$. If the population variance is known, the appropriate test statistic is $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ for tests about the mean of a population.

LOS 6.h

For two independent samples from two normally distributed populations, the difference in means can be tested with a t -statistic. When the two population variances are assumed to be equal, the denominator is based on the variance of the pooled samples.

LOS 6.i

A paired comparisons test is concerned with the mean of the differences between the paired observations of two dependent, normally distributed samples. A t -statistic,

$t = \frac{\bar{d}}{s_d}$, where $s_d = \sqrt{\frac{s^2_d}{n}}$, and \bar{d} is the average difference of the n paired observations, is used to test whether the means of two dependent normal variables are equal. Values outside the critical t -values lead us to reject equality.

LOS 6.j

The test of a hypothesis about the population variance for a normally distributed population uses a chi-square test statistic: $\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma_0^2}$, where n is the sample size, s^2 is the sample variance, and σ_0^2 is the hypothesized value for the population variance. Degrees of freedom are $n - 1$.

The test comparing two variances based on independent samples from two normally distributed populations uses an F -distributed test statistic: $F = \frac{s_1^2}{s_2^2}$, where s_1^2 is the variance of the first sample and s_2^2 is the (smaller) variance of the second sample.

LOS 6.k

Parametric tests, like the t -test, F -test, and chi-square tests, make assumptions regarding the distribution of the population from which samples are drawn. Nonparametric tests either do not consider a particular population parameter or have few assumptions about the sampled population. Nonparametric tests are used when the assumptions of parametric tests can't be supported or when the data are not suitable for parametric tests.

LOS 6.1

To test a hypothesis that a population correlation coefficient equals zero, the appropriate test statistic is a t -statistic with $n - 2$ degrees of freedom, calculated as $\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$, where r is the sample correlation coefficient.

A non-parametric test of correlation can be performed when we have only ranks (e.g., deciles of investment performance). The Spearman rank correlation test tests whether the ranks for

multiple periods are correlated. The rank correlation is $r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$, where d_i^2 is the sum of

the squared difference in pairs of ranks and n is the number of sample periods. The test statistic follows a t -distribution for sample sizes greater than 30.

LOS 6.m

A contingency table can be used to test the hypothesis that two characteristics (categories) of a sample of items are independent. The test statistic follows a chi-square distribution and is calculated as:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

where

O_{ij} = the number of observations in cell i, j , row i and column j (i.e., observed frequency)

E_{ij} = the expected number of observations for cell i, j of the contingency table with independence

r = the number of row categories and c = the number of column categories

The degrees of freedom are $[(r - 1)(c - 1)]$. If the test statistic is greater than the critical chi-square value for a given level of significance, we reject the hypothesis that the two characteristics are independent.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 6.1

1. C To test whether the population mean is greater than 20, the test would attempt to reject the null hypothesis that the mean is less than or equal to 20. The null hypothesis must always include the "equal to" condition. (LOS 6.a)
2. C Rejecting the null when it is actually true is a Type I error. A Type II error is failing to reject the null hypothesis when it is false. The significance level equals the probability of a Type I error. (LOS 6.c)

3. **C** A Type I error is rejecting the null hypothesis when it's true. The probability of rejecting a false null is $[1 - \text{Prob Type II}] = [1 - 0.60] = 40\%$, which is called the power of the test. A and B are not necessarily true, since the null may be false and the probability of rejection unknown. (LOS 6.c)
4. **A** The power of a test is $1 - P(\text{Type II error}) = 1 - 0.15 = 0.85$. (LOS 6.c)

Module Quiz 6.2

1. **C** The standard error is $\frac{5.29}{\sqrt{28}} = 1.0$. Test statistic = $1.645/1.0 = 1.645$. The critical value for t -test is greater than the critical value for a z -test at a 5% significance level (which is 1.645 for a one-tailed test), so the calculated test statistic of 1.645 must be less than the critical value for a t -test (which is 1.703 for a one-tailed test with 27 degrees of freedom) and we cannot reject the null hypothesis that mean excess return is greater than zero. (LOS 6.f)
2. **A** With a small sample size, a t -test may be used if the population is approximately normally distributed. If the population has a nonnormal distribution, no test statistic is available unless the sample size is large. (LOS 6.g)
3. **B** This is a two-tailed test with $14 - 1 = 13$ degrees of freedom. From the t -table, 2.160 is the critical value to which the analyst should compare the calculated t -statistic. (LOS 6.g)

Module Quiz 6.3

1. **C** When the variances are assumed to be unequal, we just calculate the denominator (standard error) differently and use both sample variances to calculate the t -statistic. (LOS 6.h)
2. **A** Since the observations are likely dependent (both related to market returns), a paired comparisons (mean differences) test is appropriate and is based on a t -statistic. (LOS 6.h, LOS 6.i)

Module Quiz 6.4

1. **B** The F -test is the appropriate test. (LOS 6.j)
2. **C** A test of the population variance is a chi-square test. (LOS 6.j)
3. **A** Degrees of freedom are $n - 2$ for a test of the hypothesis that correlation is equal to zero. The test statistic increases with sample size (degrees of freedom increase) and follows a t -distribution. (LOS 6.l)
4. **A** The test statistic for the Spearman rank correlation test follows a t -distribution. (LOS 6.l)
5. **B** A contingency table is used to determine whether two characteristics of a group are independent. (LOS 6.m)

READING 7

INTRODUCTION TO LINEAR REGRESSION

EXAM FOCUS

This introduction covers simple linear regression, which involves two variables: an independent and a dependent variable. Candidates should be able to construct a simple regression model and state the assumptions under which a linear model is valid. Given the estimated model parameters (coefficients), you should be able to use the model to predict the dependent variable. Finally, you may be required to interpret an ANOVA table and test the significance of estimated regression coefficients. Note that an F -test, in the context of a simple regression, is equivalent to a t -test of the significance of the estimated slope coefficient.

MODULE 7.1: LINEAR REGRESSION: INTRODUCTION



Video covering this content is available online.

LOS 7.a: Describe a simple linear regression model and the roles of the dependent and independent variables in the model.

The purpose of **simple linear regression** is to explain the variation in a dependent variable in terms of the variation in a single independent variable. Here, the term “variation” is interpreted as the degree to which a variable differs from its mean value. Don’t confuse variation with variance—they are related but are not the same.

$$\text{variation in } Y = \text{variation in } \bar{Y} = \sum_{i=1}^n (Y_i - \bar{Y})^2$$

- The **dependent variable** is the variable whose variation is explained by the independent variable. We are interested in answering the question, “What explains fluctuations in the dependent variable?” The dependent variable is also referred to as the *explained variable*, the *endogenous variable*, or the *predicted variable*.
- The **independent variable** is the variable used to explain the variation of the dependent variable. The independent variable is also referred to as the *explanatory variable*, the *exogenous variable*, or the *predicting variable*.

EXAMPLE: Dependent vs. independent variables

Suppose that we want to predict stock returns based on GDP growth. Which variable is the independent variable?

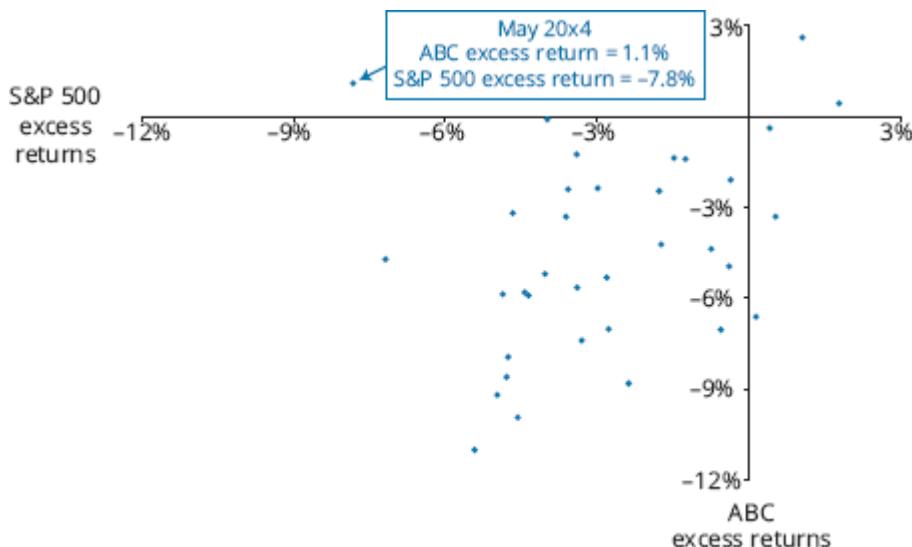
Answer:

Because GDP is going to be used as a *predictor* of stock returns, stock returns are being *explained* by GDP. Hence, stock returns are the dependent (explained) variable, and GDP is the independent (explanatory) variable.

Suppose we want to use excess returns on the S&P 500 (the independent variable) to explain the variation in excess returns on ABC common stock (the dependent variable). For this model, we define excess return as the difference between the actual return and the return on 1-month Treasury bills.

We would start by creating a scatter plot with ABC excess returns on the vertical axis and S&P 500 excess returns on the horizontal axis. Monthly excess returns for both variables from June 20x2 to May 20x5 are plotted in Figure 7.1. For example, look at the point labeled May 20x4. In that month, the excess return on the S&P 500 was -7.8% and the excess return on ABC was 1.1% .

Figure 7.1: Scatter Plot of ABC Excess Returns vs. S&P 500 Index Excess Returns



The two variables in Figure 7.1 appear to be positively correlated: excess ABC returns tended to be positive (negative) in the same month that S&P 500 excess returns were positive (negative). This is not the case for all the observations, however (for example, May 20x4). In fact, the correlation between these variables is approximately 0.40.

LOS 7.b: Describe the least squares criterion, how it is used to estimate regression coefficients, and their interpretation.

Simple Linear Regression Model

The following linear regression model is used to describe the relationship between two variables, X and Y :

$$Y_i = b_0 + b_1 X_i + \epsilon_i, i = 1, \dots, n$$

where:

Y_i = i th observation of the dependent variable, Y

X_i = i th observation of the independent variable, X

b_0 = regression intercept term

b_1 = regression slope coefficient

ϵ_i = residual for the i th observation (also referred to as the disturbance term or error term)

Based on this regression model, the regression process estimates an equation for a line through a scatter plot of the data that “best” explains the observed values for Y in terms of the observed values for X .

The linear equation, often called the line of best fit or **regression line**, takes the following form:

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i, i = 1, 2, 3, \dots, n$$

where:

\hat{Y}_i = estimated value of Y_i given X_i

\hat{b}_0 = estimated intercept term

\hat{b}_1 = estimated slope coefficient



PROFESSOR'S NOTE

The hat “^” above a variable or parameter indicates a predicted value.

The regression line is just one of the many possible lines that can be drawn through the scatter plot of X and Y . The criteria used to estimate this line is the essence of linear regression. The regression line is the line that minimizes the sum of the squared differences (vertical distances) between the Y -values predicted by the regression equation ($\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$) and the *actual* Y -values, Y_i . The sum of the squared vertical distances between the estimated and actual Y -values is referred to as the **sum of squared errors (SSE)**.

Thus, the regression line is the line that minimizes the SSE. This explains why simple linear regression is frequently referred to as **ordinary least squares (OLS)** regression, and the values determined by the estimated regression equation, \hat{Y}_i , are called least squares estimates.

The estimated **slope coefficient** (\hat{b}_1) for the regression line describes the change in Y for a one-unit change in X . It can be positive, negative, or zero, depending on the relationship between the regression variables. The slope term is calculated as:

$$\hat{b}_1 = \frac{\text{Cov}_{XY}}{\sigma_X^2}$$

The intercept term (\hat{b}_0) is the line’s intersection with the Y -axis at $X = 0$. It can be positive, negative, or zero. A property of the least squares method is that the intercept term may be expressed as:

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$$

where:

\bar{Y} = mean of Y

\bar{X} = mean of X

The intercept equation highlights the fact that the regression line passes through a point with coordinates equal to the mean of the independent and dependent variables (i.e., the point \bar{X}, \bar{Y}).

EXAMPLE: Computing the slope coefficient and intercept term

Compute the slope coefficient and intercept term for the ABC regression example using the following information:

$$\begin{array}{ll} \text{Cov(S\&P 500, ABC)} = 0.000336 & \text{Mean return, S\&P 500} = -2.70\% \\ \text{Var(S\&P 500)} = 0.000522 & \text{Mean return, ABC} = -4.05\% \end{array}$$

Answer:

The slope coefficient is calculated as $\hat{b}_1 = 0.000336 / 0.000522 = 0.64$.

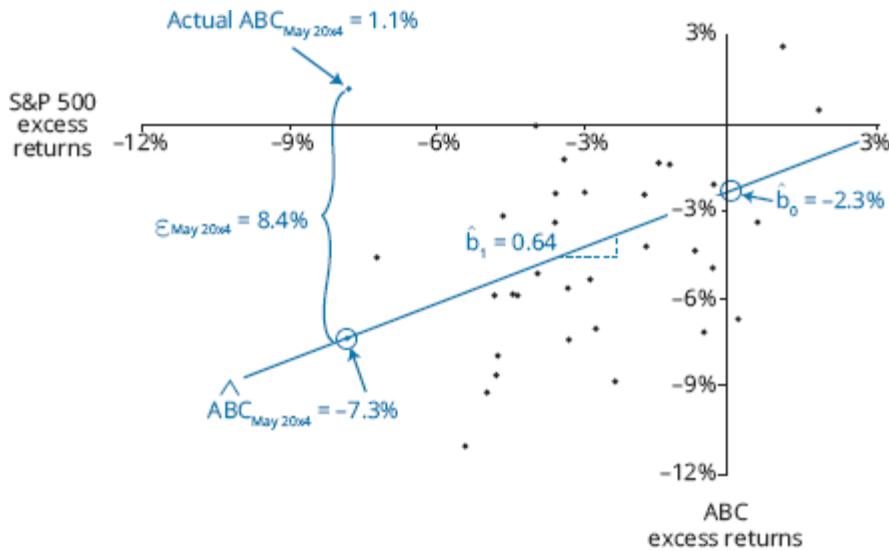
The intercept term is:

$$\hat{b}_0 = \bar{ABC} - \hat{b}_1 \bar{S\&P 500} = -4.05\% - 0.64(-2.70\%) = -2.3\%$$

The estimated regression line that minimizes the SSE in our ABC stock return example is shown in Figure 7.2.

This regression line has an intercept of -2.3% and a slope of 0.64 . The model predicts that if the S&P 500 excess return is -7.8% (May 20x4 value), then the ABC excess return would be $-2.3\% + (0.64)(-7.8\%) = -7.3\%$. The residual (error) for the May 20x4 ABC prediction is 8.4% , the difference between the actual ABC excess return of 1.1% and the predicted return of -7.3% .

Figure 7.2: Estimated Regression Equation for ABC vs. S&P 500 Excess Returns



Interpreting a Regression Coefficient

The estimated intercept represents the value of the dependent variable at the point of intersection of the regression line and the axis of the dependent variable (usually the vertical axis). In other words, the intercept is an estimate of the dependent variable when the independent variable is zero.

We also mentioned earlier that the estimated slope coefficient is interpreted as the expected change in the dependent variable for a one-unit change in the independent variable. For example, an estimated slope coefficient of 2 would indicate that the dependent variable is expected to change by two units for every one-unit change in the independent variable.

EXAMPLE: Interpreting regression coefficients

In the ABC regression example, the estimated slope coefficient was 0.64 and the estimated intercept term was -2.3%. Interpret each coefficient estimate.

Answer:

The slope coefficient of 0.64 can be interpreted to mean that when excess S&P 500 returns increase (decrease) by 1%, ABC excess return is expected to increase (decrease) by 0.64%.

The intercept term of -2.3% can be interpreted to mean that when the excess return on the S&P 500 is zero, the expected return on ABC stock is -2.3%.



PROFESSOR'S NOTE

The slope coefficient in a linear regression of the excess return of an individual security (the Y-variable) on the excess return on the market (the X-variable) is called the stock's beta, which is an estimate of systematic risk of ABC's stock. Notice that ABC is less risky than the average because its returns tend to increase or decrease by less than the change in the market returns. A stock with a beta (regression line slope) of one would have an average level of systematic risk and a stock with a beta greater than one would have more-than-average systematic risk. We will apply this concept in the Portfolio Management topic area.

Keep in mind, however, that any conclusions regarding the importance of an independent variable in explaining a dependent variable are based on the statistical significance of the slope coefficient. The magnitude of the slope coefficient tells us nothing about the strength of the linear relationship between the dependent and independent variables. A hypothesis test must be conducted, or a confidence interval must be formed, to assess the explanatory power of the independent variable. Later in this reading we will perform these hypothesis tests.

LOS 7.c: Explain the assumptions underlying the simple linear regression model, and describe how residuals and residual plots indicate if these assumptions may have been violated.

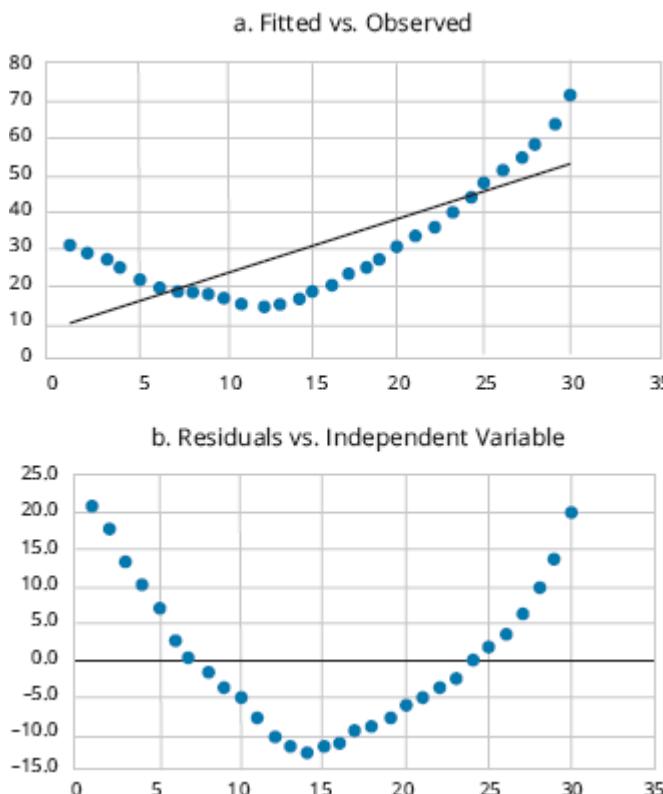
Linear regression is based on a number of assumptions. Most of the assumptions pertain to the regression model's residual term (ϵ). Linear regression assumes the following:

1. A linear relationship exists between the dependent and the independent variables.
2. The variance of the residual term is constant for all observations (homoskedasticity).
3. The residual term is independently distributed; that is, the residual for one observation is not correlated with that of another observation.
4. The residual term is normally distributed.

Linear Relationship

A linear regression model is not appropriate when the underlying relationship between X and Y is nonlinear. In Panel a of Figure 7.3, we illustrate a regression line fitted to a nonlinear relationship. Note that the prediction errors (vertical distances from the dots to the line) are positive for low values of X, then increasingly negative for higher values of X, and then turning positive for still greater values of X. One way of checking for linearity is to examine the model residuals (prediction errors) in relation to the independent regression variable. In Panel b, we show the pattern of residuals over the range of the independent variable: positive, negative, then positive.

Figure 7.3: Nonlinear Relationship

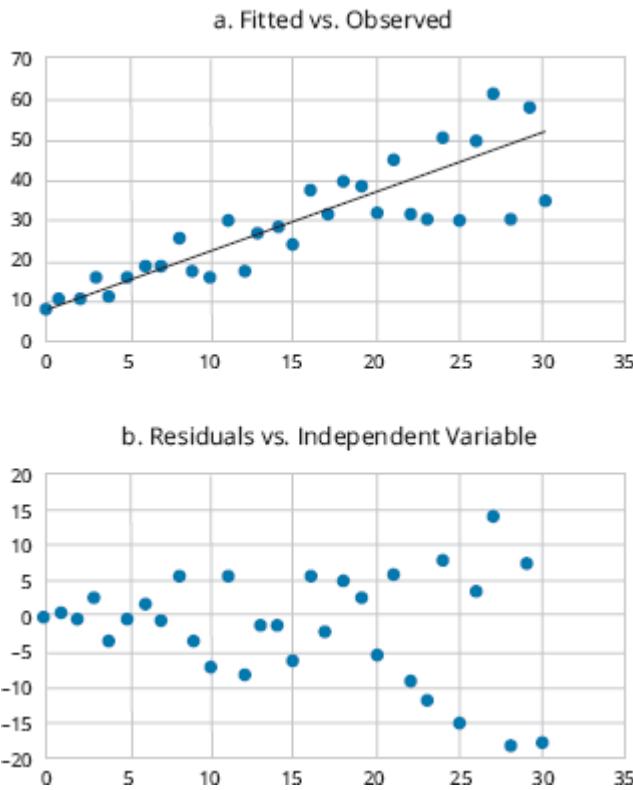


Homoskedasticity

Homoskedasticity refers to the case where prediction errors all have the same variance.

Heteroskedasticity refers to the situation when the assumption of homoskedasticity is violated. Figure 7.4 Panel a shows a scatter plot of observations around a fitted regression line where the residuals (prediction errors) increase in magnitude with larger values of the independent variable X. Panel b shows the residuals plotted versus the value of the independent variable, and also illustrates that the variance of the error terms is not likely constant for all observations.

Figure 7.4: Heteroskedasticity

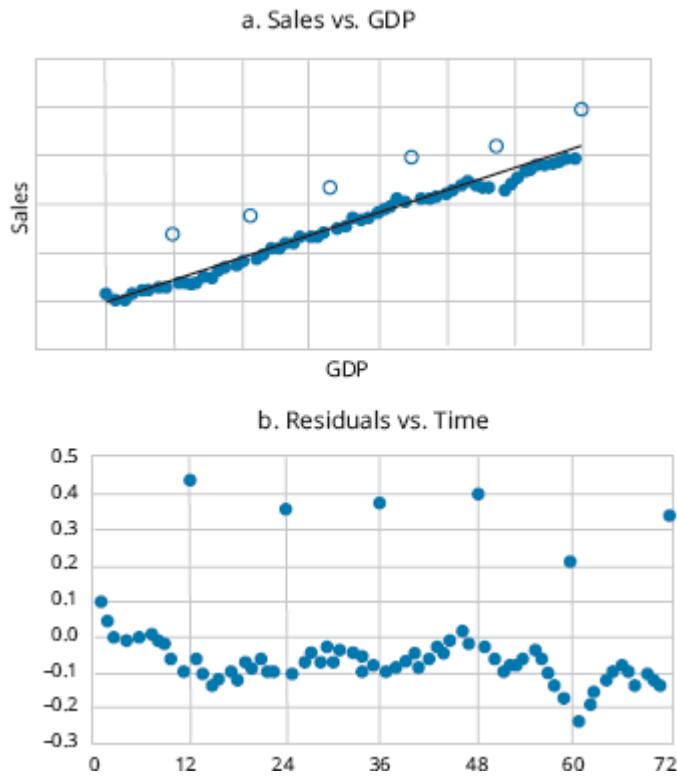


Another type of heteroskedasticity results if the variance of the error term changes over time (rather than with the magnitude of the independent variable). We could observe this by plotting the residuals from a linear regression model versus the dates of each observation and finding that the magnitude of the errors exhibits a pattern of changing over time. To illustrate this, we could plot the residuals versus a time index (as the X variable). Residuals would exhibit a pattern of increasing over time.

Independence

Suppose we collect a company's monthly sales and plot them against monthly GDP as in Figure 7.5 Panel a and observe that some prediction errors (the unfilled dots) are noticeably larger than others. To investigate this, we plot the residuals versus time, as in Panel b. The residuals plot illustrates that there are large negative prediction errors every 12 months (in December). This suggests that there is seasonality in sales such that December sales (the unfilled dots in Figure 7.5) are noticeably further from their predicted values than sales for the other months. If the relationship between X and Y is not independent, the residuals are not independent, and our estimates of variance, as well as our estimates of the model parameters, will not be correct.

Figure 7.5: Independence



Normality

When the residuals (prediction errors) are normally distributed, we can conduct hypothesis testing for evaluating the goodness of fit of the model (discussed later). With a large sample size, based on the central limit theorem, our parameter estimates may be valid, even when the residuals are not normally distributed.

Outliers are observations (one or a few) that are far from our regression line (have large prediction errors or X values that are far from the others). Outliers will influence our parameter estimates so that the OLS model will not fit the other observations well.



MODULE QUIZ 7.1

1. Which of the following is *least likely* a necessary assumption of simple linear regression analysis?
 - A. The residuals are normally distributed.
 - B. There is a constant variance of the error term.
 - C. The dependent variable is uncorrelated with the residuals.
2. What is the *most appropriate* interpretation of a slope coefficient estimate of 10.0?
 - A. The predicted value of the dependent variable when the independent variable is zero is 10.0.
 - B. For every one unit change in the independent variable, the model predicts that the dependent variable will change by 10 units.
 - C. For every 1-unit change in the independent variable, the model predicts that the dependent variable will change by 0.1 units.

MODULE 7.2: GOODNESS OF FIT AND HYPOTHESIS TESTS



Video covering
this content is
available online.

LOS 7.d: Calculate and interpret the coefficient of determination and the F-statistic in a simple linear regression.

LOS 7.e: Describe the use of analysis of variance (ANOVA) in regression analysis, interpret ANOVA results, and calculate and interpret the standard error of estimate in a simple linear regression.

Analysis of variance (ANOVA) is a statistical procedure for analyzing the total variability of the dependent variable. Let's define some terms before we move on to ANOVA tables:

- **Total sum of squares (SST)** measures the total variation in the dependent variable. SST is equal to the sum of the squared differences between the actual Y-values and the mean of Y.

$$SST = \sum_{i=1}^n (Y_i - \bar{Y})^2$$

- **Sum of squares regression (SSR)** measures the variation in the dependent variable that is explained by the independent variable. SSR is the sum of the squared distances between the predicted Y-values and the mean of Y.

$$SSR = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

- **Sum of squared errors (SSE)** measures the unexplained variation in the dependent variable. It's also known as the sum of squared residuals or the residual sum of squares. SSE is the sum of the squared vertical distances between the actual Y-values and the predicted Y-values on the regression line.

$$SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

You probably will not be surprised to learn that:

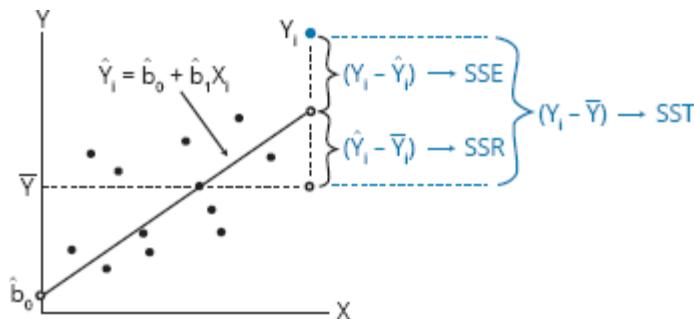
$$\text{total variation} = \text{explained variation} + \text{unexplained variation}$$

or:

$$SST = SSR + SSE$$

Figure 7.6 illustrates how the total variation in the dependent variable (SST) is composed of SSR and SSE.

Figure 7.6: Components of Total Variation



The output of the ANOVA procedure is an ANOVA table, which is a summary of the variation in the dependent variable. ANOVA tables are included in the regression output of many statistical

software packages. You can think of the ANOVA table as the source of the data for the computation of many of the regression concepts discussed in this reading. A generic ANOVA table for a simple linear regression (one independent variable) is presented in Figure 7.7.

Figure 7.7: ANOVA Table

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Regression (explained)	1	SSR	$MSR = \frac{SSR}{k} = \frac{SSR}{1} = SSR$
Error (unexplained)	$n - 2$	SSE	$MSE = \frac{SSE}{n - 2}$
Total	$n - 1$	SST	



PROFESSOR'S NOTE

In Figure 7.7, k is the number of slope parameters estimated and n is the number of observations. In general (including regressions with more than one independent variable), the regression $df = k$ and the error $df = (n - k - 1)$. Because we are limited to simple linear regressions in this reading (one independent variable), we use $k = 1$ for the regression degrees of freedom and $n - 1 - 1 = n - 2$ for the error degrees of freedom.

Standard Error of Estimate (SEE)

SEE for a regression is the standard deviation of its residuals. The lower the SEE, the better the model fit.

$$SEE = \sqrt{MSE}$$

Coefficient of Determination (R^2)

The **coefficient of determination** (R^2) is defined as the percentage of the total variation in the dependent variable explained by the independent variable. For example, an R^2 of 0.63 indicates that the variation of the independent variable explains 63% of the variation in the dependent variable.

$$R^2 = \frac{SSR}{SST}$$



PROFESSOR'S NOTE

For simple linear regression (i.e., with one independent variable), the coefficient of determination, R^2 , may be computed by simply squaring the correlation coefficient, r . In other words, $R^2 = r^2$ for a regression with one independent variable.

EXAMPLE: Using the ANOVA table

Complete the ANOVA table for the ABC regression example and calculate the R^2 and the standard error of estimate (SEE).

Partial ANOVA Table for ABC Regression Example

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Regression (explained)	?	0.00756	?
Error (unexplained)	?	0.04064	?
Total	?	?	

Answer:

Recall that the data included three years of monthly return observations, so the total number of observations (n) is 36.

Completed ANOVA Table for ABC Regression Example

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Regression (explained)	1	0.0076	0.0076
Error (unexplained)	34	0.0406	0.0012
Total	35	0.0482	

$$R^2 = \frac{\text{explained variation (SSR)}}{\text{total variation (SST)}} = \frac{0.0076}{0.0482} = 0.158 \text{ or } 15.8\%$$

$$\text{SEE} = \sqrt{\text{MSE}} = \sqrt{0.0012} = 0.035$$

The F-Statistic

An F -test assesses how well a set of independent variables, as a group, explains the variation in the dependent variable.

The F -statistic is calculated as:

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{\text{SSR}/k}{\text{SSE}/(n - k - 1)}$$

where:

MSR = mean regression sum of squares

MSE = mean squared error

Important: This is always a one-tailed test!

For simple linear regression, there is only one independent variable, so the F -test is equivalent to a t -test for statistical significance of the slope coefficient:

$$H_0: b_1 = 0 \text{ versus } H_a: b_1 \neq 0$$

To determine whether b_1 is statistically significant using the F -test, the calculated F -statistic is compared with the critical F -value, F_c , at the appropriate level of significance. The degrees of freedom for the numerator and denominator with one independent variable are:

$$df_{\text{numerator}} = k = 1$$

$$df_{\text{denominator}} = n - k - 1 = n - 2$$

where:

n = number of observations

The decision rule for the *F*-test is: reject H_0 if $F > F_c$.

Rejecting the null hypothesis that the value of the slope coefficient equals zero at a stated level of significance indicates that the independent variable and the dependent variable have a significant linear relationship.

EXAMPLE: Calculating and interpreting the *F*-statistic

Use the completed ANOVA table from the previous example to calculate and interpret the *F*-statistic. Test the null hypothesis at the 5% significance level that the slope coefficient is equal to 0.

Answer:

$$F = \frac{MSR}{MSE} = \frac{0.0076}{0.0012} = 6.33$$

$$df_{\text{numerator}} = k = 1$$

$$df_{\text{denominator}} = n - k - 1 = 36 - 1 - 1 = 34$$

The null and alternative hypotheses are: $H_0: b_1 = 0$ versus $H_a: b_1 \neq 0$. The critical *F*-value for 1 and 34 degrees of freedom at a 5% significance level is approximately 4.1. (Remember, it's a one-tail test, so we use the 5% *F*-table!) Therefore, we can reject the null hypothesis and conclude that the slope coefficient is significantly different than zero.

LOS 7.f: Formulate a null and an alternative hypothesis about a population value of a regression coefficient, and determine whether the null hypothesis is rejected at a given level of significance.

A *t*-test may also be used to test the hypothesis that the true slope coefficient, b_1 , is equal to a hypothesized value. Letting \hat{b}_1 be the point estimate for b_1 , the appropriate test statistic with $n - 2$ degrees of freedom is:

$$t_{b_1} = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}}$$

The decision rule for tests of significance for regression coefficients is:

Reject H_0 if $t > +t_{\text{critical}}$ or $t < -t_{\text{critical}}$

Rejection of the null supports the alternative hypothesis that the slope coefficient is *different* from the hypothesized value of b_1 . To test whether an independent variable explains the variation in the dependent variable (i.e., it is statistically significant), the null hypothesis is that the true slope is zero ($b_1 = 0$). The appropriate test structure for the null and alternative hypotheses is:

$H_0: b_1 = 0$ versus $H_a: b_1 \neq 0$

EXAMPLE: Hypothesis test for significance of regression coefficients

The estimated slope coefficient from the ABC example is 0.64 with a standard error equal to 0.26. Assuming that the sample has 36 observations, determine if the estimated slope coefficient is significantly different than zero at a 5% level of significance.

Answer:

The calculated test statistic is:

$$t = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}} = \frac{0.64 - 0}{0.26} = 2.46$$

The critical two-tailed t -values are ± 2.03 (from the t -table with $df = 36 - 2 = 34$). Because $t > t_{critical}$ (i.e., $2.46 > 2.03$), we reject the null hypothesis and conclude that the slope is different from zero.

Note that the t -test for a simple linear regression is equivalent to a t -test for the correlation coefficient between x and y :

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$



MODULE QUIZ 7.2

Use the following data to answer Questions 1 and 2.

An analyst is interested in predicting annual sales for XYZ Company, a maker of paper products. The following table reports a regression of the annual sales for XYZ against paper product industry sales.

Regression Output

Parameters	Coefficient	Standard Error of the Coefficient
Intercept	-94.88	32.97
Slope (industry sales)	0.2796	0.0363

The correlation between company and industry sales is 0.9757. The regression was based on five observations.

1. Which of the following is *closest* to the value and reports the *most likely* interpretation of the R^2 for this regression?
 - A. The R^2 is 0.048, indicating that the variability of industry sales explains about 4.8% of the variability of company sales.
 - B. The R^2 is 0.952, indicating that the variability of industry sales explains about 95.2% of the variability of company sales.
 - C. The R^2 is 0.952, indicating that the variability of company sales explains about 95.2% of the variability of industry sales.

2. Based on the regression results, XYZ Company's market share of any increase in industry sales is expected to be *closest* to:
- 4%.
 - 28%.
 - 45%.

Use the following information to answer Questions 3 and 4.

A study was conducted by the British Department of Transportation to estimate urban travel time between locations in London, England. Data was collected for motorcycles and passenger cars. Simple linear regression was conducted using data sets for both types of vehicles, where Y = urban travel time in minutes and X = distance between locations in kilometers. The following results were obtained:

Regression Results for Travel Times Between Distances in London		
Passenger cars:	$\hat{Y} = 1.85 + 3.86X$	$R^2 = 0.758$
Motorcycles:	$\hat{Y} = 2.50 + 1.93X$	$R^2 = 0.676$

- The estimated increase in travel time for a motorcycle commuter planning to move 8 km farther from his workplace in London is *closest* to:
 - 31 minutes.
 - 15 minutes.
 - 0.154 hours.
 - Based on the regression results, which model is more reliable?
 - The passenger car model because $3.86 > 1.93$.
 - The motorcycle model because $1.93 < 3.86$.
 - The passenger car model because $0.758 > 0.676$.
 - Consider the following statement: In a simple linear regression, the appropriate degrees of freedom for the critical t -value used to calculate a confidence interval around both a parameter estimate and a predicted Y -value is the same as the number of observations minus two. The statement is:
 - justified.
 - not justified, because the appropriate of degrees of freedom used to calculate a confidence interval around a parameter estimate is the number of observations.
 - not justified, because the appropriate of degrees of freedom used to calculate a confidence interval around a predicted Y -value is the number of observations.
 - What is the appropriate alternative hypothesis to test the statistical significance of the intercept term in the following regression?
- $$Y = a_1 + a_2(X) + \varepsilon$$
- $H_A: a_1 \neq 0$.
 - $H_A: a_1 > 0$.
 - $H_A: a_2 \neq 0$.

MODULE 7.3: PREDICTING DEPENDENT VARIABLES AND FUNCTIONAL FORMS



Video covering
this content is
available online.

LOS 7.g: Calculate and interpret the predicted value for the dependent variable, and a prediction interval for it, given an estimated linear regression model and a value for the independent variable.

Predicted values are values of the dependent variable based on the estimated regression coefficients and a prediction about the value of the independent variable. They are the values that are *predicted* by the regression equation, given an estimate of the independent variable.

For a simple regression, the predicted (or forecast) value of Y is:

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 X_p$$

where:

\hat{Y} = predicted value of the dependent variable

X_p = forecasted value of the independent variable

EXAMPLE: Predicting the dependent variable

Given the ABC regression equation:

$$\widehat{\text{ABC}} = -2.3\% + (0.64)(\widehat{\text{S\&P 500}})$$

Calculate the predicted value of ABC excess returns if forecasted S&P 500 excess returns are 10%.

Answer:

The predicted value for ABC excess returns is determined as follows:

$$\widehat{\text{ABC}} = -2.3\% + (0.64)(10\%) = 4.1\%$$

Confidence Intervals for Predicted Values

The equation for the confidence interval for a predicted value of Y is:

$$\hat{Y} \pm (t_c \times s_f) \Rightarrow [\hat{Y} - (t_c \times s_f) < Y < \hat{Y} + (t_c \times s_f)]$$

where:

t_c = two-tailed critical t -value at the desired level of significance with $df = n - 2$

s_f = standard error of the forecast

The challenge with computing a confidence interval for a predicted value is calculating s_f . On the Level I exam it's highly unlikely that you will have to calculate the standard error of the forecast (it will probably be provided if you need to compute a confidence interval for the dependent variable). However, if you do need to calculate s_f it can be done with the following formula for the variance of the forecast:

$$s_f^2 = \text{SEE}^2 \left[1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{(n - 1)s_x^2} \right]$$

where:

SEE^2 = variance of the residuals = the square of the standard error of estimate

s_x^2 = variance of the independent variable

X = value of the independent variable for which the forecast was made

EXAMPLE: Confidence interval for a predicted value

Calculate a 95% prediction interval on the predicted value of ABC excess returns from the previous example. Suppose the standard error of the forecast is 3.67, and the forecast value of S&P 500 excess returns is 10%.

Answer:

The predicted value for ABC excess returns is:

$$\widehat{ABC} = -2.3\% + (0.64)(10\%) = 4.1\%$$

The 5% two-tailed critical t -value with 34 degrees of freedom is 2.03. The prediction interval at the 95% confidence level is:

$$\widehat{ABC} \pm (t_c \times s_f) \Rightarrow [4.1\% \pm (2.03 \times 3.67\%)] = 4.1\% \pm 7.5\%$$

or -3.4% to 11.6%

We can interpret this to mean that, given a forecast value for S&P 500 excess returns of 10%, we can be 95% confident that the ABC excess returns will be between -3.4% and 11.6%.

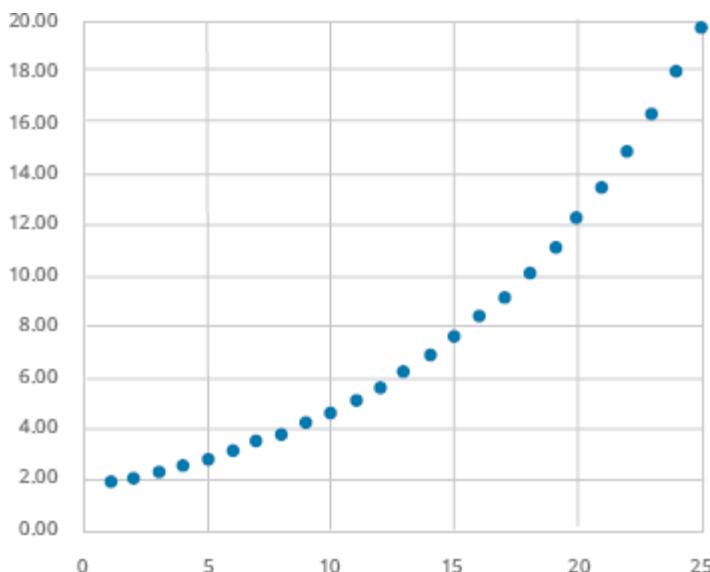
LOS 7.h: Describe different functional forms of simple linear regressions.

One of the assumptions of linear regression is that the relationship between X and Y is linear.

What if that assumption is violated? Consider $Y = \text{EPS}$ for a company and $X = \text{time index}$.

Suppose that EPS is growing at approximately 10% annually. Figure 7.8 shows the plot of actual EPS versus time.

Figure 7.8: Nonlinear Relationship



In such a situation, transforming one or both of the variables can produce a linear relationship. The appropriate transformation depends on the relationship between the two variables. One

often-used transformation is to take the natural log of one or both of the variables. Some examples are:

- **Log-lin model.** If the dependent variable is logarithmic while the independent variable is linear.
- **Lin-log model.** If the dependent variable is linear while the independent variable is logarithmic.
- **Log-log model.** Both the dependent variable and the independent variable are logarithmic.

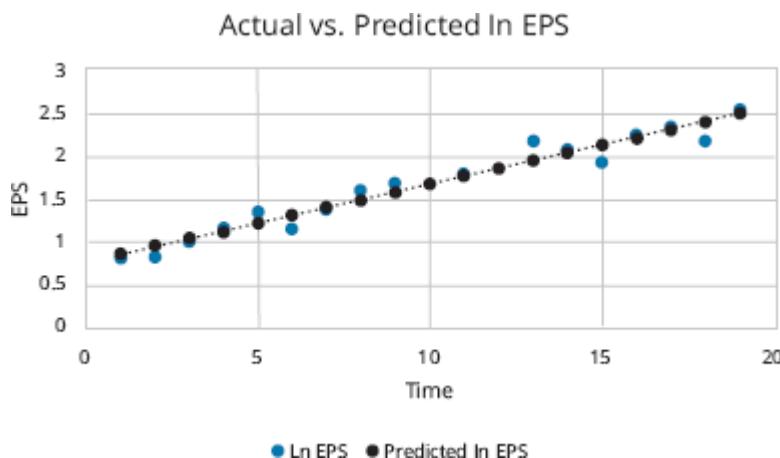
Log-Lin Model

Taking the natural logarithm of the dependent variable, our model now becomes:

$$\ln Y_i = b_0 + b_1 X_i + \varepsilon_i$$

In this model, the slope coefficient is interpreted as the *relative* change in the dependent variable for an absolute change in the independent variable. Figure 7.9 shows the results after taking the natural log of EPS, and fitting that data using a log-lin model.

Figure 7.9: Log-Lin Model, EPS Data



Lin-Log Model

Taking the natural logarithm of the independent variable, our model now becomes:

$$Y_i = b_0 + b_1 \ln(X)_i + \varepsilon_i$$

In this model, the slope coefficient is interpreted as the *absolute* change in the dependent variable for a *relative* change in the independent variable.

Log-Log Model

Taking the natural logarithm of both variables, our model now becomes:

$$\ln Y_i = b_0 + b_1 \ln(X)_i + \varepsilon_i$$

In this model, the slope coefficient is interpreted as the relative change in the dependent variable for a relative change in the independent variable.

Selection of Functional Form

Selecting the correct functional form involves determining the nature of the variables and evaluation of the goodness of fit measures (e.g., R^2 , SEE, F -stat).



MODULE QUIZ 7.3

1. The variation in the dependent variable explained by the independent variable is measured by:
 - A. the mean squared error.
 - B. the sum of squared errors.
 - C. the regression sum of squares.
2. Results from a regression analysis are presented in the following figures.

Estimated Coefficients

Coefficient	Coefficient Estimate	Standard Error
b_0	0.0023	0.0022
b_1	1.1163	0.0624

Partial ANOVA Table

Source of Variation	Sum of Squares
Regression (explained)	0.0228
Error (unexplained)	0.0024

Are the intercept term and the slope coefficient statistically significantly different from zero at the 5% significance level?

- | <u>Intercept term significant?</u> | <u>Slope coefficient significant?</u> |
|------------------------------------|---------------------------------------|
| A. Yes | Yes |
| B. Yes | No |
| C. No | Yes |

3. Partial ANOVA Table

Source of Variation	Sum of Squares
Regression (explained)	0.0228
Error (unexplained)	0.0024

To test the following hypothesis: $H_0: b_1 \leq 1$ versus $H_1: b_1 > 1$, at the 1% significance level, the calculated t -statistic and the appropriate conclusion are:

- | <u>Calculated t-statistic</u> | <u>Appropriate conclusion</u> |
|--|-------------------------------|
| A. 1.86 | Reject H_0 |
| B. 1.86 | Fail to reject H_0 |
| C. 2.44 | Reject H_0 |

4. The appropriate regression model for a linear relationship between the relative change in an independent variable and the absolute change in the dependent variable is a:
 - A. log-lin model.
 - B. lin-log model.
 - C. lin-lin model.

5. For a regression model of $Y = 5 + 3.5X$, the analysis (based on a large data sample) provides the standard error of the forecast as 2.5 and the standard error of the slope coefficient as 0.8. A 90% confidence interval for the estimate of Y when the value of the independent variable is 10 is closest to:
- 35.1 to 44.9.
 - 35.6 to 44.4.
 - 35.9 to 44.1.

KEY CONCEPTS

LOS 7.a

Linear regression provides an estimate of the linear relationship between an independent variable (the explanatory variable) and a dependent variable (the predicted variable).

LOS 7.b

The general form of a simple linear regression model is $Y_i = b_0 + b_1 X_i + \varepsilon_i$.

The least-squares model minimizes the sum of squared errors.

- $\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$ = fitted intercept
- \hat{b}_1 fitted slope coefficient = $\text{cov}(X, Y) / \text{variance of } X$

The estimated intercept, \hat{b}_0 , represents the value of the dependent variable at the point of intersection of the regression line and the axis of the dependent variable (usually the vertical axis). The estimated slope coefficient, \hat{b}_1 , is interpreted as the change in the dependent variable for a one-unit change in the independent variable.

LOS 7.c

Assumptions made regarding simple linear regression include the following:

1. A linear relationship exists between the dependent and the independent variable.
2. The variance of the residual term is constant (homoskedasticity).
3. The residual term is independently distributed (residuals are uncorrelated).
4. The residual term is normally distributed.

LOS 7.d, e

ANOVA Table for Simple Linear Regression ($k = 1$)

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Regression (explained)	1	SSR	$MSR = \frac{SSR}{k} = \frac{SSR}{1} = SSR$
Error (unexplained)	$n - 2$	SSE	$MSE = \frac{SSE}{n - 2}$
Total	$n - 1$	SST	

The standard error of the estimate in a simple linear regression is calculated as:

$$SEE = \sqrt{\frac{SSE}{n - 2}}$$

The standard error of the estimate in a simple linear regression is calculated as:

$$SEE = \sqrt{\frac{SSE}{n-2}}$$

The coefficient of determination, R^2 , is the proportion of the total variation of the dependent variable explained by the regression:

$$R^2 = \frac{SSR}{SST} = \frac{SST - SSE}{SST}$$

In simple linear regression, because there is only one independent variable ($k = 1$), the F -test tests the same null hypothesis as testing the statistical significance of b_1 using the t -test: $H_0: b_1 = 0$ versus $H_a: b_1 \neq 0$. With only one independent variable, F is calculated as:

$$F\text{-stat} = \frac{MSR}{MSE} \text{ with } 1 \text{ and } n - 2 \text{ degrees of freedom}$$

LOS 7.f

A t -test with $n - 2$ degrees of freedom is used to conduct hypothesis tests of the estimated regression parameters:

$$t = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}}$$

LOS 7.g

A predicted value of the dependent variable, \hat{Y} , is determined by inserting the predicted value of the independent variable, X_p , in the regression equation and calculating $\hat{Y}_p = \hat{b}_0 + \hat{b}_1 X_p$

The confidence interval for a predicted Y -value is $[\hat{Y} - (t_c \times s_f) < Y < \hat{Y} + (t_c \times s_f)]$, where s_f is the standard error of the forecast.

LOS 7.h

Dependent Variable	Independent Variable	Model	Slope Interpretation
Logarithmic	Linear	Log-lin	<i>Relative</i> change in dependent variable for an absolute change in the independent variable
Linear	Logarithmic	Lin-log	<i>Absolute</i> change in dependent variable for a relative change in the independent variable
Logarithmic	Logarithmic	Log-log	<i>Relative</i> change in dependent variable for a <i>relative</i> change in the independent variable

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 7.1

- C** The model does not assume that the dependent variable is uncorrelated with the residuals. It does assume that the independent variable is uncorrelated with the residuals. (LOS 7.c)
- B** The slope coefficient is best interpreted as the predicted change in the dependent variable for a 1-unit change in the independent variable. If the slope coefficient estimate is 10.0 and the independent variable changes by 1

unit, the dependent variable is expected to change by 10 units. The intercept term is best interpreted as the value of the dependent variable when the independent variable is equal to zero. (LOS 7.b)

Module Quiz 7.2

1. **B** The R^2 is computed as the correlation squared: $(0.9757)^2 = 0.952$.

The interpretation of this R^2 is that 95.2% of the variation in Company XYZ's sales is explained by the variation in industry sales. The independent variable (industry sales) explains the variation in the dependent variable (company sales). This interpretation is based on the economic reasoning used in constructing the regression model. (LOS 7.d)

2. **B** The slope coefficient of 0.2796 indicates that a \$1 million increase in industry sales will result in an increase in firm sales of approximately 28% of that amount (\$279,600). (LOS 7.b)

3. **B** The slope coefficient is 1.93, indicating that each additional kilometer increases travel time by 1.93 minutes:

$$1.93 \times 8 = 15.44$$

(LOS 7.b)

4. **C** The higher R^2 for the passenger car model indicates that regression results are more reliable. Distance is a better predictor of travel time for cars. Perhaps the aggressiveness of the driver is a bigger factor in travel time for motorcycles than it is for autos. (LOS 7.d)

5. **A** In simple linear regression, the appropriate degrees of freedom for both confidence intervals is the number of observations in the sample (n) minus two. (LOS 7.d)

6. **A** In this regression, a_1 is the intercept term. To test the statistical significance means to test the null hypothesis that a_1 is equal to zero, versus the alternative that a_1 is not equal to zero. (LOS 7.d)

Module Quiz 7.3

1. **C** The regression sum of squares measures the amount of variation in the dependent variable explained by the independent variable (i.e., the explained variation). The sum of squared errors measures the variation in the dependent variable not explained by the independent variable. The mean squared error is equal to the sum of squared errors divided by its degrees of freedom. (Module 7.2, LOS 7.e)
2. **C** The critical two-tailed 5% t-value with 34 degrees of freedom is approximately 2.03. The calculated t-statistics for the intercept term and slope coefficient are, respectively, $0.0023 / 0.0022 = 1.05$ and $1.1163 / 0.0624 = 17.9$. Therefore, the intercept term is not statistically different from zero at the 5% significance level, while the slope coefficient is. (LOS 7.g)
3. **B** Note that this is a one-tailed test. The critical one-tailed 1% t-value with 34 degrees of freedom is approximately 2.44. The calculated t-statistic for the slope coefficient is $(1.1163 - 1) / 0.0624 = 1.86$. Therefore, the slope coefficient is not statistically different from one at the 1% significance level and the analyst should fail to reject the null hypothesis. (LOS 7.g)
4. **B** The appropriate model would be a lin-log model, in which the values of the dependent variable (Y) are regressed on the natural logarithms of the independent variable (X), $Y = b_0 + b_1 \ln X$. (LOS 7.h)
5. **C** The estimate of Y, given $X = 10$ is: $Y = 5 + 3.5(10) = 40$. The critical value for a 90% confidence interval with a large sample size (z-statistic) is approximately 1.65. Given the standard error of the forecast of 2.5, the confidence interval for the estimated value of Y is $40 \pm 1.65(2.5) = 35.875$ to 44.125 . (LOS 7.g)

TOPIC QUIZ: QUANTITATIVE METHODS

You have now finished the Quantitative Methods topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Quantitative Methods. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

READING 8

TOPICS IN DEMAND AND SUPPLY ANALYSIS

EXAM FOCUS

The Level I Economics curriculum assumes candidates are familiar with concepts such as supply and demand, utility-maximizing consumers, and the product and cost curves of firms. CFA Institute has posted three assigned readings to its website as prerequisites for Level I Economics. If you have not studied economics before (or if it has been a while), you should review these readings, along with the video instruction, study notes, and review questions for each of them in your online Schweser Resource Library to get up to speed.

MODULE 8.1: ELASTICITY



LOS 8.a: Calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure.

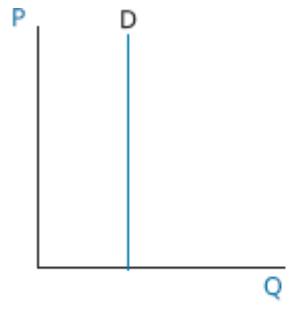
Video covering
this content is
available online.

Own-Price Elasticity of Demand

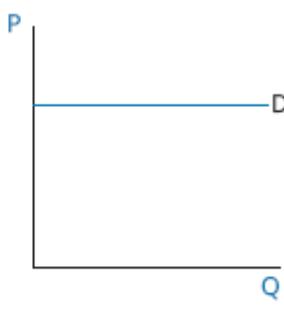
Own-price elasticity is a measure of the responsiveness of the quantity demanded to a change in price. It is calculated as the ratio of the percentage change in quantity demanded to a percentage change in price. With downward-sloping demand (i.e., an increase in price decreases quantity demanded), own-price elasticity is negative.

When the quantity demanded is very responsive to a change in price (absolute value of elasticity > 1), we say demand is elastic; when the quantity demanded is not very responsive to a change in price (absolute value of elasticity < 1), we say that demand is inelastic. In Figure 8.1, we illustrate the most extreme cases: perfectly elastic demand (at any higher price, quantity demanded decreases to zero) and perfectly inelastic demand (a change in price has no effect on quantity demanded).

Figure 8.1: Perfectly Inelastic and Perfectly Elastic Demand



(a) Perfectly inelastic demand
(elasticity = 0)



(b) Perfectly elastic demand
(elasticity = ∞)

When there are few or no good substitutes for a good, demand tends to be relatively inelastic. Consider a drug that keeps you alive by regulating your heart. If two pills per day keep you alive, you are unlikely to decrease your purchases if the price goes up and also quite unlikely to increase your purchases if price goes down.

When one or more goods are very good substitutes for the good in question, demand will tend to be very elastic. Consider two gas stations along your regular commute that offer gasoline of equal quality. A decrease in the posted price at one station may cause you to purchase all your gasoline there, while a price increase may lead you to purchase all your gasoline at the other station. Remember, we calculate demand and elasticity while holding the prices of related goods (in this case, the price of gas at the other station) constant.

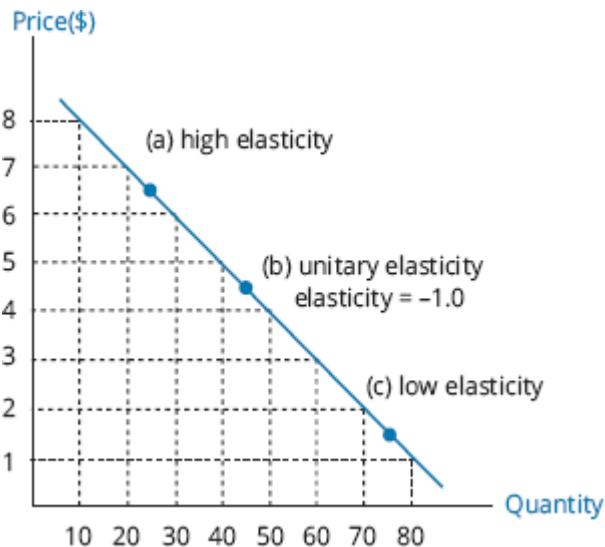
Other factors affect demand elasticity in addition to the quality and availability of substitutes:

- **Portion of income spent on a good.** The larger the proportion of income spent on a good, the more elastic an individual's demand for that good. If the price of a preferred brand of toothpaste increases, a consumer may not change brands or adjust the amount used if the customer prefers to simply pay the extra cost. When housing costs increase, however, a consumer will be much more likely to adjust consumption, because rent is a fairly large proportion of income.
- **Time.** Elasticity of demand tends to be greater the longer the time period since the price change. For example, when energy prices initially rise, some adjustments to consumption are likely made quickly. Consumers can lower the thermostat temperature. Over time, adjustments such as smaller living quarters, better insulation, more efficient windows, and installation of alternative heat sources are more easily made, and the effect of the price change on consumption of energy is greater.

It is important to understand that elasticity is not equal to the slope of a demand curve (except for the extreme examples of perfectly elastic or perfectly inelastic demand). Slope is dependent on the units that price and quantity are measured in. Elasticity is not dependent on units of measurement because it is based on percentage changes.

Figure 8.2 shows how elasticity changes along a linear demand curve. In the upper part of the demand curve, elasticity is greater (in absolute value) than 1; in other words, the percentage change in quantity demanded is greater than the percentage change in price. In the lower part of the curve, the percentage change in quantity demanded is smaller than the percentage change in price.

Figure 8.2: Price Elasticity Along a Linear Demand Curve



- At point (a), in a higher price range, the price elasticity of demand is greater than at point (c) in a lower price range.
- The elasticity at point (b) is -1.0 ; a 1% increase in price leads to a 1% decrease in quantity demanded. This is the point of greatest total revenue ($P \times Q$), which is $4.50 \times 45 = \$202.50$.
- At prices less than $\$4.50$ (inelastic range), total revenue will increase when price increases. The percentage decrease in quantity demanded will be less than the percentage increase in price.
- At prices above $\$4.50$ (elastic range), a price increase will decrease total revenue since the percentage decrease in quantity demanded will be greater than the percentage increase in price.

An important point to consider about the price and quantity combination for which price elasticity equals -1.0 (**unit or unitary elasticity**) is that total revenue (price \times quantity) is maximized at that price. An increase in price moves us to the elastic region of the curve so that the percentage decrease in quantity demanded is greater than the percentage increase in price, resulting in a decrease in total revenue. A decrease in price from the point of unitary elasticity moves us into the inelastic region of the curve so that the percentage decrease in price is more than the percentage increase in quantity demanded, resulting, again, in a decrease in total revenue.

Income Elasticity of Demand

Recall that one of the independent variables in our example of a demand function for gasoline was income. The sensitivity of quantity demanded to a change in income is termed **income elasticity**. Holding other independent variables constant, we can measure income elasticity as the ratio of the percentage change in quantity demanded to the percentage change in income.

For most goods, the sign of income elasticity is positive—an increase in income leads to an increase in quantity demanded. Goods for which this is the case are termed **normal goods**. For other goods, it may be the case that an increase in income leads to a decrease in quantity demanded. Goods for which this is true are termed **inferior goods**.

Cross-Price Elasticity of Demand

Recall that some of the independent variables in a demand function are the prices of related goods (related in the sense that their prices affect the demand for the good in question). The ratio of the percentage change in the quantity demanded of a good to the percentage change in the price of a related good is termed the **cross-price elasticity of demand**.

When an increase in the price of a related good increases demand for a good, the two goods are substitutes. If Bread A and Bread B are two brands of bread, considered good substitutes by many consumers, an increase in the price of one will lead consumers to purchase more of the other (substitute the other). When the cross-price elasticity of demand is positive (price of one is up and quantity demanded for the other is up), we say those goods are substitutes.

When an increase in the price of a related good decreases demand for a good, the two goods are **complements**. If an increase in the price of automobiles (less automobiles purchased) leads to a decrease in the demand for gasoline, they are complements. Right shoes and left shoes are perfect complements for most of us and, as a result, shoes are priced by the pair. If they were priced separately, there is little doubt that an increase in the price of left shoes would decrease the quantity demanded of right shoes. Overall, the cross-price elasticity of demand is more positive the better substitutes two goods are and more negative the better complements the two goods are.

Calculating Elasticities

The price elasticity of demand is defined as:

$$\frac{\% \Delta Q}{\% \Delta P} = \frac{\Delta Q / Q_0}{\Delta P / P_0} = \left(\frac{P_0}{Q_0} \right) \times \left(\frac{\Delta Q}{\Delta P} \right)$$

The term $\frac{\Delta Q}{\Delta P}$ is the slope of a demand *function* that (for a linear demand function) takes the form:

$$\text{quantity demanded} = A + B \times \text{price}$$

In such a function, B is the slope of the line. A demand *curve* is the inverse of the demand function, in which price is given as a function of quantity demanded.

As an example, consider a demand function with $A = 100$ and $B = -2$, so that $Q = 100 - 2P$. The slope, $\frac{\Delta Q}{\Delta P}$, of this line is -2 . The corresponding demand curve for this demand function is: $P = 100 / 2 - Q / 2 = 50 - 1/2 Q$. Therefore, given a demand curve, we can calculate the slope of the demand function as the reciprocal of slope term, $-1/2$, of the demand curve (i.e., the reciprocal of $-1/2$ is -2 , the slope of the demand function).

EXAMPLE: Calculating price elasticity of demand

A demand function for gasoline is as follows:

$$Q_{D\text{gas}} = 138,500 - 12,500P_{\text{gas}}$$

Calculate the price elasticity at a gasoline price of \$3 per gallon.

Answer:

We can calculate the quantity demanded at a price of \$3 per gallon as $138,500 - 12,500(3) = 101,000$. Substituting 3 for P_0 , 101,000 for Q_0 , and $-12,500$ for $(\frac{\Delta Q}{\Delta P})$, we can calculate the price elasticity of demand as:

$$E_{\text{Demand}} = \frac{\% \Delta Q}{\% \Delta P} = \left(\frac{3}{101,000} \right) \times (-12,500) = -0.37$$

For this demand function, at a price and quantity of \$3 per gallon and 101,000 gallons, demand is inelastic.

The techniques for calculating the income elasticity of demand and the cross-price elasticity of demand are the same, as illustrated in the following example. We assume values for all the independent variables, except the one of interest, then calculate elasticity for a given value of the variable of interest.

EXAMPLE: Calculating income elasticity and cross-price elasticity

An individual has the following demand function for gasoline:

$$Q_D \text{ gas} = 15 - 3P_{\text{gas}} + 0.02I + 0.11P_{\text{BT}} - 0.008P_{\text{auto}}$$

where income and car price are measured in thousands, and the price of bus travel is measured in average dollars per 100 miles traveled.

Assuming the average automobile price is \$22,000, income is \$40,000, the price of bus travel is \$25, and the price of gasoline is \$3, calculate and interpret the income elasticity of gasoline demand and the cross-price elasticity of gasoline demand with respect to the price of bus travel.

Answer:

Inserting the prices of gasoline, bus travel, and automobiles into our demand equation, we get:

$$Q_D \text{ gas} = 15 - 3(3) + 0.02(\text{income in thousands}) + 0.11(25) - 0.008(22)$$

and

$$Q_D \text{ gas} = 8.574 + 0.02(\text{income in thousands})$$

Our slope term on income is 0.02, and for an income of 40,000, $Q_D \text{ gas} = 9.374$ gallons.

The formula for the income elasticity of demand is:

$$\frac{\% \Delta Q}{\% \Delta I} = \frac{\Delta Q / Q_0}{\Delta I / I_0} = \left(\frac{I_0}{Q_0} \right) \times \left(\frac{\Delta Q}{\Delta I} \right)$$

Substituting our calculated values, we have:

$$\left(\frac{40}{9.374} \right) \times (0.02) = 0.085$$

This tells us that for these assumed values (at a single point on the demand curve), a 1% increase (decrease) in income will lead to an increase (decrease) of 0.085% in the quantity of gasoline demanded.

In order to calculate the cross-price elasticity of demand for bus travel and gasoline, we construct a demand function with only the price of bus travel as an independent variable:

$$Q_D \text{ gas} = 15 - 3P_{\text{gas}} + 0.02I + 0.11P_{\text{BT}} - 0.008P_{\text{auto}}$$

$$Q_D \text{ gas} = 15 - 3(3) + 0.02(40) + 0.11P_{\text{BT}} - 0.008(22)$$

$$Q_D \text{ gas} = 6.624 + 0.11P_{\text{BT}}$$

For a price of bus travel of \$25, the quantity of gasoline demanded is:

$$Q_D \text{ gas} = 6.624 + 0.11P_{\text{BT}}$$

$$Q_D \text{ gas} = 6.624 + 0.11(25) = 9.374 \text{ gallons}$$

The cross-price elasticity of the demand for gasoline with respect to the price of bus travel is:

$$\frac{\% \Delta Q}{\% \Delta P_{\text{BT}}} = \frac{\Delta Q / Q_0}{\Delta P_{\text{BT}} / P_{0 \text{ BT}}} = \left(\frac{P_{0 \text{ BT}}}{Q_0} \right) \times \left(\frac{\Delta Q}{\Delta P_{\text{BT}}} \right) = \frac{25}{9.374} \times 0.11$$

$$= 0.293$$

As noted, gasoline and bus travel are substitutes, so the cross-price elasticity of demand is positive. We can interpret this value to mean that, for our assumed values, a 1% change in the price of bus travel will lead to a 0.293% change in the quantity of gasoline demanded in the same direction, other things equal.

MODULE 8.2: DEMAND AND SUPPLY



LOS 8.b: Compare substitution and income effects.

Video covering this content is available online.

When the price of Good X decreases, there is a **substitution effect** that shifts consumption towards more of Good X. Because the total expenditure on the consumer's original bundle of goods falls when the price of Good X falls, there is also an **income effect**. The income effect can be toward more or less consumption of Good X. This is the key point here: the substitution effect always acts to increase the consumption of a good that has fallen in price, while the income effect can either increase or decrease consumption of a good that has fallen in price.

Based on this analysis, we can describe three possible outcomes of a decrease in the price of Good X:

1. The substitution effect is positive, and the income effect is also positive—consumption of Good X will *increase*.
2. The substitution effect is positive, and the income effect is negative but smaller than the substitution effect—consumption of Good X will *increase*.
3. The substitution effect is positive, and the income effect is negative and larger than the substitution effect—consumption of Good X will *decrease*.

LOS 8.c: Contrast normal goods with inferior goods.



PROFESSOR'S NOTE

Candidates who are not already familiar with profit maximization based on a firm's cost curves (e.g., average cost and marginal cost) and firm revenue (e.g., average revenue, total revenue, and marginal revenue) should study the material in the CFA curriculum prerequisite reading "Demand and Supply Analysis: The Firm" prior to their study of the following material.

Earlier, we defined normal goods and inferior goods in terms of their income elasticity of demand. A normal good is one for which the income effect is positive. An inferior good is one for which the income effect is negative.

A specific good may be an inferior good for some ranges of income and a normal good for other ranges of income. For a really poor person or population (e.g., underdeveloped country), an increase in income may lead to greater consumption of noodles or rice. Now, if incomes rise a bit (e.g., college student or developing country), more meat or seafood may become part of the diet. Over this range of incomes, noodles can be an inferior good and ground meat a normal good. If incomes rise to a higher range (e.g., graduated from college and got a job), the consumption of ground meat may fall (inferior) in favor of preferred cuts of meat (normal).

For many of us, commercial airline travel is a normal good. When our incomes rise, vacations are more likely to involve airline travel, be more frequent, and extend over longer distances so that airline travel is a normal good. For wealthy people (e.g., hedge fund manager), an increase in income may lead to travel by private jet and a decrease in the quantity of commercial airline travel demanded.

A **Giffen good** is an inferior good for which the negative income effect outweighs the positive substitution effect when price falls. A Giffen good is theoretical and would have an upward-sloping demand curve. At lower prices, a smaller quantity would be demanded as a result of the dominance of the income effect over the substitution effect. Note that the existence of a Giffen good is not ruled out by the axioms of the theory of consumer choice.

A **Veblen good** is one for which a higher price makes the good more desirable. The idea is that the consumer gets utility from being seen to consume a good that has high status (e.g., Gucci bag), and that a higher price for the good conveys more status and increases its utility. Such a good could conceivably have a positively sloped demand curve for some individuals over some range of prices. If such a good exists, there must be a limit to this process, or the price would rise without limit. Note that the existence of a Veblen good does violate the theory of consumer choice. If a Veblen good exists, it is not an inferior good, so both the substitution and income effects of a price increase are to decrease consumption of the good.

LOS 8.d: Describe the phenomenon of diminishing marginal returns.

Factors of production are the resources a firm uses to generate output. Factors of production include:

- *Land*—where the business facilities are located.
- *Labor*—includes all workers from unskilled laborers to top management.
- *Capital*—sometimes called physical capital or plant and equipment to distinguish it from financial capital. Refers to manufacturing facilities, equipment, and machinery.
- *Materials*—refers to inputs into the productive process, including raw materials, such as iron ore or water, or manufactured inputs, such as wire or microprocessors.

For economic analysis, we often consider only two inputs, capital and labor. The quantity of output that a firm can produce can be thought of as a function of the amounts of capital and labor employed. Such a function is called a **production function**.

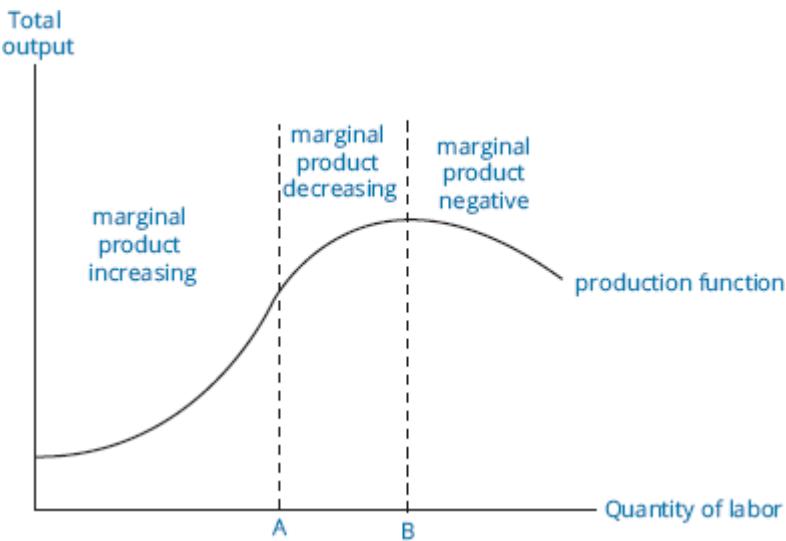
If we consider a given amount of capital (a firm's plant and equipment), we can examine the increase in production (increase in total product) that will result as we increase the amount of labor employed. The output with only one worker is considered the marginal product of the first unit of labor. The addition of a second worker will increase total product by the marginal product of the second worker. The marginal product of (additional output from) the second worker is likely greater than the marginal product of the first. This is true if we assume that two workers can produce more than twice as much output as one because of the benefits of teamwork or specialization of tasks. At this low range of labor input (remember, we are holding capital constant), we can say that the marginal product of labor is increasing.

As we continue to add additional workers to a fixed amount of capital, at some point, adding one more worker will increase total product by less than the addition of the previous worker, although total product continues to increase. When we reach the quantity of labor for which the additional output for each additional worker begins to decline, we have reached the point of **diminishing marginal productivity** of labor, or that labor has reached the point of **diminishing marginal returns**. Beyond this quantity of labor, the additional output from each additional worker continues to decline.

There is, theoretically, some quantity for labor for which the marginal product of labor is actually negative (i.e., the addition of one more worker actually decreases total output).

In Figure 8.3, we illustrate all three cases. For quantities of labor between zero and A, the marginal product of labor is increasing (slope is increasing). Beyond the inflection point in the production at quantity of labor A up to quantity B, the marginal product of labor is still positive but decreasing. The slope of the production function is positive but decreasing, and we are in a range of diminishing marginal productivity of labor. Beyond the quantity of labor B, adding additional workers decreases total output. The marginal product of labor in this range is negative, and the production function slopes downward.

Figure 8.3: Production Function—Capital Fixed, Labor Variable



LOS 8.e: Determine and interpret breakeven and shutdown points of production.

In economics, we define the **short run** for a firm as the time period over which some factors of production are fixed. Typically, we assume that capital is fixed in the short run so that a firm cannot change its scale of operations (plant and equipment) over the short run. All factors of production (costs) are variable in the **long run**. The firm can let its leases expire and sell its equipment, thereby avoiding costs that are fixed in the short run.

Shutdown and Breakeven Under Perfect Competition

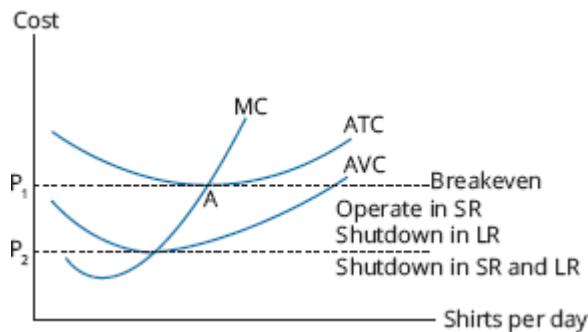
As a simple example of shutdown and breakeven analysis, consider a retail store with a 1-year lease (fixed cost) and one employee (quasi-fixed cost), so that variable costs are simply the store's cost of merchandise. If the total sales (total revenue) just covers both fixed and variable costs, price equals both average revenue and average total cost, so we are at the breakeven output quantity, and economic profit equals zero.

During the period of the lease (the short run), as long as items are being sold for more than their variable cost, the store should continue to operate to minimize losses. If items are being sold for less than their average variable cost, losses would be reduced by shutting down the business in the short run.

In the long run, a firm should shut down if the price is less than average total cost, regardless of the relation between price and average variable cost.

In the case of a firm under perfect competition, price = marginal revenue = average revenue, as we have noted. For a firm under perfect competition (a price taker), we can use a graph of cost functions to examine the profitability of the firm at different output prices. In Figure 8.4, at price P_1 , price and average revenue equal average total cost. At the output level of Point A, the firm is making an economic profit of zero. At a price above P_1 , economic profit is positive, and at prices less than P_1 , economic profit is negative (the firm has economic losses).

Figure 8.4: Shutdown and Breakeven



Because some costs are fixed in the short run, it will be better for the firm to continue production in the short run as long as average revenue is greater than average variable costs. At prices between P_1 and P_2 in Figure 8.4, the firm has losses, but the loss is less than the losses that would occur if all production were stopped. As long as total revenue is greater than total variable cost, at least some of the firm's fixed costs are covered by continuing to produce and sell its product. If the firm were to shut down, losses would be equal to the fixed costs that still must be paid. As long as price is greater than average variable costs, the firm will minimize its losses in the short run by continuing in business.

If average revenue is less average variable cost, the firm's losses are greater than its fixed costs, and it will minimize its losses by shutting down production in the short run. In this case (a price less than P_2 in Figure 8.4), the loss from continuing to operate is greater than the loss (total fixed costs) if the firm is shut down.

In the long run, all costs are variable, so a firm can avoid its (short-run) fixed costs by shutting down. For this reason, if price is expected to remain below minimum average total cost (Point A in Figure 8.4) in the long run, the firm will shut down rather than continue to generate losses.

To sum up, if average revenue is less than average variable cost in the short run, the firm should shut down. This is its **short-run shutdown point**. If average revenue is greater than average variable cost in the short run, the firm should continue to operate, even if it has losses. In the long run, the firm should shut down if average revenue is less than average total cost. This is the **long-run shutdown point**. If average revenue is just equal to average total cost, total revenue is just equal to total (economic) cost, and this is the firm's **breakeven point**.

- If $AR \geq ATC$, the firm should stay in the market in both the short and long run.
- If $AR \geq AVC$, but $AR < ATC$, the firm should stay in the market in the short run but will exit the market in the long run.
- If $AR < AVC$, the firm should shut down in the short run and exit the market in the long run.

Shutdown and Breakeven Under Imperfect Competition

For price-searcher firms (those that face downward-sloping demand curves), we could compare average revenue to ATC and AVC, just as we did for price-taker firms, to identify shutdown and breakeven points. However, marginal revenue is no longer equal to price.

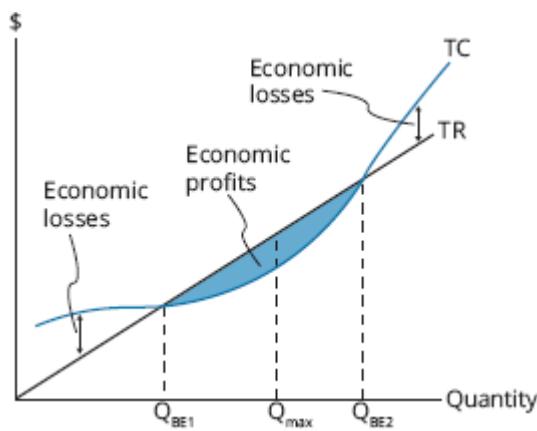
We can, however, still identify the conditions under which a firm is breaking even, should shut down in the short run, and should shut down in the long run in terms of total costs and total revenue. These conditions are:

- $TR = TC$: break even.
- $TC > TR > TVC$: firm should continue to operate in the short run but shut down in the long run.
- $TR < TVC$: firm should shut down in the short run and the long run.

Because price does not equal marginal revenue for a firm in imperfect competition, analysis based on total costs and revenues is better suited for examining breakeven and shutdown points.

The previously described relations hold for both price-taker and price-searcher firms. We illustrate these relations in Figure 8.5 for a price-taker firm (TR increases at a constant rate with quantity). Total cost equals total revenue at the breakeven quantities Q_{BE1} and Q_{BE2} . The quantity for which economic profit is maximized is shown as Q_{max} .

Figure 8.5: Breakeven Point Using the Total Revenue/Total Cost Approach



If the entire TC curve exceeds TR (i.e., no breakeven point), the firm will want to minimize the economic loss in the short run by operating at the quantity corresponding to the smallest (negative) value of $TR - TC$.

EXAMPLE: Short-run shutdown decision

For the last fiscal year, Legion Gaming reported total revenue of \$700,000, total variable costs of \$800,000, and total fixed costs of \$400,000. Should the firm continue to operate in the short run?

Answer:

The firm should shut down. Total revenue of \$700,000 is less than total costs of \$1,200,000 and also less than total variable costs of \$800,000. By shutting down, the firm will lose an amount equal to fixed costs of \$400,000. This is less than the loss of operating, which is $TR - TC = \$500,000$.

EXAMPLE: Long-run shutdown decision

Suppose instead that Legion reported total revenue of \$850,000. Should the firm continue to operate in the short run? Should it continue to operate in the long run?

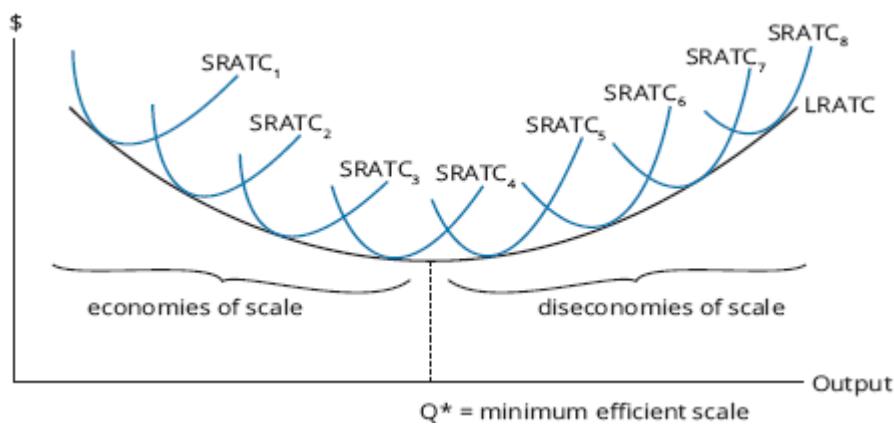
Answer:

In the short run, $TR > TVC$, and the firm should continue operating. The firm should consider exiting the market in the long run, as TR is not sufficient to cover all of the fixed costs and variable costs.

LOS 8.f: Describe how economies of scale and diseconomies of scale affect costs.

While plant size is fixed in the short run, in the long run, firms can choose their most profitable scale of operations. Because the long-run average total cost (LRATC) curve is drawn for many different plant sizes or scales of operation, each point along the curve represents the minimum ATC for a given plant size or scale of operations. In Figure 8.6, we show a firm's LRATC curve along with short-run average total cost (SRATC) curves for many different plant sizes, with $SRATC_{n+1}$ representing a larger scale of operations than $SRATC_n$.

Figure 8.6: Economies and Diseconomies of Scale



We draw the LRATC curve as U-shaped. Average total costs first decrease with larger scale and eventually increase. The lowest point on the LRATC corresponds to the scale or plant size at which the average total cost of production is at a minimum. This scale is sometimes called the **minimum efficient scale**. Under perfect competition, firms must operate at minimum efficient scale in long-run equilibrium, and LRATC will equal the market price. Recall that under perfect competition, firms earn zero economic profit in long-run equilibrium. Firms that have chosen a different scale of operations with higher average total costs will have economic losses and must either leave the industry or change to minimum efficient scale.

The downward-sloping segment of the long-run average total cost curve presented in Figure 8.6 indicates that economies of scale (or *increasing returns to scale*) are present. Economies of scale result from factors such as labor specialization, mass production, and investment in more efficient equipment and technology. In addition, the firm may be able to negotiate lower input prices with suppliers as firm size increases and more resources are purchased. A firm operating with economies of scale can increase its competitiveness by expanding production and reducing costs.

The upward-sloping segment of the LRATC curve indicates that **diseconomies of scale** are present. Diseconomies of scale may result as the increasing bureaucracy of larger firms leads to inefficiency, problems with motivating a larger workforce, and greater barriers to innovation.

and entrepreneurial activity. A firm operating under diseconomies of scale will want to decrease output and move back toward the minimum efficient scale. The U.S. auto industry is an example of an industry that has exhibited diseconomies of scale.

There may be a relatively flat portion at the bottom of the LRATC curve that exhibits *constant returns to scale*. Over a range of constant returns to scale, costs are constant for the various plant sizes.



MODULE QUIZ 8.1, 8.2

1. Total revenue is greatest in the part of a demand curve that is:

- A. elastic
- B. inelastic
- C. unit elastic.

2. A demand function for air conditioners is given by:

$$Q_D \text{air conditioner} = 10,000 - 2 P_{\text{air conditioner}} + 0.0004 \text{ income} + 30 P_{\text{electric fan}} - 4 P_{\text{electricity}}$$

At current average prices, an air conditioner costs 5,000 yen, a fan costs 200 yen, and electricity costs 1,000 yen. Average income is 4,000,000 yen. The income elasticity of demand for air conditioners is *closest* to:

- A. 0.0004.
- B. 0.444.
- C. 40,000.

3. When the price of a good decreases, and an individual's consumption of that good also decreases, it is *most likely* that:

- A. the income effect and substitution effect are both negative.
- B. the substitution effect is negative and the income effect is positive.
- C. the income effect is negative and the substitution effect is positive.

4. A good is classified as an inferior good if its:

- A. income elasticity is negative.
- B. own-price elasticity is negative.
- C. cross-price elasticity is negative.

5. Increasing the amount of one productive input while keeping the amounts of other inputs constant results in diminishing marginal returns:

- A. in all cases.
- B. when it causes total output to decrease.
- C. when the increase in total output becomes smaller.

6. A firm's average revenue is greater than its average variable cost and less than its average total cost. If this situation is expected to persist, the firm should:

- A. shut down in the short run and in the long run.
- B. shut down in the short run but operate in the long run.
- C. operate in the short run but shut down in the long run.

7. If a firm's long-run average total cost increases by 6% when output is increased by 6%, the firm is experiencing:

- A. economies of scale.
- B. diseconomies of scale.
- C. constant returns to scale.

KEY CONCEPTS

LOS 8.a

Elasticity is measured as the ratio of the percentage change in one variable to a percentage change in another. Three elasticities related to a demand function are of interest:

$$\text{own-price elasticity} = \frac{\% \text{ change in quantity demanded}}{\% \text{ change in own price}}$$

$$\text{cross-price elasticity} = \frac{\% \text{ change in quantity demanded}}{\% \text{ change in price of related good}}$$

$$\text{income elasticity} = \frac{\% \text{ change in quantity demanded}}{\% \text{ change in income}}$$

$|\text{own-price elasticity}| > 1$: demand is elastic

$|\text{own-price elasticity}| < 1$: demand is inelastic

$\text{cross-price elasticity} > 0$: related good is a substitute

$\text{cross-price elasticity} < 0$: related good is a complement

$\text{income elasticity} < 0$: good is an inferior good

$\text{income elasticity} > 0$: good is a normal good

LOS 8.b

When the price of a good decreases, the substitution effect leads a consumer to consume more of that good and less of goods for which prices have remained the same.

A decrease in the price of a good that a consumer purchases leaves her with unspent income (for the same combination of goods). The effect of this additional income on consumption of the good for which the price has decreased is termed the income effect.

LOS 8.c

For a normal good, the income effect of a price decrease is positive—income elasticity of demand is positive.

For an inferior good, the income effect of a price decrease is negative—income elasticity of demand is negative. An increase in income reduces demand for an inferior good.

A Giffen good is an inferior good for which the negative income effect of a price decrease outweighs the positive substitution effect, so that a decrease (increase) in the good's price has a net result of decreasing (increasing) the quantity consumed.

A Veblen good is also one for which an increase (decrease) in price results in an increase (decrease) in the quantity consumed. However, a Veblen good is not an inferior good and is not supported by the axioms of the theory of demand.

LOS 8.d

Marginal returns refer to the additional output that can be produced by using one more unit of a productive input while holding the quantities of other inputs constant. Marginal returns may increase as the first units of an input are added, but as input quantities increase, they reach a point at which marginal returns begin to decrease. Inputs beyond this quantity are said to produce diminishing marginal returns.

LOS 8.e

Under perfect competition:

- The breakeven quantity of production is the quantity for which price (P) = average total cost (ATC) and total revenue (TR) = total cost (TC).
- The firm should shut down in the long run if $P < ATC$ so that $TR < TC$.
- The firm should shut down in the short run (and the long run) if $P <$ average variable cost (AVC) so that $TR <$ total variable cost (TVC).

Under imperfect competition (firm faces downward sloping demand):

- Breakeven quantity is the quantity for which $TR = TC$.
- The firm should shut down in the long run if $TR < TC$.
- The firm should shut down in the short run (and the long run) if $TR < TVC$.

LOS 8.f

The long-run average total cost (LRATC) curve shows the minimum average total cost for each level of output assuming that the plant size (scale of the firm) can be adjusted. A downward-sloping segment of an LRATC curve indicates economies of scale (increasing returns to scale). Over such a segment, increasing the scale of the firm reduces ATC. An upward-sloping segment of an LRATC curve indicates diseconomies of scale, where average unit costs will rise as the scale of the business (and long-run output) increases.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 8.1, 8.2

1. **C** Total revenue is maximized at the quantity at which own-price elasticity equals -1. (Module 8.1, LOS 8.a)

2. **B** Substituting current values for the independent variables other than income, the demand function becomes:

$$\begin{aligned} QD_{\text{air conditioner}} &= 10,000 - 2(5,000) + 0.0004 \text{ income} + 30(200) - 4(1,000) \\ &= 0.0004 \text{ income} + 2,000. \end{aligned}$$

The slope of income is 0.0004, and for an income of 4,000,000 yen, $QD = 3,600$.

Income elasticity = $I_0 / Q_0 \times \Delta Q / \Delta I = 4,000,000 / 3,600 \times 0.0004 = 0.444$. (Module 8.1, LOS 8.a)

3. **C** The substitution effect of a price decrease is always positive, but the income effect can be either positive or negative. Consumption of a good will decrease when the price of that good decreases only if the income effect is both negative and greater than the substitution effect. (Module 8.2, LOS 8.b)

4. **A** An inferior good is one that has a negative income elasticity of demand. (Module 8.2, LOS 8.c)

5. **C** Productive inputs exhibit diminishing marginal returns at the level where an additional unit of input results in a smaller increase in output than the previous unit of input. (Module 8.2, LOS 8.d)

6. **C** If a firm is generating sufficient revenue to cover its variable costs and part of its fixed costs, it should continue to operate in the short run. If average revenue is likely to remain below average total costs in the long run, the firm should shut down. (Module 8.2, LOS 8.e)

7. **B** Increasing long-run average total cost as a result of increasing output demonstrates diseconomies of scale. (Module 8.2, LOS 8.f)

READING 9

THE FIRM AND MARKET STRUCTURES

EXAM FOCUS

This reading covers four market structures: perfect competition, monopolistic competition, oligopoly, and monopoly. You need to be able to compare and contrast these structures in terms of numbers of firms, firm demand elasticity and pricing power, long-run economic profits, barriers to entry, and the amount of product differentiation and advertising. Finally, know the two quantitative concentration measures, their implications for market structure and pricing power, and their limitations in this regard. We will apply all of these concepts when we analyze industry competition and pricing power of companies in the Equity Investments topic area.

MODULE 9.1: PERFECT COMPETITION



LOS 9.a: Describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly.

Video covering this content is available online.

In this reading, we examine four types of market structure: perfect competition, monopolistic competition, oligopoly, and monopoly. We can analyze where an industry falls along this spectrum by examining the following five factors:

1. Number of firms and their relative sizes.
2. Degree to which firms differentiate their products.
3. Bargaining power of firms with respect to pricing.
4. Barriers to entry into or exit from the industry.
5. Degree to which firms compete on factors other than price.

At one end of the spectrum is **perfect competition**, in which many firms produce identical products, and competition forces them all to sell at the market price. At the other extreme, we have **monopoly**, where only one firm is producing the product. In between are **monopolistic competition** (many sellers and differentiated products) and **oligopoly** (few firms that compete in a variety of ways). Each market structure has its own characteristics and implications for firm strategy, and we will examine each in turn.

Perfect competition refers to a market in which many firms produce identical products, barriers to entry into the market are very low, and firms compete for sales only on the basis of price. Firms face perfectly elastic (horizontal) demand curves at the price determined in the market because no firm is large enough to affect the market price. The market for wheat in a region is a

good approximation of such a market. Overall market supply and demand determine the price of wheat.

Monopolistic competition differs from perfect competition in that products are not identical. Each firm differentiates its product(s) from those of other firms through some combination of differences in product quality, product features, and marketing. The demand curve faced by each firm is downward sloping; while demand is elastic, it is not perfectly elastic. Prices are not identical because of perceived differences among competing products, and barriers to entry are low. The market for toothpaste is a good example of monopolistic competition. Firms differentiate their products through features and marketing with claims of more attractiveness, whiter teeth, fresher breath, and even of actually cleaning your teeth and preventing decay. If the price of your personal favorite increases, you are not likely to immediately switch to another brand as under perfect competition. Some customers would switch in response to a 10% increase in price and some would not. This is why firm demand is downward sloping.

The most important characteristic of an *oligopoly* market is that there are only a few firms competing. In such a market, each firm must consider the actions and responses of other firms in setting price and business strategy. We say that such firms are interdependent. While products are typically good substitutes for each other, they may be either quite similar or differentiated through features, branding, marketing, and quality. Barriers to entry are high, often because economies of scale in production or marketing lead to very large firms. Demand can be more or less elastic than for firms in monopolistic competition. The automobile market is dominated by a few very large firms and can be characterized as an oligopoly. The product and pricing decisions of Toyota certainly affect those of Ford and vice versa. Automobile makers compete based on price, but also through marketing, product features, and quality, which is often signaled strongly through brand name. The oil industry also has a few dominant firms but their products are very good substitutes for each other.

A *monopoly* market is characterized by a single seller of a product with no close substitutes. This fact alone means that the firm faces a downward-sloping demand curve (the market demand curve) and has the power to choose the price at which it sells its product. High barriers to entry protect a monopoly producer from competition. One source of monopoly power is the protection offered by copyrights and patents. Another possible source of monopoly power is control over a resource specifically needed to produce the product. Most frequently, monopoly power is supported by government. A **natural monopoly** refers to a situation where the average cost of production is falling over the relevant range of consumer demand. In this case, having two (or more) producers would result in a significantly higher cost of production and be detrimental to consumers. Examples of natural monopolies include the electric power and distribution business and other public utilities. When privately owned companies are granted such monopoly power, the price they charge is often regulated by government as well.

Sometimes market power is the result of *network effects* or *synergies* that make it very difficult to compete with a company once it has reached a critical level of market penetration. eBay gained such a large share of the online auction market that its information on buyers and sellers and the number of buyers who visit eBay essentially precluded others from establishing competing businesses. While it may have competition to some degree, its market share is such that it has negatively sloped demand and a good deal of pricing power. Sometimes we refer to such companies as having a moat around them that protects them from competition. It is best to

remember, however, that changes in technology and consumer tastes can, and usually do, reduce market power over time. Polaroid had a monopoly on instant photos for years, but the introduction of digital photography forced the firm into bankruptcy in 2001.

The table in Figure 9.1 shows the key features of each market structure.

Figure 9.1: Characteristics of Market Structures

	Perfect Competition	Monopolistic Competition	Oligopoly	Monopoly
Number of sellers	Many firms	Many firms	Few firms	Single firm
Barriers to entry	Very low	Low	High	Very high
Nature of substitute products	Very good substitutes	Good substitutes but differentiated	Very good substitutes or differentiated	No good substitutes
Nature of competition	Price only	Price, marketing, features	Price, marketing, features	Advertising
Pricing power	None	Some	Some to significant	Significant

LOS 9.b: Explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure.

LOS 9.d: Describe and determine the optimal price and output for firms under each market structure.

LOS 9.f: Explain factors affecting long-run equilibrium under each market structure.

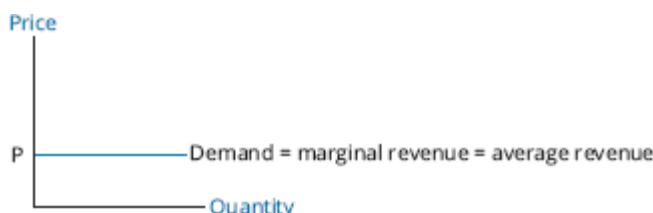


PROFESSOR'S NOTE

We cover these LOS together and slightly out of curriculum order so that we can present the complete analysis of each market structure to better help candidates understand the economics of each type of market structure.

Producer firms in perfect competition have no influence over market price. Market supply and demand determine price. As illustrated in Figure 9.2, the *individual firm's* demand schedule is *perfectly elastic* (horizontal).

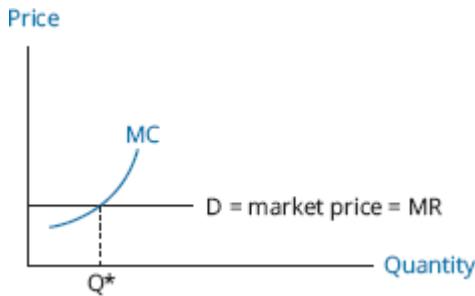
Figure 9.2: Price-Taker Demand



In a perfectly competitive market, a firm will continue to expand production until marginal revenue (MR) equals marginal cost (MC). Marginal revenue is the increase in total revenue from selling one more unit of a good or service. For a price taker, marginal revenue is simply the price

because all additional units are assumed to be sold at the same (market) price. In *pure competition*, a firm's marginal revenue is equal to the market price, and a firm's MR curve, presented in Figure 9.3, is identical to its demand curve. A profit maximizing firm will produce the quantity, Q^* , when $MC = MR$.

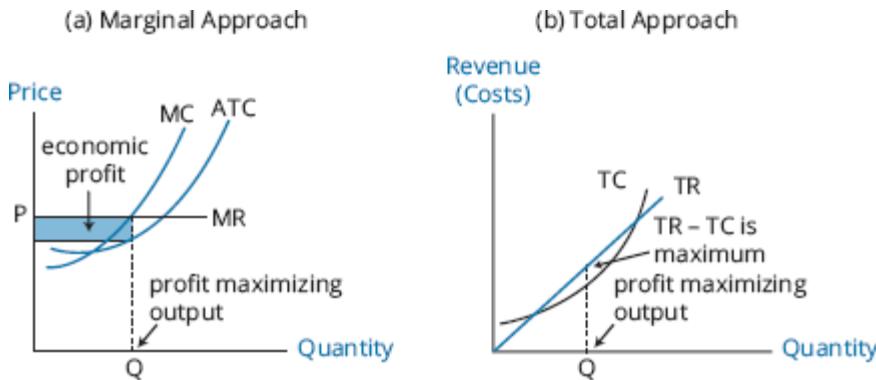
Figure 9.3: Profit Maximizing Output For A Price Taker



All firms maximize (economic) profit by producing and selling the quantity for which marginal revenue equals marginal cost. For a firm in a perfectly competitive market, this is the same as producing and selling the quantity for which marginal cost equals (market) price. Economic profit equals total revenues less the opportunity cost of production, which includes the cost of a normal return to all factors of production, including invested capital.

Panel (a) of Figure 9.4 illustrates that in the *short run*, economic profit is maximized at the quantity for which $marginal\ revenue = marginal\ cost$. As shown in Panel (b), profit maximization also occurs when total revenue exceeds total cost by the maximum amount.

Figure 9.4: Short-Run Profit Maximization



An *economic loss* occurs on any units for which marginal revenue is less than marginal cost. At any output above the quantity where $MR = MC$, the firm will be generating losses on its marginal production and will maximize profits by reducing output to where $MR = MC$.

In a perfectly competitive market, firms will not earn economic profits for any significant period of time. The assumption is that new firms (with average and marginal cost curves identical to those of existing firms) will enter the industry to earn economic profits, increasing market supply and eventually reducing market price so that it just equals firms' average total cost (ATC). In equilibrium, each firm is producing the quantity for which $P = MR = MC = ATC$, so that no firm earns economic profits and each firm is producing the quantity for which ATC is a

minimum (the quantity for which $ATC = MC$). This equilibrium situation is illustrated in Figure 9.5.

Figure 9.5: Equilibrium in a Perfectly Competitive Market

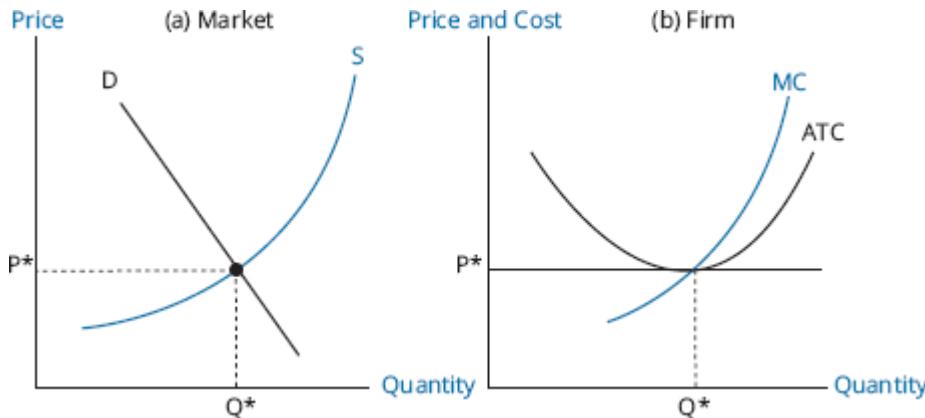
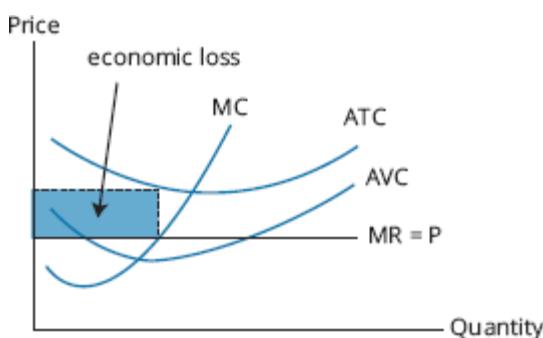


Figure 9.6 illustrates that firms will experience economic losses when price is below average total cost ($P < ATC$). In this case, the firm must decide whether to continue operating. A firm will minimize its losses in the short run by continuing to operate when price is less than ATC but greater than AVC . As long as the firm is covering its variable costs and some of its fixed costs, its loss will be less than its fixed (in the short run) costs. If the firm is only just covering its variable costs ($P = AVC$), the firm is operating at its **shutdown point**. If the firm is not covering its variable costs ($P < AVC$) by continuing to operate, its losses will be greater than its fixed costs. In this case, the firm will shut down (zero output) and lay off its workers. This will limit its losses to its fixed costs (e.g., its building lease and debt payments). If the firm does not believe price will ever exceed ATC in the future, going out of business is the only way to eliminate fixed costs.

Figure 9.6: Short-Run Loss

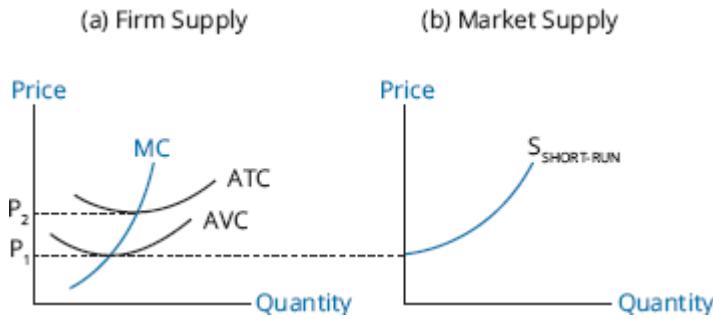


The *long-run equilibrium output* level for perfectly competitive firms is where $MR = MC = ATC$, which is where ATC is at a minimum. At this output, economic profit is zero and only a normal return is realized.

Recall that price takers should produce where $P = MC$. Referring to Panel (a) in Figure 9.7, a firm will shut down at a price below P_1 . Between P_1 and P_2 , a firm will continue to operate in the short run. At P_2 , the firm is earning a normal profit—economic profit equals zero. At prices above P_2 , a firm is making economic profits and will expand its production along the MC line.

Thus, the **short-run supply curve for a firm** is its MC line above the average variable cost curve, AVC. The supply curve shown in Panel (b) is the **short-run market supply curve**, which is the horizontal sum (add up the quantities from all firms at each price) of the MC curves for all firms in a given industry. Because firms will supply more units at higher prices, the short-run market supply curve slopes upward to the right.

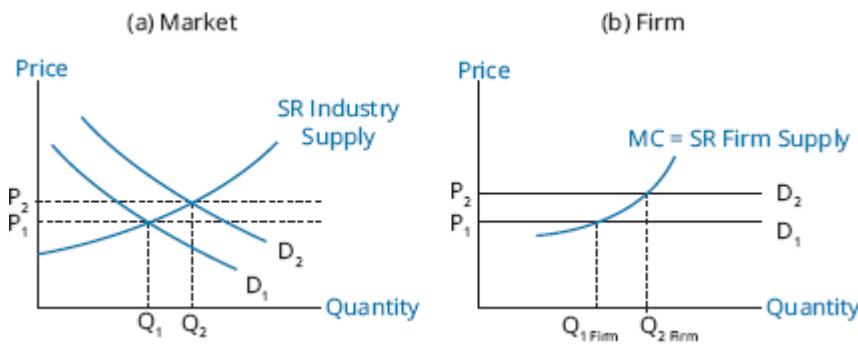
Figure 9.7: Short-Run Supply Curves



Changes in Demand, Entry and Exit, and Changes in Plant Size

In the short run, an increase in market demand (a shift of the market demand curve to the right) will increase both equilibrium price and quantity, while a decrease in market demand will reduce both equilibrium price and quantity. The change in equilibrium price will change the (horizontal) demand curve faced by each individual firm and the profit-maximizing output of a firm. These effects for an increase in demand are illustrated in Figure 9.8. An increase in market demand from D_1 to D_2 increases the short-run equilibrium price from P_1 to P_2 and equilibrium output from Q_1 to Q_2 . In Panel (b) of Figure 9.8, we see the short-run effect of the increased market price on the output of an individual firm. The higher price leads to a greater profit-maximizing output, $Q_{2\text{ Firm}}$. At the higher output level, a firm will earn an economic profit in the short run. In the long run, some firms will increase their scale of operations in response to the increase in demand, and new firms will likely enter the industry. In response to a decrease in demand, the short-run equilibrium price and quantity will fall, and in the long run, firms will decrease their scale of operations or exit the market.

Figure 9.8: Short-Run Adjustment to an Increase in Demand Under Perfect Competition



A firm's long-run adjustment to a shift in industry demand and the resulting change in price may be either to alter the size of its plant or leave the market entirely. The marketplace abounds with examples of firms that have increased their plant sizes (or added additional production facilities) to increase output in response to increasing market demand. Other firms, such as Ford

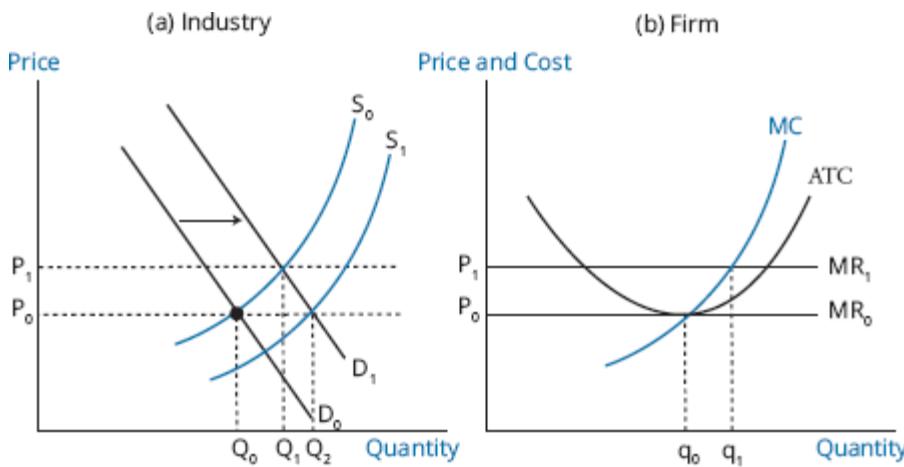
and GM, have decreased plant size to reduce economic losses. This strategy is commonly referred to as *downsizing*.

If an industry is characterized by firms earning economic profits, new firms will enter the market. This will cause industry supply to increase (the industry supply curve shifts downward and to the right), increasing equilibrium output and decreasing equilibrium price. Even though industry output increases, however, individual firms will produce less because as price falls, each individual firm will move down its own supply curve. The end result is that a firm's total revenue and economic profit will decrease.

If firms in an industry are experiencing economic losses, some of these firms will exit the market. This will decrease industry supply and increase equilibrium price. Each remaining firm in the industry will move up its individual supply curve and increase production at the higher market price. This will cause total revenues to increase, reducing any economic losses the remaining firms had been experiencing.

A *permanent change in demand* leads to the entry of firms to, or exit of firms from, an industry. Let's consider the permanent increase in demand illustrated in Figure 9.9. The initial long-run industry equilibrium condition shown in Panel (a) is at the intersection of demand curve D_0 and supply curve S_0 , at price P_0 and quantity Q_0 . As indicated in Panel (b) of Figure 9.9, at the market price of P_0 each firm will produce q_0 . At this price and output, each firm earns a normal profit, and economic profit is zero. That is, $MC = MR = P$, and ATC is at its minimum. Now, suppose industry demand permanently increases such that the industry demand curve in Panel (a) shifts to D_1 . The new market price will be P_1 and industry output will increase to Q_1 . At the new price P_1 , existing firms will produce q_1 and realize an economic profit because $P_1 > ATC$. Positive economic profits will cause new firms to enter the market. As these new firms increase total industry supply, the industry supply curve will gradually shift to S_1 , and the market price will decline back to P_0 . At the market price of P_0 , the industry will now produce Q_2 , with an increased number of firms in the industry, each producing at the original quantity, q_0 . The individual firms will no longer enjoy an economic profit because $ATC = P_0$ at q_0 .

Figure 9.9: Effects of a Permanent Increase in Demand



MODULE QUIZ 9.1

1. When a firm operates under conditions of pure competition, marginal revenue always equals:
 - A. price.
 - B. average cost.
 - C. marginal cost.
2. In which market structure(s) can a firm's supply function be described as its marginal cost curve above its average variable cost curve?
 - A. Oligopoly or monopoly.
 - B. Perfect competition only.
 - C. Perfect competition or monopolistic competition.
3. In a purely competitive market, economic losses indicate that:
 - A. price is below average total costs.
 - B. collusion is occurring in the market place.
 - C. firms need to expand output to reduce costs.
4. A purely competitive firm will tend to expand its output so long as:
 - A. marginal revenue is positive.
 - B. marginal revenue is greater than price.
 - C. market price is greater than marginal cost.
5. A firm is likely to operate in the short run as long as price is at least as great as:
 - A. marginal cost.
 - B. average total cost.
 - C. average variable cost.

MODULE 9.2: MONOPOLISTIC COMPETITION



Monopolistic competition has the following market characteristics:

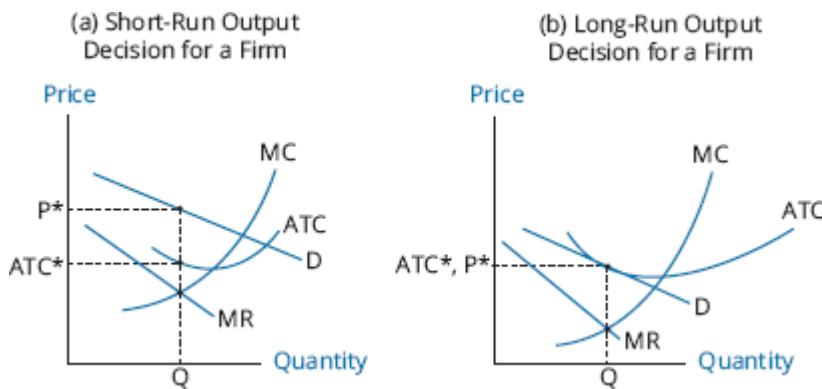
- *A large number of independent sellers:* (1) Each firm has a relatively small market share, so no individual firm has any significant power over price. (2) Firms need only pay attention to average market price, not the price of individual competitors. (3) There are too many firms in the industry for collusion (price fixing) to be possible.
- *Differentiated products:* Each producer has a product that is slightly different from its competitors (at least in the minds of consumers). The competing products are close substitutes for one another.
- *Firms compete on price, quality, and marketing* as a result of product differentiation. *Quality* is a significant product-differentiating characteristic. *Price* and output can be set by firms because they face downward-sloping demand curves, but there is usually a strong correlation between quality and the price that firms can charge. *Marketing* is a must to inform the market about a product's differentiating characteristics.
- *Low barriers to entry* so that firms are free to enter and exit the market. If firms in the industry are earning economic profits, new firms can be expected to enter the industry.

Firms in monopolistic competition face *downward-sloping demand* curves (they are price searchers). Their demand curves are highly *elastic* because competing products are perceived by consumers as close substitutes. Think about the market for toothpaste. All toothpaste is quite similar, but differentiation occurs due to taste preferences, influential advertising, and the reputation of the seller.

Video covering
this content is
available online.

The price/output decision for monopolistic competition is illustrated in Figure 9.10. Panel (a) of Figure 9.10 illustrates the short-run price/output characteristics of monopolistic competition for a single firm. As indicated, firms in monopolistic competition maximize economic profits by producing where marginal revenue (MR) equals marginal cost (MC), and by charging the price for that quantity from the demand curve, D . Here the firm earns positive economic profits because price, P^* , exceeds average total cost, ATC^* . Due to low barriers to entry, competitors will enter the market in pursuit of these economic profits.

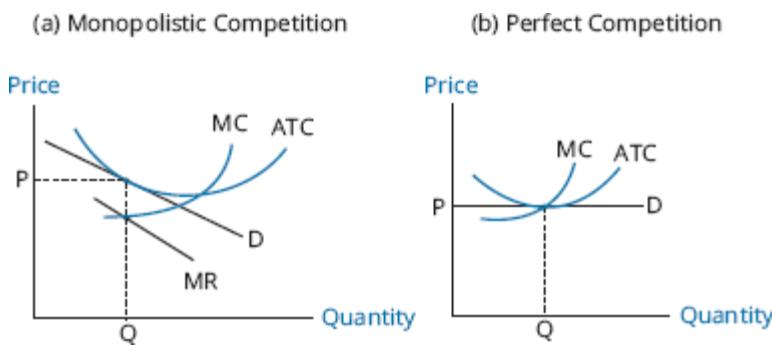
Figure 9.10: Short-Run and Long-Run Output Under Monopolistic Competition



Panel (b) of Figure 9.10 illustrates long-run equilibrium for a *representative firm* after new firms have entered the market. As indicated, the entry of new firms shifts the demand curve faced by each individual firm down to the point where price equals average total cost ($P^* = ATC^*$), such that economic profit is zero. At this point, there is no longer an incentive for new firms to enter the market, and long-run equilibrium is established. The firm in monopolistic competition continues to produce at the quantity where $MR = MC$ but no longer earns positive economic profits.

Figure 9.11 illustrates the differences between long-run equilibrium in markets with monopolistic competition and markets with perfect competition. Note that with monopolistic competition, price is greater than marginal cost (i.e., producers can realize a **markup**), average total cost is not at a minimum for the quantity produced (suggesting **excess capacity**, or an inefficient scale of production), and the price is slightly higher than under perfect competition. The point to consider here, however, is that perfect competition is characterized by no product differentiation. The question of the efficiency of monopolistic competition becomes, “Is there an economically efficient amount of product differentiation?”

Figure 9.11: Firm Output Under Monopolistic and Perfect Competition



In a world with only one brand of toothpaste, clearly average production costs would be lower. That fact alone probably does not mean that a world with only one brand/type of toothpaste would be a better world. While product differentiation has costs, it also has benefits to consumers.

Consumers definitely benefit from brand name promotion and advertising because they receive information about the nature of a product. This often enables consumers to make better purchasing decisions. Convincing consumers that a particular brand of deodorant will actually increase their confidence in a business meeting or make them more attractive to the opposite sex is not easy or inexpensive. Whether the perception of increased confidence or attractiveness from using a particular product is worth the additional cost of advertising is a question probably better left to consumers of the products. Some would argue that the increased cost of advertising and sales is not justified by the benefits of these activities.

Product innovation is a necessary activity as firms in monopolistic competition pursue economic profits. Firms that bring new and innovative products to the market are confronted with less-elastic demand curves, enabling them to increase price and earn economic profits. However, close substitutes and imitations will eventually erode the initial economic profit from an innovative product. Thus, firms in monopolistic competition must continually look for innovative product features that will make their products relatively more desirable to some consumers than those of the competition.

Innovation does not come without costs. The costs of product innovation must be weighed against the extra revenue that it produces. A firm is considered to be spending the optimal amount on innovation when the marginal cost of (additional) innovation just equals the marginal revenue (marginal benefit) of additional innovation.

Advertising expenses are high for firms in monopolistic competition. This is to inform consumers about the unique features of their products and to create or increase a perception of differences between products that are actually quite similar. We just note here that advertising costs for firms in monopolistic competition are greater than those for firms in perfect competition and those that are monopolies.

As you might expect, advertising costs increase the average total cost curve for a firm in monopolistic competition. The increase to average total cost attributable to advertising decreases as output increases, because more fixed advertising dollars are being averaged over a larger quantity. In fact, if advertising leads to enough of an increase in output (sales), it can actually decrease a firm's average total cost.

Brand names provide information to consumers by providing them with signals about the quality of the branded product. Many firms spend a significant portion of their advertising budget on brand name promotion. Seeing the brand name BMW likely tells a consumer more about the quality of a newly introduced automobile than an inspection of the vehicle itself would reveal. At the same time, the reputation BMW has for high quality is so valuable that the firm has an added incentive not to damage it by producing vehicles of low quality.



MODULE QUIZ 9.2

1. The demand for products from monopolistic competitors is relatively elastic due to:
 - A. high barriers to entry.

- B. the availability of many close substitutes.
 - C. the availability of many complementary goods.
2. Compared to a perfectly competitive industry, in an industry characterized by monopolistic competition:
- A. both price and quantity are likely to be lower.
 - B. price is likely to be higher and quantity is likely to be lower.
 - C. quantity is likely to be higher and price is likely to be lower.
3. A firm will *most likely* maximize profits at the quantity of output for which:
- A. price equals marginal cost.
 - B. price equals marginal revenue.
 - C. marginal cost equals marginal revenue.

MODULE 9.3: OLIGOPOLY



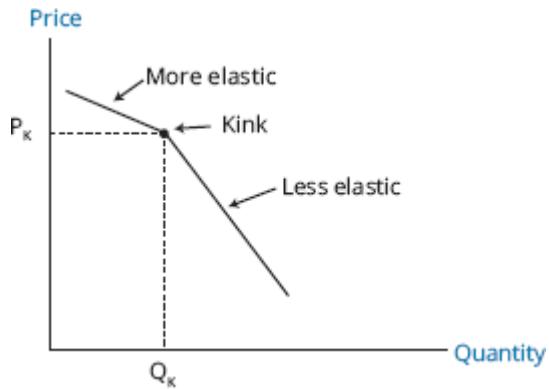
Compared to monopolistic competition, an oligopoly market has higher barriers to entry and fewer firms. The other key difference is that the firms are interdependent, so a price change by one firm can be expected to be met by a price change by its competitors. This means that the actions of another firm will directly affect a given firm's demand curve for the product. Given this complicating fact, models of oligopoly pricing and profits must make a number of important assumptions. In the following, we describe four of these models and their implications for price and quantity:

Video covering
this content is
available online.

1. Kinked demand curve model.
2. Cournot duopoly model.
3. Nash equilibrium model (prisoner's dilemma).
4. Stackelberg dominant firm model.

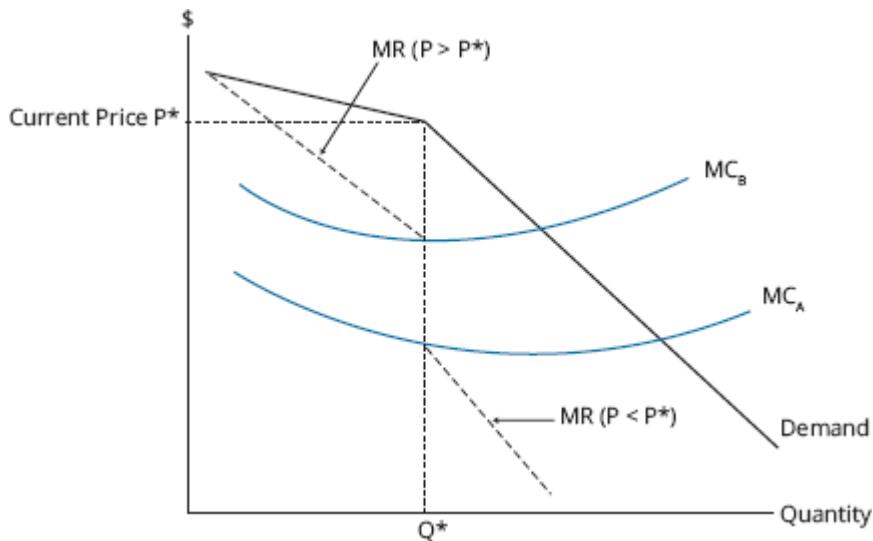
One traditional model of oligopoly, the **kinked demand curve model**, is based on the assumption that an increase in a firm's product price will not be followed by its competitors, but a decrease in price will. According to the kinked demand curve model, each firm believes that it faces a demand curve that is more elastic (flatter) above a given price (the kink in the demand curve) than it is below the given price. The kinked demand curve model is illustrated in Figure 9.12. The kink price is at price P_K , where a firm produces Q_K . A firm believes that if it raises its price above P_K , its competitors will remain at P_K , and it will lose market share because it has the highest price. Above P_K , the demand curve is considered to be relatively elastic, where a small price increase will result in a large decrease in demand. On the other hand, if a firm decreases its price below P_K , other firms will match the price cut, and all firms will experience a relatively small increase in sales relative to any price reduction. Therefore, Q_K is the profit-maximizing level of output.

Figure 9.12: Kinked Demand Curve Model



It is worth noting that with a kink in the market demand curve, we also get a gap in the associated marginal revenue curve, as shown in Figure 9.13. For any firm with a marginal cost curve passing through this gap, the price at which the kink is located is the firm's profit maximizing price.

Figure 9.13: Gap in Marginal Revenue Curve



We say that the decisions of firms in an oligopoly are interdependent; that is, the pricing decision of one firm depends on the pricing decision of another firm. Some models of market price equilibrium have a set of rules for the actions of oligopolists. These rules assume they choose prices based on the choices of the other firms. By specifying the decision rules that each firm follows, we can design a model that allows us to determine the equilibrium prices and quantities for firms operating in an oligopoly market. An early model of oligopoly pricing decisions is the **Cournot model**. In Cournot's model, two firms with identical marginal cost curves each choose their preferred selling price based on the price the other firm chose in the previous period. The equilibrium for an oligopoly with two firms (duopoly), in the Cournot model, is for both firms to sell the same amounts and same quantities, splitting the market equally at the equilibrium price. The equilibrium price is less than the price a single monopolist would charge, but greater than the equilibrium price that would result under perfect competition.

Another model, the **Stackelberg model**, uses a different set of rules. One firm is the "leader" and chooses its price first, and the other firm chooses a price based on the leader's price. In

equilibrium, under these rules, the leader charges a higher price and receives a greater proportion of the firms' total profits.

Firms determine their quantities simultaneously each period and, under the assumptions of the Cournot model, these quantities will change each period until they are equal. When each firm selects the same quantity, there is no longer any additional profit to be gained by changing quantity, and we have a stable equilibrium. The resulting market price is less than the profit maximizing price that a monopolist would charge, but higher than marginal cost, the price that would result from perfect competition. Additional analysis shows that as more firms are added to the model, the equilibrium market price falls towards marginal cost, which is the equilibrium price in the limit as the number of firms gets large.

These rules-based models are early versions of what are called *strategic games*, decision models in which the best choice for a firm depends on the actions (reactions) of other firms. A more general model of this strategic game was developed by Nobel Prize winner John Nash, who developed the concept of a **Nash equilibrium**. A Nash equilibrium is reached when the choices of all firms are such that there is no other choice that makes any firm better off (increases profits or decreases losses).

The concept of a Nash equilibrium can be applied to the situation presented in Figure 9.14, which shows the choices and resulting profits for two firms. Each firm can charge either a high price or a low price. If both firms charge a high price, Firm A earns 1,000 and Firm B earns 600. While Firm A would not charge the low price (it would earn less regardless of Firm B's decision), Firm B can increase profits to 700 by charging a low price. With Firm A charging a high price and Firm B charging a low price, neither firm can increase profits by changing its price strategy. Thus, we can identify the Nash equilibrium in this scenario as Firm B charging a low price and Firm A charging a high price.

Figure 9.14: Nash Equilibrium

	Firm B High Price	Firm B Low Price
Firm A High Price	A earns 1,000 B earns 600	A earns 600 B earns 700
Firm A Low Price	A earns 160 B earns 0	A earns 100 B earns 140

The firms could, however, collude. The greatest joint profits (1,600) are earned when both firms charge a high price. If Firm A offers to pay Firm B 200 for charging a high price, Firm A's profits increase from 600 to 1,000. After paying 200 to Firm B, Firm A still gains 200. Firm B's profits (including the payment of 200) increase from 700 to 800. Collusion, in this case, increases the profits of both firms, compared to the Nash equilibrium.

If firms can enter into and enforce an agreement regarding pricing and output, often they can all benefit. Such agreements among producers are illegal in many countries because they reduce competition.

An example of a collusive agreement is the OPEC **cartel**. Cartel member countries agree to restrict their oil production in order to increase the world price of oil. Members sometimes choose to "cheat" on the cartel agreement by producing more than the amount of oil they have

agreed to produce. If members of a cartel do not adhere to the agreement, taking advantage of the higher market price but failing to restrict output to the agreed-upon amount, the agreement can quickly break down.

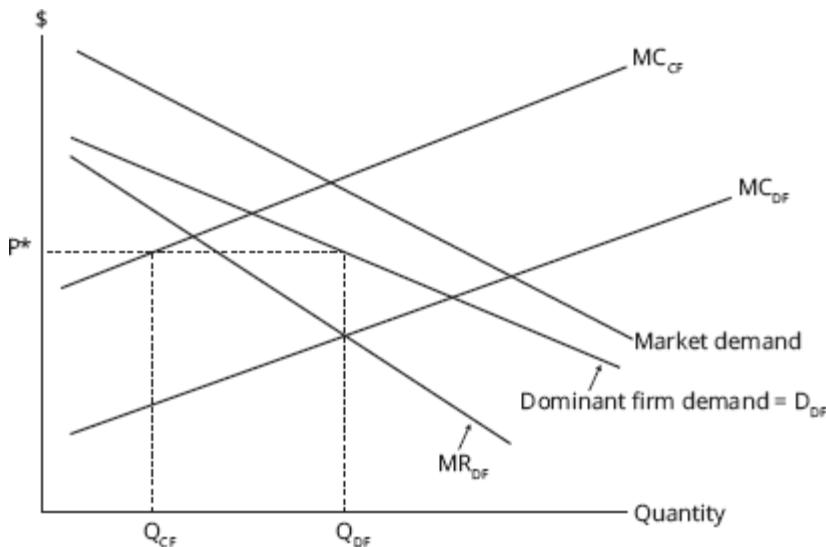
In general, collusive agreements to increase price in an oligopoly market will be more successful (have less cheating) when:

- There are fewer firms.
- Products are more similar (less differentiated).
- Cost structures are more similar.
- Purchases are relatively small and frequent.
- Retaliation by other firms for cheating is more certain and more severe.
- There is less actual or potential competition from firms outside the cartel.

A final model of oligopoly behavior to consider is the **dominant firm model**. In this model, there is a single firm that has a significantly large market share because of its greater scale and lower cost structure—the dominant firm (DF). In such a model, the market price is essentially determined by the dominant firm, and the other competitive firms (CF) take this market price as given.

The dominant firm believes that the quantity supplied by the other firms decreases at lower prices, so that the dominant firm's demand curve is related to the market demand curve as shown in Figure 9.15. Based on this demand curve (D_{DF}) and its associated marginal revenue (MR_{DF}) curve, the firm will maximize profits at a price of P^* . The competitive firms maximize profits by producing the quantity for which their marginal cost (MC_{CF}) equals P^* , quantity Q_{CF} .

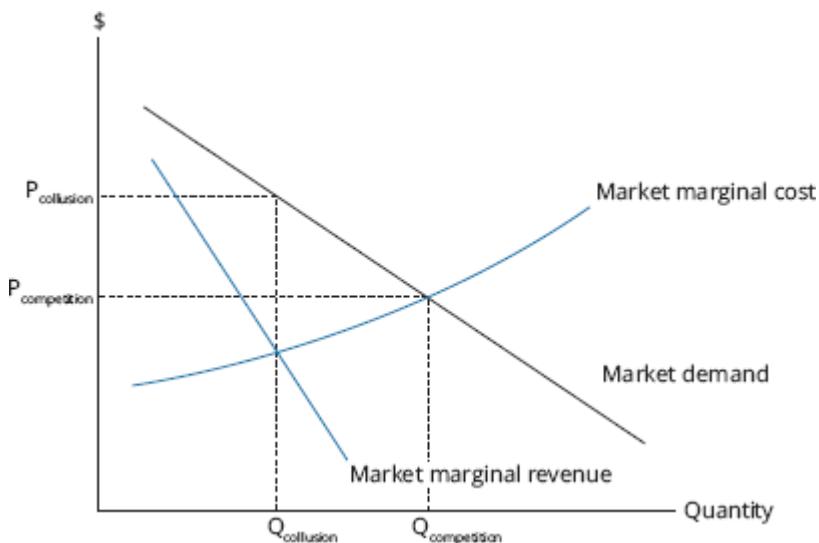
Figure 9.15: Dominant Firm Oligopoly



A price decrease by one of the competitive firms, which increases Q_{CF} in the short run, will lead to a decrease in price by the dominant firm, and competitive firms will decrease output and/or exit the industry in the long run. The long-run result of such a price decrease by competitors below P^* would then be to decrease the overall market share of competitor firms and increase the market share of the dominant firm.

Clearly, there are many possible outcomes in oligopoly markets that depend on the characteristics of the firms and the market itself. The important point is that the firms' decisions are interdependent so that the expected reaction of other firms is an important consideration. Overall, the resulting price will be somewhere between the price based on perfect collusion that would maximize total profits to all firms in the market (actually the monopoly price, which is addressed next) and the price that would result from perfect competition and generate zero economic profits in the long run. These two limiting outcomes are illustrated in Figure 9.16 as $P_{\text{collusion}}$ with $Q_{\text{collusion}}$ for perfect collusion and $P_{\text{competition}}$ and $Q_{\text{competition}}$ for perfect competition.

Figure 9.16: Collusion vs. Perfect Competition



MODULE QUIZ 9.3

1. An oligopolistic industry has:
 - few barriers to entry.
 - few economies of scale.
 - a great deal of interdependence among firms.
2. Consider a firm in an oligopoly market that believes the demand curve for its product is more elastic above a certain price than below this price. This belief fits *most closely* to which of the following models?
 - Cournot model.
 - Dominant firm model.
 - Kinked demand model.
3. Consider an agreement between France and Germany that will restrict wine production so that maximum economic profit can be realized. The possible outcomes of the agreement are presented in the table below.

	Germany complies	Germany defaults
France complies	France gets €8 billion Germany gets €8 billion	France gets €2 billion Germany gets €10 billion
France defaults	France gets €10 billion Germany gets €2 billion	France gets €4 billion Germany gets €4 billion

Based on the concept of a Nash equilibrium, the *most likely* strategy followed by the two countries with respect to whether they comply with or default on the agreement will be:

- A. both countries will default.
- B. both countries will comply.
- C. one country will default and the other will comply.

MODULE 9.4: MONOPOLY AND CONCENTRATION



A monopoly faces a downward-sloping demand curve for its product, so profit maximization involves a trade-off between price and quantity sold if the firm sells at the same price to all buyers. Assuming a single selling price, a monopoly firm must lower its price in order to sell a greater quantity. Unlike a firm in perfect competition, a firm facing a downward-sloping demand curve must determine what price to charge, hoping to find the price and output combination that will bring the maximum profit to the firm.

Video covering this content is available online.

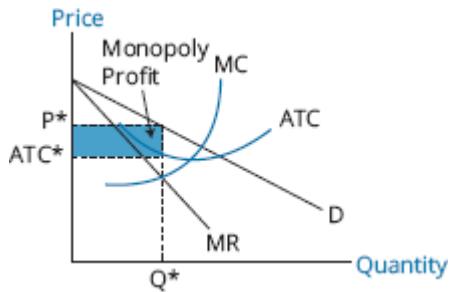
Two pricing strategies that are possible for a monopoly firm are *single-price* and *price discrimination*. If the monopoly's customers cannot resell the product to each other, the monopoly can maximize profits by charging different prices to different groups of customers. When price discrimination isn't possible, the monopoly will charge a single price. Price discrimination is described in more detail after we address single-price profit maximization.

To maximize profit, monopolists will expand output until marginal revenue (MR) equals marginal cost (MC). Due to high entry barriers, monopolist profits do not attract new market entrants. Therefore, long-run positive economic profits can exist. Do monopolists charge the highest possible price? The answer is no, because monopolists want to maximize profits, not price.

One way to calculate marginal revenue for a firm that faces a downward-sloping demand curve and sells all units for the same price is $MR = P \left(1 - \frac{1}{E_p}\right)$, where MR is marginal revenue, P is the current price, and E_p is the absolute value of the price elasticity of demand at price = P. Therefore, we can also express the single-price profit-maximizing output as that output for which $MC = P \left(1 - \frac{1}{E_p}\right)$.

Figure 9.17 shows the revenue-cost structure facing the monopolist. Note that production will expand until $MR = MC$ at optimal output Q^* . To find the price at which it will sell Q^* units, you must go to the demand curve. The demand curve itself does not determine the optimal behavior of the monopolist. Just like the perfect competition model, the profit maximizing output for a monopolist is where $MR = MC$. To ensure a profit, the demand curve must lie above the firm's average total cost (ATC) curve at the optimal quantity so that $price > ATC$. The optimal quantity will be in the elastic range of the demand curve.

Figure 9.17: Monopoly Short-Run Costs and Revenues



Once again, the *profit maximizing* output for a monopolistic firm is the one for which $MR = MC$. As shown in Figure 9.17, the profit maximizing output is Q^* , with a price of P^* , and an economic profit equal to $(P^* - ATC^*) \times Q^*$.

Monopolists are *price searchers* and have *imperfect information* regarding market demand. They must experiment with different prices to find the one that maximizes profit.

Price discrimination is the practice of charging different consumers different prices for the same product or service. Examples are different prices for airline tickets based on whether a Saturday-night stay is involved (separates business travelers and leisure travelers) and different prices for movie tickets based on age.

The motivation for a monopolist is to capture more consumer surplus as economic profit than is possible by charging a single price.

For price discrimination to work, the seller must:

- Face a downward-sloping demand curve.
- Have at least two identifiable groups of customers with *different price elasticities of demand* for the product.
- Be able to prevent the customers paying the lower price from reselling the product to the customers paying the higher price.

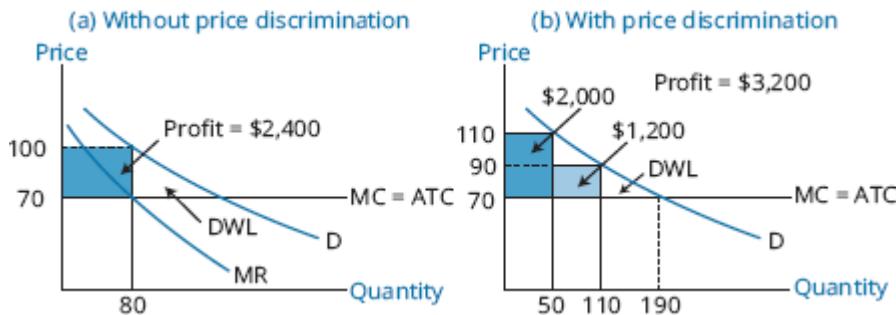
As long as these conditions are met, firm profits can be increased through price discrimination.

Figure 9.18 illustrates how price discrimination can increase the total quantity supplied and increase economic profits compared to a single-price pricing strategy. For simplicity, we have assumed no fixed costs and constant variable costs so that $MC = ATC$. In Panel (a), the single profit-maximizing price is \$100 at a quantity of 80 (where $MC = MR$), which generates a profit of \$2,400. In Panel (b), the firm is able to separate consumers, charges one group \$110 and sells them 50 units, and sells an additional 60 units to another group (with more elastic demand) at a price of \$90. Total profit is increased to \$3,200, and total output is increased from 80 units to 110 units.

Compared to the quantity produced under perfect competition, the quantity produced by a monopolist reduces the sum of consumer and producer surplus by an amount represented by the triangle labeled *deadweight loss* (DWL) in Panel (a) of Figure 9.18. Consumer surplus is reduced not only by the decrease in quantity but also by the increase in price relative to perfect competition. Monopoly is considered inefficient because the reduction in output compared to perfect competition reduces the sum of consumer and producer surplus. Because marginal benefit is greater than marginal cost, less than the efficient quantity of resources are allocated to the production of the good. Price discrimination reduces this inefficiency by increasing

output toward the quantity where marginal benefit equals marginal cost. Note that the deadweight loss is smaller in Panel (b). The firm gains from those customers with inelastic demand while still providing goods to customers with more elastic demand. This may even cause production to take place when it would not otherwise.

Figure 9.18: Effect of Price Discrimination on Output and Operating Profit



An extreme (and largely theoretical) case of price discrimination is perfect price discrimination. If it were possible for the monopolist to charge each consumer the maximum they are willing to pay for each unit, there would be no deadweight loss because a monopolist would produce the same quantity as under perfect competition. With perfect price discrimination, there would be no consumer surplus. It would all be captured by the monopolist.

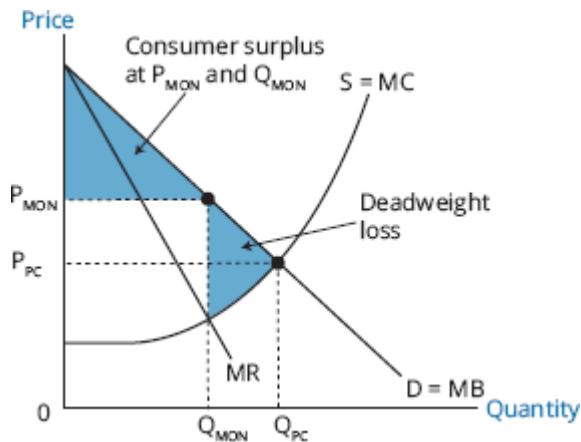
Figure 9.19 illustrates the difference in allocative efficiency between monopoly and perfect competition. Under perfect competition, the industry supply curve, S , is the sum of the supply curves of the many competing firms in the industry. The perfect competition equilibrium price and quantity are at the intersection of the industry supply curve and the market demand curve, D . The quantity produced is Q_{PC} at an equilibrium price P_{PC} . Because each firm is small relative to the industry, there is nothing to be gained by attempting to decrease output in an effort to increase price.

A monopolist facing the same demand curve, and with the same marginal cost curve, MC , will maximize profit by producing Q_{MON} (where $MC = MR$) and charging a price of P_{MON} .

The important thing to note here is that when compared to a perfectly competitive industry, the monopoly firm will produce less total output and charge a higher price.

Recall from our review of perfect competition that the efficient quantity is the one for which the sum of consumer surplus and producer surplus is maximized. In Figure 9.19, this quantity is where $S = D$, or equivalently, where marginal cost (MC) = marginal benefit (MB). *Monopoly creates a deadweight loss* relative to perfect competition because monopolies produce a quantity that does not maximize the sum of consumer surplus and producer surplus. A further loss of efficiency results from **rent seeking** when producers spend time and resources to try to acquire or establish a monopoly.

Figure 9.19: Perfect Competition vs. Monopoly

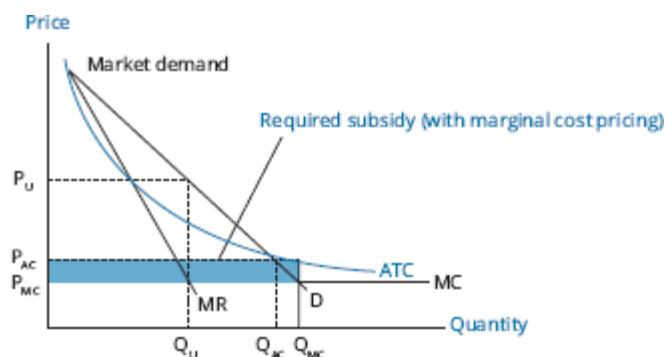


Natural Monopoly

In some industries, the economics of production lead to a single firm supplying the entire market demand for the product. When there are large economies of scale, it means that the average cost of production decreases as a single firm produces greater and greater output. An example is an electric utility. The fixed costs of producing electricity and building the power lines and related equipment to deliver it to homes are quite high. The marginal cost of providing electricity to an additional home or of providing more electricity to a home is, however, quite low. The more electricity provided, the lower the average cost per kilowatt hour. When the average cost of production for a single firm is falling throughout the relevant range of consumer demand, we say that the industry is a **natural monopoly**. The entry of another firm into the industry would divide the production between two firms and result in a higher average cost of production than for a single producer. Thus, large economies of scale in an industry present significant barriers to entry.

We illustrate the case of a natural monopoly in Figure 9.20. Left unregulated, a single-price monopolist will maximize profits by producing where $MR = MC$, producing quantity Q_U and charging P_U . Given the economies of scale, having another firm in the market would increase the ATC significantly. Note in Figure 9.20 that if two firms each produced approximately one-half of output Q_{AC} , average cost for each firm would be much higher than for a single producer producing Q_{AC} . Thus, there is a potential gain from monopoly because of lower average cost production when LRAC is decreasing so that economies of scale lead to a single supplier.

Figure 9.20: Natural Monopoly—Average Cost and Marginal Cost Pricing



Regulators often attempt to increase competition and efficiency through efforts to reduce artificial barriers to trade, such as licensing requirements, quotas, and tariffs.

Because monopolists produce less than the optimal quantity (do not achieve efficient resource allocation), government regulation may be aimed at improving resource allocation by regulating the prices monopolies may charge. This may be done through average cost pricing or marginal cost pricing.

Average cost pricing is the most common form of regulation. This would result in a price of P_{AC} and an output of Q_{AC} as illustrated in Figure 9.20. It forces monopolists to reduce price to where the firm's ATC intersects the market demand curve. This will:

- Increase output and decrease price.
- Increase social welfare (allocative efficiency).
- Ensure the monopolist a *normal* profit because price = ATC.

Marginal cost pricing, which is also referred to as *efficient regulation*, forces the monopolist to reduce price to the point where the firm's MC curve intersects the market demand curve. This increases output and reduces price, but causes the monopolist to incur a loss because price is below ATC, as illustrated in Figure 9.20. Such a solution requires a government subsidy in order to provide the firm with a normal profit and prevent it from leaving the market entirely.

Another way of regulating a monopoly is for the government to sell the monopoly right to the highest bidder. The right to build a gasoline station and food court on a tollway is one example. In theory, the winning bidder will be an efficient supplier that bids an amount equal to the value of expected economic profit and sets prices equal to long-run average cost.

LOS 9.c: Describe a firm's supply function under each market structure.

The short-run supply function for a firm under perfect competition is its marginal cost curve above its average variable cost curve, as described earlier. The short-run market supply curve is constructed simply by summing the quantities supplied at each price across all firms in the market.

In markets characterized as monopolistic competition, oligopoly, and monopoly, there is no well-defined supply function. This is because under all three of these market structures, firms face downward-sloping demand curves. In each case, the quantity supplied is determined by the intersection of marginal cost and marginal revenue, and the price charged is then determined by the demand curve the firm faces. We cannot construct a function of quantity supplied as a function of price as we can under perfect competition, where price equals marginal revenue. The quantity supplied depends not only on a firm's marginal cost, but on demand and marginal revenue (which change with quantity) as well.

LOS 9.e: Describe pricing strategy under each market structure.

We have covered each market structure separately in detail, so we will simply summarize optimal pricing strategies.

Perfect competition: Profits are maximized by producing the quantity for which marginal cost equals marginal revenue. Note that marginal revenue and price are equal so price also equals marginal cost at the profit-maximizing quantity.

Monopoly: Profits are also maximized by producing the quantity for which marginal revenue equals marginal cost. Because the firm's demand curve is downward sloping, price is greater than marginal revenue and greater than marginal cost.

Monopolistic competition: Profits are maximized when a firm produces the quantity for which marginal revenue equals marginal cost. Similar to a monopoly structure, the firm faces a downward sloping demand curve and price will be greater than marginal cost and marginal revenue.

Oligopoly: Because one of the key characteristics of oligopoly is the interdependence of firms' pricing and output decisions, the optimal pricing strategy depends on our assumptions about the reactions of other firms to each firm's actions. Here we note different possible assumptions and the strategy that is implied by each.

1. Kinked demand curve: This assumes competitors will match a price decrease but not a price increase. Firms produce the quantity for which marginal revenue equals marginal cost. However, the marginal revenue curve is discontinuous (there's a gap in it), so for many cost structures the optimal quantity is the same, given they face the same kinked demand curve.
2. Collusion: If all producers agree to share the market to maximize total industry profits, they will produce a total quantity for which marginal cost equals marginal revenue and charge the price from the industry demand curve at which that quantity can be sold. This is the same overall price and quantity as for a profit maximizing monopoly firm, but the oligopoly firms must agree to share this total output among themselves and share the economic profits as a result.
3. Dominant firm model: In this case, we assume one firm has the lowest cost structure and a large market share as a result. The dominant firm will maximize profits by producing the quantity for which its marginal cost equals its marginal revenue and charge the price on its firm demand curve for that quantity. Other firms in the market will essentially take that price as given and produce the quantity for which their marginal cost equals that price.
4. Game theory: Because of the interdependence of oligopoly firms' decisions, assumptions about how a competitor will react to a particular price and output decision by a competitor can determine the optimal output and pricing strategy. Given the variety of models and assumptions about competitor reactions, the long-run outcome is indeterminate. We can only say that the price will be between the monopoly price (if firms successfully collude) and the perfect competition price which equals marginal cost (if potential competition rules out prices above that level).

LOS 9.g: Describe the use and limitations of concentration measures in identifying market structure.

When examining the pricing power of firms in an industry, we would like to be able to measure elasticity of demand directly, but that is very difficult. Regulators often use percentage of

market sales (market share) to measure the degree of monopoly or market power of a firm. Often, mergers or acquisitions of companies in the same industry or market are not permitted by government authorities when they determine the market share of the combined firms will be too high and, therefore, detrimental to the economy.

Rather than estimate elasticity of demand, **concentration measures** for a market or industry are very often used as an indicator of market power. One concentration measure is the **N-firm concentration ratio**, which is calculated as the sum of the percentage market shares of the largest N firms in a market. While this measure is simple to calculate and understand, it does not directly measure market power or elasticity of demand.

One limitation of the N-firm concentration ratio is that it may be relatively insensitive to mergers of two firms with large market shares. This problem is reduced by using an alternative measure of market concentration, the **Herfindahl-Hirschman Index (HHI)**. The HHI is calculated as the sum of the squares of the market shares of the largest firms in the market. The following example illustrates this difference between the two measures and their calculation.

EXAMPLE: 4-firm concentration ratios

Given the market shares of the following firms, calculate the 4-firm concentration ratio and the 4-firm HHI, both before and after a merger of Acme and Blake.

Firm	Sales/Total Market Sales
Acme	25%
Blake	15%
Curtis	15%
Dent	10%
Erie	5%
Federal	5%

Answer:

Prior to the merger, the 4-firm concentration ratio for the market is $25 + 15 + 15 + 10 = 65\%$. After the merger, the Acme + Blake firm has 40% of the market, and the 4-firm concentration ratio is $40 + 15 + 10 + 5 = 70\%$. Although the 4-firm concentration ratio has only increased slightly, the market power of the largest firm in the industry has increased significantly from 25% to 40%.

Prior to the merger, the 4-firm HHI is $0.25^2 + 0.15^2 + 0.15^2 + 0.10^2 = 0.1175$.

After the merger, the 4-firm HHI is $0.40^2 + 0.15^2 + 0.10^2 + 0.05^2 = 0.1950$, a significant increase.

A second limitation that applies to both of our simple concentration measures is that barriers to entry are not considered in either case. Even a firm with high market share may not have much pricing power if barriers to entry are low and there is *potential competition*. With low barriers to entry, it may be the case that other firms stand ready to enter the market if firms currently in the market attempt to increase prices significantly. In this case, the elasticity of demand for

existing firms may be high even though they have relatively high market shares and industry concentration measures.

LOS 9.h: Identify the type of market structure within which a firm operates.

The identification of the type of market structure within which a firm is operating is based on the characteristics we outlined earlier. Our earlier table is repeated here in Figure 9.21. Because the analyst is attempting to determine the degree of pricing power firms in the industry have, the focus is on number of firms in the industry, the importance of barriers to entry, the nature of substitute products, and the nature of industry competition. Significant interdependence among firm pricing and output decisions is always a characteristic of an oligopoly market, although some interdependence is present under monopolistic competition, even with many more firms than for an oligopoly structure.

The following table illustrates the differences in characteristics among the various market structures.

Figure 9.21: Characteristics of Market Structures

	Perfect Competition	Monopolistic Competition	Oligopoly	Monopoly
Number of sellers	Many firms	Many firms	Few firms	Single firm
Barriers to entry	Very low	Low	High	Very high
Nature of substitute products	Very good substitutes	Good substitutes but differentiated	Very good substitutes or differentiated	No good substitutes
Nature of competition	Price only	Price, marketing, features	Price, marketing, features	Advertising
Pricing power	None	Some	Some to significant	Significant



MODULE QUIZ 9.4

1. Which of the following statements *most accurately* describes a significant difference between a monopoly firm and a perfectly competitive firm? A perfectly competitive firm:
 - A. minimizes costs; a monopolistic firm maximizes profit.
 - B. maximizes profit; a monopolistic firm maximizes price.
 - C. takes price as given; a monopolistic firm must search for the best price.
2. A monopolist will expand production until $MR = MC$ and charge a price determined by:
 - A. the demand curve.
 - B. the marginal cost curve.
 - C. the average total cost curve.
3. When a regulatory agency requires a monopolist to use average cost pricing, the intent is to price the product where:
 - A. the ATC curve intersects the MR curve.
 - B. the MR curve intersects the demand curve.
 - C. the ATC curve intersects the demand curve.

4. Which of the following is *most likely* an advantage of the Herfindahl-Hirschman Index relative to the N -firm concentration ratio? The Herfindahl-Hirschman Index:
 - A. is simpler to calculate.
 - B. considers barriers to entry.
 - C. is more sensitive to mergers.
5. A market characterized by low barriers to entry, good substitutes, limited pricing power, and marketing of product features is *best* characterized as:
 - A. oligopoly.
 - B. perfect competition.
 - C. monopolistic competition.

KEY CONCEPTS

LOS 9.a

Perfect competition is characterized by:

- Many firms, each small relative to the market.
- Very low barriers to entry into or exit from the industry.
- Homogeneous products that are perfect substitutes, no advertising or branding.
- No pricing power.

Monopolistic competition is characterized by:

- Many firms.
- Low barriers to entry into or exit from the industry.
- Differentiated products, heavy advertising and marketing expenditure.
- Some pricing power.

Oligopoly markets are characterized by:

- Few sellers.
- High barriers to entry into or exit from the industry.
- Products that may be homogeneous or differentiated by branding and advertising.
- Firms that may have significant pricing power.

Monopoly is characterized by:

- A single firm that comprises the whole market.
- Very high barriers to entry into or exit from the industry.
- Advertising used to compete with substitute products.
- Significant pricing power.

LOS 9.b

Perfect competition:

- Price = marginal revenue = marginal cost (in equilibrium).
- Perfectly elastic demand, zero economic profit in equilibrium.

Monopolistic competition:

- Price > marginal revenue = marginal cost (in equilibrium).

- Zero economic profit in long-run equilibrium.

Oligopoly:

- Price > marginal revenue = marginal cost (in equilibrium).
- May have positive economic profit in long-run equilibrium, but moves toward zero economic profit over time.

Monopoly:

- Price > marginal revenue = marginal cost (in equilibrium).
- May have positive economic profit in long-run equilibrium, profits may be zero because of expenditures to preserve monopoly.

LOS 9.c

Under perfect competition, a firm's short-run supply curve is the portion of the firm's short-run marginal cost curve above average variable cost. A firm's long-run supply curve is the portion of the firm's long-run marginal cost curve above average total cost.

Firms operating under monopolistic competition, oligopoly, and monopoly do not have well-defined supply functions, so neither marginal cost curves nor average cost curves are supply curves in these cases.

LOS 9.d

All firms maximize profits by producing the quantity of output for which marginal cost equals marginal revenue. Under perfect competition (perfectly elastic demand), marginal revenue also equals price.

Firms in monopolistic competition or that operate in oligopoly or monopoly markets all face downward-sloping demand curves. Selling price is determined from the price on the demand curve for the profit maximizing quantity of output.

LOS 9.e

Whether a firm operates in perfect competition, monopolistic competition, or is a monopoly, profits are maximized by producing and selling the quantity for which marginal revenue equals marginal cost. Under perfect competition, price equals marginal revenue. Under monopolistic competition or monopoly, firms face downward-sloping demand curves so that marginal revenue is less than price, and the price charged at the profit-maximizing quantity is the price from the firm's demand curve at the optimal (profit-maximizing) level of output.

Under oligopoly, the pricing strategy is not clear. Because firm decisions are interdependent, the optimal pricing and output strategy depends on the assumptions made about other firms' cost structures and about competitors' responses to a firm's price changes.

LOS 9.f

An increase (decrease) in demand will increase (decrease) economic profits in the short run under all market structures. Positive economic profits result in entry of firms into the industry unless barriers to entry are high. Negative economic profits result in exit of firms from the industry unless barriers to exit are high. When firms enter (exit) an industry, market supply increases (decreases), resulting in a decrease (increase) in market price and an increase (decrease) in the equilibrium quantity traded in the market.

LOS 9.g

A concentration ratio for N firms is calculated as the percentage of market sales accounted for by the N largest firms in the industry and is used as a simple measure of market structure and market power.

The Herfindahl-Hirschman Index measure of concentration is calculated as the sum of the squared market shares of the largest N firms in an industry and better reflects the effect of mergers on industry concentration.

Neither measure actually measures market power directly. Both can be misleading measures of market power when potential competition restricts pricing power.

LOS 9.h

To identify the market structure in which a firm is operating, we need to examine the number of firms in its industry, whether products are differentiated or other types of non-price competition exist, and barriers to entry, and compare these to the characteristics that define each market structure.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 9.1

1. **A** When a firm operates under conditions of pure competition, MR always equals price. This is because, in pure competition, demand is perfectly elastic (a horizontal line), so MR is constant and equal to price. (LOS 9.a)
2. **B** The supply function is not well-defined in markets other than those that can be characterized as perfect competition. (LOS 9.c)
3. **A** In a purely competitive market, economic losses indicate that firms are overproducing, causing prices to fall below average total costs. This can occur in the short run. In the long run, however, market supply will decrease as firms exit the industry, and prices will rise to the point where economic profits are zero. (LOS 9.b)
4. **C** A purely competitive firm will tend to expand its output so long as the market price is greater than MC. In the short run and long run, profit is maximized when $P = MC$. (LOS 9.d)
5. **C** If price is greater than average variable cost, a firm will continue to operate in the short run because it is covering at least some of its fixed costs. (LOS 9.d)

Module Quiz 9.2

1. **B** The demand for products from firms competing in monopolistic competition is relatively elastic due to the availability of many close substitutes. If a firm increases its product price, it will lose customers to firms selling substitute products at lower prices. (LOS 9.b)
2. **B** Monopolistic competition is likely to result in a higher price and lower quantity of output compared to perfect competition. (LOS 9.d)
3. **C** The profit-maximizing output is the quantity at which marginal revenue equals marginal cost. In a price-searcher industry structure (i.e., any structure that is not perfect competition), price is greater than marginal revenue. (LOS 9.d, 9.e, 9.f)

Module Quiz 9.3

1. **C** An oligopolistic industry has a great deal of interdependence among firms. One firm's pricing decisions or advertising activities will affect the other firms. (LOS 9.a)
2. **C** The kinked demand model assumes that each firm in a market believes that at some price, demand is more elastic for a price increase than for a price decrease. (LOS 9.b)
3. **A** The Nash equilibrium results when each nation pursues the strategy that is best, given the strategy that is pursued by the other nation.
 - Given that Germany complies with the agreement: France will get €8 billion if it complies, but €10 billion if it defaults. Therefore, France should default.
 - Given that Germany defaults: France will get €2 billion if it complies, but €4 billion if it defaults. Therefore, France should default.
 - Because France is better off in either case by defaulting, France will default.
 - Germany will follow the same logic and reach the same conclusion.(LOS 9.f)

Module Quiz 9.4

1. **C** Monopolists must search for the profit maximizing price (and output) because they do not have perfect information regarding demand. Firms under perfect competition take the market price as given and only determine the profit maximizing quantity. (LOS 9.b)
2. **A** A monopolist will expand production until $MR = MC$, and the price of the product will be determined by the demand curve. (LOS 9.d)
3. **C** When a regulatory agency requires a monopolist to use average cost pricing, the intent is to price the product where the ATC curve intersects the market demand curve. A problem in using this method is actually determining exactly what the ATC is. (LOS 9.f)
4. **C** Although the N -firm concentration ratio is simple to calculate, it can be relatively insensitive to mergers between companies with large market shares. Neither the HHI nor the N -firm concentration ratio consider barriers to entry. (LOS 9.g)
5. **C** These characteristics are associated with a market structure of monopolistic competition. Firms in perfect competition do not compete on product features. Oligopolistic markets have high barriers to entry. (LOS 9.h)

READING 10

AGGREGATE OUTPUT, PRICES, AND ECONOMIC GROWTH

EXAM FOCUS

This reading introduces macroeconomics and the measurement of aggregate economic output. The crucial concepts to grasp here are aggregate demand, short-run aggregate supply, and long-run aggregate supply. Know the factors that cause the aggregate demand and supply curves to shift and the sources of long-run economic growth. Understand the various measures of aggregate income (nominal and real GDP, national income, personal income, and personal disposable income). The interaction among saving, investment, the fiscal balance, and the trade balance will be built on in later readings on international trade and foreign exchange.

MODULE 10.1: GDP, INCOME, AND EXPENDITURES



Video covering this content is available online.

LOS 10.a: Calculate and explain gross domestic product (GDP) using expenditure and income approaches.

Gross domestic product (GDP) is the total market value of the goods and services produced in a country within a certain time period. GDP is the most widely used measure of the size of a nation's economy. GDP includes only purchases of newly produced goods and services. The sale or resale of goods produced in previous periods is excluded. Transfer payments made by the government (e.g., unemployment, retirement, and welfare benefits) are not economic output and are not included in the calculation of GDP.

The values used in calculating GDP are *market values of final goods* and services—that is, goods and services that will not be resold or used in the production of other goods and services. The value of the computer chips that Intel makes is not explicitly included in GDP; their value is included in the final prices of computers that use the chips. The value of a Rembrandt painting that sells for 10 million euros is not included in the calculation of GDP, as it was not produced during the period.

Goods and services provided by government are included in GDP even though they are not explicitly priced in markets. For example, the services provided by police and the judiciary, and goods such as roads and infrastructure improvements, are included. Because these goods and services are not sold at market prices, they are valued at their cost to the government.

GDP also includes the value of owner-occupied housing, just as it includes the value of rental housing services. Because the value of owner-occupied housing is not revealed in market transactions, the value is estimated for inclusion in GDP. The value of labor not sold, such as a homeowner's repairs to his own home, is not included in GDP. By-products of production, such as environmental damage, are not included in GDP.

GDP can be calculated as the sum of all the spending on newly produced goods and services, or as the sum of the income received as a result of producing these goods and services. Under the **expenditure approach**, GDP is calculated by summing the amounts spent on goods and services produced during the period. Under the **income approach**, GDP is calculated by summing the amounts earned by households and companies during the period, including wage income, interest income, and business profits.

For the whole economy, total expenditures and total income must be equal, so the two approaches should produce the same result. In practice, measurement issues result in different values under the two methods.

LOS 10.b: Compare the sum-of-value-added and value-of-final-output methods of calculating GDP.

So far, we have described the calculation of GDP under the expenditure approach as summing the values of all final goods and services produced. This expenditure method is termed the **value-of-final-output method**.

Under the **sum-of-value-added method**, GDP is calculated by summing the additions to value created at each stage of production and distribution. An example of the calculation for a specific product is presented in Figure 10.1.

Figure 10.1: Value Added at Stages of Production

Stage of Production	Sales Value (\$)	Value Added (\$)
Raw materials/components	\$100	\$100
Manufacturing	\$350	\$250
Retail	\$400	\$50
Sum of value added		\$400

The intuition is clear. The prices of final goods and services include, and are equal to, the additions to value at each stage of production (e.g., from mining iron ore and making steel to assembling an automobile that contains machined steel parts).

LOS 10.c: Compare nominal and real GDP and calculate and interpret the GDP deflator.

Nominal GDP is simply GDP as we have described it under the expenditures approach: the total value of all goods and services produced by an economy, valued at current market prices. For an economy with N different goods and services, we can express nominal GDP as:

$$\begin{aligned}
 \text{nominal GDP}_t \text{ for year } t &= \sum_{i=1}^N P_{i,t} Q_{i,t} \\
 &= \sum_{i=1}^N (\text{price of good } i \text{ in year } t) \\
 &\quad \times (\text{quantity of good } i \text{ produced in year } t)
 \end{aligned}$$

Because nominal GDP is based on current prices, inflation will increase nominal GDP even if the physical output of goods and services remains constant from one year to the next. **Real GDP** measures the output of the economy using prices from a base year, removing the effect of changes in prices so that inflation is not counted as economic growth.

Real GDP is calculated relative to a *base year*. By using base-year prices and current-year output quantities, real GDP growth reflects only increases in total output, not simply increases (or decreases) in the money value of total output.

Assuming the base year prices are those for five years ago, real GDP can be calculated as:

$$\begin{aligned}
 \text{real GDP for year } t &= \sum_{i=1}^N P_{i,t-5} Q_{i,t} \\
 &= \sum_{i=1}^N (\text{price of good } i \text{ in year } t - 5) \\
 &\quad \times (\text{quantity of good } i \text{ produced in year } t)
 \end{aligned}$$

The **GDP deflator** is a price index that can be used to convert nominal GDP into real GDP, taking out the effects of changes in the overall price level. The GDP deflator is based on the current mix of goods and services, using prices at the beginning and end of the period. The GDP deflator is calculated as:

$$\begin{aligned}
 \text{GDP deflator for year } t &= \frac{\sum_{i=1}^N P_{i,t} Q_{i,t}}{\sum_{i=1}^N P_{i, \text{base year}} Q_{i,t}} \times 100 \\
 &= \frac{\text{nominal GDP in year } t}{\text{value of year } t \text{ output at base year prices}} \times 100
 \end{aligned}$$

Per-capita real GDP is defined as real GDP divided by population and is often used as a measure of the economic well-being of a country's residents.

EXAMPLE: Calculating and using the GDP deflator

1. GDP in 20X2 is \$1.80 billion at 20X2 prices and \$1.65 billion when calculated using 20X1 prices. Calculate the GDP deflator using 20X1 as the base period.
2. Nominal GDP was \$213 billion in 20X6 and \$150 billion in 20X1. The 20X6 GDP deflator relative to the base year 20X1 is 122.3. Calculate real GDP for 20X6 and the compound annual real growth rate of economic output from 20X1 to 20X6.

Answer:

1. GDP deflator = $1.80 / 1.65 \times 100 = 109.1$, reflecting a 9.1% increase in the price level.
2. Real GDP 20X6 = $\$213 / 1.223 = \174.16 .

Noting that real and nominal GDP are the same for the base year, the compound real annual growth rate of economic output over the 5-year period is:

$$\left(\frac{174.16}{150}\right)^{\frac{1}{5}} - 1 = 3.03\%$$

LOS 10.d: Compare GDP, national income, personal income, and personal disposable income.

Using the expenditure approach, the major components of real GDP are consumption, investment, government spending, and **net exports** (exports minus imports). These components are summarized in the equation:

$$GDP = C + I + G + (X - M)$$

where:

C = consumption spending

I = business investment (capital equipment, inventories)

G = government purchases

X = exports

M = imports

We may also express this equation as:

$$GDP = (C + GC) + (I + GI) + (X - M)$$

where:

GC = government consumption

GI = government investment (capital goods, inventories)

Under the income approach, we have the following equation for GDP, or **gross domestic income (GDI)**:

$$GDP = \text{national income} + \text{capital consumption allowance} + \text{statistical discrepancy}$$

A **capital consumption allowance (CCA)** measures the depreciation (i.e., wear) of physical capital from the production of goods and services over a period. CCA can be thought of as the amount that would have to be reinvested to maintain the productivity of physical capital from one period to the next. The *statistical discrepancy* is an adjustment for the difference between GDP measured under the income approach and the expenditure approach because they use different data.

National income is the sum of the income received by all factors of production that go into the creation of final output:

$$\begin{aligned} \text{national income} &= \text{compensation of employees (wages and benefits)} \\ &+ \text{corporate and government enterprise profits before taxes} \\ &+ \text{interest income} \\ &+ \text{unincorporated business net income (business owners' incomes)} \\ &+ \text{rent} \\ &+ \text{indirect business taxes} - \text{subsidies (taxes and subsidies that are included in final prices)} \end{aligned}$$



PROFESSOR'S NOTE

Candidates should be aware that different countries' economic reporting bureaus may use their own terminology. For example, Statistics Canada defines *gross domestic income* as "net domestic income + consumption of fixed capital + statistical discrepancy," where net domestic income is "compensation of employees + gross operating surplus + gross mixed income + taxes less subsidies on production + taxes less subsidies on products and imports."

Personal income is a measure of the pretax income received by households and is one determinant of consumer purchasing power and consumption. Personal income differs from national income in that personal income includes all income that households receive, including government transfer payments such as unemployment or disability benefits.

Household disposable income or personal disposable income is personal income after taxes. Disposable income measures the amount that households have available to either save or spend on goods and services and is an important economic indicator of the ability of consumers to spend and save.

LOS 10.e: Explain the fundamental relationship among saving, investment, the fiscal balance, and the trade balance.

To show how private savings are related to investment, the government sector, and foreign trade, we will combine the income and expenditure approaches to measuring GDP.

As we have seen, total expenditures can be stated as $GDP = C + I + G + (X - M)$. Total income, which must equal total expenditures, can be stated as:

$$GDP = C + S + T$$

where:

C = consumption spending

S = household and business savings

T = net taxes (taxes paid minus transfer payments received)

Because total income equals total expenditures, we have the equality:

$$C + I + G + (X - M) = C + S + T$$

Rearranging this equation and solving for S (household and business savings), we get the following fundamental relationship:

$$S = I + (G - T) + (X - M)$$

Note that $(G - T)$ is the **fiscal balance**, or the difference between government spending and tax receipts. Recall that $(X - M)$ is net exports, or the **trade balance**. This equation shows that private savings must equal private investment, plus government borrowing or minus government savings, and minus the trade deficit or plus the trade surplus.



PROFESSOR'S NOTE

In this equation and the ones we will derive from it, a positive value for $(G - T)$ is a government budget deficit and a negative value for $(G - T)$ is a budget surplus. On the other hand, a positive value for $(X - M)$ is a trade surplus and a negative value for $(X - M)$ is a trade deficit.

If we solve this equation for the fiscal balance, we get:

$$(G - T) = (S - I) - (X - M)$$

From this equation, we can see that a government deficit ($G - T > 0$) must be financed by some combination of a trade deficit ($X - M < 0$) or an excess of private saving over private investment ($S - I > 0$).



PROFESSOR'S NOTE

In the reading on International Trade and Capital Flows, we will see that a trade deficit (current account deficit) must be associated with an inflow of foreign investment (capital account surplus). So we can interpret this equation as saying a fiscal deficit must be financed by a combination of domestic and foreign capital.



MODULE QUIZ 10.1

1. The *least appropriate* approach to calculating a country's gross domestic product (GDP) is summing for a given time period:
 - A. the value of all purchases and sales that took place within the country.
 - B. the amount spent on final goods and services produced within the country.
 - C. the income generated in producing all final goods and services produced within the country.
2. Gross domestic product does not include the value of:
 - A. transfer payments.
 - B. government services.
 - C. owner-occupied housing.
3. When GDP is calculated by the sum-of-value-added method, what is the value of a manufactured product in GDP?
 - A. The sum of the product's value at each stage of production and distribution.
 - B. The sum of the increases in the product's value at each stage of production and distribution.
 - C. The product's retail price less the value added at each stage of production and distribution.
4. Real GDP is *best* described as the value of:
 - A. current output measured at current prices.
 - B. current output measured at base-year prices.
 - C. base-year output measured at current prices.
5. The GDP deflator is calculated as 100 times the ratio of:
 - A. nominal GDP to real GDP.
 - B. base year prices to current year prices.
 - C. current year nominal GDP to base year nominal GDP.
6. Which of the following measures of income is the sum of wages and benefits, pretax profits, interest income, owners' income from unincorporated businesses, rent, and taxes net of subsidies?
 - A. Personal income.
 - B. National income.
 - C. Disposable income.
7. If a government budget deficit increases, net exports must:
 - A. increase, or the excess of private saving over private investment must decrease.
 - B. decrease, or the excess of private saving over private investment must increase.

C. decrease, or the excess of private saving over private investment must decrease.

MODULE 10.2: AGGREGATE DEMAND AND SUPPLY



Video covering
this content is
available online.

LOS 10.f: Explain how the aggregate demand curve is generated.

The **aggregate demand curve (AD curve)** illustrates the negative relationship between the price level and the level of real output demanded by consumers, businesses, and government. Points on the AD are combinations of the price level and real output for which the following two conditions hold:

1. *The goods market is in equilibrium:* Aggregate income equals aggregate expenditure, as we saw earlier in the fundamental relationship among saving, investment, and the fiscal and trade balances.
2. *The money market is in equilibrium:* Individuals and businesses are willing to hold the real money supply (nominal money supply adjusted for the price level).

Three effects explain why the AD curve slopes downward: a wealth effect, an interest rate effect, and a real exchange rate effect.

The **wealth effect** results from how changes in the price level affect consumers' purchasing power. For any given amount of nominal wealth (wealth stated in currency units), an increase in the price level reduces the amount of goods and services that amount of nominal wealth will purchase. A decrease in the price level increases the purchasing power of existing nominal wealth so that consumers will demand more goods and services. (People feel richer with lower prices and increase their consumption of goods and services.)

The **interest rate effect** is related to the demand for money. When the price level increases, to purchase the same amount of real goods and services, consumers need to hold greater nominal money balances. However, as we are holding the nominal money supply constant, only an increase in interest rates will restore the equilibrium between the money supply and the nominal balances people desire to hold (interest is the opportunity cost of holding money rather than investing it in interest-bearing securities). Higher interest rates will tend to decrease both demand for consumption goods (especially goods that are typically purchased on credit, such as cars and appliances) and business investment demand (a higher cost of capital reduces the number of profitable investment projects, as we explain in detail in the Corporate Issuers topic area).

The **real exchange rate effect** refers to the effect of an increase in the domestic price level on the net exports ($X - M$) component of GDP. When the domestic price level increases relative to the price level in a foreign country, the real price (the amount of foreign goods they must give up) of the domestic country's goods increases for foreigners, which reduces demand for exports. At the same time, the real price of imported goods to domestic consumers and businesses falls, increasing the quantity of imports demanded by domestic consumers. Both the decrease in

exports and the increase in imports reduce net exports ($X - M$), reducing aggregate demand for real goods and services.

To summarize, the aggregate demand curve slopes downward (a lower price level is associated with higher output) because lower price levels increase real wealth, decrease equilibrium real interest rates, make domestic goods less expensive to foreign consumers, and make foreign goods more expensive to domestic consumers and businesses, all of which increase the quantity of domestic output demanded.

LOS 10.g: Explain the aggregate supply curve in the short run and long run.

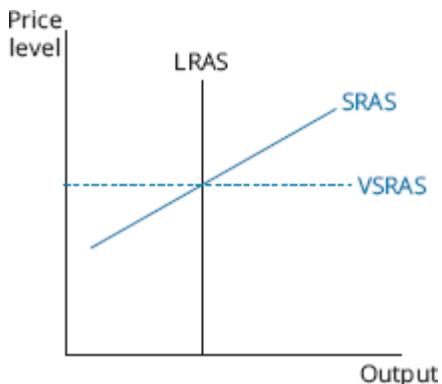
The Aggregate Supply Curve

The **aggregate supply (AS) curve** describes the relationship between the price level and the quantity of real GDP supplied, when all other factors are kept constant. That is, it represents the amount of output that firms will produce at different price levels.

We need to consider three aggregate supply curves with different time frames: the very short-run aggregate supply (VSRAS) curve, the short-run aggregate supply (SRAS) curve, and the long-run aggregate supply (LRAS) curve.

In the very short run, firms will adjust output without changing price by adjusting labor hours and intensity of use of plant and equipment in response to changes in demand. We represent this with the perfectly elastic very short run aggregate supply (VSRAS) curve in Figure 10.2.

Figure 10.2: Aggregate Supply Curves



In the short run, the SRAS curve slopes upward because some input prices will change as production is increased or decreased. We assume in the short run that *output prices* will change proportionally to the price level but that at least some *input prices* are sticky, meaning that they do not adjust to changes in the price level in the short run. When output prices increase, the price level increases, but firms see no change in input prices in the short run. Firms respond by increasing output in anticipation of greater profits from higher output prices. The result is an upward-sloping SRAS curve.

All input costs can vary in the long run, and the LRAS curve in Figure 10.2 is perfectly inelastic. In the long run, wages and other input prices change proportionally to the price level, so the

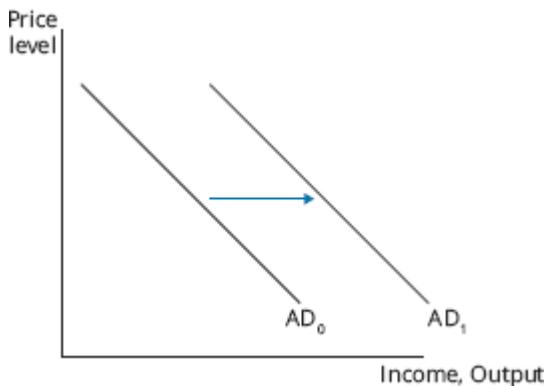
price level has no long-run effect on aggregate supply. We refer to this level of output as **potential GDP** or **full-employment GDP**.

LOS 10.h: Explain causes of movements along and shifts in aggregate demand and supply curves.

Shifts in the Aggregate Demand Curve

The aggregate demand (AD) curve reflects the total level of expenditures in an economy by consumers, businesses, governments, and foreigners. A number of factors can affect this level of expenditures and cause the AD curve to shift. Note that a *change in the price level* is represented as a *movement along the AD curve*, not a shift in the AD curve. In Figure 10.3, an increase in aggregate demand is shown by a shift to the right, indicating that the quantity of goods and services demanded is greater at any given price level.

Figure 10.3: Increase in Aggregate Demand



In trying to understand and remember the factors that affect aggregate demand, it may help to recall that, from the expenditure point of view, $GDP = C + I + G + \text{net } X$. For changes in each of the following factors that increase aggregate demand (shift AD to the right), we identify which component of expenditures is increased.

1. **Increase in consumers' wealth:** As the value of households' wealth increases (real estate, stocks, and other financial securities), the proportion of income saved decreases and spending increases, increasing aggregate demand (C increases).
2. **Business expectations:** When businesses are more optimistic about future sales, they tend to increase their investment in plant, equipment, and inventory, which increases aggregate demand (I increases).
3. **Consumer expectations of future income:** When consumers expect higher future incomes, due to a belief in greater job stability or expectations of rising wage income, they save less for the future and increase spending now, increasing aggregate demand (C increases).
4. **High capacity utilization:** When companies produce at a high percentage¹ of their capacity, they tend to invest in more plant and equipment, increasing aggregate demand (I increases).
5. **Expansionary monetary policy:** When the rate of growth of the money supply is increased, banks have more funds to lend, which puts downward pressure on interest rates. Lower

interest rates increase investment in plant and equipment because the cost of financing these investments declines. Lower interest rates and greater availability of credit will also increase consumers' spending on consumer durables (e.g., automobiles, large appliances) that are typically purchased on credit. Thus, the effect of expansionary monetary policy is to increase aggregate demand (C and I increase).

Note that if the economy is operating at potential GDP (LRAS) when the monetary expansion takes place, the increase in real output will be only for the short run. In the long run, subsequent increases in input prices decrease SRAS and return output to potential GDP.

6. **Expansionary fiscal policy:** Expansionary fiscal policy refers to a decreasing government budget surplus (or an increasing budget deficit) from decreasing taxes, increasing government expenditures, or both. A decrease in taxes increases disposable income and consumption, while an increase in government spending increases aggregate demand directly (C increases for tax cut, G increases for spending increase).



PROFESSOR'S NOTE

A complete analysis of monetary and fiscal policy as they relate to overall expenditures and GDP is presented in our reading on Monetary and Fiscal Policy.

7. **Exchange rates:** A decrease in the relative value of a country's currency will increase exports and decrease imports. Both of these effects tend to increase domestic aggregate demand (net X increases).



PROFESSOR'S NOTE

We will analyze the effect of exchange rates on exports and imports in our reading on Currency Exchange Rates.

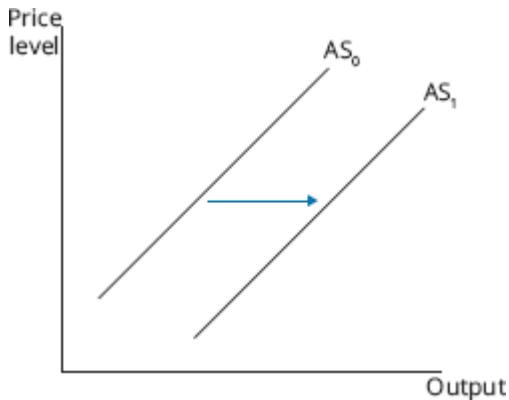
8. **Global economic growth:** GDP growth in foreign economies tends to increase the quantity of imports (domestic exports) foreigners demand. By increasing domestic export demand, this will increase aggregate demand (net X increases).

Note that for each factor, a change in the opposite direction will tend to decrease aggregate demand.

Shifts in the Short-Run Aggregate Supply Curve

The **short-run aggregate supply (SRAS) curve** reflects the relationship between output and the price level when wages and other input prices are held constant (or are slow to adjust to higher output prices). The curve shows the total level of output that businesses are willing to supply at different price levels. A number of factors can affect this level of output and cause the SRAS curve to shift. In Figure 10.4, an increase in aggregate supply is shown by a shift to the right, as the quantity supplied at each price level increases.

Figure 10.4: Increase in Aggregate Supply



In addition to changes in potential GDP (shifts in long-run aggregate supply), a number of factors can cause the SRAS curve to shift to the right:

1. **Labor productivity:** Holding the wage rate constant, an increase in labor productivity (output per hour worked) will decrease unit costs to producers. Producers will increase output as a result, increasing SRAS (shifting it to the right).
2. **Input prices:** A decrease in nominal wages or the prices of other important productive inputs will decrease production costs and cause firms to increase production, increasing SRAS. Wages are often the largest contributor to a producer's costs and have the greatest impact on SRAS.
3. **Expectations of future output prices:** When businesses expect the price of their output to increase in the future, they will expand production, increasing SRAS.
4. **Taxes and government subsidies:** Either a decrease in business taxes or an increase in government subsidies for a product will decrease the costs of production. Firms will increase output as a result, increasing SRAS.
5. **Exchange rates:** Appreciation of a country's currency in the foreign exchange market will decrease the cost of imports. To the extent that productive inputs are purchased from foreign countries, the resulting decrease in production costs will cause firms to increase output, increasing SRAS.

Again, an opposite change in any of these factors will tend to decrease SRAS.

Shifts in the Long-Run Aggregate Supply Curve

The **long-run aggregate supply (LRAS) curve** is vertical (perfectly inelastic) at the potential (full-employment) level of real GDP. Changes in factors that affect the real output that an economy can produce at full employment will shift the LRAS curve.

Factors that will shift the LRAS curve are:

1. **Increase in the supply and quality of labor:** Because LRAS reflects output at full employment, an increase in the labor force will increase full-employment output and the LRAS. An increase in the skills of the workforce, through training and education, will increase the productivity of a labor force of a given size, increasing potential real output and increasing LRAS.

- Increase in the supply of natural resources:** Just as with an increase in the labor force, increases in the available amounts of other important productive inputs will increase potential real GDP and LRAS.
- Increase in the stock of physical capital:** For a labor force of a given size, an increase in an economy's accumulated stock of capital equipment will increase potential output and LRAS.
- Technology:** In general, improvements in technology increase labor productivity (output per unit of labor) and thereby increase the real output that can be produced from a given amount of productive inputs, increasing LRAS.

Decreases in labor quality, labor supply, the supply of natural resources, or the stock of physical capital will all decrease LRAS (move the curve to the left). Technology does not really retreat, but a law prohibiting the use of an improved technology could decrease LRAS.

Movement Along Aggregate Demand and Supply Curves

In contrast with *shifts* in the aggregate demand and aggregate supply curves, *movements along* these curves reflect the impact of a change in the price level on the quantity demanded and the quantity supplied. Changes in the price level alone do not cause shifts in the AD and AS curves, although we have allowed that changes in expected future prices can.



MODULE QUIZ 10.2

- The aggregate demand curve illustrates which of the following relationships?
 - Direct relationship between aggregate income and the price level.
 - Inverse relationship between aggregate income and the price level.
 - Direct relationship between aggregate income and the real interest rate.
- An economy's potential output is *best* represented by:
 - long-run aggregate supply.
 - short-run aggregate supply.
 - long-run aggregate demand.
- A stronger domestic currency relative to foreign currencies is *most likely* to result in:
 - a shift in the aggregate supply curve toward lower supply.
 - a shift in the aggregate demand curve toward lower demand.
 - a movement along the aggregate demand curve towards higher prices.
- Which of the following factors would be *least likely* to shift the aggregate demand curve?
 - The price level increases.
 - The federal deficit expands.
 - Expected inflation decreases.

MODULE 10.3: MACROECONOMIC EQUILIBRIUM AND GROWTH



Video covering this content is available online.

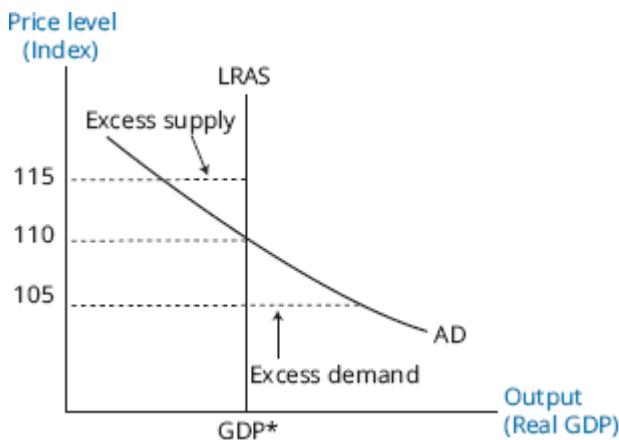
LOS 10.i: Describe how fluctuations in aggregate demand and aggregate supply cause short-run changes in the economy and the business cycle.

LOS 10.j: Distinguish among the following types of macroeconomic equilibria: long-run full employment, short-run recessionary gap, short-run inflationary gap, and short-run stagflation.

LOS 10.k: Explain how a short-run macroeconomic equilibrium may occur at a level above or below full employment.

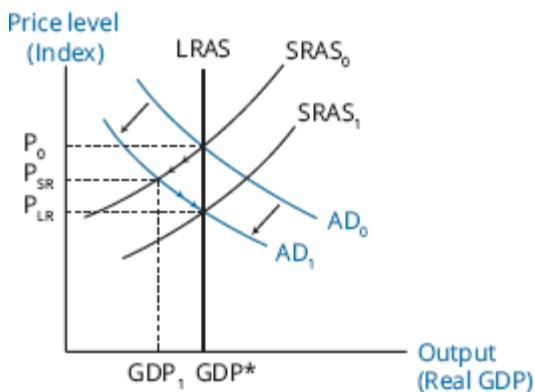
Having explained the factors that cause shifts in the aggregate demand and aggregate supply curves, we now turn our attention to the effects of fluctuations in aggregate supply and demand on real GDP and the business cycle. Our starting point is an economy that is in *long-run full-employment equilibrium*, as illustrated in Figure 10.5.

Figure 10.5: Long-Run Equilibrium Real Output



First consider a decrease in aggregate demand, which can result from a decrease in the growth rate of the money supply, an increase in taxes, a decrease in government spending, lower equity and house prices, or a decrease in the expectations of consumers and businesses for future economic growth. As illustrated in Figure 10.6, a decrease in aggregate demand will reduce both real output and the price level in the short run. The new short-run equilibrium output, GDP_1 , is less than full employment (potential) GDP. The decrease in aggregate demand has resulted in both lower real output and a lower price level.

Figure 10.6: Adjustment to a Decrease in Aggregate Demand



Because real GDP is less than full employment GDP, we say there is a **recessionary gap**. A recession is a period of declining GDP and rising unemployment. Classical economists believed that unemployment would drive down wages, as workers compete for available jobs, which in turn would increase SRAS and return the economy to its full employment level of real GDP. Keynesian economists, on the other hand, believe that this might be a slow and economically

painful process and that increasing aggregate demand through government action is the preferred alternative. Both expansionary fiscal policy (increasing government spending or decreasing taxes) and expansionary monetary policy (increasing the growth rate of the money supply to reduce interest rates) are methods to increase aggregate demand and return real GDP to its full employment (potential) level.

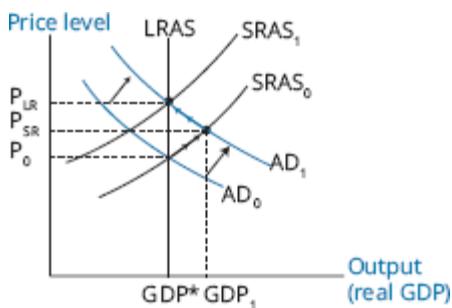


PROFESSOR'S NOTE

We will describe Classical, Keynesian, and other business cycle theories in the reading on Understanding Business Cycles.

A second case to consider is an increase in aggregate demand that results in an equilibrium at a level of GDP greater than full-employment GDP in the short run, as illustrated in Figure 10.7. Note that both GDP and the price level are increased. The economy can operate at a level of GDP greater than full-employment GDP in the short run, as workers work overtime and maintenance of productive equipment is delayed, but output greater than full-employment GDP cannot be maintained in the long run. In the long run, the economy always returns to full-employment GDP along the LRAS curve.

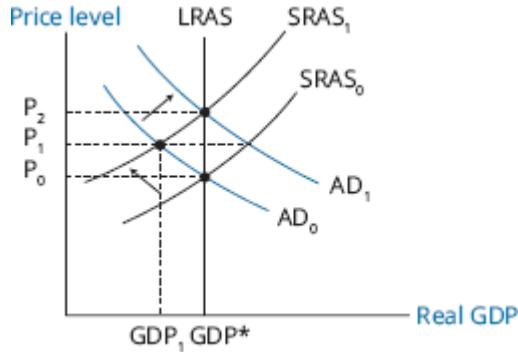
Figure 10.7: Adjustment to an Increase in Aggregate Demand



We term the difference between GDP_1 and full-employment GDP in Figure 10.7 an **inflationary gap** because the increase in aggregate demand from its previous level causes upward pressure on the price level. Competition among producers for workers, raw materials, and energy may shift the SRAS curve to the left, returning the economy to full-employment GDP but at a price level that is higher still. Alternatively, government policy makers can reduce aggregate demand by decreasing government spending, increasing taxes, or slowing the growth rate of the money supply, in order to move the economy back to the initial long run equilibrium at full-employment GDP.

Changes in wages or the prices of other important productive inputs can shift the SRAS curve, affecting real GDP and the price level in the short run. An important case to consider is a decrease in SRAS caused by an increase in the prices of raw materials or energy. As illustrated in Figure 10.8, the new short-run equilibrium is at lower GDP and a higher overall price level for goods and services compared to the initial long-run equilibrium. This combination of declining economic output and higher prices is termed **stagflation** (stagnant economy with inflation).

Figure 10.8: Stagflation

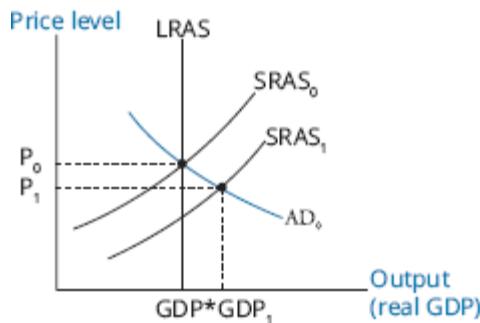


A subsequent decrease in input prices can return the economy to its long-run equilibrium output. An increase in aggregate demand from either expansionary fiscal or monetary policy can also return the economy to its full employment level, but at a price level that is higher still compared to the initial equilibrium.

Stagflation is an especially difficult situation for policy makers because actions to increase aggregate demand to restore full employment will also increase the price level even more. Conversely, a decision by policy makers to fight inflation by decreasing aggregate demand will decrease GDP even further. A decrease in wages and the prices of other productive inputs may be expected to increase SRAS and restore full-employment equilibrium. However, this process may be quite slow and doing nothing may be a very risky strategy for a government when voters expect action to restore economic growth or stem inflationary pressures.

The fourth case to consider is an increase in SRAS due to a decrease in the price of important productive inputs. As illustrated in Figure 10.9, the resulting new short-run equilibrium is at a level of GDP greater than full-employment GDP and a lower overall price level.

Figure 10.9: Decrease in Input Prices



In Figure 10.10, we present a summary of the short-run effects of shifts in aggregate demand and in aggregate supply on real GDP, unemployment, and the price level.

Figure 10.10: Short-Run Macroeconomic Effects

Type of Change	Real GDP	Unemployment	Price Level
Increase in AD	Increase	Decrease	Increase
Decrease in AD	Decrease	Increase	Decrease
Increase in AS	Increase	Decrease	Decrease
Decrease in AS	Decrease	Increase	Increase

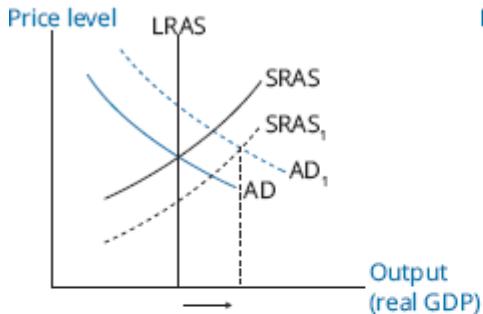
LOS 10.l: Analyze the effect of combined changes in aggregate supply and demand on the economy.

When both aggregate supply and aggregate demand change, the effects on equilibrium output and the price level may be clear when the effects on the variable are in the same direction (or ambiguous when the effects on the variable are in opposite directions). We summarize the effects of combined changes in demand and supply in Figure 10.11.

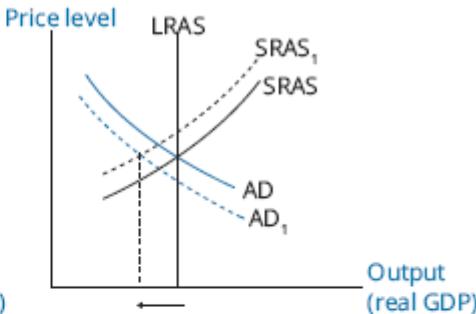
- When *aggregate demand and aggregate supply both increase*, real GDP increases but the effect on the price level depends on the relative magnitudes of the changes because their price effects are in opposite directions [Panel (a) of Figure 10.11].
- When *aggregate demand and aggregate supply both decrease*, real GDP decreases but the effect on the price level depends on the relative magnitudes of the changes because their price effects are in opposite directions [Panel (b) of Figure 10.11].
- When *aggregate demand increases and aggregate supply decreases*, the price level will increase but the effect on real GDP depends on the relative magnitudes of the changes because their effects on economic output are in opposite directions [Panel (c) of Figure 10.11].
- When *aggregate demand decreases and aggregate supply increases*, the price level will decrease but the effect on real GDP depends on the relative magnitudes of the changes because their effects on economic output are in opposite directions [Panel (d) of Figure 10.11].

Figure 10.11: Changes in Aggregate Supply and Aggregate Demand

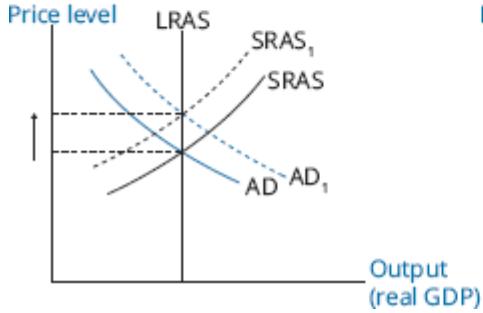
(a) AD and SRAS increase



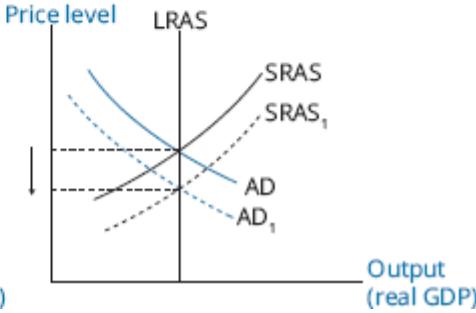
(b) AD and SRAS decrease



(c) AD increases, SRAS decreases



(d) AD decreases, SRAS increases



LOS 10.m: Describe sources, measurement, and sustainability of economic growth.

Economic growth can best be explained by examining five important **sources of economic growth**:

1. **Labor supply.** The **labor force** is the number of people over the age of 16 who are either working or available for work but currently unemployed. It is affected by population growth, net immigration, and the labor force participation rate (described in our reading on Understanding Business Cycles). Growth of the labor force is an important source of economic growth.
2. **Human capital.** The education and skill level of a country's labor force can be just as important a determinant of economic output as the size of the labor force. Because workers who are skilled and well-educated (possess more human capital) are more productive and better able to take advantage of advances in technology, investment in human capital leads to greater economic growth.
3. **Physical capital stock.** A high rate of investment increases a country's stock of physical capital. As noted earlier, a larger capital stock increases labor productivity and potential GDP. An increased rate of investment in physical capital can increase economic growth.
4. **Technology.** As noted previously, improvements in technology increase productivity and potential GDP. More rapid improvements in technology lead to greater rates of economic growth.
5. **Natural resources.** Raw material inputs, such as oil and land, are necessary to produce economic output. These resources may be *renewable* (e.g., forests) or *non-renewable* (e.g., coal). Countries with large amounts of productive natural resources can achieve greater rates of economic growth.

Sustainability of Economic Growth

One way to view potential GDP is with the following equation:

$$\text{potential GDP} = \text{aggregate hours worked} \times \text{labor productivity}$$

Or, stated in terms of economic growth:

$$\text{growth in potential GDP} = \text{growth in labor force} + \text{growth in labor productivity}$$

An economy's sustainable growth rate can be estimated by estimating the growth rate of labor productivity and the growth rate of the labor force. For example, if Japan's labor force is projected to shrink by 1%, while its labor productivity is expected to grow by 2%, then we would estimate the growth in potential GDP as: $-1\% + 2\% = 1\%$.

The **sustainable rate of economic growth** is important because long-term equity returns are highly dependent on economic growth over time. A country's sustainable rate of economic growth is the rate of increase in the economy's productive capacity (potential GDP).

LOS 10.n: Describe the production function approach to analyzing the sources of economic growth.

A **production function** describes the relationship of output to the size of the labor force, the capital stock, and productivity.

Economic output can be thought of as a function of the amounts of labor and capital that are available and their productivity, which depends on the level of technology available. That is:

$$Y = A \times f(L, K)$$

where:

Y = aggregate economic output

L = size of labor force

K = amount of capital available

A = total factor productivity

The multiplier, A, is referred to as **total factor productivity** and quantifies the amount of output growth that is not explained by increases in the size of the labor force and capital. Total factor productivity is closely related to technological advances. Generally, total factor productivity cannot be observed directly and must be inferred based on the other factors.

The production function can be stated on a per-worker basis by dividing by L:

$$\frac{Y}{L} = A \times f\left(\frac{K}{L}\right)$$

where:

$\frac{Y}{L}$ = output per worker (labor productivity)

$\frac{K}{L}$ = physical capital per worker

This relationship suggests that labor productivity can be increased by either improving technology or increasing physical capital per worker.

We assume that the production function exhibits **diminishing marginal productivity** for each individual input, meaning the amount of additional output produced by each additional unit of input declines (holding the quantities of other inputs constant). For this reason, sustainable long-term growth cannot necessarily be achieved simply by **capital deepening investment**—that is to say, increasing physical capital per worker over time. Productivity gains and growth of the labor force are also necessary for long-term sustainable growth.

LOS 10.0: Define and contrast input growth with growth of total factor productivity as components of economic growth.

A well-known model (the *Solow model* or *neoclassical model*) of the contributions of technology, labor, and capital to economic growth is:

$$\text{growth in potential GDP} = \text{growth in technology} + W_L(\text{growth in labor}) + W_C(\text{growth in capital})$$

where W_L and W_C are labor's percentage share of national income and capital's percentage share of national income. Like the multiplier, A, in a production function, the additional growth in potential GDP from "growth in technology" represents the change in total factor productivity, the growth of output that is not explained by the growth of labor and capital. Growth in technology is the primary driver of the growth in total factor productivity.

Consider a developed country where $W_L = 0.7$ and $W_C = 0.3$. For that country, a 1% increase in the labor force will lead to a much greater increase in economic output than a 1% increase in the capital stock. Similarly, sustained growth of the labor force will result in greater economic growth over time than sustained growth of the capital stock of an equal magnitude.

Sometimes the relationship between potential GDP, improvements in technology, and capital growth is written on a per-capita basis² as:

$$\text{growth in per-capita potential GDP} = \text{growth in technology} + W_C (\text{growth in the capital-to-labor ratio})$$

With $W_C = 0.25$, for example, each 1% increase in capital per worker will increase GDP per worker by 0.25%. In developed economies, where capital per worker is already relatively high, growth of technology will be the primary source of growth in GDP per worker. At higher levels of capital per worker, an economy will experience diminishing marginal productivity of capital and must look to advances in technology for strong economic growth.



MODULE QUIZ 10.3

1. Starting from short-run equilibrium, if aggregate demand is increasing faster than long-run aggregate supply:
 - A. the price level is likely to increase.
 - B. downward pressure on wages should ensue.
 - C. supply will increase to meet the additional demand.
2. A short-run macroeconomic equilibrium in which output must decrease to restore long-run equilibrium is most accurately characterized as:
 - A. stagflation.
 - B. a recessionary gap.
 - C. an inflationary gap.
3. Which of the following combinations of changes in aggregate demand and aggregate supply is *most likely* to result in decreasing prices? Aggregate demand:
 - A. decreases while aggregate supply increases.
 - B. decreases while aggregate supply decreases.
 - C. increases while aggregate supply decreases.
4. Labor productivity is *most likely* to increase as a result of:
 - A. an increase in physical capital.
 - B. a decrease in net immigration.
 - C. an increase in the labor force participation rate.
5. Long-term sustainable growth of an economy is *least likely* to result from growth in:
 - A. the supply of labor.
 - B. capital per unit of labor.
 - C. output per unit of labor.
6. In a production function model of economic output, total factor productivity represents the output growth that can be accounted for by:
 - A. capital growth but not labor growth.
 - B. neither labor growth nor capital growth.
 - C. the combined effects of labor growth and capital growth.
7. In a developed economy, the primary source of growth in potential GDP is:
 - A. capital investment.
 - B. labor supply growth.
 - C. technology advances.

KEY CONCEPTS

LOS 10.a

Gross domestic product (GDP) is the market value of all final goods and services produced within a country during a certain time period.

Using the expenditure approach, GDP is calculated as the total amount spent on goods and services produced in the country during a time period.

Using the income approach, GDP is calculated as the total income earned by households and businesses in the country during a time period.

LOS 10.b

The expenditure approach to measuring GDP can use the sum-of-value-added method or the value-of-final-output method.

- Sum-of-value-added: GDP is calculated by summing the additions to value created at each stage of production and distribution.
- Value-of-final-output: GDP is calculated by summing the values of all final goods and services produced during the period.

LOS 10.c

Nominal GDP values goods and services at their current prices. Real GDP measures current year output using prices from a base year.

The GDP deflator is a price index that can be used to convert nominal GDP into real GDP by removing the effects of changes in prices.

LOS 10.d

The four components of gross domestic product are consumption spending, business investment, government spending, and net exports.

$$GDP = C + I + G + (X - M).$$

National income is the income received by all factors of production used in the creation of final output.

Personal income is the pretax income received by households.

Household disposable income is personal income after taxes.

LOS 10.e

Private saving and investment are related to the fiscal balance and the trade balance. A fiscal deficit must be financed by some combination of a trade deficit or an excess of private saving over private investment.

$$(G - T) = (S - I) - (X - M).$$

LOS 10.f

The aggregate demand curve slopes downward because higher price levels reduce real wealth, increase real interest rates, and make domestically produced goods more expensive compared to goods produced abroad, all of which reduce the quantity of domestic output demanded.

LOS 10.g

The short-run aggregate supply curve shows the positive relationship between real GDP supplied and the price level, when other factors are held constant. Holding some input costs such as wages fixed in the short run, the curve slopes upward because higher output prices result in greater output (real wages fall).

Because all input prices are assumed to be flexible in the long run, the long-run aggregate supply curve is perfectly inelastic (vertical). Long-run aggregate supply represents potential GDP, the full employment level of economic output.

LOS 10.h

Changes in the price level cause movement along the aggregate demand or aggregate supply curves.

Shifts in the aggregate demand curve are caused by changes in household wealth, business and consumer expectations, capacity utilization, fiscal policy, monetary policy, currency exchange rates, and global economic growth rates.

Shifts in the short-run aggregate supply curve are caused by changes in nominal wages or other input prices, expectations of future prices, business taxes, business subsidies, and currency exchange rates, as well as by the factors that affect long-run aggregate supply.

Shifts in the long-run aggregate supply curve are caused by changes in labor supply and quality, the supply of physical capital, the availability of natural resources, and the level of technology.

LOS 10.i

The short-run effects of changes in aggregate demand and in aggregate supply are summarized in the following table:

Type of Change	Real GDP	Unemployment	Price Level
Increase in AD	Increase	Decrease	Increase
Decrease in AD	Decrease	Increase	Decrease
Increase in AS	Increase	Decrease	Decrease
Decrease in AS	Decrease	Increase	Increase

LOS 10.j

In long-run equilibrium, real GDP is equal to full-employment (potential) GDP. An increase in aggregate demand can result in a short-run equilibrium with GDP greater than full-employment GDP, termed an inflationary gap. A decrease in aggregate demand can result in a short-run equilibrium with GDP less than full-employment, termed a recessionary gap. When short-run aggregate supply decreases, the resulting short-run equilibrium is with GDP reduced to less than full-employment GDP but with an increase in the price level, termed stagflation.

LOS 10.k

From a situation of long-run equilibrium: an increase in either aggregate demand or aggregate supply can result in a short-run equilibrium with real GDP greater than full employment GDP; a decrease in either aggregate demand or aggregate supply can result in a short-run equilibrium with real GDP less than full-employment GDP.

LOS 10.l

Short-run effects of shifts in both aggregate demand and aggregate supply on the price level and real GDP:

Aggregate Demand	Aggregate Supply	Change in Real GDP	Change in Price Level
Increase	Increase	Increase	May increase or decrease
Decrease	Decrease	Decrease	May increase or decrease
Increase	Decrease	May increase or decrease	Increase
Decrease	Increase	May increase or decrease	Decrease

LOS 10.m

Sources of economic growth include increases in the supply of labor, increases in human capital, increases in the supply of physical capital, increasing availability of natural resources, and advances in technology.

The sustainable rate of economic growth is determined by the rate of increase in the labor force and the rate of increase in labor productivity.

LOS 10.n

A production function relates economic output to the supply of labor, the supply of capital, and total factor productivity. Total factor productivity is a residual factor, which represents that part of economic growth not accounted for by increases in the supply of labor and capital. Increases in total factor productivity can be attributed to advances in technology.

LOS 10.o

In developed countries, where a high level of capital per worker is available and capital inputs experience diminishing marginal productivity, technological advances that increase total factor productivity are the main source of sustainable economic growth.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 10.1

- A** Adding all purchases and sales is not appropriate because these would include goods that were produced before the time period in question. All purchases and sales could also result in double-counting intermediate goods. GDP is the market value of all final goods and services produced in a country in a certain period of time. GDP can be calculated either by totaling the amount spent on goods and services produced in the economy (the expenditure approach), or the income generated in producing these goods and services (the income approach). (LOS 10.b)
- A** Owner-occupied housing and government services are included in GDP at imputed (estimated) values. Transfer payments are excluded from the calculation of GDP. (LOS 10.a)
- B** Using the sum-of-value-added method, GDP can be calculated by summing the value added at each stage in the production and distribution process. Summing the value of the product at each stage of production would count the value added at earlier stages multiple times. The value added at earlier stages would not be included in GDP if it was deducted from the retail price. (LOS 10.b)
- B** Real GDP is the value of current period output calculated using prices from a base year. (LOS 10.c)

5. **A** The GDP deflator is the ratio of nominal GDP to real GDP, or equivalently the ratio of current year prices to base year prices. (LOS 10.c)
6. **B** National income is the income received by all factors of production used in the generation of final output. Personal income measures the pretax income that households receive. Disposable income is personal income after taxes. (LOS 10.d)
7. **B** The fundamental relationship among saving, investment, the fiscal balance, and the trade balance is described by the following equation: $(G - T) = (S - I) - (X - M)$. If the government budget deficit ($G - T$) increases, the larger budget deficit must be financed by some combination of an increase in the excess of private saving over private investment ($S - I$) or a decrease in net exports ($X - M$). (LOS 10.e)

Module Quiz 10.2

1. **B** The inverse relationship between aggregate income (or output) and the price level is the aggregate demand curve. (LOS 10.f)
2. **A** The LRAS curve is vertical at the level of potential GDP. (LOS 10.g)
3. **B** Strengthening of the domestic currency should cause exports to decrease and imports to increase, causing the AD curve to shift to the left (lower demand). At the same time, the cost of raw material inputs should decrease in domestic currency terms, causing the SRAS curve to shift to the right (greater supply). Changes in the price level cause movement along the AD and AS curves; in this case, any shifts along these curves will be towards lower prices. (LOS 10.h)
4. **A** Since the y-axis of the aggregate supply/demand model is the price level, a change in the price level is a movement along the AD curve. As long as inflation expectations are unchanged, an increase in the price level will not shift the aggregate demand curve. (LOS 10.h)

Module Quiz 10.3

1. **A** If AD is increasing faster than LRAS, the economy is expanding faster than its full-employment rate of output. This will cause pressure on wages and resource prices and lead to an increase in the price level. The SRAS curve will shift to the left—a decrease in supply for any given price level—until the rate of output growth slows to its full-employment potential. (LOS 10.i)
2. **C** If output must decrease to restore long-run equilibrium, the short-run equilibrium must be at an output level greater than long-run aggregate supply. This describes an inflationary gap. (LOS 10.j, 10.k)
3. **A** Decreasing aggregate demand combined with increasing aggregate supply will result in decreasing prices. Increasing aggregate demand combined with decreasing aggregate supply will result in increasing prices. A decrease or an increase in both aggregate demand and aggregate supply may either increase or decrease prices. (LOS 10.l)
4. **A** Increased investment in physical capital can increase labor productivity. Labor force participation rates and net immigration affect the size of the labor force and the aggregate number of hours worked, but do not necessarily affect labor productivity. (LOS 10.m)
5. **B** The sustainable rate of economic growth is a measurement of the rate of increase in the economy's productive capacity. An economy's sustainable rate of growth depends on the growth rate of the labor supply and the growth rate of labor productivity. Due to diminishing marginal productivity, an economy generally cannot achieve long-term sustainable growth through continually increasing the stock of capital relative to labor (i.e., capital deepening). (LOS 10.m)
6. **B** Total factor productivity represents output growth in excess of that resulting from the growth in labor and capital. (LOS 10.n)

7. C For developed economies, advances in technology are likely to be the primary source of growth in potential GDP because capital per worker is already high enough to experience diminishing marginal productivity of capital. (LOS 10.o)
-

¹ According to the Federal Reserve, "Industrial plants usually operate at capacity utilization rates that are well below 100 percent... For total industry and total manufacturing, utilization rates have exceeded 90 percent only in wartime." (Federal Reserve Statistical Release G.17, "Industrial Production and Capacity Utilization," www.federalreserve.gov/releases/g17/current/g17.pdf)

² Paul R. Kutasovic, CFA, and Richard G. Fritz, *Aggregate Output, Prices, and Economic Growth*, CFA[®] Program Level I 2022 Curriculum, Volume 2.

READING 11

UNDERSTANDING BUSINESS CYCLES

EXAM FOCUS

The phase of the business cycle is the starting point for top-down financial analysis. Candidates need to know how to interpret the many economic indicators that are available and why various indicators tend to lead, coincide with, or lag behind changes in economic activity. Indicators of unemployment and inflation are crucial for understanding fiscal and monetary policy actions.

MODULE 11.1: BUSINESS CYCLE PHASES



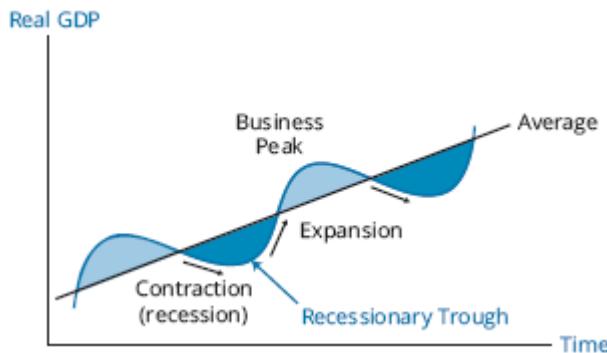
LOS 11.a: Describe the business cycle and its phases.

Video covering this content is available online.

The **business cycle** is characterized by fluctuations in economic activity. Real gross domestic product (GDP) and the rate of unemployment are the key variables used to determine the current phase of the cycle.

The business cycle has four phases: **expansion** (real GDP is increasing), **peak** (real GDP stops increasing and begins decreasing), **contraction or recession** (real GDP is decreasing), and **trough** (real GDP stops decreasing and begins increasing). The phases are illustrated in Figure 11.1.

Figure 11.1: Business Cycle



An expansion features growth in most sectors of the economy, with increasing employment, consumer spending, and business investment. As an expansion approaches its peak, the rates of increase in spending, investment, and employment slow but remains positive, while inflation accelerates.

A contraction or recession is associated with declines in most sectors, with inflation typically decreasing. When the contraction reaches a trough and the economy begins a new expansion or **recovery**, economic growth becomes positive again and inflation is typically moderate, but employment growth may not start to increase until the expansion has taken hold convincingly.

A common rule of thumb is to consider two consecutive quarters of growth in real GDP as the beginning of an expansion and two consecutive quarters of declining real GDP as indicating the beginning of a contraction. Statistical agencies that date expansions and recessions, such as the National Bureau of Economic Research in the United States, look at a wider variety of economic data such as employment, industrial production, and real personal income to identify turning points in the business cycle.

A key aspect of business cycles is that they recur, but not at regular intervals. Past business cycles have been as short as a year or longer than a decade.

The idea of a business cycle applies to economies that consist mainly of businesses. For economies that are mostly subsistence agriculture or dominated by state planning, fluctuations in activity are not really “business cycles” in the sense we are discussing here.

LOS 11.b: Describe credit cycles.

Credit cycles refer to cyclical fluctuations in interest rates and the availability of loans (credit). Typically, lenders are more willing to lend and tend to offer lower interest rates during economic expansions and are less willing to lend and require higher interest rates when the economy is slowing (contracting).

Credit cycles may amplify business cycles. Widely available or “loose” credit conditions during expansions can lead to “bubbles” (prices based on implausible expectations) in the markets for some assets, such as subprime mortgages in the period leading up to the financial crisis of 2007–2009. Some research suggests that expansions tend to be stronger, and contractions deeper and longer lasting, when they coincide with credit cycles. They do not always coincide, however, as historical data suggests credit cycles have been longer in duration than business cycles on average.

LOS 11.c: Describe how resource use, consumer and business activity, housing sector activity, and external trade sector activity vary as an economy moves through the business cycle.

Business Activity and Resource Use Fluctuation

Inventories are an important business cycle indicator. Firms try to keep enough inventory on hand to meet sales demand but do not want to keep too much of their capital tied up in inventory. As a result, the ratio of inventory to sales in many industries trends toward a normal level in times of steady economic growth.

When an expansion is approaching its peak, sales growth begins to slow, and unsold inventories accumulate. This can be seen in an increase in the **inventory-sales ratio** above its normal level.

Firms respond to an unplanned increase in inventory by reducing production, which is one of the causes of the subsequent contraction in the economy. An increase in inventories is counted in the GDP statistics as economic output, whether the increase is planned or unplanned. An analyst who looks only at GDP growth, rather than the inventory-sales ratio, might see economic strength rather than the beginning of weakness.

The opposite occurs when a contraction reaches its trough. Having reduced their production levels to adjust for lower sales demand, firms find their inventories becoming depleted more quickly once sales growth begins to accelerate. This causes the inventory-sales ratio to decrease below its normal level. To meet the increase in demand, firms will increase output, and the inventory-sales ratio will increase toward normal levels.

One of the ways firms react to fluctuations in business activity is by adjusting their utilization of labor and physical capital. Adding and subtracting workers in lockstep with changes in economic growth would be costly for firms, in terms of both direct expenses and the damage it would do to employee morale and loyalty. Instead, firms typically begin by changing how they utilize their current workers, producing less or more output per hour or adjusting the hours they work by adding or removing overtime. Only when an expansion or contraction appears likely to persist will they hire or lay off workers.

Similarly, because it is costly to adjust production levels by frequently buying and selling plant and equipment, firms first adjust their production levels by using their existing physical capital more or less intensively. As an expansion persists, firms will increase their production capacity by investing more in plant and equipment. During contractions, however, firms will not necessarily sell plant and equipment outright. They can reduce their physical capacity by spending less on maintenance or by delaying the replacement of equipment that is near the end of its useful life.

Consumer Sector Activity

Consumer spending, the largest component of gross domestic product, depends on the level of consumers' current incomes and their expectations about their future incomes. As a result, consumer spending increases during expansions and decreases during contractions.

Consumer spending in some sectors is more sensitive to business cycle phases than spending in other sectors. Spending on **durable goods** is highly cyclical because they are often higher-value purchases. Consumers are more willing to purchase high-value durable goods (e.g., appliances, furniture, automobiles) during expansions, when incomes are increasing and economic confidence is high. During contractions (and sometimes extending into the early stages of expansions), consumers often postpone durable goods purchases until they are more confident about their employment status and prospects for income growth.

Consumer spending on **services** is also positively correlated with business cycle phases, but not to the same extent as durable goods spending. Services include spending that is more discretionary, such as for lodging or food away from home, but also includes spending that is less discretionary, such as spending on telecommunications, health care, and insurance. The more discretionary a service is, the more cyclical consumer spending on it tends to be. Spending on **nondurable goods**, such as food at home or household products for everyday use, remains relatively stable over the business cycle.

Housing Sector Activity

Although the housing sector is a small part of the economy relative to overall consumer spending, cyclical swings in activity in the housing market can be large so that the effect on overall economic activity is greater than it otherwise would be. Important determinants of the level of economic activity in the housing sector are:

1. **Mortgage rates:** Low interest rates tend to increase home buying and construction while high interest rates tend to reduce home buying and construction.
2. **Housing costs relative to income:** When incomes are cyclically high (low) relative to home costs, including mortgage financing costs, home buying and construction tend to increase (decrease). Housing activity can decrease even when incomes are rising late in a cycle if home prices are rising faster than incomes, leading to decreases in purchase and construction activity in the housing sector.
3. **Speculative activity:** As we saw in the housing sector in 2007 and 2008 in many economies, rising home prices can lead to purchases based on expectations of further gains. Higher prices led to more construction and eventually excess building. This resulted in falling prices that decreased or eliminated speculative demand and led to dramatic decreases in housing activity overall.
4. **Demographic factors:** The proportion of the population in the 25- to 40-year-old segment is positively related to activity in the housing sector because these are the ages of greatest household formation. In China, a strong population shift from rural areas to cities as manufacturing activity has grown has required large increases in construction of new housing to accommodate those needs.

External Trade Sector Activity

The most important factors determining the level of a country's imports and exports are domestic GDP growth, GDP growth of trading partners, and currency exchange rates. Increasing growth of domestic GDP leads to increases in purchases of foreign goods (imports), while decreasing domestic GDP growth reduces imports. Exports depend on the growth rates of GDP of other economies (especially those of important trading partners). Increasing foreign incomes increase sales to foreigners (exports) and decreasing economic growth in foreign countries decreases domestic exports.

An increase in the value of a country's currency makes its goods more expensive to foreign buyers and foreign goods less expensive to domestic buyers, which tends to decrease exports and increase imports. A decrease in the value of a country's currency has the opposite effect, increasing exports and decreasing imports. Currencies affect import and export volumes over time in response to persistent trends in foreign exchange rates, rather than in response to short-term changes which can be quite volatile.

Currency effects can differ in direction from GDP growth effects and change in response to a complex set of variables. The effects of changes in GDP levels and growth rates are more direct and immediate.

Typical business cycle characteristics may be summarized as follows:

Trough:

- GDP growth rate changes from negative to positive.
- High unemployment rate, increasing use of overtime and temporary workers.
- Spending on consumer durable goods and housing may increase.
- Moderate or decreasing inflation rate.

Expansion:

- GDP growth rate increases.
- Unemployment rate decreases as hiring accelerates.
- Investment increases in producers' equipment and home construction.
- Inflation rate may increase.
- Imports increase as domestic income growth accelerates.

Peak:

- GDP growth rate decreases.
- Unemployment rate decreases but hiring slows.
- Consumer spending and business investment grow at slower rates.
- Inflation rate increases.

Contraction/recession:

- GDP growth rate is negative.
 - Hours worked decrease, unemployment rate increases.
 - Consumer spending, home construction, and business investment decrease.
 - Inflation rate decreases with a lag.
 - Imports decrease as domestic income growth slows.
-

LOS 11.d: Describe theories of the business cycle.

The causes of business cycles are a subject of considerable debate among economists.

Neoclassical school economists believe shifts in both aggregate demand and aggregate supply are primarily *driven by changes in technology* over time. They also believe that the economy has a strong tendency toward full-employment equilibrium, as recession puts downward pressure on the money wage rate, or as over-full employment puts upward pressure on the money wage rate. They conclude that business cycles result from *temporary deviations from long-run equilibrium*.

The Great Depression of the 1930s did not support the beliefs of the neoclassical economists. The economy in the United States operated significantly below its full-employment level for many years. Additionally, business cycles in general have been more severe and more prolonged than the neoclassical model would suggest.

British economist John Maynard Keynes attempted to explain the Depression and the nature of business cycles. He provided policy recommendations for moving the economy toward full-employment GDP and reducing the severity and duration of business cycles. Keynes believed

that *shifts in aggregate demand due to changes in expectations* were the primary cause of business cycles. **Keynesian school** economists believe these fluctuations are primarily due to swings in the level of optimism of those who run businesses. They overinvest and overproduce when they are too optimistic about future growth in potential GDP, and they underinvest and underproduce when they are too pessimistic or fearful about the future growth in potential GDP.

Keynesians argue that wages are “downward sticky,” reducing the ability of a decrease in money wages to increase short-run aggregate supply and move the economy from recession (or depression) back toward full employment. The policy prescription of Keynesian economists is to increase aggregate demand directly, through monetary policy (increasing the money supply) or through fiscal policy (increasing government spending, decreasing taxes, or both).

The **New Keynesian school** added the assertion that the prices of productive inputs other than labor are also “downward sticky,” presenting additional barriers to the restoration of full-employment equilibrium.

A third view of macroeconomic equilibrium is that held by the **Monetarist school**. Monetarists believe the variations in aggregate demand that cause business cycles are due to variations in the rate of growth of the money supply, likely from *inappropriate decisions by the monetary authorities*. Monetarists believe that recessions can be caused by external shocks or by inappropriate decreases in the money supply. They suggest that to keep aggregate demand stable and growing, the central bank should follow a policy of steady and predictable increases in the money supply.

Economists of the **Austrian school** believe business cycles are caused by *government intervention in the economy*. When policymakers force interest rates down to artificially low levels, firms invest too much capital in long-term and speculative lines of production, compared to actual consumer demand. When these investments turn out poorly, firms must decrease output in those lines, which causes a contraction.



PROFESSOR'S NOTE

Austrian school economists refer to this misdirection of capital as “malinvestment.”

The theory is related closely to the credit cycles discussed earlier.

New Classical school economists introduced **real business cycle theory (RBC)**. RBC emphasizes the effect of real economic variables such as *changes in technology and external shocks*, as opposed to monetary variables, as the cause of business cycles. RBC applies utility theory, which we described in the readings on microeconomic analysis, to macroeconomics. Based on a model in which individuals and firms maximize expected utility, New Classical economists argue that policymakers should not try to counteract business cycles because expansions and contractions are efficient market responses to real external shocks.



MODULE QUIZ 11.1

1. In the early part of an economic expansion, inventory-sales ratios are *most likely* to:
 - A. increase because sales are unexpectedly low.
 - B. increase because businesses plan for expansion.
 - C. decrease because of unexpected increases in sales.
2. The contraction phase of the business cycle is *least likely* accompanied by decreasing:
 - A. unemployment.

- B. inflation pressure.
 - C. economic output.
3. According to which business cycle theory should expansionary monetary policy be used to fight a recession?
- A. Keynesian school.
 - B. Monetarist school.
 - C. New Classical school.

MODULE 11.2: INFLATION AND INDICATORS



LOS 11.e: Interpret a set of economic indicators and describe their uses and limitations.

Video covering this content is available online.

Economic indicators can be classified into three categories: **leading indicators** that have been known to change direction before peaks or troughs in the business cycle, **coincident indicators** that change direction at roughly the same time as peaks or troughs, and **lagging indicators** that don't tend to change direction until after expansions or contractions are already underway.

The Conference Board publishes indexes of leading, coincident, and lagging indicators for several countries. Their indexes for the United States include the following components:

- *Leading indicators*: Average weekly hours in manufacturing; initial claims for unemployment insurance; manufacturers' new orders for consumer goods; manufacturers' new orders for non-defense capital goods ex-aircraft; Institute for Supply Management new orders index; building permits for new houses; S&P 500 equity price index; Leading Credit Index; 10-year Treasury to Fed funds interest rate spread; and consumer expectations.
- *Coincident indicators*: Employees on nonfarm payrolls; real personal income; index of industrial production; manufacturing and trade sales.
- *Lagging indicators*: Average duration of unemployment; inventory-sales ratio; change in unit labor costs; average prime lending rate; commercial and industrial loans; ratio of consumer installment debt to income; change in consumer price index.

Other sources, such as the Organization for Economic Cooperation and Development (OECD) and the Economic Cycle Research Institute (ECRI), also publish indexes of economic indicators for the world's major economies.

Analysts should use leading, coincident, and lagging indicators together to determine the phase of the business cycle. They should also use the composite indexes to confirm what is indicated by individual indicators. If a widely followed leading indicator, such as stock prices or initial claims for unemployment insurance, changes direction, but most other leading indicators have not, an analyst should not yet conclude that a peak or trough is imminent.

EXAMPLE: Interpreting economic indicators

Karen Trumbull, CFA, gathers the following economic reports for the United States in the most recent two months:

	Latest Month	Prior Month
Building permits	+1.8%	+0.7%
Commercial and industrial loans	-0.9%	-1.6%
Consumer price index	-0.1%	-0.2%
Index of industrial production	+0.2%	0.0%
New orders for consumer goods	+2.2%	+1.6%
Real personal income	0.0%	-0.4%

Based on these indicators, what should Trumbull conclude about the phase of the business cycle?

Answer:

Commercial and industrial loans and the consumer price index are lagging indicators. Industrial production and real personal income are coincident indicators. These indicators suggest the business cycle has been in the contraction phase.

Building permits and orders for consumer goods are leading indicators. Increases in both of these in the latest two months suggest an economic expansion may be emerging.

Taken together, these data indicate that the business cycle may be at or just past its trough.

Analysts should be aware that the classifications *leading*, *coincident*, and *lagging* indicators reflect tendencies in the timing of their turning points, not exact relationships with the business cycle. Not all changes in direction of leading indicator indexes have been followed by corresponding changes in the business cycle, and even when they have, the lead time has varied. This common criticism is summed up in the often repeated comment, “Declines in stock prices have predicted nine of the last four recessions.”



PROFESSOR'S NOTE

Analysts who use economic indicators in forecasting models must guard against look-ahead bias. The data are not available immediately. For example, data for May are typically first released in mid- to late June and may be revised in July and August.

LOS 11.f: Describe types of unemployment and compare measures of unemployment.

Unemployment can be divided into three categories:

1. **Frictional unemployment** results from the time lag necessary to match employees who seek work with employers needing their skills. Frictional unemployment is always with us as employers expand or contract their businesses and workers move, are fired, or quit to seek other opportunities.
2. **Structural unemployment** is caused by long-run changes in the economy that eliminate some jobs while generating others for which unemployed workers are not qualified. Structural unemployment differs from frictional unemployment in that the unemployed workers do not currently have the skills needed to perform the jobs that are available.

3. **Cyclical unemployment** is caused by changes in the general level of economic activity.

Cyclical unemployment is positive when the economy is operating at less than full capacity and can be negative when an expansion leads to employment temporarily over the full employment level.

A person who is not working is considered to be **unemployed** if he is actively searching for work.¹ One who has been seeking work unsuccessfully for several months is referred to as *long-term unemployed*.

The **unemployment rate** is the percentage of people in the labor force who are unemployed. The **labor force** includes all people who are either employed or unemployed. People who choose not to be in the labor force are said to be *voluntarily unemployed* and are not included in the calculation of the unemployment rate.

A person who is employed part time but would prefer to work full time or is employed at a low-paying job despite being qualified for a significantly higher-paying one is said to be **underemployed**. Identification of the number of underemployed is somewhat subjective and not easily discernible from employment statistics.

The **participation ratio** (also referred to as the *activity ratio* or *labor force participation rate*) is the percentage of the working-age population who are either employed or actively seeking employment.

Short-term fluctuations in the participation ratio can occur because of changes in the number of **discouraged workers**, those who are available for work but are neither employed nor actively seeking employment. The participation rate tends to increase when the economy expands and decrease during recessions. Discouraged workers who stopped seeking jobs during a recession are motivated to seek work again once the expansion takes hold and they believe their prospects of finding work are better.

This movement of discouraged workers out of and back into the labor force causes the unemployment rate to be a lagging indicator of the business cycle. Early in an expansion when hiring prospects begin to improve, the number of discouraged workers who re-enter the labor force is greater than the number who are hired immediately. This causes the unemployment rate to increase even though employment is expanding. To gauge the current state of the labor market, analysts should include other widely available indicators such as the number of employees on payrolls.

Earlier, we noted that firms tend to be slow to hire or lay off workers at business cycle turning points. This also causes the unemployment rate to lag the business cycle. The effect can also be seen in data on **productivity**, or output per hour worked. Productivity declines early in contractions as firms try to keep employees on despite producing less output. Productivity increases early in expansions as firms try to produce more output but are not yet ready to hire new workers.

When comparing unemployment rates across countries, analysts should note that different reporting agencies may use somewhat dissimilar methods for calculating the statistics. Also, all of the employment indicators mentioned here apply only to legal employment. Participants in illegal sectors of the economy are not reflected in employment data.

LOS 11.g: Explain inflation, hyperinflation, disinflation, and deflation.

Inflation is a persistent increase in the price level over time. If the price level increases in a single jump but does not continue rising, the economy is not experiencing inflation. An increase in the price of a single good, or in relative prices of some goods, is not inflation. If inflation is present, the prices of almost all goods and services are increasing.

Inflation erodes the purchasing power of a currency. Inflation favors borrowers at the expense of lenders because when the borrower returns the principal to the lender, it is worth less in terms of goods and services (in real terms) than it was worth when it was borrowed. Inflation that accelerates out of control is referred to as **hyperinflation**, which can destroy a country's monetary system and bring about social and political upheavals.

The **inflation rate** is the percentage increase in the price level, typically compared to the prior year. Analysts can use the inflation rate as a business cycle indicator and to anticipate changes in central bank monetary policy. As we will see in the reading on Monetary and Fiscal Policy, an objective of central banks is to keep inflation within some target range. **Disinflation** refers to an inflation rate that is decreasing over time but remains greater than zero.

A persistently decreasing price level (i.e., a negative inflation rate) is called **deflation**. Deflation is commonly associated with deep recessions. When most prices are decreasing, consumers delay purchases because they believe they can buy the same goods more cheaply in the future. For firms, deflation results in decreasing revenue and increasing real fixed costs.



PROFESSOR'S NOTE

Values stated as "real" are adjusted for inflation over some defined period. This makes values at different points in time comparable in terms of purchasing power.

LOS 11.h: Explain the construction of indexes used to measure inflation.

To calculate a rate of inflation, we need to use a **price index** as a proxy for the price level. A price index measures the average price for a defined basket of goods and services. The **consumer price index (CPI)** is the best-known indicator of U.S. inflation. Many countries use indexes similar to the CPI.

The CPI basket represents the purchasing patterns of a typical urban household. Weights for the major categories in the CPI are shown in Figure 11.2.

Figure 11.2: Relative Importance in the CPI as of April 2016

Category	Percent of Index
Food	13.9%
Energy	6.6%
All items less food and energy	79.5%
Commodities less food and energy commodities:	
Apparel	3.2%
New vehicles	3.8%
Used cars and trucks	2.1%
Medical care commodities	1.8%
Alcoholic beverages	1.0%
Tobacco and smoking products	0.7%
Services less energy services:	
Shelter	33.3%
Medical care services	6.6%
Transportation services	5.9%

Source: Bureau of Labor Statistics, U.S. Department of Labor (stats.bls.gov)

To calculate the CPI, the Bureau of Labor Statistics compares the cost of the CPI basket today with the cost of the basket in an earlier *base period*. The value of the index is as follows:

$$\text{CPI} = \frac{\text{cost of basket at current prices}}{\text{cost of basket at base period prices}} \times 100$$

EXAMPLE: Calculating a price index

The following table shows price information for a simplified basket of goods:

Item	Quantity	Price in Base Period	Current Price
Cheeseburgers	200	2.50	3.00
Movie tickets	50	7.00	10.00
Gasoline (in gallons)	300	1.50	3.00
Digital watches	100	12.00	9.00

Calculate the change in the price index for this basket from the base period to the current period.

Answer:

Reference base period:

Cheeseburgers	$200 \times 2.50 =$	500
Movie tickets	$50 \times 7.00 =$	350
Gasoline	$300 \times 1.50 =$	450
Watches	$100 \times 12.00 =$	1,200
Cost of basket		2,500

Current period:

Cheeseburgers	$200 \times 3.00 =$	600
Movie tickets	$50 \times 10.00 =$	500
Gasoline	$300 \times 3.00 =$	900
Watches	$100 \times 9.00 =$	900
Cost of basket		2,900

$$\text{price index} = \frac{2,900}{2,500} \times 100 = 116$$

The price index is up $\frac{116}{100} - 1 = 16\%$ over the period.



PROFESSOR'S NOTE

The LOS requires you to “explain the construction of” price indexes but does not require you to calculate them.

Analysts who compare price indexes for different countries should be aware of differences in their composition. The weights assigned to each good and service reflect the typical consumer’s purchasing patterns, which are likely to be significantly different across countries and regions. There can also be differences in how the data are collected. In the United States, for example, the most frequently cited CPI measure is based on the purchases typical of “all urban consumers.” Other countries may survey a different set of consumers and consequently use different baskets of goods.

An alternative measure of consumer price inflation is the *price index for personal consumption expenditures*. In the United States, this index is created by surveying businesses rather than consumers. The *GDP deflator*, which we described in an earlier reading, is another widely used inflation measure.

Analysts who look for emerging trends in consumer prices are often interested in the prices of goods in process. Widespread price increases for producers’ goods may be passed along to consumers. For most major economies, a **producer price index (PPI)** or **wholesale price index (WPI)** is available. Analysts can observe the PPI for different stages of processing (raw materials, intermediate goods, and finished goods) to watch for emerging price pressure. Sub-indexes of the PPI are also useful for identifying changes in relative prices of producers’ inputs, which may indicate shifts in demand among industries.

For both consumer and producer prices, analysts and policymakers often distinguish between **headline inflation** and **core inflation**. Headline inflation refers to price indexes for all goods. Core inflation refers to price indexes that exclude food and energy. Food and energy prices are typically more volatile than those of most other goods. Thus, core inflation can sometimes be a more useful measure of the underlying trend in prices.

LOS 11.i: Compare inflation measures, including their uses and limitations.

The price index we calculated in our example is a **Laspeyres index**, which uses a constant basket of goods and services. Most countries calculate consumer price inflation this way.

Three factors cause a Laspeyres index of consumer prices to be biased upward as a measure of the cost of living:

- *New goods.* Older products are often replaced by newer, but initially more expensive, products. New goods are periodically added to the market basket, and the older goods they replace are reduced in weight in the index. This biases the index upward.
- *Quality changes.* If the price of a product increases because the product has improved, the price increase is not due to inflation but still increases the price index.
- *Substitution.* Even in an inflation-free economy, prices of goods relative to each other change all the time. When two goods are substitutes for each other, consumers increase their purchases of the relatively cheaper good and buy less of the relatively more expensive good. Over time, such changes can make a Laspeyres index's fixed basket of goods a less accurate measure of typical household spending.

A technique known as **hedonic pricing** can be used to adjust a price index for product quality. To address the bias from substitution, reporting agencies can use a *chained* or *chain-weighted* price index such as a **Fisher index**. A Fisher index is the geometric mean of a Laspeyres index and a **Paasche index**. A Paasche index uses the current consumption weights, prices from the base period, and prices in the current period.

EXAMPLE: Paasche index

Continuing the example we presented earlier, assume the basket of goods has changed as follows:

Item	Quantity in base period	Price in base period	Quantity in current period	Current price
Cheeseburgers	200	2.50	205	3.00
Movie tickets	50	7.00	45	10.00
Gasoline (in gallons)	300	1.50	295	3.00
Digital watches	100	12.00	105	9.00

Calculate a Paasche index for the current period, compare it to the Laspeyres index (previously calculated as 116), and explain the difference.

Answer:

Reference base period:

Cheeseburgers	$205 \times 2.50 =$	512.50
Movie tickets	$45 \times 7.00 =$	315.00
Gasoline	$295 \times 1.50 =$	442.50
Watches	$105 \times 12.00 =$	<u>1,260.00</u>
Cost of basket		2,530.00

Current period:

Cheeseburgers	$205 \times 3.00 =$	615.00
Movie tickets	$45 \times 10.00 =$	450.00
Gasoline	$295 \times 3.00 =$	885.00
Watches	$105 \times 9.00 =$	<u>945.00</u>
Cost of basket		2,895.00

$$\text{Paasche index} = \frac{2,895}{2,530} \times 100 = 114.43$$

The Paasche index is less than 116 because, compared to the base period, consumers have substituted away from the two goods with the largest percentage price increases (gasoline and movie tickets).



PROFESSOR'S NOTE

The LOS does not require you to calculate these indexes. We show these examples to illustrate how substitution of goods by consumers can affect index values.

LOS 11.j: Contrast cost-push and demand-pull inflation.

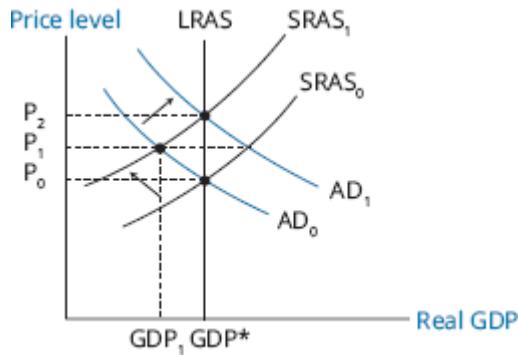
The two types of inflation are cost-push and demand-pull. **Cost-push inflation** results from a decrease in aggregate supply, while **demand-pull inflation** results from an increase in aggregate demand.

Cost-Push Inflation

Inflation can result from an initial decrease in aggregate supply caused by an increase in the real price of an important factor of production, such as wages or energy. Figure 11.3 illustrates the effect on output and the price level of a decrease in aggregate supply. The reduction from $SRAS_0$ to $SRAS_1$ increases the price level to P_1 , and with no initial change in aggregate demand, reduces output to GDP_1 .

If the decline in GDP brings a policy response that stimulates aggregate demand so output returns to its long-run potential, the result would be a further increase in the price level to P_2 .

Figure 11.3: Cost-Push Inflation



Because labor is the most important cost of production, wage pressure can be a source of cost-push inflation (sometimes called *wage-push inflation* when it occurs). Upward pressure on wages is more likely to emerge when cyclical unemployment is low, but it can occur even when cyclical unemployment is present. Because every individual provides a different type and quality of labor, some segments of the economy may have trouble finding enough qualified workers even during a contraction. As a result, the **non-accelerating inflation rate of unemployment (NAIRU)**, also called the **natural rate of unemployment (NARU)**, can be higher than the rate associated with the absence of cyclical unemployment. NARU or NAIRU can vary over time and is likely different across countries.

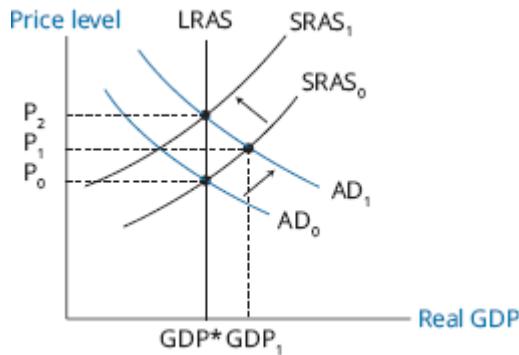
Analysts can use publicly available data on hourly and weekly earnings and labor productivity to identify signs of potential wage pressure. Wage increases are not inflationary as long as they remain in line with gains in productivity. A useful indicator of wages and benefits in terms of productivity is **unit labor costs**, the ratio of total labor compensation per hour to output units per hour.

An additional source of wage pressure is **expected inflation**. If workers expect inflation to increase, they will increase their wage demands accordingly. One indicator analysts use to gauge expected inflation is the difference in yield between inflation-indexed bonds, such as Treasury Inflation-Protected Securities, and otherwise similar non-indexed Treasury bonds.

Demand-Pull Inflation

Demand-pull inflation can result from an increase in the money supply, increased government spending, or any other change that increases aggregate demand. Figure 11.4 shows the effect on the price level when the economy is at full employment and aggregate demand increases (shifts to the right). In Figure 11.4, the economy is initially at full-employment equilibrium, with output at GDP^* and the price level at P_0 , so that the aggregate demand and short-run aggregate supply curves are AD_0 and $SRAS_0$. Real GDP is equal to potential GDP, which is represented by the long-run aggregate supply curve LRAS.

Figure 11.4: Demand-Pull Inflation



Now suppose the central bank increases the money supply, which increases aggregate demand to AD_1 . With no initial change in aggregate supply, output increases to GDP_1 , and the price level increases to P_1 . Prices rise, and real GDP is above potential (full-employment) GDP.

With real GDP above its full-employment level, the increase in GDP is not sustainable. Unemployment falls below its natural rate, which puts upward pressure on real wages. Rising real wages result in a decrease in short-run aggregate supply (the curve shifts left to $SRAS_1$) until real GDP reverts back to full-employment GDP. Output falls back to GDP^* , and the price level increases further to P_2 .

In the absence of other changes, the economy would reach a new equilibrium price level at P_2 . But what would happen if the central bank tried to keep GDP above the full-employment level with further increases in the money supply? The same results would occur repeatedly. Output cannot remain above its potential in the long run, but the induced increase in aggregate demand and the resulting pressure on wages would keep the price level rising even higher. Demand-pull inflation would persist until the central bank reduced the growth rate of the money supply and allowed the economy to return to full-employment equilibrium at a level of real GDP equal to potential GDP.

Economists often use the capacity utilization rate of industry to indicate the potential for demand-pull inflation. High rates of capacity utilization suggest the economy is producing at or above potential GDP and may experience inflationary pressure.

The impact on output is the key difference between the demand-pull and cost-push effects. The demand-pull effect increases GDP above full-employment GDP, while with cost-push inflation, a decrease in aggregate supply initially decreases GDP.



MODULE QUIZ 11.2

- An economic indicator that has turning points which tend to occur after the turning points in the business cycle is classified as:
 - a lagging indicator.
 - a leading indicator.
 - a trailing indicator.
- The unemployment rate is defined as the number of unemployed as a percentage of:
 - the labor force.
 - the number of employed.
 - the working-age population.
- A country's year-end consumer price index over a 5-year period is as follows:

Year 1 106.5

Year 2 114.2

Year 3 119.9

Year 4 124.8

Year 5 128.1

The behavior of inflation as measured by this index is *best* described as:

- A. deflation.
 - B. disinflation.
 - C. hyperinflation.
4. Core inflation is *best* described as an inflation rate:
 - A. for producers' raw materials.
 - B. the central bank views as acceptable.
 - C. that excludes certain volatile goods prices.
 5. Which of the following is *least likely* to reduce substitution bias in a consumer price index?
 - A. Use a chained index.
 - B. Use a Paasche index.
 - C. Adjust for the bias directly using hedonic pricing.
 6. In which of the following inflation scenarios does short-run aggregate supply decrease due to increasing wage demands?
 - A. Cost-push inflation.
 - B. Demand-pull inflation.
 - C. Both cost-push and demand-pull inflation.

KEY CONCEPTS

LOS 11.a

The business cycle has four phases:

1. Expansion: Real GDP is increasing.
2. Peak: Real GDP stops increasing and begins decreasing.
3. Contraction: Real GDP is decreasing.
4. Trough: Real GDP stops decreasing and begins increasing.

Expansions feature increasing output, employment, consumption, investment, and inflation. Contractions are characterized by decreases in these indicators.

Business cycles are recurring but do not occur at regular intervals, can differ in strength or severity, and do not persist for specific lengths of time.

LOS 11.b

Credit cycles are cyclical fluctuations in interest rates and credit availability. Credit cycles may amplify business cycles and cause bubbles in the markets for some assets.

LOS 11.c

Inventory to sales ratios typically increase late in expansions when sales slow and decrease near the end of contractions when sales begin to accelerate. Firms decrease or increase production to restore their inventory-sales ratios to their desired levels.

Because hiring and laying off employees have high costs, firms prefer to adjust their utilization of current employees. As a result, firms are slow to lay off employees early in contractions and slow to add employees early in expansions.

Firms use their physical capital more intensively during expansions, investing in new capacity only if they believe the expansion is likely to continue. They use physical capital less intensively during contractions, but they are more likely to reduce capacity by deferring maintenance and not replacing equipment than by selling their physical capital.

Consumer spending fluctuates with the business cycle. Durable goods spending is highly sensitive to business cycles and spending on services is somewhat sensitive, but spending on nondurable goods is relatively insensitive to business cycles.

The level of activity in the housing sector is affected by mortgage rates, demographic changes, the ratio of income to housing prices, and investment or speculative demand for homes resulting from recent price trends.

Domestic imports tend to rise with increases in GDP growth and domestic currency appreciation, while increases in foreign incomes and domestic currency depreciation tend to increase domestic export volumes.

LOS 11.d

Neoclassical economists believe business cycles are temporary and driven by changes in technology, and that rapid adjustments of wages and other input prices cause the economy to move to full-employment equilibrium.

Keynesian economists believe excessive optimism or pessimism among business managers causes business cycles and that contractions can persist because wages are slow to move downward. New Keynesians believe input prices other than wages are also slow to move downward.

Monetarists believe inappropriate changes in the rate of money supply growth cause business cycles, and that money supply growth should be maintained at a moderate and predictable rate to support the growth of real GDP.

Austrian-school economists believe business cycles are initiated by government intervention that drives interest rates to artificially low levels.

Real business cycle theory holds that business cycles can be explained by utility-maximizing actors responding to real economic forces such as external shocks and changes in technology, and that policymakers should not intervene in business cycles.

LOS 11.e

Leading indicators have turning points that tend to precede those of the business cycle.

Coincident indicators have turning points that tend to coincide with those of the business cycle.

Lagging indicators have turning points that tend to occur after those of the business cycle.

A limitation of using economic indicators to predict business cycles is that their relationships with the business cycle are inexact and can vary over time.

LOS 11.f

Frictional unemployment results from the time it takes for employers looking to fill jobs and employees seeking those jobs to find each other. Structural unemployment results from long-term economic changes that require workers to learn new skills to fill available jobs. Cyclical unemployment is positive (negative) when the economy is producing less (more) than its potential real GDP.

A person is considered unemployed if he is not working, is available for work, and is actively seeking work. The labor force includes all people who are either employed or unemployed. The unemployment rate is the percentage of labor force participants who are unemployed.

LOS 11.g

Inflation is a persistent increase in the price level over time. An inflation rate is a percentage increase in the price level from one period to the next.

Disinflation is a decrease in the inflation rate over time. Deflation refers to a persistent decrease in the price level (i.e., a negative inflation rate).

LOS 11.h

A price index measures the cost of a specific basket of goods and services relative to its cost in a prior (base) period. The inflation rate is most often calculated as the annual percentage change in a price index.

The most widely followed price index is the consumer price index (CPI), which is based on the purchasing patterns of a typical household. The GDP deflator and the producer or wholesale price index are also used as measures of inflation.

Headline inflation is a percentage change in a price index for all goods. Core inflation is calculated by excluding food and energy prices from a price index because of their high short-term volatility.

LOS 11.i

A Laspeyres price index is based on the cost of a specific basket of goods and services that represents actual consumption in a base period. New goods, quality improvements, and consumers' substitution of lower-priced goods for higher-priced goods over time cause a Laspeyres index to be biased upward.

A Paasche price index uses current consumption weights for the basket of goods and services for both periods and thereby reduces substitution bias. A Fisher price index is the geometric mean of a Laspeyres and a Paasche index.

LOS 11.j

Cost-push inflation results from a decrease in aggregate supply caused by an increase in the real price of an important factor of production, such as labor or energy.

Demand-pull inflation results from persistent increases in aggregate demand that increase the price level and temporarily increase economic output above its potential or full-employment level.

The non-accelerating inflation rate of unemployment (NAIRU) represents the unemployment rate below which upward pressure on wages is likely to develop.

Wage demands reflect inflation expectations.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 11.1

1. **C** Early in an expansion, inventory-sales ratios typically decrease below their normal levels as accelerating sales draw down inventories of produced goods. (LOS 11.c)
2. **A** An economic contraction is likely to feature increasing unemployment (i.e., decreasing employment), along with declining economic output and decreasing inflation pressure. (LOS 11.a)
3. **A** Keynesian school economists recommend monetary or fiscal policy action to stimulate aggregate demand and restore full employment. Monetarists believe the rate of money supply growth should be kept stable and predictable. The New Classical school recommends against monetary or fiscal policy intervention because recessions reflect individuals' and firms' utility-maximizing response to real factors in the economy. (LOS 11.d)

Module Quiz 11.2

1. **A** Lagging indicators have turning points that occur after business cycle turning points. (LOS 11.e)

2. **A** The unemployment rate is the number of unemployed as a percentage of the labor force. (LOS 11.f)

3. **B** The yearly inflation rate is as follows:

$$\text{Year 2 } (114.2 - 106.5) / 106.5 = 7.2\%$$

$$\text{Year 3 } (119.9 - 114.2) / 114.2 = 5.0\%$$

$$\text{Year 4 } (124.8 - 119.9) / 119.9 = 4.1\%$$

$$\text{Year 5 } (128.1 - 124.8) / 124.8 = 2.6\%$$

The inflation rate is decreasing, but the price level is still increasing. This is best described as disinflation. (LOS 11.g)

4. **C** Core inflation is measured using a price index that excludes food and energy prices. (LOS 11.h)

5. **C** Adopting a chained price index method addresses substitution bias, as does using a Paasche index. Hedonic pricing adjusts for improvements in the quality of products over time, not substitution bias. (LOS 11.i)

6. **C** Both inflation scenarios can involve a decrease in short-run aggregate supply due to increasing wage demands. In a wage-push scenario, which is a form of cost-push inflation, the decrease in aggregate supply causes real GDP to fall below full employment. In a demand-pull inflation scenario, an increase in aggregate demand causes real GDP to increase beyond full employment, which creates wage pressure that results in a decrease in short-run aggregate supply. (LOS 11.j)

READING 12

MONETARY AND FISCAL POLICY

EXAM FOCUS

This reading covers the supply and demand for money, as well as fiscal and monetary policy. This is a lot of material, but you really need to get it all down to be prepared for the exam. Concentrate initially on all the definitions and the basics of expansionary and contractionary fiscal and monetary policy. When you read it the second time, try to understand every cause-and-effect relationship so you can trace the effects of a policy change through the economy. In this way, you will be able to answer questions about the effect of, for example, open market purchases of securities by the central bank on interest rates, consumption, saving, private investment, and, of course, real GDP in the short and long run. You should understand the role of the central bank in a developed economy, including its limitations in achieving its stated objectives.

MODULE 12.1: MONEY AND INFLATION



LOS 12.a: Compare monetary and fiscal policy.

Video covering this content is available online.

Fiscal policy refers to a government's use of spending and taxation to influence economic activity. The budget is said to be *balanced* when tax revenues equal government expenditures. A **budget surplus** occurs when government tax revenues exceed expenditures, and a **budget deficit** occurs when government expenditures exceed tax revenues.

Monetary policy refers to the central bank's actions that affect the quantity of money and credit in an economy in order to influence economic activity. Monetary policy is said to be **expansionary** (or *accommodative* or *easy*) when the central bank increases the quantity of money and credit in an economy. Conversely, when the central bank is reducing the quantity of money and credit in an economy, the monetary policy is said to be **contractionary** (or *restrictive* or *tight*).

Both monetary and fiscal policies are used by policymakers with the goals of maintaining stable prices and producing positive economic growth. Fiscal policy can also be used as a tool for redistribution of income and wealth.

LOS 12.b: Describe functions and definitions of money.

Money is most commonly defined as a generally accepted medium of exchange. Rather than exchanging goods and services directly (bartering), using money facilitates indirect exchange.

Money has three primary functions:

- Money serves as a **medium of exchange** or **means of payment** because it is accepted as payment for goods and services.
- Money also serves as a **unit of account** because prices of all goods and services are expressed in units of money: dollars, yen, rupees, pesos, and so forth. This allows us to determine how much of any good we are foregoing when consuming another.
- Money provides a **store of value** because money received for work or goods now can be saved to purchase goods later.

Narrow money is the amount of notes (currency) and coins in circulation in an economy plus balances in checkable bank deposits. **Broad money** includes narrow money plus any amount available in liquid assets, which can be used to make purchases.

Measures of money differ among monetary authorities, but there is consistency in that broad measures of money include money that is less liquid (immediately spendable) than that included in narrow money measures. We have included definitions of narrow and broad monetary aggregates used by the U.S. Federal Reserve and by the European Central Bank as examples.

According to the Federal Reserve Bank of New York:

The money supply measures reflect the different degrees of liquidity—or spendability—that different types of money have. The narrowest measure, M1, is restricted to the most liquid forms of money; it consists of currency in the hands of the public; travelers checks; demand deposits, and other deposits against which checks can be written. M2 includes M1, plus savings accounts, time deposits of under \$100,000, and balances in retail money market mutual funds.

The European Central Bank describes their monetary aggregates as follows:

	M1	M2	M3
Currency in circulation	X	X	X
Overnight deposits	X	X	X
Deposits with an agreed maturity of up to 2 years		X	X
Deposits redeemable at notice of up to 3 months		X	X
Repurchase agreements			X
Money market fund shares/units			X
Debt securities issued with a maturity of up to 2 years			X

LOS 12.c: Explain the money creation process.

In the early stages of money development, **promissory notes** were developed. When customers deposited gold (or other precious metal) with early bankers, they were issued a promissory note, which was a promise by the banker to return that gold on demand from the depositor. Promissory notes themselves then became a medium of exchange. Bankers, recognizing that all the deposits would never be withdrawn at the same time, started lending a portion of deposits to earn interest. This led to what is called **fractional reserve banking**.

In a fractional reserve banking system, a bank holds a proportion of deposits in reserve. In most countries, banks are required to hold a minimum percentage of deposits as reserves.

When cash is deposited in a bank, the portion that is not required to be held in reserve can be loaned out. When a bank makes a cash loan and the borrower spends the money, the sellers who receive this cash may deposit it in banks as well. These funds can now be loaned out by these banks, except for the portion that must be held as reserves by each bank. This process of lending, spending, and depositing can continue until deposits are some multiple of the original cash amount.

Consider a bank that has \$1,000 in **excess reserves** (cash not needed for reserves) that it lends. Assume the required reserve ratio is 25%. If the borrower of the \$1,000 deposits the cash in a second bank, the second bank will be able to lend its excess reserves of \$750 ($0.75 \times \$1,000$). Those funds may be deposited in a third bank, which can then lend its excess reserve of \$563 ($0.75 \times \750). If this lending and depositing continues, the money supply can expand to \$4,000 [$(1 / 0.25) \times \$1,000$]. One dollar of excess reserves can generate a \$4 increase in the money supply.

The total amount of money that can be created is calculated as:

$$\text{money created} = \frac{\text{new deposit}}{\text{reserve requirement}} = \frac{1,000}{0.25} = \$4,000$$

With 25% of deposits held as reserves, the original deposit can result in total deposits four times as large, and we say that the **money multiplier** is four.

$$\text{money multiplier} = \frac{1}{\text{reserve requirement}} = \frac{1}{0.25} = 4$$

If the required reserve percentage is decreased, the money multiplier increases, and the quantity of money that can be created increases. If the reserve requirement was reduced from 25% to 10%, the money multiplier would increase from 4 to 10.

Relationship of Money and the Price Level

The **quantity theory of money** states that quantity of money is some proportion of the total spending in an economy and implies the **quantity equation of exchange**:

$$\text{money supply} \times \text{velocity} = \text{price} \times \text{real output} (\text{MV} = \text{PY})$$

Price multiplied by real output is total spending so that **velocity** is the average number of times per year each unit of money is used to buy goods or services. The equation of exchange must hold with velocity defined in this way.

Monetarists believe that velocity and the real output of the economy change only slowly. Assuming that velocity and real output remain constant, any increase in the money supply will lead to a proportionate increase in the price level. For example, a 5% increase in the money supply will increase average prices by 5%. For this reason, monetarists argue that monetary policy can be used to control and regulate inflation. The belief that real variables (real GDP and velocity) are not affected by monetary variables (money supply and prices) is referred to as **money neutrality**.

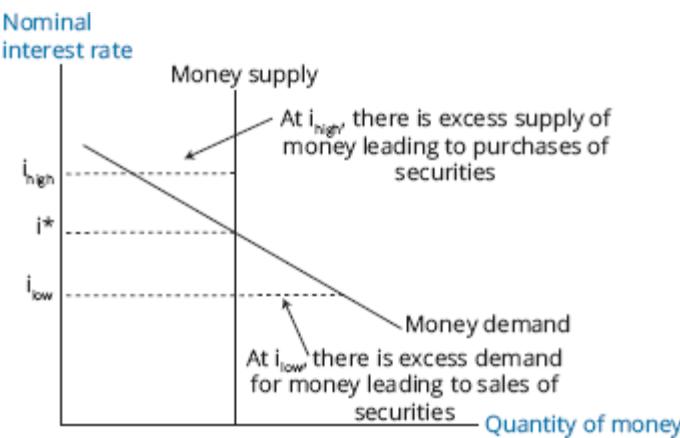
LOS 12.d: Describe theories of the demand for and supply of money.

The amount of wealth that households and firms in an economy choose to hold in the form of money is known as **demand for money**. There are three reasons for holding money:

1. *Transaction demand*: Money held to meet the need for undertaking transactions. As the level of real GDP increases, the size and number of transactions will increase, and the demand for money to carry out transactions increases.
2. *Precautionary demand*: Money held for unforeseen future needs. The demand for money for precautionary reasons is higher for large firms. In the aggregate, the total amount of precautionary demand for money increases with the size of the economy.
3. *Speculative demand*: Money that is available to take advantage of investment opportunities that arise in the future. It is inversely related to returns available in the market. As bonds and other financial instruments provide higher returns, investors would rather invest their money now than hold speculative money balances. Conversely, the demand for money for speculative reasons is positively related to perceived risk in other financial instruments. If the risk is perceived to be higher, people choose to hold money rather than invest it.

The relation between short-term interest rates and the quantity of money that firms and households demand to hold is illustrated in Figure 12.1. At lower interest rates, firms and households choose to hold more money. At higher interest rates, the opportunity cost of holding money increases, and firms and households will desire to hold less money and more interest-bearing financial assets.

Figure 12.1: The Supply and Demand for Money



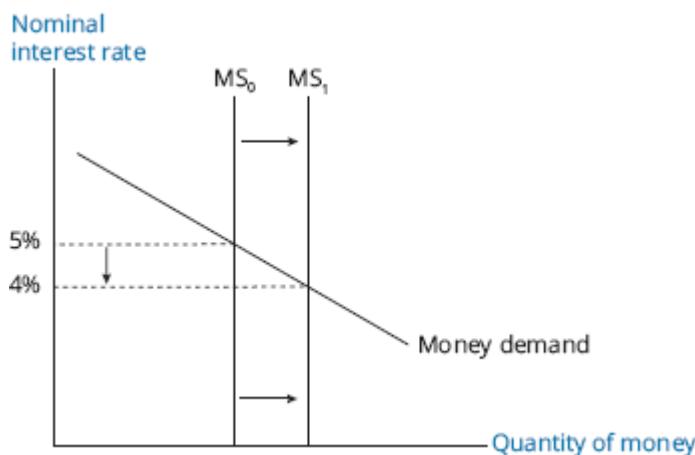
The **supply of money** is determined by the central bank (the Fed in the United States) and is independent of the interest rate. This accounts for the vertical (perfectly inelastic) supply curve in Figure 12.1.

Short-term interest rates are determined by the equilibrium between money supply and money demand. As illustrated in Figure 12.1, if the interest rate is above the equilibrium rate (i_{high}), there is excess supply of real money. Firms and households are holding more real money balances than they desire to, given the opportunity cost of holding money balances. They will

purchase securities to reduce their money balances, which will decrease the interest rate as securities prices are bid up. If interest rates are below equilibrium (i_{low}), there is excess demand for real money balances, as illustrated in Figure 12.1. Firms and households will sell securities to increase their money holdings to the desired level, decreasing securities prices and increasing the interest rate.

A central bank can affect short-term interest rates by increasing or decreasing the money supply. An increase in the money supply (shift of the money supply curve to the right) will put downward pressure on interest rates, as illustrated in Figure 12.2. With an increase in the money supply, there is excess supply of money at the previous rate of 5%. To reduce their money holdings, firms and households buy securities, increasing securities prices and decreasing the interest rate until the new equilibrium interest rate of 4% is achieved. If the central bank decreases the money supply, excess demand for money balances results in sales of securities and an increase in the interest rate.

Figure 12.2: Increase in the Money Supply



LOS 12.e: Describe the Fisher effect.

The **Fisher effect** states that the nominal interest rate is simply the sum of the real interest rate and expected inflation.

$$R_{\text{Nom}} = R_{\text{Real}} + E[I]$$

where:

R_{Nom} = nominal interest rate

R_{Real} = real interest rate

$E[I]$ = expected inflation

The idea behind the Fisher effect is that real rates are relatively stable, and changes in interest rates are driven by changes in expected inflation. This is consistent with money neutrality.

Investors are exposed to the risk that inflation and other future outcomes may be different than expected. Investors require additional return (a risk premium) for bearing this risk, which we can consider a third component of a nominal interest rate.

$$R_{\text{Nom}} = R_{\text{Real}} + E[I] + RP$$

where:

RP = risk premium for uncertainty

LOS 12.f: Describe roles and objectives of central banks.

There are several key **roles of central banks**:

1. *Sole supplier of currency*: Central banks have the sole authority to supply money. Traditionally, such money was backed by gold; the central bank stood ready to convert the money into a pre-specified quantity of gold. Later on, the gold backing was removed, and money supplied by the central bank was deemed **legal tender** by law. Money not backed by any tangible value is termed **fiat money**. As long as fiat money holds its value over time and is acceptable for transactions, it can continue to serve as a medium of exchange.
2. *Banker to the government and other banks*: Central banks provide banking services to the government and other banks in the economy.
3. *Regulator and supervisor of payments system*: In many countries, central banks may regulate the banking system by imposing standards of risk-taking allowed and reserve requirements of banks under its jurisdiction. Central banks also oversee the payments system to ensure smooth operations of the clearing system domestically and in conjunction with other central banks for international transactions.
4. *Lender of last resort*: Central banks' ability to print money allows them to supply money to banks with shortages, and this government backing tends to prevent runs on banks (i.e., large scale withdrawals) by assuring depositors their funds are secure.
5. *Holder of gold and foreign exchange reserves*: Central banks are often the repositories of the nation's gold and reserves of foreign currencies.
6. *Conductor of monetary policy*: Central banks control or influence the quantity of money supplied in an economy and growth of money supply over time.

The primary **objective of a central bank** is to *control inflation* so as to promote price stability. High inflation is not conducive to a stable economic environment. High inflation leads to **menu costs** (i.e., cost to businesses of constantly having to change their prices) and **shoe leather costs** (i.e., costs to individuals of making frequent trips to the bank so as to minimize their holdings of cash that are depreciating in value due to inflation).

In addition to price stability, some central banks have other stated goals, such as:

- Stability in exchange rates with foreign currencies.
- Full employment.
- Sustainable positive economic growth.
- Moderate long-term interest rates.

The target inflation rate in most developed countries is a range around 2% to 3%. A target of zero inflation is not used because that increases the risk of deflation, which can be very disruptive for an economy.

While most developed countries have an explicit target inflation rate, the U.S. Fed and the Bank of Japan do not. In the United States, this is because the Fed has the additional goals of maximum employment and moderate long-term interest rates. In Japan, it is because deflation, rather than inflation, has been a persistent problem in recent years.

Some developed countries, and several developing countries, choose a target level for the exchange rate of their currency with that of another country, primarily the U.S. dollar. This is referred to as **pegging** their exchange rate with the dollar. If their currency appreciates (i.e., becomes relatively more valuable), they can sell their domestic currency reserves for dollars to reduce the exchange rate. While such actions may be effective in the short run, for stability of the exchange rate over time, the monetary authorities in the pegging country must manage interest rates and economic activity to achieve their goal. This can lead to increased volatility of their money supply and interest rates. The pegging country essentially commits to a policy intended to make its inflation rate equal to the inflation rate of the country to which they peg their currency.

LOS 12.g: Contrast the costs of expected and unexpected inflation.

We turn our attention now to the costs to an economy of inflation, why central banks' target inflation rates are low, and why they care about volatility of inflation rates. At any point in time, economic agents have an expected rate of future inflation in the aggregate. The costs of inflation that is equal to the expected rate are different from the costs of inflation that differs from expectations, with the costs imposed on an economy of unanticipated inflation greater than those of perfectly anticipated inflation.

Consider an economy for which expected inflation is 6% and actual inflation will be 6% with certainty, so that inflation is *perfectly anticipated* (i.e., there is no unexpected inflation). The prices of all goods and wages could be indexed to this inflation rate so each month both wages and prices are increased approximately one-half percent. Increased demand for a product would result in monthly price increases of more than one-half percent and decreased demand would be reflected in prices that increased less than one-half percent per month.

One effect of high inflation—even when perfectly anticipated—is that the cost of holding money rather than interest-bearing securities is higher because its purchasing power decreases steadily. This will decrease the quantity of money that people willingly hold and impose some costs of more frequent movement of money from interest-bearing securities to cash or non-interest-bearing deposit accounts to facilitate transactions. To some extent, technology and the Internet have decreased these costs as movement of money between accounts has become much easier.

Much more important are the costs imposed on an economy by *unanticipated inflation*, inflation that is higher or lower than the expected rate of inflation. When inflation is higher than expected, borrowers gain at the expense of lenders as loan payments in the future are made with currency that has less value in real terms. Conversely, inflation that is less than expected will benefit lenders at the expense of borrowers. In an economy with volatile (rather than certain) inflation rates, lenders will require higher interest rates to compensate for the additional risk

they face from unexpected changes in inflation. Higher borrowing rates slow business investment and reduce the level of economic activity.

A second cost of unexpected inflation is that information about supply and demand from changes in prices becomes less reliable. Suppose that when expected inflation is 5%, a manufacturer sees that prices for his product have increased 10%. If this is interpreted as an increase in demand for the product, the manufacturer will increase capacity and production in response to the perceived increase in demand. If, in fact, general price inflation is 10% rather than the expected 5% over the recent period, the price increase in the manufacturer's product did not result from an increase in demand. The expansion of production will result in excess inventory and capacity, and the firm will decrease production, laying off workers and reducing or eliminating expenditures on increased capacity for some time. Because of these effects, unexpected inflation can increase the magnitude or frequency of business cycles. The destabilizing effects of inflation, either higher than expected or lower than expected, because of reduced information content of price changes impose real costs on an economy.



MODULE QUIZ 12.1

1. Both monetary and fiscal policy are used to:
 - A. balance the budget.
 - B. achieve economic targets.
 - C. redistribute income and wealth.
2. Which of the following statements is *least accurate*? The existence and use of money:
 - A. permits individuals to perform economic transactions.
 - B. requires the central bank to control the supply of currency.
 - C. increases the efficiency of transactions compared to a barter system.
3. If money neutrality holds, the effect of an increase in the money supply is:
 - A. higher prices.
 - B. higher output.
 - C. lower unemployment.
4. If the money supply is increasing and velocity is decreasing:
 - A. prices will decrease.
 - B. real GDP will increase.
 - C. the impact on prices and real GDP is uncertain.
5. The money supply curve is perfectly inelastic because the money:
 - A. supply is independent of interest rates.
 - B. demand schedule is downward-sloping.
 - C. supply is dependent upon interest rates.
6. The Fisher effect states that the nominal interest rate is equal to the real rate plus:
 - A. actual inflation.
 - B. average inflation.
 - C. expected inflation.
7. A central bank's policy goals *least likely* include:
 - A. price stability.
 - B. minimizing long-term interest rates.
 - C. maximizing the sustainable growth rate of the economy.
8. A country that targets a stable exchange rate with another country's currency *least likely*:
 - A. accepts the inflation rate of the other country.
 - B. will sell its currency if its foreign exchange value rises.
 - C. must also match the money supply growth rate of the other country.

MODULE 12.2: MONETARY POLICY



LOS 12.h: Describe tools used to implement monetary policy.

Video covering
this content is
available online.

Monetary policy is implemented using the **monetary policy tools** of the central bank. The three main policy tools of central banks are as follows:

1. *Policy rate*: In the United States, banks can borrow funds from the Fed if they have temporary shortfalls in reserves. The rate at which banks can borrow reserves from the Fed is termed the *discount rate*. For the European Central Bank (ECB), it is called the *refinancing rate*.

One way to lend money to banks is through a *repurchase agreement*. The central bank purchases securities from banks that, in turn, agree to repurchase the securities at a higher price in the future. The percentage difference between the purchase price and the repurchase price is effectively the rate at which the central bank is lending to member banks. The Bank of England uses this method, and its policy rate is called the *two-week repo (repurchase) rate*. A lower rate reduces banks' cost of funds, encourages lending, and tends to decrease interest rates overall. A higher policy rate has the opposite effect, decreasing lending and increasing interest rates.

In the United States, the *federal funds rate* is the rate that banks charge each other on overnight loans of reserves. The Fed sets a target for this market-determined rate and uses open market operations to move it to the target rate.

2. *Reserve requirements*: By increasing the reserve requirement (the percentage of deposits banks are required to retain as reserves), the central bank effectively decreases the funds that are available for lending and the money supply, which will tend to increase interest rates. A decrease in the reserve requirement will increase the funds available for lending and the money supply, which will tend to decrease interest rates. This tool only works well to increase the money supply if banks are willing to lend and customers are willing to borrow.
3. *Open market operations*: Buying and selling of securities by the central bank is referred to as open market operations. When the central bank buys securities, cash replaces securities in investor accounts, banks have excess reserves, more funds are available for lending, the money supply increases, and interest rates decrease. Sales of securities by the central bank have the opposite effect, reducing cash in investor accounts, excess reserves, funds available for lending, and the money supply, which will tend to cause interest rates to increase. In the United States, open market operations are the Fed's most commonly used tool and are important in achieving the federal funds target rate.

LOS 12.i: Describe the monetary transmission mechanism.

The **monetary transmission mechanism** refers to the ways in which a change in monetary policy, specifically the central bank's policy rate, affects the price level and inflation. There are four channels through which a change in the policy rates the monetary authorities control directly are transmitted to prices. They are transmitted through their effect on other short-term

rates, asset values, currency exchange rates, and expectations. We can examine the transmission mechanism in more detail by considering the effects of a change to a contractionary monetary policy implemented through an increase in the policy rate.

- Banks' *short-term lending rates will increase* in line with the increase in the policy rate. The higher rates will decrease aggregate demand as consumers reduce credit purchases and businesses cut back on investment in new projects.
- Bond prices, equity prices, and *asset prices in general will decrease* as the discount rates applied to future expected cash flows are increased. This may have a wealth effect because a decrease in the value of households' assets may increase the savings rate and decrease consumption.
- Both consumers and businesses may decrease their expenditures because their *expectations for future economic growth decrease*.
- The increase in interest rates may attract foreign investment in debt securities, leading to an *appreciation of the domestic currency relative to foreign currencies*. An appreciation of the domestic currency increases the foreign currency prices of exports and can reduce demand for the country's export goods.

Taken together, these effects act to decrease aggregate demand and put downward pressure on the price level. A decrease in the policy rate would affect the price level through the same channels, but in the opposite direction.

LOS 12.j: Explain the relationships between monetary policy and economic growth, inflation, interest, and exchange rates.

If money neutrality holds, changes in monetary policy and the policy rate will have no effect on real output. In the short run, however, changes in monetary policy can affect real economic growth as well as interest rates, inflation, and foreign exchange rates. The effects of a change to a more expansionary monetary policy may include any or all of the following:

- The central bank buys securities, which increases bank reserves.
- The interbank lending rate decreases as banks are more willing to lend each other reserves.
- Other short-term rates decrease as the increase in the supply of loanable funds decreases the equilibrium rate for loans.
- Longer-term interest rates also decrease.
- The decrease in real interest rates causes the currency to depreciate in the foreign exchange market.
- The decrease in long-term interest rates increases business investment in plant and equipment.
- Lower interest rates cause consumers to increase their purchases of houses, autos, and durable goods.
- Depreciation of the currency increases foreign demand for domestic goods.
- These increases in consumption, investment, and net exports all increase aggregate demand.
- The increase in aggregate demand increases inflation, employment, and real GDP.

The transmission mechanism for a decrease in interbank lending rates affects four things simultaneously:

1. Market rates decrease due to banks adjusting their lending rates for the short and long term.
2. Asset prices increase because lower discount rates are used for computing present values.
3. Firms and individuals raise their expectations for economic growth and profitability. They may also expect the central bank to follow up with further interest rate decreases.
4. The domestic currency depreciates due to an outflow of foreign money as real interest rates decline.

Together, these four factors increase domestic demand as people consume more (they have less incentive to save given lower interest rates) and increase net external demand (exports minus imports) because depreciation of the domestic currency makes exports less expensive to foreigners and imports more expensive in the domestic economy. The increase in overall demand and import prices tends to increase aggregate demand and domestic inflation.

LOS 12.k: Describe qualities of effective central banks.

For a central bank to succeed in its inflation-targeting policies, it should have **three essential qualities**:

1. *Independence*: For a central bank to be effective in achieving its goals, it should be free from political interference. Reducing the money supply to reduce inflation can also be expected to decrease economic growth and employment. The political party in power has an incentive to boost economic activity and reduce unemployment prior to elections. For this reason, politicians may interfere with the central bank's activities, compromising its ability to manage inflation. Independence should be thought of in relative terms (degrees of independence) rather than absolute terms. Even in the case of relatively independent central banks, the heads of the banks may be appointed by politicians.

Independence can be evaluated based on both **operational independence** and **target independence**. Operational independence means that the central bank is allowed to independently determine the policy rate. Target independence means the central bank also defines how inflation is computed, sets the target inflation level, and determines the horizon over which the target is to be achieved. The ECB has both target and operational independence, while most other central banks have only operational independence.

2. *Credibility*: To be effective, central banks should follow through on their stated intentions. If a government with large debts, instead of a central bank, set an inflation target, the target would not be credible because the government has an incentive to allow inflation to exceed the target level. On the other hand, a credible central bank's targets can become self-fulfilling prophecies. If the market believes that a central bank is serious about achieving a target inflation rate of 3%, wages and other nominal contracts will be based on 3% inflation, and actual inflation will then be close to that level.
3. *Transparency*: Transparency on the part of central banks aids their credibility. Transparency means central banks periodically disclose the state of the economic environment by issuing

inflation reports. Transparent central banks periodically report their views on the economic indicators and other factors they consider in their interest rate setting policy. When a central bank makes clear the economic indicators that it uses in establishing monetary policy and how they will be used, it not only gains credibility but makes policy changes easier to anticipate and implement.

LOS 12.l: Contrast the use of inflation, interest rate, and exchange rate targeting by central banks.

Central banks have used various economic variables and indicators over the years to make monetary policy decisions. In the past, some have used **interest rate targeting**, increasing the money supply when specific interest rates rose above the target band and decreasing the money supply (or the rate of money supply growth) when rates fell below the target band. Currently, **inflation targeting** is the most widely used tool for making monetary policy decisions and is, in fact, the method required by law in some countries. Central banks that currently use inflation targeting include the U.K., Brazil, Canada, Australia, Mexico, and the European Central Bank.

The most common inflation rate target is 2%, with a permitted deviation of $\pm 1\%$ so the target band is 1% to 3%. The reason the inflation target is not 0% is that variations around that rate would allow for negative inflation (i.e., deflation), which is considered disruptive to the smooth functioning of an economy. Central banks are not necessarily targeting current inflation, which is the result of prior policy and events, but inflation in the range of two years in the future.

Some countries, especially developing countries, use **exchange rate targeting**. That is, they target a foreign exchange rate between their currency and another (often the U.S. dollar), rather than targeting inflation. As an example, consider a country that has targeted an exchange rate for its currency versus the U.S. dollar. If the foreign exchange value of the domestic currency falls relative to the U.S. dollar, the monetary authority must use foreign reserves to purchase their domestic currency (which will reduce money supply growth and increase interest rates) in order to reach the target exchange rate. Conversely, an increase in the foreign exchange value of the domestic currency above the target rate will require sale of the domestic currency in currency markets to reduce its value (increasing the domestic money supply and decreasing interest rates) to move towards the target exchange rate. One result of exchange rate targeting may be greater volatility of the money supply because domestic monetary policy must adapt to the necessity of maintaining a stable foreign exchange rate.

Over the short term, the targeting country can purchase or sell its currency in the foreign exchange markets to influence the exchange rate. There are limits, however, on how much influence currency purchases or sales can have on exchange rates over time. For example, a country may run out of foreign reserves with which to purchase its currency when the exchange value of its currency is still below the target exchange rate.

The net effect of exchange rate targeting is that the targeting country will have the same inflation rate as the targeted currency and the targeting country will need to follow monetary policy and accept interest rates that are consistent with this goal, regardless of domestic economic circumstances.

LOS 12.m: Determine whether a monetary policy is expansionary or contractionary.

An economy's long-term sustainable real growth rate is called the **real trend rate** or, simply, the trend rate. The trend rate is not directly observable and must be estimated. The trend rate also changes over time as structural conditions of the economy change. For example, after a prolonged period of heavy debt use, consumers may increase saving and reduce consumption in order to reduce their levels of debt. This structural shift in the economy would reduce the trend growth rate.

The **neutral interest rate** of an economy is the growth rate of the money supply that neither increases nor decreases the economic growth rate:

$$\text{neutral interest rate} = \text{real trend rate of economic growth} + \text{inflation target}$$

When the policy rate is above (below) the neutral rate, the monetary policy is said to be **contractionary (expansionary)**. In general, contractionary policy is associated with a decrease in the *growth rate* of money supply, while expansionary policy increases its growth rate.

Monetary policy is often adjusted to reflect the source of inflation. For example, if inflation is above target due to higher aggregate demand (consumer and business spending), then contractionary monetary policy may be an appropriate response to reduce inflation. Suppose, however, that inflation is higher due to supply shocks, such as higher food or energy prices, and the economy is already operating below full employment. In such a situation, a contractionary monetary policy may make a bad situation worse.



PROFESSOR'S NOTE

In the United States, the Federal Reserve focuses on core inflation (i.e., excluding volatile food and energy prices) for this reason.

LOS 12.n: Describe limitations of monetary policy.

This transmission mechanism for monetary policy previously described does not always produce the intended results. In particular, long-term rates may not rise and fall with short-term rates because of the effect of monetary policy changes on expected inflation.

If individuals and businesses believe that a decrease in the money supply intended to reduce inflation will be successful, they will expect lower future inflation rates. Because long-term bond yields include a premium for expected inflation, long-term rates could fall (tending to increase economic growth), even while the central bank has increased short-term rates in order to slow economic activity. Conversely, increasing the money supply to stimulate economic activity could lead to an increase in expected inflation rates and long-term bond yields, even as short-term rates fall.

From a different perspective, monetary tightening may be viewed as too extreme, increasing the probability of a recession, making long-term bonds more attractive and reducing long-term interest rates. If money supply growth is seen as inflationary, higher expected future asset prices will make long-term bonds relatively less attractive and will increase long-term interest rates.

Bond market participants that act in this way have been called **bond market vigilantes**. When the central bank's policy is credible and investors believe that the inflation target rate will be maintained over time, this effect on long-term rates will be small.

Another situation in which the transmission mechanism may not perform as expected is if demand for money becomes very elastic and individuals willingly hold more money even without a decrease in short-term rates. Such a situation is called a **liquidity trap**. Increasing growth of the money supply will not decrease short-term rates under these conditions because individuals hold the money in cash balances instead of investing in interest-bearing securities. If an economy is experiencing deflation even though money supply policy has been expansionary, liquidity trap conditions may be present.

Compared to inflation, deflation is more difficult for central banks to reverse. In a deflationary environment, monetary policy needs to be expansionary. However, the central bank is limited to reducing the nominal policy rate to zero. Once it reaches zero, the central bank has limited ability to further stimulate the economy.

Another reason standard tools for increasing the money supply might not increase economic activity is that even with increasing excess reserves, banks may not be willing to lend. When what has become known as the *credit bubble* collapsed in 2008, banks around the world lost equity capital and desired to rebuild it. For this reason, they decreased their lending, even as money supplies were increased and short-term rates fell. With short-term rates near zero, economic growth still poor, and a real threat of deflation, central banks began a policy termed **quantitative easing**.

In the United Kingdom, quantitative easing entailed large purchases of British government bonds in the maturity range of three to five years. The intent was to reduce interest rates to encourage borrowing and to generate excess reserves in the banking system to encourage lending. Uncertainty about the economy's future caused banks to behave quite conservatively and willingly hold more excess reserves, rather than make loans.

In the United States, billions of dollars were made available for the Fed to buy assets other than short-term Treasury securities. Large amounts of mortgage securities were purchased from banks to encourage bank lending and to reduce mortgage rates in an attempt to revive the housing market, which had collapsed. When this program did not have the desired effect, a second round of quantitative easing (QE2) was initiated. The Fed purchased long-term Treasury bonds in large quantities (hundreds of billions of dollars) with the goal of bringing down longer-term interest rates and generating excess reserves to increase lending and economic growth. The Fed has also purchased securities with credit risk as part of its quantitative easing, improving banks' balance sheets but perhaps just shifting risk from the private sector to the public sector.

Monetary Policy in Developing Economies

Developing countries face problems in successfully implementing monetary policy. Without a liquid market in their government debt interest rate, information may be distorted and open market operations difficult to implement. In a very rapidly developing economy it may be quite difficult to determine the neutral rate of interest for policy purposes. Rapid financial innovation may change the demand to hold monetary aggregates. Central banks may lack credibility

because of past failure to maintain inflation rates in a target band and may not be given independence by the political authority.



MODULE QUIZ 12.2

1. A central bank conducts monetary policy primarily by altering:
 - A. the policy rate.
 - B. the inflation rate.
 - C. the long-term interest rate.
2. Purchases of securities in the open market by the monetary authorities are *least likely* to increase:
 - A. excess reserves.
 - B. cash in investor accounts.
 - C. the interbank lending rate.
3. An increase in the policy rate will *most likely* lead to an increase in:
 - A. business investment in fixed assets.
 - B. consumer spending on durable goods.
 - C. the foreign exchange value of the domestic currency.
4. Qualities of effective central banks include:
 - A. credibility and verifiability.
 - B. comparability and relevance.
 - C. independence and transparency.
5. If a country's inflation rate is below the central bank's target rate, the central bank is *most likely* to:
 - A. sell government securities.
 - B. increase the reserve requirement.
 - C. decrease the overnight lending rate.
6. Monetary policy is likely to be *least* responsive to domestic economic conditions if policymakers employ:
 - A. inflation targeting.
 - B. interest rate targeting.
 - C. exchange rate targeting.
7. Suppose an economy has a real trend rate of 2%. The central bank has set an inflation target of 4.5%. To achieve the target, the central bank has set the policy rate at 6%. Monetary policy is *most likely*:
 - A. balanced.
 - B. expansionary.
 - C. contractionary.
8. Monetary policy is *most likely* to fail to achieve its objectives when the economy is:
 - A. growing rapidly.
 - B. experiencing deflation.
 - C. experiencing disinflation.

MODULE 12.3: FISCAL POLICY



LOS 12.o: Describe roles and objectives of fiscal policy.

Video covering
this content is
available online.

Fiscal policy refers to a government's use of spending and taxation to meet macroeconomic goals. A government budget is said to be *balanced* when tax revenues equal

government expenditures. A *budget surplus* occurs when government tax revenues exceed expenditures, and a *budget deficit* occurs when government expenditures exceed tax revenues.

In general, decreased taxes and increased government spending both *increase* a budget deficit, overall demand, economic growth, and employment. Increased taxes and decreased government spending *decrease* a budget deficit, overall demand, economic growth, and employment. Budget deficits are increased in response to recessions, and budget deficits are decreased to slow growth when inflation is too high.

Keynesian economists believe that fiscal policy, through its effect on aggregate demand, can have a strong effect on economic growth when the economy is operating at less than full employment. Monetarists believe that the effect of fiscal stimulus is only temporary and that monetary policy should be used to increase or decrease inflationary pressures over time. Monetarists do not believe that monetary policy should be used in an attempt to influence aggregate demand to counter cyclical movements in the economy.

Discretionary fiscal policy refers to the spending and taxing decisions of a national government that are intended to stabilize the economy. In contrast, **automatic stabilizers** are built-in fiscal devices triggered by the state of the economy. For example, during a recession, tax receipts will fall, and government expenditures on unemployment insurance payments will increase. Both of these tend to increase budget deficits and are expansionary. Similarly, during boom times, higher tax revenues coupled with lower outflows for social programs tend to decrease budget deficits and are contractionary.

Objectives of fiscal policy may include:

- Influencing the level of economic activity and aggregate demand.
 - Redistributing wealth and income among segments of the population.
 - Allocating resources among economic agents and sectors in the economy.
-

LOS 12.p: Describe the arguments about whether the size of a national debt relative to GDP matters.

When a government runs fiscal deficits, it incurs debt that needs to be repaid as well as ongoing interest expense. Total deficits, annual deficits, and interest expense can all be evaluated relative to annual GDP. When these ratios increase beyond certain levels, it may be a cause for concern, and the solvency of the country may be questioned.

A country's **debt ratio** is the ratio of aggregate debt to GDP. Because taxes are linked to GDP, when an economy grows in real terms, tax revenues will also grow in real terms. If the real interest rate on the government's debt is higher than the real growth rate of the economy, then the debt ratio will increase over time (keeping tax rates constant). Similarly, if the real interest rate on government's debt is lower than real growth in GDP, the debt ratio will decrease (i.e., improve) over time.

Arguments *for* being concerned with the size of fiscal deficit:

- Higher deficits lead to higher future taxes. Higher future taxes will lead to disincentives to work and entrepreneurship. This leads to lower long-term economic growth.

- If markets lose confidence in the government, investors may not be willing to refinance the debt. This can lead to the government defaulting (if debt is in a foreign currency) or having to simply print money (if the debt is in local currency). Printing money would ultimately lead to higher inflation.
- Increased government borrowing will tend to increase interest rates, and firms may reduce their borrowing and investment spending as a result, decreasing the impact on aggregate demand of deficit spending. This is referred to as the **crowding-out effect** because government borrowing is taking the place of private sector borrowing.

Arguments *against* being concerned with the size of fiscal deficit:

- If the debt is primarily being held by domestic citizens, the scale of the problem is overstated.
 - If the debt is used to finance productive capital investment, future economic gains will be sufficient to repay the debt.
 - Fiscal deficits may prompt needed tax reform.
 - Deficits would not matter if private sector savings in anticipation of future tax liabilities just offsets the government deficit (Ricardian equivalence holds).
 - If the economy is operating at less than full capacity, deficits do not divert capital away from productive uses. On the contrary, deficits can aid in increasing GDP and employment.
-

LOS 12.q: Describe tools of fiscal policy, including their advantages and disadvantages.

Fiscal policy tools include spending tools and revenue tools.

Spending Tools

Transfer payments, also known as entitlement programs, redistribute wealth, taxing some and making payments to others. Examples include Social Security and unemployment insurance benefits. Transfer payments are not included in GDP computations.

Current spending refers to government purchases of goods and services on an ongoing and routine basis.

Capital spending refers to government spending on infrastructure, such as roads, schools, bridges, and hospitals. Capital spending is expected to boost future productivity of the economy.

Justification for spending tools:

- Provide services such as national defense that benefit all the residents in a country.
- Invest in infrastructure to enhance economic growth.
- Support the country's growth and unemployment targets by directly affecting aggregate demand.
- Provide a minimum standard of living.
- Subsidize investment in research and development for certain high-risk ventures consistent with future economic growth or other goals (e.g., green technology).

Revenue Tools

Direct taxes are levied on income or wealth. These include income taxes, taxes on income for national insurance, wealth taxes, estate taxes, corporate taxes, capital gains taxes, and Social Security taxes. Some progressive taxes (such as income and wealth taxes) generate revenue for wealth and income redistributing.

Indirect taxes are levied on goods and services. These include sales taxes, value-added taxes (VATs), and excise taxes. Indirect taxes can be used to reduce consumption of some goods and services (e.g., alcohol, tobacco, gambling).

Desirable attributes of tax policy:

- Simplicity to use and enforce.
- Efficiency; having the least interference with market forces and not acting as a deterrent to working.
- Fairness is quite subjective, but two commonly held beliefs are:
 - Horizontal equality: people in similar situations should pay similar taxes.
 - Vertical equality: richer people should pay more in taxes.
- Sufficiency, in that taxes should generate sufficient revenues to meet the spending needs of the government.

Advantages of fiscal policy tools:

- Social policies, such as discouraging tobacco use, can be implemented very quickly via indirect taxes.
- Quick implementation of indirect taxes also means that government revenues can be increased without significant additional costs.

Disadvantages of fiscal policy tools:

- Direct taxes and transfer payments take time to implement, delaying the impact of fiscal policy.
- Capital spending also takes a long time to implement. The economy may have recovered by the time its impact is felt.

Announcing a change in fiscal policy may have significant effects on expectations. For example, an announcement of future increase in taxes may immediately reduce current consumption, rapidly producing the desired goal of reducing aggregate demand. Note that not all fiscal policy tools affect economic activity equally. Spending tools are most effective in increasing aggregate demand. Tax reductions are somewhat less effective, as people may not spend the entire amount of the tax savings. Tax reductions for those with low incomes will be more effective in increasing aggregate demand, as those with lower incomes tend to spend a larger proportion of income on consumption; that is, they save a smaller proportion of income and have a higher marginal propensity to consume.

Fiscal Multiplier

Changes in government spending have magnified effects on aggregate demand because those whose incomes increase from increased government spending will in turn increase their spending, which increases the incomes and spending of others. The magnitude of the *multiplier effect* depends on the tax rate and on the marginal propensity to consume.

To understand the calculation of the multiplier effect, consider an increase in government spending of \$100 when the MPC is 80%, and the tax rate is 25%. The increase in spending increases incomes by \$100, but \$25 (100×0.25) of that will be paid in taxes. **Disposable income** is equal to income after taxes, so disposable income increases by $\$100 \times (1 - 0.25) = \75 . With an MPC of 80%, additional spending by those who receive the original \$100 increase is $\$75 \times 0.8 = \60 .

This additional spending will increase others' incomes by \$60 and disposable incomes by $\$60 \times 0.75 = \45 , from which they will spend $\$45 \times 0.8 = \36 .

Because each iteration of this process reduces the amount of additional spending, the effect reaches a limit. The **fiscal multiplier** determines the potential increase in aggregate demand resulting from an increase in government spending:

$$\text{fiscal multiplier} = \frac{1}{1 - \text{MPC}(1 - t)}$$

Here, with a tax rate of 25% and an MPC of 80%, the fiscal multiplier is $1 / [1 - 0.8(1 - 0.25)] = 2.5$, and the increase of \$100 in government spending has the potential to increase aggregate demand by \$250.

The fiscal multiplier is inversely related to the tax rate (higher tax rate decreases the multiplier) and directly related to the marginal propensity to consume (higher MPC increases the multiplier).

Balanced Budget Multiplier

In order to balance the budget, the government could increase taxes by \$100 to just offset a \$100 increase in spending. Changes in taxes also have a magnified effect on aggregate demand. An increase in taxes will decrease disposable income and consumption expenditures, thereby decreasing aggregate demand. The initial decrease in spending from a tax increase of \$100 is $100 \times \text{MPC} = 100 \times 0.8 = \80 ; beyond that, the multiplier effect is the same as we described for a direct increase in government spending, and the overall decrease in aggregate demand for a \$100 tax increase is $100(\text{MPC}) \times \text{fiscal multiplier}$, or, for our example, $100(0.8)(2.5) = \$200$.

Combining the total increase in aggregate demand from a \$100 increase in government spending with the total decrease in aggregate demand from a \$100 tax increase shows that the net effect on aggregate demand of both is an increase of $\$250 - \$200 = \$50$, so we can say that the balanced budget multiplier is positive.

If instead of a \$100 increase in taxes, we increased taxes by $100 / \text{MPC} = 100 / 0.8 = \125 and increased government spending by \$100, the net effect on aggregate demand would be zero.

Ricardian Equivalence

Increases in the current deficit mean greater taxes in the future. To maintain their preferred pattern of consumption over time, taxpayers may increase current savings (reduce current consumption) in order to offset the expected cost of higher future taxes. If taxpayers reduce current consumption and increase current saving by just enough to repay the principal and interest on the debt the government issued to fund the increased deficit, there is no effect on aggregate demand. This is known as **Ricardian equivalence** after economist David Ricardo. If taxpayers underestimate their future liability for servicing and repaying the debt, so that

aggregate demand is increased by equal spending and tax increases, Ricardian equivalence does not hold. Whether it does is an open question.

LOS 12.r: Explain the implementation of fiscal policy and difficulties of implementation.

Fiscal policy is implemented through changes in taxes and spending. This is called **discretionary fiscal policy** (as opposed to automatic stabilizers discussed previously). Discretionary fiscal policy would be designed to be expansionary when the economy is operating below full employment. Fiscal policy aims to stabilize aggregate demand. During recessions, actions can be taken to increase government spending or decrease taxes. Either change tends to strengthen the economy by increasing aggregate demand, putting more money in the hands of corporations and consumers to invest and spend. During inflationary economic booms, actions can be taken to decrease government spending or increase taxes. Either change tends to slow the economy by decreasing aggregate demand, taking money out of the hands of corporations and consumers, causing both investment and consumption spending to fall.

Discretionary fiscal policy is not an exact science. First, economic forecasts might be wrong, leading to incorrect policy decisions. Second, complications arise in practice that delay both the implementation of discretionary fiscal policy and the impact of policy changes on the economy. The lag between recessionary or inflationary conditions in the economy and the impact on the economy of fiscal policy changes can be divided into three types:

- **Recognition lag:** Discretionary fiscal policy decisions are made by a political process. The state of the economy is complex, and it may take policymakers time to recognize the nature and extent of the economic problems.
- **Action lag:** The time governments take to discuss, vote on, and enact fiscal policy changes.
- **Impact lag:** The time between the enactment of fiscal policy changes and when the impact of the changes on the economy actually takes place. It takes time for corporations and individuals to act on the fiscal policy changes, and fiscal multiplier effects occur only over time as well.

These lags can actually make fiscal policy counterproductive. For example, if the economy is in a recession phase, fiscal stimulus may be deemed appropriate. However, by the time fiscal stimulus is implemented and has its full impact, the economy may already be on a path to a recovery driven by the private sector.

Additional macroeconomic issues may hinder usefulness of fiscal policy:

- *Misreading economic statistics:* The full employment level for an economy is not precisely measurable. If the government relies on expansionary fiscal policy mistakenly at a time when the economy is already at full capacity, it will simply drive inflation higher.
- *Crowding-out effect:* Expansionary fiscal policy may crowd out private investment, reducing the impact on aggregate demand.
- *Supply shortages:* If economic activity is slow due to resource constraints (low availability of labor or other resources) and not due to low demand, expansionary fiscal policy will fail to achieve its objective and will probably lead to higher inflation.

- *Limits to deficits:* There is a limit to expansionary fiscal policy. If the markets perceive that the deficit is already too high as a proportion of GDP, funding the deficit will be problematic. This could lead to higher interest rates and actually make the situation worse.
 - *Multiple targets:* If the economy has high unemployment coupled with high inflation, fiscal policy cannot address both problems simultaneously.
-

LOS 12.s: Determine whether a fiscal policy is expansionary or contractionary.

Fiscal policy entails setting taxes and spending. A budget surplus (deficit) occurs when tax revenues exceed (fall short of) spending. Economists often focus on *changes* in the surplus or deficit to determine if the fiscal policy is expansionary or contractionary. An increase (decrease) in surplus is indicative of a contractionary (expansionary) fiscal policy. Similarly, an increase (decrease) in deficit is indicative of an expansionary (contractionary) fiscal policy.



PROFESSOR'S NOTE

For the exam, an increase (decrease) in a revenue item (e.g., sales tax) should be considered contractionary (expansionary), and an increase (decrease) in a spending item (e.g., construction of highways) should be considered expansionary (contractionary).

A government's intended fiscal policy is not necessarily obvious from just examining the current deficit. Consider an economy that is in recession so that transfer payments are increased and tax revenue is decreased, leading to a deficit. This does not necessarily indicate that fiscal policy is expansionary as, at least to some extent, the deficit is a natural outcome of the recession without any explicit action of the government. Economists often use a measure called the **structural (or cyclically adjusted) budget deficit** to gauge fiscal policy. This is the deficit that would occur based on current policies if the economy were at full employment.

LOS 12.t: Explain the interaction of monetary and fiscal policy.

Monetary policy and fiscal policy may each be either expansionary or contractionary, so there are four possible scenarios:

1. **Expansionary fiscal and monetary policy:** In this case, the impact will be highly expansionary taken together. Interest rates will usually be lower (due to monetary policy), and the private and public sectors will both expand.
2. **Contractionary fiscal and monetary policy:** In this case, aggregate demand and GDP would be lower, and interest rates would be higher due to tight monetary policy. Both the private and public sectors would contract.
3. **Expansionary fiscal policy + contractionary monetary policy:** In this case, aggregate demand will likely be higher (due to fiscal policy), while interest rates will be higher (due to increased government borrowing and tight monetary policy). Government spending as a proportion of GDP will increase.

4. Contractionary fiscal policy + expansionary monetary policy: In this case, interest rates will fall from decreased government borrowing and from the expansion of the money supply, increasing both private consumption and output. Government spending as a proportion of GDP will decrease due to contractionary fiscal policy. The private sector would grow as a result of lower interest rates.

Not surprisingly, the fiscal multipliers for different types of fiscal stimulus differ, and the effects of expansionary fiscal policy are greater when it is combined with expansionary monetary policy. The fiscal multiplier for direct government spending increases has been much higher than the fiscal multiplier for increases in transfers to individuals or tax reductions for workers. Within this latter category, government transfer payments to the poor have the greatest relative impact, followed by tax cuts for workers, and broader-based transfers to individuals (not targeted). For all types of fiscal stimulus, the impact is greater when the fiscal actions are combined with expansionary monetary policy. This may reflect the impact of greater inflation, falling real interest rates, and the resulting increase in business investment.



MODULE QUIZ 12.3

1. Roles and objectives of fiscal policy *most likely* include:
 - A. controlling the money supply to limit inflation.
 - B. adjusting tax rates to influence aggregate demand.
 - C. using government spending to control interest rates.
2. A government enacts a program to subsidize farmers with an expansive spending program of \$10 billion. At the same time, the government enacts a \$10 billion tax increase over the same period. Which of the following statements *best* describes the impact on aggregate demand?
 - A. Lower growth because the tax increase will have a greater effect.
 - B. No effect because the tax and spending effects just offset each other.
 - C. Higher growth because the spending increase will have a greater effect.
3. A government reduces spending by \$50 million. The tax rate is 30%, and consumers exhibit a marginal propensity to consume of 80%. The change in aggregate demand caused by the change in government spending is *closest to*:
 - A. -\$66 million.
 - B. -\$114 million.
 - C. -\$250 million.
4. The size of a national debt is *most likely* to be a concern for policymakers if:
 - A. Ricardian equivalence holds.
 - B. a crowding-out effect occurs.
 - C. debt is used to finance capital growth.
5. Sales in the retail sector have been sluggish, and consumer confidence has recently declined, indicating fewer planned purchases. In response, the president sends an expansionary government spending plan to the legislature. The plan is submitted on March 30, and the legislature refines and approves the terms of the spending plan on June 30. What type of fiscal plan is being considered, and what type of delay did the plan experience between March 30 and June 30?

Fiscal plan

- A. Discretionary
- B. Automatic
- C. Discretionary

Type of lag

- Recognition
- Action
- Action

6. A government is concerned about the timing of the impact of fiscal policy changes and is considering requiring the compilation and reporting of economic statistics weekly, rather than

quarterly. The new reporting frequency is intended to decrease:

- A. the action lag.
 - B. the impact lag.
 - C. the recognition lag.
7. Fiscal policy is *most likely* to be expansionary if tax rates:
- A. and government spending both decrease.
 - B. decrease and government spending increases.
 - C. increase and government spending decreases.
8. In the presence of tight monetary policy and loose fiscal policy, the *most likely* effect on interest rates and the private sector share in GDP are:

<u>Interest rate</u>	<u>Share of private sector</u>
A. lower	lower
B. higher	higher
C. higher	lower

KEY CONCEPTS

LOS 12.a

Fiscal policy is a government's use of taxation and spending to influence the economy. Monetary policy deals with determining the quantity of money supplied by the central bank. Both policies aim to achieve economic growth with price level stability, although governments use fiscal policy for social and political reasons as well.

LOS 12.b

Money is defined as a widely accepted medium of exchange. Functions of money include a medium of exchange, a store of value, and a unit of account.

LOS 12.c

In a fractional reserve system, new money created is a multiple of new excess reserves available for lending by banks. The potential multiplier is equal to the reciprocal of the reserve requirement and, therefore, is inversely related to the reserve requirement.

LOS 12.d

Three factors influence money demand:

- Transaction demand, for buying goods and services.
- Precautionary demand, to meet unforeseen future needs.
- Speculative demand, to take advantage of investment opportunities.

Money supply is determined by central banks with the goal of managing inflation and other economic objectives.

LOS 12.e

The Fisher effect states that a nominal interest rate is equal to the real interest rate plus the expected inflation rate.

LOS 12.f

Central bank roles include supplying currency, acting as banker to the government and to other banks, regulating and supervising the payments system, acting as a lender of last resort, holding the nation's gold and foreign currency reserves, and conducting monetary policy.

Central banks have the objective of controlling inflation, and some have additional goals of maintaining currency stability, full employment, positive sustainable economic growth, or moderate interest rates.

LOS 12.g

High inflation, even when it is perfectly anticipated, imposes costs on the economy as people reduce cash balances because of the higher opportunity cost of holding cash. More significant costs are imposed by unexpected inflation, which reduces the information value of price changes, can make economic cycles worse, and shifts wealth from lenders to borrowers. Uncertainty about the future rate of inflation increases risk, resulting in decreased business investment.

LOS 12.h

Policy tools available to central banks include the policy rate, reserve requirements, and open market operations. The policy rate is called the discount rate in the United States, the refinancing rate by the ECB, and the 2-week repo rate in the United Kingdom.

Decreasing the policy rate, decreasing reserve requirements, and making open market purchases of securities are all expansionary. Increasing the policy rate, increasing reserve requirements, and making open market sales of securities are all contractionary.

LOS 12.i

The transmission mechanism for changes in the central bank's policy rate through to prices and inflation includes one or more of the following:

- Short-term bank lending rates.
- Asset prices.
- Expectations for economic activity and future policy rate changes.
- Exchange rates with foreign currencies.

LOS 12.j

A contractionary monetary policy (increase in policy rate) will tend to decrease economic growth, increase market interest rates, decrease inflation, and lead to appreciation of the domestic currency in foreign exchange markets. An expansionary monetary policy (decrease in policy rate) will have opposite effects, tending to increase economic growth, decrease market interest rates, increase inflation, and reduce the value of the currency in foreign exchange markets.

LOS 12.k

Effective central banks exhibit independence, credibility, and transparency.

- Independence: The central bank is free from political interference.
- Credibility: The central bank follows through on its stated policy intentions.
- Transparency: The central bank makes it clear what economic indicators it uses and reports on the state of those indicators.

LOS 12.l

Most central banks set target inflation rates, typically 2% to 3%, rather than targeting interest rates as was once common. When inflation is expected to rise above (fall below) the target band,

the money supply is decreased (increased) to reduce (increase) economic activity.

Developing economies sometimes target a stable exchange rate for their currency relative to that of a developed economy, selling their currency when its value rises above the target rate and buying their currency with foreign reserves when the rate falls below the target. The developing country must follow a monetary policy that supports the target exchange rate and essentially commits to having the same inflation rate as the developed country.

LOS 12.m

The real trend rate is the long-term sustainable real growth rate of an economy. The neutral interest rate is the sum of the real trend rate and the target inflation rate. Monetary policy is said to be contractionary when the policy rate is above the neutral rate and expansionary when the policy rate is below the neutral rate.

LOS 12.n

Reasons that monetary policy may not work as intended:

- Monetary policy changes may affect inflation expectations to such an extent that long-term interest rates move opposite to short-term interest rates.
- Individuals may be willing to hold greater cash balances without a change in short-term rates (liquidity trap).
- Banks may be unwilling to lend greater amounts, even when they have increased excess reserves.
- Short-term rates cannot be reduced below zero.
- Developing economies face unique challenges in utilizing monetary policy due to undeveloped financial markets, rapid financial innovation, and lack of credibility of the monetary authority.

LOS 12.o

Fiscal policy refers to the taxing and spending policies of the government. Objectives of fiscal policy can include (1) influencing the level of economic activity, (2) redistributing wealth or income, and (3) allocating resources among industries.

LOS 12.p

Arguments for being concerned with the size of fiscal deficit:

- Higher future taxes lead to disincentives to work, negatively affecting long-term economic growth.
- Fiscal deficits may not be financed by the market when debt levels are high.
- Crowding-out effect as government borrowing increases interest rates and decreases private sector investment.

Arguments against being concerned with the size of fiscal deficit:

- Debt may be financed by domestic citizens.
- Deficits for capital spending can boost the productive capacity of the economy.
- Fiscal deficits may prompt needed tax reform.
- Ricardian equivalence may prevail: private savings rise in anticipation of the need to repay principal on government debt.

- When the economy is operating below full employment, deficits do not crowd out private investment.

LOS 12.q

Fiscal policy tools include spending tools and revenue tools. Spending tools include transfer payments, current spending (goods and services used by government), and capital spending (investment projects funded by government). Revenue tools include direct and indirect taxation.

An advantage of fiscal policy is that indirect taxes can be used to quickly implement social policies and can also be used to quickly raise revenues at a low cost.

Disadvantages of fiscal policy include time lags for implementing changes in direct taxes and time lags for capital spending changes to have an impact.

LOS 12.r

Fiscal policy is implemented by governmental changes in taxing and spending policies. Delays in realizing the effects of fiscal policy changes limit their usefulness. Delays can be caused by:

- Recognition lag: Policymakers may not immediately recognize when fiscal policy changes are needed.
- Action lag: Governments take time to enact needed fiscal policy changes.
- Impact lag: Fiscal policy changes take time to affect economic activity.

LOS 12.s

A government has a budget surplus when tax revenues exceed government spending and a deficit when spending exceeds tax revenue.

An increase (decrease) in a government budget surplus is indicative of a contractionary (expansionary) fiscal policy. Similarly, an increase (decrease) in a government budget deficit is indicative of an expansionary (contractionary) fiscal policy.

LOS 12.t

Interaction of monetary and fiscal policies:

Monetary Policy	Fiscal Policy	Interest Rates	Output	Private Sector Spending	Public Sector Spending
Tight	Tight	higher	lower	lower	lower
Easy	Easy	lower	higher	higher	higher
Tight	Easy	higher	higher	lower	higher
Easy	Tight	lower	varies	higher	lower

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 12.1

- B** Both monetary and fiscal policies primarily strive to achieve economic targets such as inflation and GDP growth. Balancing the budget is not a goal for monetary policy and is a potential outcome of fiscal policy. Fiscal policy (but not monetary policy) may secondarily be used as a tool to redistribute income and wealth. (LOS 12.a)

2. **B** Money functions as a unit of account, a medium of exchange, and a store of value. Money existed long before the idea of central banking was conceived. (LOS 12.b)
3. **A** Money neutrality is the theory that changes in the money supply do not affect real output or the velocity of money. Therefore, an increase in the money supply can only increase the price level. (LOS 12.c)
4. **C** Given the equation of exchange, $MV = PY$, an increase in the money supply is consistent with an increase in nominal GDP (PY). However, a decrease in velocity is consistent with a decrease in nominal GDP. Unless we know the size of the changes in the two variables, there is no way to tell what the net impact is on real GDP (Y) and prices (P). (LOS 12.c)
5. **A** The money supply schedule is vertical because the money supply is independent of interest rates. Central banks control the money supply. (LOS 12.d)
6. **C** The Fisher effect states that nominal interest rates are equal to the real interest rate plus the expected inflation rate. (LOS 12.e)
7. **B** Central bank goals often include maximum employment, which is interpreted as the maximum sustainable growth rate of the economy; stable prices; and *moderate* (not minimum) long-term interest rates. (LOS 12.f)
8. **C** The money supply growth rate may need to be adjusted to keep the exchange rate within acceptable bounds, but is not necessarily the same as that of the other country. The other two statements are true. (LOS 12.f)

Module Quiz 12.2

1. **A** The primary method by which a central bank conducts monetary policy is through changes in the target short-term rate or policy rate. (LOS 12.h)
2. **C** Open market purchases by monetary authorities *decrease* the interbank lending rate by increasing excess reserves that banks can lend to one another and therefore increasing their willingness to lend. (LOS 12.i)
3. **C** An increase in the policy rate is likely to increase longer-term interest rates, causing decreases in consumption spending on durable goods and business investment in plant and equipment. The increase in rates, however, makes investment in the domestic economy more attractive to foreign investors, increasing demand for the domestic currency and causing the currency to appreciate. (LOS 12.i)
4. **C** The three qualities of effective central banks are independence, credibility, and transparency. (LOS 12.k)
5. **C** Decreasing the overnight lending rate would add reserves to the banking system, which would encourage bank lending, expand the money supply, reduce interest rates, and allow GDP growth and the rate of inflation to increase. Selling government securities or increasing the reserve requirement would have the opposite effect, reducing the money supply and decreasing the inflation rate. (LOS 12.j)
6. **C** Exchange rate targeting requires monetary policy to be consistent with the goal of a stable exchange rate with the targeted currency, regardless of domestic economic conditions. (LOS 12.l)
7. **B** neutral rate = trend rate + inflation target = $2\% + 4.5\% = 6.5\%$
Because the policy rate is less than the neutral rate, monetary policy is expansionary. (LOS 12.m)
8. **B** Monetary policy has limited ability to act effectively against deflation because the policy rate cannot be reduced below zero and demand for money may be highly elastic (liquidity trap). (LOS 12.n)

Module Quiz 12.3

1. **B** Influencing the level of aggregate demand through taxation and government spending is an objective of fiscal policy. Controlling inflation and interest rates are typical objectives of monetary policy. (LOS 12.o)
2. **C** The amount of the spending program exactly offsets the amount of the tax increase, leaving the budget unaffected. The multiplier for government spending is greater than the multiplier for a tax increase. Therefore,

the balanced budget multiplier is positive. All of the government spending enters the economy as increased expenditure, whereas spending is reduced by only a portion of the tax increase. (LOS 12.q)

3. **B** fiscal multiplier = $1 / [1 - MPC(1 - T)] = 1 / [1 - 0.80(1 - 0.3)] = 2.27$

change in government spending = -\$50 million

change in aggregate demand = $-(50 \times 2.27) = -\$113.64$ million (LOS 12.q)

4. **B** Crowding out refers to the possibility that government borrowing causes interest rates to increase and private investment to decrease. If government debt is financing the growth of productive capital, this should increase future economic growth and tax receipts to repay the debt. Ricardian equivalence is the theory that if government debt increases, private citizens will increase savings in anticipation of higher future taxes, and it is an argument against being concerned about the size of government debt and budget deficits. (LOS 12.p)

5. **C** The expansionary plan initiated by the president and approved by the legislature is an example of discretionary fiscal policy. The lag from the time of the submission (March 30) through time of the vote (June 30) is known as action lag. It took the legislature three months to write and pass the necessary laws. (LOS 12.r)

6. **C** More frequent and current economic data would make it easier for authorities to monitor the economy and to recognize problems. The reduction in the time between economic reports should reduce the recognition lag. (LOS 12.r)

7. **B** Increases in government spending and decreases in taxes are expansionary fiscal policy. Decreases in spending and increases in taxes are contractionary fiscal policy. (LOS 12.s)

8. **C** Tight monetary policy and loose fiscal policy both lead to higher interest rates. Tight monetary policy decreases private sector growth, while loose fiscal policy expands the public sector, reducing the overall share of private sector in the GDP. (LOS 12.t)

READING 13

INTRODUCTION TO GEOPOLITICS

EXAM FOCUS

Candidates should be familiar with the terminology from this reading as well as the framework for analysis of cooperation versus competition and the difference between globalization and nationalism. Finally, be able to demonstrate the relation between geopolitical risks and investment risks.

MODULE 13.1: GEOPOLITICS AND GEOPOLITICAL RISK



Video covering this content is available online.

LOS 13.a: Describe geopolitics from a cooperation versus competition perspective.

Geopolitics refers to interactions among nations, including the actions of **state actors** (national governments) and **non-state actors** (corporations, non-government organizations, and individuals).

Geopolitics also refers to the study of how geography affects interactions among nations and their citizens. For example, firms located in coastal countries naturally tend to be the dominant participants in international shipping.

One way to examine geopolitics is through analysis of the extent to which individual countries cooperate with one another. Potential areas for cooperation include diplomatic and military matters and economic and cultural interactions. In terms of economics, areas of cooperation include freedom of movement across borders for goods, services, and capital; agreements to harmonize tariffs; international standardization of rules; and transfers of information and technology.

While a country that engages with other countries on these matters may be considered **cooperative** and one that does not may be considered **non-cooperative**, the extent of cooperation actually varies along a spectrum. A country might be more cooperative on some issues and less cooperative on others, and its degree of cooperation can change over time or with the outcomes of the country's domestic politics. A country's current decision makers and the length of its political cycle are factors to consider when analyzing geopolitics.

A country will typically cooperate with other countries when doing so advances its national interests. For example, a country may cooperate with its neighbors in a military alliance if doing so will further its interests in protecting its citizens from foreign invaders.

We can analyze a country's national interests as a hierarchy, with its top priorities being those that ensure its survival. A country's **geophysical resource endowment** may influence its priorities. For example, a country that has mineral resources but lacks arable land needs to trade minerals for food, and therefore has an interest in cooperating with other countries to keep international trade lanes open.

Non-state actors often have interests in cooperating across borders. Individuals and firms seek to direct their resources to their highest-valued uses, and some of those uses may be in other countries. To facilitate the flow of resources, state and non-state actors may cooperate on **standardization** of regulations and processes. One key example of standardization among countries is International Financial Reporting Standards for firms presenting their accounting data to the public, which we will examine in the Financial Statement Analysis topic area.

Cultural factors, such as historical emigration patterns or a shared language, can be another influence on a country's level of cooperation. Among these cultural factors are a country's formal and informal **institutions**, such as laws, public and private organizations, or distinct customs and habits. Strong and stable institutions can make cooperation easier for state and non-state actors. For example, countries that produce and export large amounts of cultural content tend to be those with legal and ethical institutions that protect intellectual property. Cultural exchange is one means through which a country may exercise **soft power**, the ability to influence other countries without using or threatening force.

LOS 13.b: Describe geopolitics and its relationship with globalization.

Globalization refers to the long-term trend toward worldwide integration of economic activity and cultures. Data from the World Bank suggest economic openness, as measured by international trade as a percentage of total output, increased steadily from about 25% in the early 1970s to about 60% before the 2008 financial crisis, and has remained near that level since then. We may contrast globalization with **nationalism**, which in this context refers to a nation pursuing its own economic interests independently of, or in competition with, the economic interests of other countries.



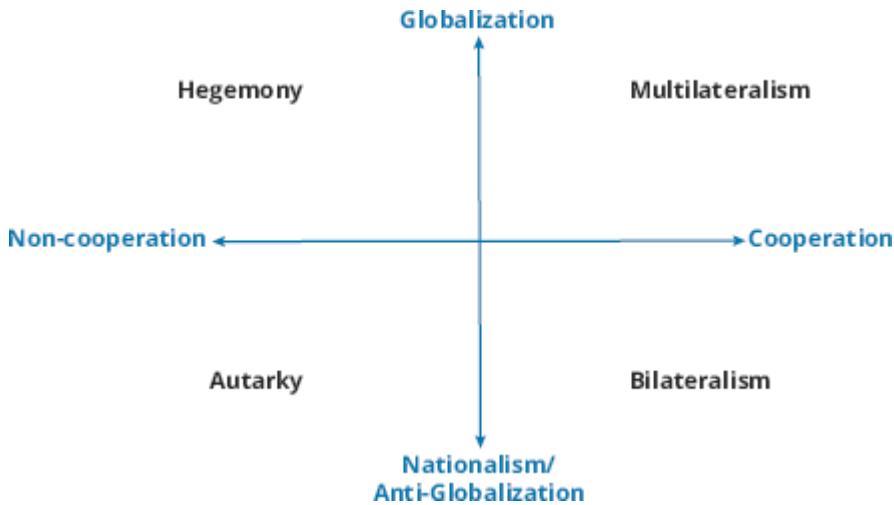
PROFESSOR'S NOTE

Debate about what the word "nationalism" means is beyond the scope of the CFA curriculum. We use it here only in the sense of opposition or resistance to globalization.

As we did with cooperation versus non-cooperation, we can think of countries' actions along a spectrum from globalization to nationalism. In general, countries that are closer to the globalization end of the spectrum are those that more actively import and export goods and services, permit freer movement of capital across borders and exchange of currencies, and are more open to cultural interaction.

In Figure 13.1 we draw each spectrum as an axis. This creates four quadrants, each of which we can associate with a type of behavior by countries. While individual countries rarely fit neatly into one of these categories, this gives us a general framework within which we can describe geopolitical actions.

Figure 13.1: Archetypes of Globalization and Cooperation



Reproduced from Level I CFA Curriculum learning module “Introduction to Geopolitics,” Exhibit 6, with permission from CFA Institute.

Characteristics we may associate with each of these categories are as follows.

- **Autarky** (non-cooperation and nationalism) refers to a goal of national self-reliance, including producing most or all necessary goods and services domestically. Autarky is often associated with a state-dominated society in general, with attributes such as government control of industry and media.
- **Hegemony** (non-cooperation and globalization) refers to countries that are open to globalization but have the size and scale to influence other countries without necessarily cooperating.
- **Bilateralism** (cooperation and nationalism) refers to cooperation between two countries. A country that engages in bilateralism may have many such relationships with other countries while tending not to involve itself in multi-country arrangements.
- **Multilateralism** (cooperation and globalization) refers to countries that engage extensively in international trade and other forms of cooperation with many other countries. Some countries may exhibit **regionalism**, cooperating multilaterally with nearby countries but less so with the world at large.

Some of the non-state actors within a country may be more oriented toward globalization than their governments. Businesses may look outside their home country for opportunities to increase profits, reduce costs, and sell to new markets. Investors may seek higher returns or diversification by investing outside their home country. Non-state actors might buy and sell foreign securities (**portfolio investment flows**) or own physical production capacity in other countries (**foreign direct investment**).

LOS 13.c: Describe tools of geopolitics and their impact on regions and economies.

We can consider **tools of geopolitics**, the means by which (primarily) state actors advance their interests in the world, as falling into three broad categories of national security, economic, and financial.

National security tools may include armed conflict, espionage, or bilateral or multilateral agreements designed to reinforce or prevent armed conflict. We can say a national security tool is *active* if a country is currently using it or *threatened* if a country is not currently using it but appears likely to do so. Armed conflict affects regions and economies by destroying productive capital and causing migration away from areas of conflict.

Economic tools can be cooperative or non-cooperative. Examples of cooperative economic tools include free trade areas, common markets, and economic and monetary unions (each of which we describe in our reading on International Trade and Capital Flows). Examples of non-cooperative economic tools include domestic content requirements, voluntary export restraints, and nationalization (i.e., the state taking control) of companies or industries.

Financial tools include foreign investment and the exchange of currencies. We can view countries as using these tools cooperatively if they allow foreign investment and the free exchange of currencies, or non-cooperatively when they restrict these activities. **Sanctions**, or restrictions on a specific geopolitical actor's financial interests, are a financial tool that state actors may use alongside national security tools.

LOS 13.d: Describe geopolitical risk and its impact on investments.

Geopolitical risk is the possibility of events that interrupt peaceful international relations. We can classify geopolitical risk into three types:

- **Event risk** refers to events about which we know the timing but not the outcome, such as national elections.
- **Exogenous risk** refers to unanticipated events, such as outbreaks of war or rebellion.
- **Thematic risk** refers to known factors that have effects over long periods, such as human migration patterns or cyber risks.

Geopolitical risk affects investment values by increasing or decreasing the risk premium investors require to hold assets in a country or region. To forecast the effect on investments of a geopolitical risk, we need to consider its probability (*likelihood*), the magnitude of its effects on investment outcomes (*impact*), and how quickly investment values would reflect these effects (*velocity*).

We can use our framework of cooperation and globalization to help estimate the **likelihood of geopolitical risk**. Countries that are more cooperative and globalized tend to have less likelihood of some geopolitical risks, such as armed conflict, but may have greater likelihood of other risks, such as the supply chain disruptions that followed the COVID-19 pandemic in 2020–2021.

To analyze the **velocity of geopolitical risk** we can classify risks as high velocity (short term), medium velocity, or low velocity (long term). Exogenous risks often have high-velocity effects on financial markets and investment values. **Black swan risk** is a term for the risk of low-likelihood exogenous events that have substantial short-term effects. Investors with longer time horizons typically do not need to react to these kinds of events, but investors with shorter horizons might find it necessary to react.

Medium-velocity risks can potentially damage specific companies or industries by increasing their costs or disrupting their production processes, while low-velocity risks tend to affect them in the “environmental, social, and governance” realm. Analyzing these kinds of risk is important for investors with long time horizons.

Because analyzing geopolitical risks requires effort, time, and resources, investors should consider whether the **impact of geopolitical risk** is likely to be high or low, and focus their analysis on risks that could have a high impact. With regard to those risks, investors should determine whether they are likely to have *discrete impacts* on a company or industry, or *broad impacts* on a country, a region, or the world. Business cycles can affect the impact of geopolitical risk, in that these risks may have greater impacts on investment values when an economy is in recession than they would have during an expansion.

Investors can use qualitative or quantitative **scenario analysis** to gauge the potential effects of geopolitical risks on their portfolios. To help identify geopolitical risks over time, investors may identify **signposts**, or data that can signal when the likelihood of an event is increasing or decreasing, such as volatility indicators in financial markets.



MODULE QUIZ 13.1

1. A state actor that is generally cooperative with other countries and primarily nationalist in pursuing its objectives is *most* accurately said to exhibit:
 - A. autarky.
 - B. hegemony.
 - C. bilateralism.
2. Which of the following tools of geopolitics is *best* described as a non-cooperative economic tool?
 - A. Voluntary export restraints.
 - B. Regional free trade agreements.
 - C. Restrictions on conversion of currencies.
3. When investing for a long time horizon, a portfolio manager should *most likely* devote resources to analyzing:
 - A. event risks.
 - B. thematic risks.
 - C. exogenous risks.

KEY CONCEPTS

LOS 13.a

Geopolitics refers to interactions among nations. On various issues ranging from diplomacy and military force to economic or cultural openness, countries lie along a spectrum from cooperative to non-cooperative.

LOS 13.b

Globalization refers to integration of economic activity and cultures among countries, and can be contrasted with nationalism, which refers to a country pursuing its own interests independently of other countries. Analysts should view geopolitical actions as being on a spectrum from nationalism to globalization.

We may describe geopolitics and its relationship with globalization using the following four broad categories: autarky (non-cooperation and nationalism); hegemony (non-cooperation and

globalization); bilateralism (cooperation and nationalism); and multilateralism (cooperation and globalization).

LOS 13.c

Tools of geopolitics include national security tools, economic tools, and financial tools.

National security tools may include armed conflict, espionage, or bilateral or multilateral national security agreements.

Cooperative economic tools include free trade areas, common markets, and economic and monetary unions. Non-cooperative economic tools include domestic content requirements, voluntary export restraints, and nationalization.

Financial tools include foreign investment, exchange of currencies, and sanctions.

LOS 13.d

Categories of geopolitical risk are event risk (when the timing is known), exogenous risk (unanticipated events), and thematic risk (known factors that have long-term effects).

Investors should analyze the likelihood of a geopolitical risk, the impact on investment values of an event if it occurs, and the velocity with which it would affect investment values.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 13.1

1. **C** Bilateralism is characterized by nationalism (as contrasted with globalization) and cooperation. Both autarky and hegemony are characterized by non-cooperation. (LOS 13.a, 13.b)
2. **A** Voluntary export restraints (exporting less of a good than the global market demands) are an example of a non-cooperative economic tool. Restrictions on the exchange of currencies are a financial tool. Free trade agreements are a cooperative economic tool. (LOS 13.c)
3. **B** Thematic risks are those that have effects over the long term. Event risks and exogenous risks are more likely to have high-velocity impacts on investment values but are less of a focus for investors with longer time horizons. (LOS 13.d)

READING 14

INTERNATIONAL TRADE AND CAPITAL FLOWS

EXAM FOCUS

International trade and currency exchange rates are key topics for both Level I and Level II. First, learn how comparative advantage results in a welfare gain from international trade and the two models of the sources of comparative advantage. Learn the types of trade restrictions and their effects on domestic price and quantity. For the balance of payments, focus on how a surplus or deficit in the broadly defined capital account must offset a deficit or surplus in the merchandise trade account. Finally, focus on how the difference between domestic income and expenditures and the difference between domestic savings and investment are related to a country's balance of trade.

MODULE 14.1: INTERNATIONAL TRADE BENEFITS



Video covering this content is available online.

Before we address specific topics and learning outcomes, it will help to define some terms as follows.

Imports: Goods and services that firms, individuals, and governments purchase from producers in other countries.

Exports: Goods and services that firms, individuals, and governments from other countries purchase from domestic producers.

Autarky or closed economy: A country that does not trade with other countries.

Free trade: A government places no restrictions or charges on import and export activity.

Trade protection: A government places restrictions, limits, or charges on exports or imports.

World price: The price of a good or service in world markets for those to whom trade is not restricted.

Domestic price: The price of a good or service in the domestic country, which may be equal to the world price if free trade is permitted or different from the world price when the domestic country restricts trade.

Net exports: The value of a country's exports minus the value of its imports over some period.

Trade surplus: Net exports are positive; the value of the goods and services a country exports are greater than the value of the goods and services it imports.

Trade deficit: Net exports are negative; the value of the goods and services a country exports is less than the value of the goods and services it imports.

Terms of trade: The ratio of an index of the prices of a country's exports to an index of the prices of its imports expressed relative to a base value of 100. If a country's terms of trade are currently 102, the prices of the goods it exports have risen relative to the prices of the goods it imports since the base period.

Foreign direct investment: Ownership of productive resources (land, factories, natural resources) in a foreign country.

Multinational corporation: A firm that has made foreign direct investment in one or more foreign countries, operating production facilities and subsidiary companies in foreign countries.

LOS 14.a: Compare gross domestic product and gross national product.

Gross domestic product over a period, typically a year, is the total value of goods and services produced within a country's borders. **Gross national product** is similar but measures the total value of goods and services produced by the labor and capital of a country's citizens. The difference is due to non-citizen incomes of foreigners working within a country, the income of citizens who work in other countries, the income of foreign capital invested within a country, and the income of capital supplied by its citizens to foreign countries. The income to capital owned by foreigners invested within a country is included in the domestic country's GDP but not in its GNP. The income of a country's citizens working abroad is included in its GNP but not in its GDP.

GDP is more closely related to economic activity within a country and so to its employment and growth.

LOS 14.b: Describe benefits and costs of international trade.

The benefits of trade are not hard to understand. As an example, consider China, and really Asia as a whole, which has had rapidly growing exports to the United States and other countries. The benefit to the importing countries has been lower-cost goods, from textiles to electronics. The benefits to the Chinese economy have been in increasing employment, increasing wages for workers, and the profits from its export products.

The costs of trade are primarily borne by those in domestic industries that compete with imported goods. Textile workers who have lost their jobs in the United States, as more and more textiles are imported, are certainly worse off in the short run. As other industries, such as health care, have grown, these workers have had to retrain to qualify for the new jobs in those fields. At the same time, U.S. firms that produce textile products using capital and technology intensive production methods have expanded. We address the reasons for this and the underlying economic theory in this reading.

Overall, economics tells us that the benefits of trade are greater than the costs for economies as a whole, so that the winners could conceivably compensate the losers and still be better off. We

now turn to the economic theory that supports this view.

LOS 14.c: Contrast comparative advantage and absolute advantage.

A country is said to have an **absolute advantage** in the production of a good if it can produce the good at a lower resource cost than another country. A country is said to have a **comparative advantage** in the production of a good if it has a lower **opportunity cost** in the production of that good, expressed as the amount of another good that could have been produced instead. Economic analysis tells us that, regardless of which country has an absolute advantage, there are potential gains from trade as long as the countries' opportunity costs of one good in terms of another are different.

This analysis is credited to David Ricardo who presented it in 1817. He used the example of the production of cloth and wine in England and Portugal. A hypothetical example of the amounts of cloth and wine these countries can produce per day of labor is presented in Figure 14.1.

Figure 14.1: Output per Unit of Labor

	Yards of Cloth	Bottles of Wine
Portugal	100	110
England	90	80

Ricardo argued that, in the absence of trading costs, England could trade cloth for wine, and Portugal could trade wine for cloth, and both countries could have more of both wine and cloth as a result. Because in Portugal a worker-day can be used to produce either 100 yards of cloth or 110 bottles of wine, its opportunity cost of a yard of cloth is $110 / 100 = 1.1$ bottles of wine and its opportunity cost of a bottle of wine is $100 / 110 = 0.91$ yards of cloth. England's opportunity cost of a yard of cloth is $80 / 90 = 0.89$ bottles of wine and its opportunity cost of a bottle of wine is $90 / 80 = 1.125$ yards of cloth.

Portugal has a comparative advantage in the production of wine as its opportunity cost is 0.91 yards of cloth compared to England's opportunity cost of 1.125 yards of cloth. As must be the case, if Portugal has a comparative advantage in wine production, England has a comparative advantage in cloth production.

To illustrate the benefits of trade, consider the output change if Portugal shifts 8 worker-days from cloth production to wine production and England shifts 10 worker-days from wine production to cloth production.

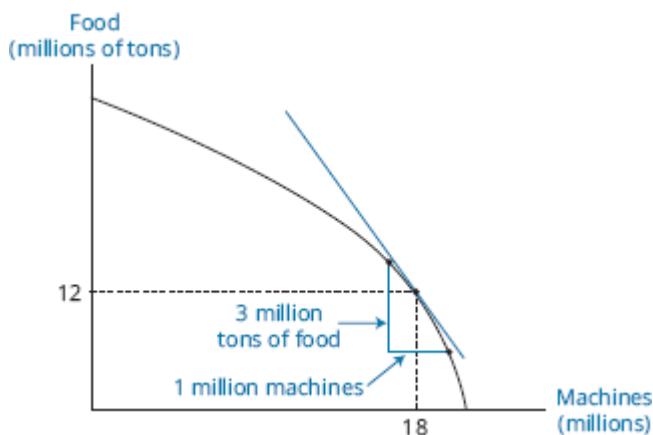
- The change in Portugal's production is $8 \times 110 = +880$ wine and $-8 \times 100 = -800$ cloth.
- The change in England's production is $-10 \times 80 = -800$ wine and $10 \times 90 = +900$ cloth.

Total production by the two countries will have increased by 80 bottles of wine and 100 yards of cloth; these are the gains from trade. The negotiated terms of trade will determine how the two countries share these gains. The important result is that *total output has increased* through trade, there's greater specialization by Portugal in wine production, and there's greater specialization by England in cloth production.

Note that Portugal has an absolute advantage in the production of both goods. However, because the countries' opportunity costs of production differ, each has a comparative advantage in one of the goods, and trade can make both countries better off.

In our simple example, we assume constant opportunity costs. As a country specializes and increases the production of an export good, increasing costs (e.g., using more marginal land for grape growing) will increase the opportunity cost of the export good. The **production possibility frontier** shown in Figure 14.2 illustrates such a situation and shows all combinations of food and machinery that an economy can produce. The slope of the frontier measures the opportunity cost of machinery in terms of food at each possible combination of food and machinery. Over a range of possible output choices around 12 million tons of food and 18 million machines, we show the slope is -3 and the opportunity cost of each million machines is 3 million tons of food. If the country were to increase the production of machinery, the amount of food production foregone would increase, as shown by the increasingly negative slope of the frontier.

Figure 14.2: A Production Possibility Frontier



LOS 14.d: Compare the Ricardian and Heckscher-Ohlin models of trade and the source(s) of comparative advantage in each model.

The **Ricardian model of trade** has only one factor of production—labor. The source of differences in production costs in Ricardo's model is *differences in labor productivity* due to differences in technology.

Heckscher and Ohlin presented a model in which there are two factors of production—capital and labor. The source of comparative advantage (differences in opportunity costs) in this model is *differences in the relative amounts of each factor* the countries possess. We can view the England and Portugal example in these terms by assuming that England has more capital (machinery) compared to labor than Portugal. Additionally, we need to assume that cloth production is more capital intensive than wine production. The result of their analysis is that the country that has more capital will specialize in the capital intensive good and trade for the less capital intensive good with the country that has relatively more labor and less capital.

In the **Heckscher-Ohlin model**, there is a redistribution of wealth within each country between labor and the owners of capital. The price of the relatively less scarce (more available)

factor of production in each country will increase so that owners of capital will earn more in England, and workers will earn more in Portugal compared to what they were without trade. This is easy to understand in the context of prices of the two goods. The good that a country imports will fall in price (that is why they import it), and the good that a country exports will rise in price. In our example, this means that the price of wine falls, and the price of cloth rises in England. Because with trade, more of the capital-intensive good, cloth, is produced in England, demand for capital and the price of capital will increase in England. As a result, capital receives more income at the expense of labor in England. In Portugal, increasing the production of wine (which is labor intensive) increases the demand for and price of labor, and workers gain at the expense of the owners of capital.



PROFESSOR'S NOTE

Remember that the model named after one economist has one factor of production, and the model named after two economists has two factors of production.



MODULE QUIZ 14.1

1. The income from a financial investment in Country P by a citizen of Country Q is *most likely* included in:
 - A. Country P's GDP but not its GNP.
 - B. Country Q's GNP and GDP.
 - C. Country P's GDP and GNP.
2. Which of the following effects is *most likely* to occur in a country that increases its openness to international trade?
 - A. Increased prices of consumer goods.
 - B. Greater specialization in domestic output.
 - C. Decreased employment in exporting industries.
3. Which of the following statements about international trade is *least accurate*? If two countries have different opportunity costs of production for two goods, by engaging in trade:
 - A. each country gains by importing the good for which it has a comparative advantage.
 - B. each country can achieve a level of consumption outside its domestic production possibility frontier.
 - C. the low opportunity cost producer of each good will export to the high opportunity cost producer of that good.
4. With regard to the Ricardian and Heckscher-Ohlin models of international trade, the amount of capital relative to labor within a country is a factor in:
 - A. both of these models.
 - B. neither of these models.
 - C. only one of these models.

MODULE 14.2: TRADE RESTRICTIONS



LOS 14.e: Compare types of trade and capital restrictions and their economic implications.

Video covering this content is available online.

There are many reasons (at least stated reasons) why governments impose trade restrictions. Some have support among economists as conceivably valid in terms of increasing a country's

welfare, while others have little or no support from economic theory. Some of the reasons for trade restrictions that have support from economists are:

- *Infant industry.* Protection from foreign competition is given to new industries to give them an opportunity to grow to an internationally competitive scale and get up the learning curve in terms of efficient production methods.
- *National security.* Even if imports are cheaper, it may be in the country's best interest to protect producers of goods crucial to the country's national defense so that those goods are available domestically in the event of conflict.

Other arguments for trade restrictions that have little support in theory are:

- *Protecting domestic jobs.* While some jobs are certainly lost, and some groups and regions are negatively affected by free trade, other jobs (in export industries or growing domestic goods and services industries) will be created, and prices for domestic consumers will be less without import restrictions.
- *Protecting domestic industries.* Industry firms often use political influence to get protection from foreign competition, usually to the detriment of consumers, who pay higher prices.

Other arguments include retaliation for foreign trade restrictions; government collection of tariffs (like taxes on imported goods); countering the effects of government subsidies paid to foreign producers; and preventing foreign exports at less than their cost of production (*dumping*).

Types of trade restrictions include:

- **Tariffs:** Taxes on imported good collected by the government.
- **Quotas:** Limits on the amount of imports allowed over some period.
- **Export subsidies:** Government payments to firms that export goods.
- **Minimum domestic content:** Requirement that some percentage of product content must be from the domestic country.
- **Voluntary export restraint:** A country voluntarily restricts the amount of a good that can be exported, often in the hope of avoiding tariffs or quotas imposed by their trading partners.

Economic Implications of Trade Restrictions

We will now examine the effects of the primary types of trade restrictions, tariffs, and subsidies.

A **tariff** placed on an imported good increases the domestic price, decreases the quantity imported, and increases the quantity supplied domestically. Domestic producers gain, foreign exporters lose, and the domestic government gains by the amount of the tariff revenues.

A **quota** restricts the quantity of a good imported to the quota amount. Domestic producers gain, and domestic consumers lose from an increase in the domestic price. The right to export a specific quantity to the domestic country is granted by the domestic government, which may or may not charge for the import licenses to foreign countries. If the import licenses are sold, the domestic government gains the revenue.

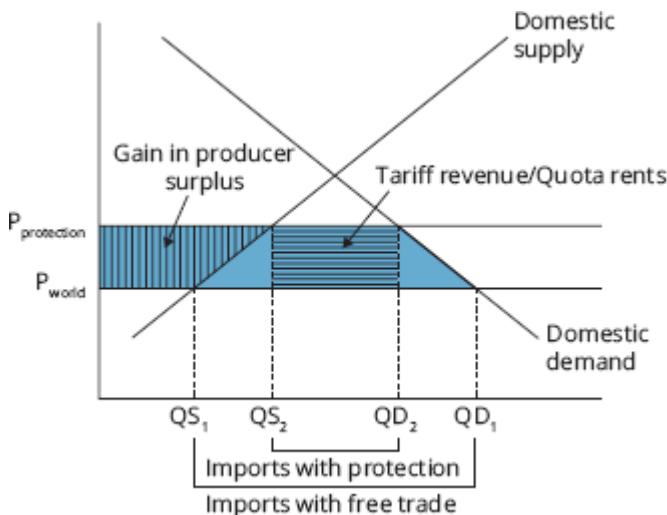
We illustrate the overall welfare effects of quotas and tariffs for a small country in Figure 14.3. We define a quota that is equivalent to a given tariff as a quota that will result in the same

decrease in the quantity of a good imported as the tariff. Defined this way, a tariff and an equivalent quota both increase the domestic price from P_{world} , the price that prevails with no trade restriction, to $P_{\text{protection}}$.

At P_{world} , prior to any restriction, the domestic quantity supplied is QS_1 , and the domestic quantity demanded is QD_1 , with the difference equal to the quantity imported, $QD_1 - QS_1$. Placing a tariff on imports increases the domestic price to $P_{\text{protection}}$, increases the domestic quantity supplied to QS_2 , and decreases the domestic quantity demanded to QD_2 . The difference is the new quantity imported. An equivalent quota will have the same effect, decreasing the quantity imported to $QD_2 - QS_2$.

The entire shaded area in Figure 14.3 represents the loss of consumer surplus in the domestic economy. The portion with vertical lines, the area to the left of the domestic supply curve between $P_{\text{protection}}$ and P_{world} , represents the gain in the producer surplus of domestic producers. The portion with horizontal lines, the area bounded by $QD_2 - QS_2$ and $P_{\text{protection}} - P_{\text{world}}$, represents the gain to the domestic government from tariff revenue. The two remaining triangular areas are the deadweight loss from the restriction on free trade.

Figure 14.3: Effects of Tariffs and Quotas



In the case of a quota, if the domestic government collects the full value of the import licenses, the result is the same as for a tariff. If the domestic government does not charge for the import licenses, this amount is a gain to those foreign exporters who receive the import licenses under the quota and are termed **quota rents**.

In terms of overall economic gains from trade, the deadweight loss is the amount of lost welfare from the imposition of the quota or tariff. From the viewpoint of the domestic country, the loss in consumer surplus is only partially offset by the gains in domestic producer surplus and the collection of tariff revenue.

If none of the quota rents are captured by the domestic government, the overall welfare loss to the domestic economy is greater by the amount of the quota rents. It is the entire difference between the gain in producer surplus and the loss of consumer surplus.

A **voluntary export restraint (VER)** is just as it sounds. It refers to a voluntary agreement by a government to limit the quantity of a good that can be exported. VERs are another way of protecting the domestic producers in the importing country. They result in a welfare loss to the importing country equal to that of an equivalent quota with no government charge for the import licenses; that is, no capture of the quota rents.

Export subsidies are payments by a government to its country's exporters. Export subsidies benefit producers (exporters) of the good but increase prices and reduce consumer surplus in the exporting country. In a small country, the price will increase by the amount of the subsidy to equal the world price plus the subsidy. In the case of a large exporter of the good, the world price decreases and some benefits from the subsidy accrue to foreign consumers, while foreign producers are negatively affected.

Most of the effects of all four of these protectionist policies are the same. With respect to the domestic (importing) country, import quotas, tariffs, and VERs all:

- Reduce imports.
- Increase price.
- Decrease consumer surplus.
- Increase domestic quantity supplied.
- Increase producer surplus.

With one exception, all will decrease national welfare. Quotas and tariffs in a large country could increase national welfare under a specific set of assumptions, primarily because for a country that imports a large amount of the good, setting a quota or tariff could reduce the world price for the good.

Capital Restrictions

Some countries impose **capital restrictions** on the flow of financial capital across borders. Restrictions include outright prohibition of investment in the domestic country by foreigners, prohibition of or taxes on the income earned on foreign investments by domestic citizens, prohibition of foreign investment in certain domestic industries, and restrictions on repatriation of earnings of foreign entities operating in a country.

Overall, capital restrictions are thought to decrease economic welfare. However, over the short term, they have helped developing countries avoid the impact of great inflows of foreign capital during periods of optimistic expansion and the impact of large outflows of foreign capital during periods of correction and market unease or outright panic. Even these short-term benefits may not offset longer-term costs if the country is excluded from international markets for financial capital flows.

LOS 14.f: Explain motivations for and advantages of trading blocs, common markets, and economic unions.

There are various types of agreements among countries with respect to trade policy. The essence of all of them is to reduce trade barriers among the countries. Reductions in trade restrictions among countries have some, by now familiar, positive and negative effects on

economic welfare. The positive effects result from increased trade according to comparative advantage, as well as increased competition among firms in member countries. The negative effects result because some firms, some industries, and some groups of workers will see their wealth and incomes decrease. Workers in affected industries may need to learn new skills to get new jobs.

On balance, economic welfare is improved by reducing or eliminating trade restrictions. Note, however, that to the extent that a trade agreement increases trade restrictions on imports from non-member countries, economic welfare gains are reduced and, in an extreme case, could be outweighed by the costs such restrictions impose. This could result if restrictions on trade with non-member countries increases a country's (unrestricted) imports from a member that has higher prices than the country's previous imports from a non-member.

We list these types of agreements, generally referred to as **trading blocs** or **regional trading agreements (RTA)**, in order of their degrees of integration.

Free Trade Areas

1. All barriers to import and export of goods and services among member countries are removed.

Customs Union

1. All barriers to import and export of goods and services among member countries are removed.
2. All countries adopt a common set of trade restrictions with non-members.

Common Market

1. All barriers to import and export of goods and services among the countries are removed.
2. All countries adopt a common set of trade restrictions with non-members.
3. All barriers to the movement of labor and capital goods among member countries are removed.

Economic Union

1. All barriers to import and export of goods and services among the countries are removed.
2. All countries adopt a common set of trade restrictions with non-members.
3. All barriers to the movement of labor and capital goods among member countries are removed.
4. Member countries establish common institutions and economic policy for the union.

Monetary Union

1. All barriers to import and export of goods and services among the countries are removed.
2. All countries adopt a common set of trade restrictions with non-members.

3. All barriers to the movement of labor and capital goods among member countries are removed.
4. Member countries establish common institutions and economic policy for the union.
5. Member countries adopt a single currency.

The North American Free Trade Agreement (NAFTA) is an example of a free trade area, the European Union (EU) is an example of an economic union, and the euro zone is an example of a monetary union.

LOS 14.g: Describe common objectives of capital restrictions imposed by governments.

Governments sometimes place restrictions on the flow of investment capital into their country, out of their country, or both. Commonly cited objectives of capital flow restrictions include the following:

- *Reduce the volatility of domestic asset prices.* In times of macroeconomic crisis, capital flows out of the country can drive down asset prices drastically, especially prices of liquid assets such as stocks and bonds. With no restrictions on inflows or outflows of foreign investment capital, the asset markets of countries with economies that are small relative to the amount of foreign investment can be quite volatile over a country's economic cycle.
- *Maintain fixed exchange rates.* For countries with fixed exchange rate targets, limiting flows of foreign investment capital makes it easier to meet the exchange rate target and, therefore, to be able to use monetary and fiscal policy to pursue only the economic goals for the domestic economy.
- *Keep domestic interest rates low.* By restricting the outflow of investment capital, countries can keep their domestic interest rates low and manage the domestic economy with monetary policy, as investors cannot pursue higher rates in foreign countries. China is an example of a country with a fixed exchange rate regime where restrictions on capital flows allow policymakers to maintain the target exchange rate as well as to pursue a monetary policy independent of concerns about its effect on currency exchange rates.
- *Protect strategic industries.* Governments sometimes prohibit investment by foreign entities in industries considered to be important for national security, such as the telecommunications and defense industries.

LOS 14.h: Describe the balance of payments accounts including their components.

When a country's firms and individuals pay for their purchases of foreign goods, services, and financial assets, they must buy the currencies of the foreign countries in order to accomplish those transactions. Similarly, payment for sales of goods, services, and financial assets to foreigners requires them to purchase the currency of the domestic country. With adjustment for changes in foreign debt to the domestic country and domestic debt to foreign countries, these amounts must balance each other.

According to the U.S. Federal Reserve, “The BOP [**balance of payments**] includes the **current account**, which mainly measures the flows of goods and services; the **capital account**, which consists of capital transfers and the acquisition and disposal of non-produced, non-financial assets; and the **financial account**, which records investment flows.”¹

Drawing on the N.Y. Fed’s explanation, the items recorded in each account are as follows.

Current Account

The current account comprises three sub-accounts:

- **Merchandise and services.** Merchandise consists of all raw materials and manufactured goods bought, sold, or given away. Services include tourism, transportation, and business and engineering services, as well as fees from patents and copyrights on new technology, software, books, and movies.
- **Income receipts** include foreign income from dividends on stock holdings and interest on debt securities.
- **Unilateral transfers** are one-way transfers of assets, such as money received from those working abroad and direct foreign aid. In the case of foreign aid and gifts, the capital account of the donor nation is debited.

Capital Account

The capital account comprises two sub-accounts:

- **Capital transfers** include debt forgiveness and goods and financial assets that migrants bring when they come to a country or take with them when they leave. Capital transfers also include the transfer of title to fixed assets and of funds linked to the purchase or sale of fixed assets, gift and inheritance taxes, death duties, and uninsured damage to fixed assets.
- **Sales and purchases of non-financial assets** that are not produced assets include rights to natural resources and intangible assets, such as patents, copyrights, trademarks, franchises, and leases.

Financial Account

The financial account comprises two sub-accounts:

- **Government-owned assets abroad** include gold, foreign currencies, foreign securities, reserve position in the International Monetary Fund, credits and other long-term assets, direct foreign investment, and claims against foreign banks.
- **Foreign-owned assets in the country** are divided into foreign official assets and other foreign assets in the domestic country. These assets include domestic government and corporate securities, direct investment in the domestic country, domestic country currency, and domestic liabilities to foreigners reported by domestic banks.

A country that has imports valued more than its exports is said to have a *current account (trade) deficit*, while countries with more exports than imports are said to have a *current account surplus*. For a country with a trade deficit, it must be balanced by a net surplus in the capital and financial accounts. As a result, investment analysts often think of all financing flows as a single capital account that combines items in the capital and financial accounts. Thinking in this way, any deficit in the current account must be made up by a surplus in the combined

capital account. That is, the excess of imports over exports must be offset by sales of assets and debt incurred to foreign entities. A current account surplus is similarly offset by purchases of foreign physical or financial assets.

LOS 14.i: Explain how decisions by consumers, firms, and governments affect the balance of payments.

The primary influences referred to here are on the current account deficit or surplus. If a country's net savings (both government savings and private savings) are less than the amount of investment in domestic capital, this investment must be financed by foreign borrowing. Foreign borrowing results in a capital account surplus, which means there is a trade deficit.

We can write the relation between the trade deficit, saving, and domestic investment as:

$$X - M = \text{private savings} + \text{government savings} - \text{investment}$$

Lower levels of private saving, larger government deficits, and high rates of domestic investment all tend to result in or increase a current account deficit. The intuition here is that low private or government savings in relation to private investment in domestic capital requires foreign investment in domestic capital.

We can make a distinction, however, between a trade deficit resulting from high government or private consumption and one resulting from high private investment in capital. In the first case, borrowing from foreign countries to finance high consumption (low savings) increases the domestic country's liabilities without any increase to its future productive power. In the second case, borrowing from foreign countries to finance a high level of private investment in domestic capital, the added liability is accompanied by an increase in future productive power because of the investment in capital.

LOS 14.j: Describe functions and objectives of the international organizations that facilitate trade, including the World Bank, the International Monetary Fund, and the World Trade Organization.

Perhaps the best way to understand the roles of the organizations designed to facilitate trade is to examine their own statements.

According to the **International Monetary Fund** (IMF; more available at www.IMF.org):

Article I of the Articles of Agreement sets out the IMF's main goals:

- promoting international monetary cooperation;
- facilitating the expansion and balanced growth of international trade;
- promoting exchange stability;
- assisting in the establishment of a multilateral system of payments; and
- making resources available (with adequate safeguards) to members experiencing balance of payments difficulties.

According to the **World Bank** (more available at www.WorldBank.org):

The World Bank is a vital source of financial and technical assistance to developing countries around the world. Our mission is to fight poverty with passion and professionalism for lasting results and to help people help themselves and their environment by providing resources, sharing knowledge, building capacity and forging partnerships in the public and private sectors.

We are not a bank in the common sense; we are made up of two unique development institutions owned by 187 member countries: the International Bank for Reconstruction and Development (IBRD) and the International Development Association (IDA).

Each institution plays a different but collaborative role in advancing the vision of inclusive and sustainable globalization. The IBRD aims to reduce poverty in middle-income and creditworthy poorer countries, while IDA focuses on the world's poorest countries.

...Together, we provide low-interest loans, interest-free credits and grants to developing countries for a wide array of purposes that include investments in education, health, public administration, infrastructure, financial and private sector development, agriculture and environmental and natural resource management.

According to the **World Trade Organization** (WTO; more available at www.WTO.org):

The World Trade Organization (WTO) is the only international organization dealing with the global rules of trade between nations. Its main function is to ensure that trade flows as smoothly, predictably and freely as possible.

...Trade friction is channeled into the WTO's dispute settlement process where the focus is on interpreting agreements and commitments, and how to ensure that countries' trade policies conform with them. That way, the risk of disputes spilling over into political or military conflict is reduced.

...At the heart of the system—known as the multilateral trading system—are the WTO's agreements, negotiated and signed by a large majority of the world's trading nations, and ratified in their parliaments. These agreements are the legal ground-rules for international commerce. Essentially, they are contracts, guaranteeing member countries important trade rights. They also bind governments to keep their trade policies within agreed limits to everybody's benefit.



MODULE QUIZ 14.2

1. An agreement with another country to limit the volume of goods and services sold to them is *best* described as:
 - A. a quota.
 - B. a voluntary export restraint.
 - C. a minimum domestic content rule.
2. Which of the following groups would be *most likely* to suffer losses from the imposition of a tariff on steel imports?
 - A. Domestic steel producers.
 - B. Workers in the domestic auto industry.
 - C. Workers in the domestic steel industry.
3. The *most likely* motivation for establishing a trading bloc is to:
 - A. increase economic welfare in the member countries.
 - B. increase tariff revenue for the member governments.
 - C. protect domestic industries in the member economies.
4. In which type of regional trade agreement are economic policies conducted independently by the member countries, while labor and capital are free to move among member countries?
 - A. Free trade area.
 - B. Common market.
 - C. Economic union.
5. The goal of a government that imposes restrictions on foreign capital flows is *most likely* to:
 - A. stimulate domestic interest rates.

- B. decrease domestic asset price volatility.
 - C. encourage competition with domestic industries.
6. Which of the following is *least likely* a component of the current account?
- A. Unilateral transfers.
 - B. Payments for fixed assets.
 - C. Payments for goods and services.
7. A current account deficit is *most likely* to decrease as a result of an increase in:
- A. domestic savings.
 - B. private investment.
 - C. the fiscal budget deficit.
8. Which international organization is primarily concerned with providing economic assistance to developing countries?
- A. World Bank.
 - B. World Trade Organization.
 - C. International Monetary Fund.

KEY CONCEPTS

LOS 14.a

Gross domestic product is the total value of goods and services produced within a country's borders. Gross national product measures the total value of goods and services produced by the labor and capital supplied by a country's citizens, regardless of where the production takes place.

LOS 14.b

Free trade among countries increases overall economic welfare. Countries can benefit from trade because one country can specialize in the production of an export good and benefit from economies of scale. Economic welfare can also be increased by greater product variety, more competition, and a more efficient allocation of resources.

Costs of free trade are primarily losses to those in domestic industries that lose business to foreign competition, especially less efficient producers who leave an industry. While other domestic industries will benefit from freer trade policies, unemployment may increase over the period in which workers are retrained for jobs in the expanding industries. Some argue that greater income inequality may result, but overall the gains from liberalization of trade policies are thought to exceed the costs, so that the winners could conceivably compensate the losers and still be better off.

LOS 14.c

A country is said to have an absolute advantage in the production of a good if it can produce the good at lower cost in terms of resources relative to another country.

A country is said to have a comparative advantage in the production of a good if its opportunity cost in terms of other goods that could be produced instead is lower than that of another country.

LOS 14.d

The Ricardian model of trade has only one factor of production—labor. The source of differences in production costs and comparative advantage in Ricardo's model is differences in labor productivity due to differences in technology.

Heckscher and Ohlin presented a model in which there are two factors of production—capital and labor. The source of comparative advantage (differences in opportunity costs) in this model is differences in the relative amounts of each factor that countries possess.

LOS 14.e

Types of trade restrictions include:

- Tariffs: Taxes on imported good collected by the government.
- Quotas: Limits on the amount of imports allowed over some period.
- Minimum domestic content: Requirement that some percentage of product content must be from the domestic country.
- Voluntary export restraints: A country voluntarily restricts the amount of a good that can be exported, often in the hope of avoiding tariffs or quotas imposed by their trading partners.

Within each importing country, all of these restrictions will tend to:

- Increase prices of imports and decrease quantities of imports.
- Increase demand for and quantity supplied of domestically produced goods.
- Increase producer's surplus and decrease consumer surplus.

Export subsidies decrease export prices and benefit importing countries at the expense of the government of the exporting country.

Restrictions on the flow of financial capital across borders include outright prohibition of investment in the domestic country by foreigners, prohibition of or taxes on the income earned on foreign investments by domestic citizens, prohibition of foreign investment in certain domestic industries, and restrictions on repatriation of earnings of foreign entities operating in a country.

LOS 14.f

Trade agreements, which increase economic welfare by facilitating trade among member countries, take the following forms:

- Free trade area: All barriers to the import and export of goods and services among member countries are removed.
- Customs union: Member countries *also* adopt a common set of trade restrictions with non-members.
- Common market: Member countries *also* remove all barriers to the movement of labor and capital goods among members.
- Economic union: Member countries *also* establish common institutions and economic policy for the union.
- Monetary union: Member countries *also* adopt a single currency.

LOS 14.g

Commonly cited objectives of capital flow restrictions include:

- Reducing the volatility of domestic asset prices.
- Maintaining fixed exchange rates.

- Keeping domestic interest rates low and enabling greater independence regarding monetary policy.
- Protecting strategic industries from foreign ownership.

LOS 14.h

The balance of payments refers to the fact that increases in a country's assets and decreases in its liabilities must equal (balance with) decreases in its assets and increases in its liabilities. These financial flows are classified into three types:

- The current account includes imports and exports of merchandise and services, foreign income from dividends on stock holdings and interest on debt securities, and unilateral transfers such as money received from those working abroad and direct foreign aid.
- The capital account includes debt forgiveness, assets that migrants bring to or take away from a country, transfer of funds for the purchase or sale of fixed assets, and purchases of non-financial assets, including rights to natural resources, patents, copyrights, trademarks, franchises, and leases.
- The financial account includes government-owned assets abroad such as gold, foreign currencies and securities, and direct foreign investment and claims against foreign banks. The financial account also includes foreign-owned assets in the country, domestic government and corporate securities, direct investment in the domestic country, and domestic country currency.

Overall, any surplus (deficit) in the current account must be offset by a deficit (surplus) in the capital and financial accounts.

LOS 14.i

In equilibrium, we have the relationship:

$$\text{exports} - \text{imports} = \text{private savings} + \text{government savings} - \text{domestic investment}$$

When total savings is less than domestic investment, exports must be less than imports so that there is a deficit in the current account. Lower levels of private saving, larger government deficits, and high rates of domestic investment all tend to result in or increase a current account deficit. The intuition here is that low private or government savings in relation to private investment in domestic capital requires foreign investment in domestic capital.

LOS 14.j

The International Monetary Fund facilitates trade by promoting international monetary cooperation and exchange rate stability, assists in setting up international payments systems, and makes resources available to member countries with balance of payments problems.

The World Bank provides low-interest loans, interest-free credits, and grants to developing countries for many specific purposes. It also provides resources and knowledge and helps form private/public partnerships with the overall goal of fighting poverty.

The World Trade Organization has the goal of ensuring that trade flows freely and works smoothly. Its main focus is on instituting, interpreting, and enforcing a number of multilateral trade agreements that detail global trade policies for a large majority of the world's trading nations.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 14.1

1. **A** The income from a financial investment in Country P of a citizen of Country Q is included in Country P's GDP but not its GNP. It is included in Country Q's GNP but not its GDP. (LOS 14.a)
2. **B** Openness to international trade increases specialization as production shifts to those products in which domestic producers have a comparative advantage. Greater competition from imports will tend to decrease prices for consumer goods. Increasing international trade is likely to increase profitability and employment in exporting industries but may decrease profitability and employment in industries that compete with imported goods. (LOS 14.b)
3. **A** Each country gains by *exporting* the good for which it has a comparative advantage. (LOS 14.c)
4. **C** In the Ricardian model, labor is the only factor of production considered. In the Heckscher-Ohlin model, comparative advantage results from the relative amounts of labor and capital available in different countries. (LOS 14.d)

Module Quiz 14.2

1. **B** Voluntary export restraints are agreements to limit the volume of goods and services exported to another country. Minimum domestic content rules are limitations imposed by a government on its domestic firms. Import quotas are limitations on imports, not on exports. (LOS 14.e)
2. **B** Imposing a tariff on steel imports benefits domestic steel producers and workers by increasing the domestic price of steel and benefits the national government by increasing tax (tariff) revenue. However, the increase in the domestic price of steel would increase costs in industries that use significant amounts of steel, such as the automobile industry. The resulting increase in the price of automobiles reduces the quantity of automobiles demanded and ultimately reduces employment in that industry. (LOS 14.e)
3. **A** The motivation for trading blocs is to increase economic welfare in the member countries by eliminating barriers to trade. Joining a trading bloc may have negative consequences for some domestic industries and may decrease tariff revenue for the government. (LOS 14.f)
4. **B** These characteristics describe a common market. In a free trade area, member countries remove restrictions on goods and services trade with one another but may still restrict movement of labor and capital among member countries. In an economic union, member countries also coordinate their economic policies and institutions. (LOS 14.f)
5. **B** Decreasing the volatility of domestic asset prices may be a goal of a government that imposes capital restrictions. Other typical goals include keeping domestic interest rates low and protecting certain domestic industries, such as the defense industry. (LOS 14.g)
6. **B** Purchases and sales of fixed assets are recorded in the capital account. Goods and services trade and unilateral transfers are components of the current account. (LOS 14.h)
7. **A** Other things equal, an increase in domestic savings would tend to decrease the current account deficit, while an increase in private investment or an increase in the fiscal budget deficit would tend to increase the current account deficit. (LOS 14.i)
8. **A** The World Bank provides technical and financial assistance to economically developing countries. The World Trade Organization is primarily concerned with settling disputes among countries concerning international trade. The International Monetary Fund promotes international trade and exchange rate stability and assists member countries that experience balance of payments trouble. (LOS 14.j)

READING 15

CURRENCY EXCHANGE RATES

EXAM FOCUS

Candidates must understand spot exchange rates, forward exchange rates, and all the calculations having to do with currency appreciation and depreciation. Additionally, candidates should understand the steps a country can take to decrease a trade deficit and the requirements for these to be effective under both the elasticities and absorption approaches. Finally, candidates should make sure to know the terms for and definitions of the various exchange rate regimes countries may adopt.

MODULE 15.1: FOREIGN EXCHANGE RATES



LOS 15.a: Define an exchange rate and distinguish between nominal and real exchange rates and spot and forward exchange rates.

Video covering this content is available online.

An **exchange rate** is simply the price or cost of units of one currency in terms of another. For the purposes of this book we will write 1.416 USD/EUR to mean that each euro costs \$1.416. If you read the “/” as *per*, you will have no trouble with the notation. We say the exchange rate is \$1.416 per euro.



PROFESSOR'S NOTE

There are alternative notations for foreign exchange quotes, but expressing them as the price of the denominator currency in terms of the numerator currency is what we will use and what you can expect on the Level I exam.

In a foreign currency quotation we have the price of one currency in units of another currency. These are often referred to as the **base currency** and the **price currency**. In the quotation 1.25 USD/EUR, the USD is the price currency and the EUR is the base currency. The price of one euro (base currency) is 1.25 USD (the price currency) so 1.25 is the price of one unit of the base currency in terms of the other. It may help to remember that the euro in this example is in the bottom or “base” of the exchange rate given in terms of USD/EUR.

Sometimes an exchange rate expressed as price currency/base currency is referred to as a **direct quote** from the point of view of an investor in the price currency country and an **indirect quote** from the point of view of an investor in the base currency country. For example, a quote of 1.17 USD/EUR would be a direct exchange rate quote for a USD-based investor and an indirect quote for a EUR-based investor. Conversely, a quote of $1 / 1.17 = 0.845$ EUR/USD would

be a direct exchange rate quote for a EUR-based investor and an indirect quote for a USD-based investor.

The exchange rate at a point in time is referred to as a **nominal exchange rate**. If the nominal exchange rate (price/base) increases, the cost of a unit of the base currency in terms of the price currency has increased, so that the purchasing power of the price currency has decreased. If the USD/EUR exchange rate increases from 1.10 to 1.15, the cost of 100 euros increases from \$110 to \$115. The purchasing power of the dollar has decreased relative to the euro because the cost of 100 euros worth of goods to a consumer in the U.S. has increased over the period.

The purchasing power of one currency relative to another is also affected by changes in the price levels of the two countries. The **real exchange rate** between two currencies refers to the purchasing power of one currency in terms of the amount of goods priced in another currency, relative to an earlier (base) period.

Consider a situation in which the nominal USD/EUR exchange rate is unchanged at 1:1 over a period and the price level in the U.S. is unchanged, while prices in the Eurozone have increased by 5%. Eurozone goods that cost 100 euros at the beginning of the period cost 105 euros at the end of the period. With the nominal exchange rate unchanged, the purchasing power of the USD in the Eurozone has decreased, because exchanging 100 USD for 100 EUR will now buy only $100/105 = 95.2\%$ of the goods 100 EUR could buy at the beginning of the period.

To summarize:

- An increase in the *nominal* USD/EUR rate decreases the purchasing power of the USD in the Eurozone (and increases the purchasing power of the EUR in the U.S.); the *real* USD/EUR exchange rate has increased.
- A decrease in the *nominal* USD/EUR rate increases the purchasing power of the USD in the Eurozone (and decreases the purchasing power of the EUR in the U.S.); the *real* USD/EUR exchange rate has decreased.
- An increase in the Eurozone price level, relative to the price level in the U.S., will increase the *real* USD/EUR exchange rate, decreasing the purchasing power of the USD in the Eurozone (and increasing the purchasing power of the EUR in the U.S.).
- A decrease in the Eurozone price level, relative to the price level in the U.S., will decrease the *real* USD/EUR exchange rate, increasing the purchasing power of the USD in the Eurozone (and decreasing the purchasing power of the EUR in the U.S.).

The end-of-period real P/B exchange rate can be calculated as:

$$\text{real P/B exchange rate} = \text{nominal P/B exchange rate} \times \frac{\text{CPI}_{\text{base currency}}}{\text{CPI}_{\text{price currency}}}$$

where the CPI values are relative to base period values of 100.

We can see from the formula that:

- An increase (decrease) in the nominal exchange rate over the period increases (decreases) the end-of-period real exchange rate and the purchasing power of the price currency decreases (increases).
- An increase in the price level in the price currency country relative to the price level in the base currency country will decrease the real exchange rate, increasing the purchasing power

- of the price currency in terms of base country goods.
- Conversely, a decrease in the price level in the price currency country relative to the price level in the base currency country will increase the real exchange rate, decreasing the purchasing power of the price currency in terms of base country goods.

In the following example we calculate the end-of-period real \$/£ exchange rate when the nominal \$/£ exchange rate has decreased over the period, (which tends to decrease the real exchange rate and increase the purchasing power of the price currency), and when the price level in the U.K. has increased by more than the price level in the U.S. over the period (which tends to increase the real exchange rate and decrease the purchasing power of the price currency). The relative increase in U.K. prices has reduced the effects of the decrease in the nominal exchange rate on the increase in the purchasing power of the USD.

EXAMPLE: Real exchange rate

At a base period, the CPIs of the U.S. and U.K. are both 100, and the exchange rate is \$1.70/£. Three years later, the exchange rate is \$1.60/£, and the CPI has risen to 110 in the United States and 112 in the U.K. What is the real exchange rate at the end of the three-year period?

Answer:

The real exchange rate is $\$1.60/\text{£} \times 112 / 110 = \$1.629/\text{£}$ which means that U.S. goods and services that cost \$1.70 at the base period now cost only \$1.629 (in real terms) if purchased in the U.K. and the real exchange rate, \$/£, has fallen. The decrease in the real exchange rate (and the increase in the purchasing power of the USD in terms of U.K. goods) over the period is less than it would have been if the relative prices between the two countries had not changed.

A **spot exchange rate** is the currency exchange rate for immediate delivery, which for most currencies means the exchange of currencies takes place two days after the trade.

A **forward exchange rate** is a currency exchange rate for an exchange to be done in the future. Forward rates are quoted for various future dates (e.g., 30 days, 60 days, 90 days, or one year). A forward is actually an agreement to exchange a specific amount of one currency for a specific amount of another on a future date specified in the forward agreement.

A French firm that will receive 10 million GBP from a British firm six months from now has uncertainty about the amount of euros that payment will be equivalent to six months from now. By entering into a forward agreement covering 10 million GBP at the 6-month forward rate of 1.192 EUR/GBP, the French firm has agreed to exchange 10 million GBP for 11.92 million euros in six months.

LOS 15.b: Calculate and interpret the percentage change in a currency relative to another currency.

Consider a USD/EUR exchange rate that has changed from 1.42 to 1.39 USD/EUR. The percentage change in the dollar price of a euro is simply $1.39 / 1.42 - 1 = -0.0211 = -2.11\%$. Because the dollar price of a euro has fallen, the euro has *depreciated* relative to the dollar, and a

euro now buys 2.11% fewer U.S. dollars. It is correct to say that the euro has depreciated by 2.11% relative to the dollar.

On the other hand, it is *not* correct to say that the dollar has appreciated by 2.11%. To calculate the percentage appreciation of the dollar, we need to convert the quotes to EUR/USD. So our beginning quote of 1.42 USD/EUR becomes $1 / 1.42 = 0.7042$ EUR/USD, and our ending quote of 1.39 USD/EUR becomes $1 / 1.39 = 0.7194$ EUR/USD. Using these exchange rates, we can calculate the change in the euro price of a dollar as $0.7194 / 0.7042 - 1 = 0.0216 = 2.16\%$. In this case, it is correct to say that the dollar has appreciated 2.16% with respect to the euro. For the same quotes, the percentage appreciation of the dollar is not the same as the percentage depreciation in the euro.

The key point to remember is that we can correctly calculate the percentage change of the *base currency* in a foreign exchange quotation.

LOS 15.c: Describe functions of and participants in the foreign exchange market.

Foreign currency markets serve companies and individuals that purchase or sell foreign goods and services denominated in foreign currencies. An even larger market, however, exists for capital flows. Foreign currencies are needed to purchase foreign physical assets as well as foreign financial securities.

Many companies have foreign exchange risk arising from their cross-border transactions. A Japanese company that expects to receive 10 million euros when a transaction is completed in 90 days has yen/euro exchange rate risk as a result. By entering into a **forward currency contract** to sell 10 million euros in 90 days for a specific quantity of yen, the firm can reduce or eliminate its foreign exchange risk associated with the transaction. When a firm takes a position in the foreign exchange market to reduce an existing risk, we say the firm is **hedging** its risk.

Alternatively, when a transaction in the foreign exchange markets increases currency risk, we term it a **speculative** transaction or position. Investors, companies, and financial institutions, such as banks and investment funds, all regularly enter into speculative foreign currency transactions.

The primary dealers in currencies and originators of forward foreign exchange (FX) contracts are large multinational banks. This part of the FX market is often called the **sell side**. On the other hand, the **buy side** consists of the many buyers of foreign currencies and forward FX contracts. These buyers include the following:

- **Corporations** regularly engage in cross-border transactions, purchase and sell foreign currencies as a result, and enter into FX forward contracts to hedge the risk of expected future receipts and payments denominated in foreign currencies.
- **Investment accounts** of many types transact in foreign currencies, hold foreign securities, and may both speculate and hedge with currency derivatives. **Real money accounts** refer to mutual funds, pension funds, insurance companies, and other institutional accounts that do not use derivatives. **Leveraged accounts** refer to the various types of investment firms that do use derivatives, including hedge funds, firms that trade for their own accounts, and other trading firms of various types.

- **Governments** and **government entities**, including **sovereign wealth funds** and pension funds, acquire foreign exchange for transactional needs, investment, or speculation. Central banks sometimes engage in FX transactions to affect exchange rates in the short term in accordance with government policy.
- The **retail market** refers to FX transactions by households and relatively small institutions and may be for tourism, cross-border investment, or speculative trading.

LOS 15.d: Calculate and interpret currency cross-rates.

The **cross rate** is the exchange rate between two currencies implied by their exchange rates with a common third currency. Cross rates are necessary when there is no active FX market in the currency pair. The rate must be computed from the exchange rates between each of these two currencies and a third currency, usually the USD or EUR.

Let's assume that we have the following quotations for Mexican pesos and Australian dollars: MXN/USD = 10.70 and USD/AUD = 0.60. The cross rate between Australian dollars and pesos (MXN/AUD) is:

$$\text{MXN/AUD} = \text{USD/AUD} \times \text{MXN/USD} = 0.60 \times 10.70 = 6.42$$

So our MXN/AUD cross rate is 6.42 pesos per Australian dollar. The key to calculating cross rates is to note that the basis of the quotations must be such that we get the desired result algebraically. If we had started with an AUD/USD quotation of 1.67, we would have taken the inverse to get the quotation into USD/AUD terms. Another approach is to divide through, as is illustrated in the following example.

EXAMPLE: Cross rate calculation

The spot exchange rate between the Swiss franc (CHF) and the USD is CHF/USD = 1.7799, and the spot exchange rate between the New Zealand dollar (NZD) and the U.S. dollar is NZD/USD = 2.2529. Calculate the CHF/NZD spot rate.

Answer:

The CHF/NZD cross rate is:

$$(\text{CHF/USD}) / (\text{NZD/USD}) = 1.7799 / 2.2529 = 0.7900$$



MODULE QUIZ 15.1

1. One year ago, the nominal exchange rate for USD/EUR was 1.300. Since then, the real exchange rate has increased by 3%. This *most likely* implies that:
 - A. the nominal exchange rate is less than USD/EUR 1.235.
 - B. the purchasing power of the euro has increased approximately 3% in terms of U.S. goods.
 - C. inflation in the euro zone was approximately 3% higher than inflation in the United States.
2. Sell-side participants in the foreign exchange market are *most likely* to include:
 - A. banks.
 - B. hedge funds.
 - C. insurance companies.

3. Suppose that the quote for British pounds (GBP) in New York is USD/GBP 1.3110. What is the quote for U.S. dollars (USD) in London (GBP/USD)?
 - A. 0.3110.
 - B. 0.7628.
 - C. 1.3110.
4. The Canadian dollar (CAD) exchange rate with the Japanese yen (JPY) changes from JPY/CAD 75 to JPY/CAD 78. The CAD has:
 - A. depreciated by 3.8%, and the JPY has appreciated by 4.0%.
 - B. appreciated by 3.8%, and the JPY has depreciated by 4.0%.
 - C. appreciated by 4.0%, and the JPY has depreciated by 3.8%.
5. Today's spot rate for the Indonesian rupiah (IDR) is IDR/USD 2,400.00, and the New Zealand dollar trades at NZD/USD 1.6000. The NZD/IDR cross rate is:
 - A. 0.00067.
 - B. 1,492.53.
 - C. 3,840.00.
6. The NZD is trading at USD/NZD 0.3500, and the SEK is trading at NZD/SEK 0.3100. The USD/SEK cross rate is:
 - A. 0.1085.
 - B. 8.8573.
 - C. 9.2166.

MODULE 15.2: FORWARD EXCHANGE RATES



LOS 15.e: Calculate an outright forward quotation from forward quotations expressed on a points basis or in percentage terms.

Video covering this content is available online.

A forward exchange rate quote typically differs from the spot quotation and is expressed in terms of the difference between the spot exchange rate and the forward exchange rate. One way to indicate this is with points. The unit of points is the last decimal place in the spot rate quote. For a spot currency quote to four decimal places, such as 2.3481, each point is 0.0001 or 1/10,000th. A quote of +18.3 points for a 90-day forward exchange rate means that the forward rate is 0.00183 more than the spot exchange rate.

EXAMPLE: Forward exchange rates in points

The AUD/EUR spot exchange rate is 0.7313 with the 1-year forward rate quoted at +3.5 points. What is the 1-year forward AUD/EUR exchange rate?

Answer:

The forward exchange rate is $0.7313 + 0.00035 = 0.73165$.

EXAMPLE: Forward exchange rates in percent

The AUD/EUR spot rate is quoted at 0.7313, and the 120-day forward exchange rate is given as -0.062%. What is the 120-day forward AUD/EUR exchange rate?

Answer:

The forward exchange rate is $0.7313 (1 - 0.00062) = 0.7308$.

LOS 15.f: Explain the arbitrage relationship between spot rates, forward rates, and interest rates.

When currencies are freely traded and forward currency contracts exist, the percentage difference between forward and spot exchange rates is approximately equal to the difference between the two countries' interest rates. This is because there is an arbitrage trade with a riskless profit to be made when this relation does not hold.

We call this a no-arbitrage condition because if it doesn't hold there is an opportunity to make a profit without risk. The possible arbitrage is as follows: borrow Currency A at interest rate A, convert it to Currency B at the spot rate and invest it to earn interest rate B, and sell the proceeds from this investment forward at the forward rate to turn it back into Currency A. If the forward rate does not correctly reflect the difference between interest rates, such an arbitrage could generate a profit to the extent that the return from investing Currency B and converting it back to Currency A with a forward contract is greater than the cost of borrowing Currency A for the period. We consider a numerical analysis of such an arbitrage later in this reading.

For spot and forward rates expressed as price currency/base currency, the no-arbitrage relation (commonly referred to as *interest rate parity*) is:

$$\frac{\text{forward}}{\text{spot}} = \frac{(1 + \text{interest rate}_{\text{price currency}})}{(1 + \text{interest rate}_{\text{base currency}})}$$

This formula can be rearranged as necessary in order to solve for specific values of the relevant terms.

LOS 15.g: Calculate and interpret a forward discount or premium.

The **forward discount** or **forward premium** for a currency is calculated relative to the spot exchange rate. The forward discount or premium *for the base currency* is the percentage difference between the forward price and the spot price.

Consider the following spot and forward exchange rates as the price in U.S. dollars of one euro.

$$\text{USD/EUR spot} = \$1.312 \quad \text{USD/EUR 90-day forward} = \$1.320$$

The (90-day) forward premium or discount on the euro = forward/spot - 1 = $1.320 / 1.312 - 1 = 0.609\%$. Because this is positive, it is interpreted as a forward premium on the euro of 0.609%. Since we have the forward rate for 3 months, we could annualize the discount simply by multiplying by 4 ($= 12 / 3$).

Because the forward quote is greater than the spot quote, it will take more dollars to buy one euro 90 days from now, so the euro is expected to appreciate versus the dollar, and the dollar is expected to depreciate relative to the euro.

If the forward quote were less than the spot quote, the calculated amount would be negative and we would interpret that as a forward discount for the euro relative to the U.S. dollar.

LOS 15.h: Calculate and interpret the forward rate consistent with the spot rate and the interest rate in each currency.

EXAMPLE: Calculating the arbitrage-free forward exchange rate

Consider two currencies, the ABE and the DUB. The spot ABE/DUB exchange rate is 4.5671, the 1-year riskless ABE rate is 5%, and the 1-year riskless DUB rate is 3%. What is the 1-year forward exchange rate that will prevent arbitrage profits?

Answer:

Rearranging our formula, we have:

$$\text{forward} = \text{spot} \left(\frac{1 + I_{\text{ABE}}}{1 + I_{\text{DUB}}} \right) \text{ and we can calculate the forward rate as}$$

$$\text{forward} = 4.5671 \left(\frac{1.05}{1.03} \right) = 4.6558 \text{ ABE/DUB}$$

Note that the forward rate is greater than the spot rate by $4.6558 / 4.5671 - 1 = 1.94\%$. This is approximately equal to the interest rate differential of $5\% - 3\% = 2\%$. The currency with the higher interest rate should depreciate over time by approximately the amount of the interest rate differential.

If we are calculating a 90-day or 180-day forward exchange rate, we need to use interest rates for 90-day or 180-day periods rather than annual rates. Note that these shorter-term rates are quoted as annualized money market yields.

EXAMPLE: Calculating the arbitrage-free forward exchange rate with 90-day interest rates

The spot ABE/DUB exchange rate is 4.5671, the 90-day riskless ABE rate is 5%, and the 90-day riskless DUB rate is 3%. What is the 90-day forward exchange rate that will prevent arbitrage profits?

Answer:

$$\text{forward} = 4.5671 \left[\frac{1 + \frac{0.05}{4}}{1 + \frac{0.03}{4}} \right] = 4.5671 \left(\frac{1.0125}{1.0075} \right) = 4.5898 \text{ ABE/DUB}$$

In our previous example, we calculated the no-arbitrage one-year forward ABE/DUB exchange rate as 4.6558. Here, we illustrate the arbitrage profit that could be gained if the forward exchange rate differs from this no-arbitrage rate. Consider a forward rate of 4.6000 so that the depreciation in the ABE is less than that implied by interest rate parity. This makes the ABE attractive to a DUB investor who can earn a riskless profit as follows:

- Borrow 1,000 DUB for one year at 3% to purchase ABE and get 4,567.1 ABE.
- Invest the 4,567.1 ABE at the ABE rate of 5% to have $1.05(4,567.1) = 4,795.45$ ABE at the end of one year.
- Enter into a currency forward contract to exchange 4,795.45 ABE in one year at the forward rate of 4.6000 ABE/DUB in order to receive $4,795.45 / 4.6000 = 1,042.49$ DUB.

The investor has ended the year with a 4.249% return on his 1,000 DUB investment, which is higher than the 3% 1-year DUB interest rate. After repaying the 1,000 DUB loan plus interest (1,030 DUB), the investor has a profit of $1,042.49 - 1,030 = 12.49$ DUB with no risk and no initial out-of-pocket investment (i.e., a pure arbitrage profit).

Arbitrageurs will pursue this opportunity, buying ABE (driving down the spot ABE/DUB exchange rate) and selling ABE forward (driving up the forward ABE/DUB exchange rate), until the interest rate parity relation is restored and arbitrage profits are no longer available.



MODULE QUIZ 15.2

1. The spot CHF/GBP exchange rate is 1.3050. In the 180-day forward market, the CHF/GBP exchange rate is –42.5 points. The 180-day forward CHF/GBP exchange rate is *closest* to:
 - A. 1.2625.
 - B. 1.3008.
 - C. 1.3093.
2. The spot rate on the New Zealand dollar (NZD) is NZD/USD 1.4286, and the 180-day forward rate is NZD/USD 1.3889. This difference means:
 - A. interest rates are lower in the United States than in New Zealand.
 - B. interest rates are higher in the United States than in New Zealand.
 - C. it takes more NZD to buy one USD in the forward market than in the spot market.
3. The current spot rate for the British pound in terms of U.S. dollars is \$1.533 and the 180-day forward rate is \$1.508. Relative to the pound, the dollar is trading *closest* to a 180-day forward:
 - A. discount of 1.63%.
 - B. premium of 1.66%.
 - C. discount of 1.66%.
4. The annual interest rates in the United States (USD) and Sweden (SEK) are 4% and 7% per year, respectively. If the current spot rate is SEK/USD 9.5238, then the 1-year forward rate in SEK/USD is:
 - A. 9.2568.
 - B. 9.7985.
 - C. 10.2884.
5. The annual risk-free interest rate is 10% in the United States (USD) and 4% in Switzerland (CHF), and the 1-year forward rate is USD/CHF 0.80. Today's USD/CHF spot rate is *closest* to:
 - A. 0.7564.
 - B. 0.8462.
 - C. 0.8888.

MODULE 15.3: MANAGING EXCHANGE RATES



LOS 15.i: Describe exchange rate regimes.

Video covering
this content is
available online.

The IMF categorizes **exchange rate regimes** into the following types, two for countries that do not issue their own currencies and seven for countries that issue their own currencies.

Countries That Do Not Have Their Own Currency

- A country can use the currency of another country (**formal dollarization**). The country cannot have its own monetary policy, as it does not create money/currency.
- A country can be a member of a **monetary union** in which several countries use a common currency. Within the European Union, for example, most countries use the euro. While individual countries give up the ability to set domestic monetary policy, they all participate in determining the monetary policy of the European Central Bank.

Countries That Have Their Own Currency

- A **currency board arrangement** is an explicit commitment to exchange domestic currency for a specified foreign currency at a fixed exchange rate. A notable example of such an arrangement is Hong Kong. In Hong Kong, currency is (and may be) only issued when fully backed by holdings of an equivalent amount of U.S. dollars. The Hong Kong Monetary Authority can earn interest on its U.S. dollar balances. With dollarization, there is no such income, as the income is earned by the U.S. Federal Reserve when it buys interest-bearing assets with the U.S. currency it issues. While the monetary authority gives up the ability to conduct independent monetary policy and essentially imports the inflation rate of the outside currency, there may be some latitude to affect interest rates over the short term.
- In a **conventional fixed peg arrangement**, a country pegs its currency within margins of $\pm 1\%$ versus another currency or a basket that includes the currencies of its major trading or financial partners. The monetary authority can maintain exchange rates within the band by purchasing or selling foreign currencies in the foreign exchange markets (*direct intervention*). In addition, the country can use *indirect intervention*, including changes in interest rate policy, regulation of foreign exchange transactions, and convincing people to constrain foreign exchange activity. The monetary authority retains more flexibility to conduct monetary policy than with dollarization, a monetary union, or a currency board. However, changes in policy are constrained by the requirements of the peg.
- In a system of pegged exchange rates within horizontal bands or a **target zone**, the permitted fluctuations in currency value relative to another currency or basket of currencies are wider (e.g., $\pm 2\%$). Compared to a conventional peg, the monetary authority has more policy discretion because the bands are wider.
- With a **crawling peg**, the exchange rate is adjusted periodically, typically to adjust for higher inflation versus the currency used in the peg. This is termed a *passive crawling peg*, as opposed to an *active crawling peg* in which a series of exchange rate adjustments over time is announced and implemented. An active crawling peg can influence inflation expectations, adding some predictability to domestic inflation. Monetary policy is restricted in much the same way it is with a fixed peg arrangement.
- With **management of exchange rates within crawling bands**, the width of the bands that identify permissible exchange rates is increased over time. This method can be used to transition from a fixed peg to a floating rate when the monetary authority's lack of credibility

makes an immediate change to floating rates impractical. Again, the degree of monetary policy flexibility increases with the width of the bands.

- With a system of **managed floating exchange rates**, the monetary authority attempts to influence the exchange rate in response to specific indicators such as the balance of payments, inflation rates, or employment without any specific target exchange rate or predetermined exchange rate path. Intervention may be direct or indirect. Such management of exchange rates may induce trading partners to respond in ways that reduce stability.
 - When a currency is **independently floating**, the exchange rate is market-determined, and foreign exchange market intervention is used only to slow the rate of change and reduce short-term fluctuations, not to keep exchange rates at a target level.
-

LOS 15.j: Explain the effects of exchange rates on countries' international trade and capital flows.

We address the question of how a change in exchange rates affects a country's balance of trade using two approaches. The **elasticities approach** focuses on the impact of exchange rate changes on the total value of imports and on the total value of exports. Because a trade deficit (surplus) must be offset by a surplus (deficit) in the capital account, we can also view the effects of a change in exchange rates on capital flows rather than on goods flows. The **absorption approach** to analyzing the effect of a change in exchange rates focuses on capital flows.

The relation between the balance of trade and capital flows is expressed by the identity we presented in the reading on Aggregate Output, Prices, and Economic Growth. This identity is:

$$\text{exports} - \text{imports} \equiv (\text{private savings} - \text{investment in physical capital}) + (\text{tax revenue} - \text{government spending})$$

or

$$X - M \equiv (S - I) + (T - G)$$

The intuition is that a trade deficit ($X - M < 0$) means that the right-hand side must also be negative so that the total savings (private savings + government savings) is less than domestic investment in physical capital. The additional amount to fund domestic investment must come from foreigners, so there is a surplus in the capital account to offset the deficit in the trade account. Another thing we can see from this identity is that any government deficit not funded by an excess of domestic saving over domestic investment is consistent with a trade deficit (imports > exports) which is offset by an inflow of foreign capital (a surplus in the capital account).

Elasticities Approach

This approach to understanding the impact of exchange rate changes on the balance of trade focuses on how exchange rate changes affect total expenditures on imports and exports.

Consider an initial situation in which a country has a merchandise trade deficit (i.e., its imports exceed its exports). Depreciation of the domestic currency will make imports more expensive in domestic currency terms and exports less expensive in foreign currency terms. Thus, depreciation of the domestic currency will increase exports and decrease imports and would seem to unambiguously reduce the trade deficit. However, it is not the *quantity* of imports and

exports, but the total *expenditures* on imports and exports that must change in order to affect the trade deficit. Thus, the elasticity of demand for export goods and import goods is a crucial part of the analysis.

The condition under which a depreciation of the domestic currency will decrease a trade deficit are given in what is called the generalized **Marshall-Lerner condition**:

$$W_X \varepsilon_X + W_M (\varepsilon_M - 1) > 0$$

where:

W_X = proportion of total trade that is exports

W_M = proportion of total trade that is imports

ε_X = absolute value of price elasticity of demand for exports

ε_M = absolute value of price elasticity of demand for imports

In the case where import expenditures and export revenues are equal, $W_X = W_M$, this condition reduces to $\varepsilon_X + \varepsilon_M > 1$, which is most often cited as the classic Marshall-Lerner condition.

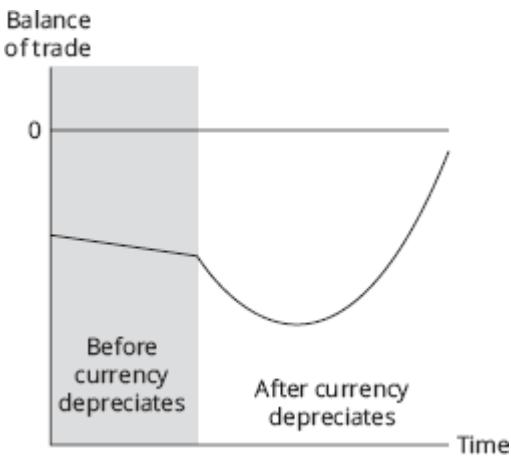
The elasticities approach tells us that currency depreciation will result in a greater improvement in the trade deficit when either import or export demand is elastic. For this reason, the compositions of export goods and import goods are an important determinant of the success of currency depreciation in reducing a trade deficit. In general, elasticity of demand is greater for goods with close substitutes, goods that represent a high proportion of consumer spending, and luxury goods in general. Goods that are necessities, have few or no good substitutes, or represent a small proportion of overall expenditures tend to have less elastic demand. Thus, currency depreciation will have a greater effect on the balance of trade when import or export goods are primarily luxury goods, goods with close substitutes, and goods that represent a large proportion of overall spending.

The J-Curve

Because import and export contracts for the delivery of goods most often require delivery and payment in the future, import and export quantities may be relatively insensitive to currency depreciation in the short run. This means that a currency depreciation may worsen a trade deficit initially. Importers adjust over time by reducing quantities. The Marshall-Lerner conditions take effect and the currency depreciation begins to improve the trade balance.

This short-term increase in the deficit followed by a decrease when the Marshall-Lerner condition is met is referred to as the **J-curve** and is illustrated in Figure 15.1.

Figure 15.1: J-Curve Effect



The Absorption Approach

One shortcoming of the elasticities approach is that it only considers the microeconomic relationship between exchange rates and trade balances. It ignores capital flows, which must also change as a result of a currency depreciation that improves the balance of trade. The absorption approach is a macroeconomic technique that focuses on the capital account and can be represented as:

$$BT = Y - E$$

where:

Y = domestic production of goods and services or national income

E = domestic absorption of goods and services, which is total expenditure

BT = balance of trade

Viewed in this way, we can see that income relative to expenditure must increase (domestic absorption must fall) for the balance of trade to improve in response to a currency depreciation. For the balance of trade to improve, domestic saving must increase relative to domestic investment in physical capital (which is a component of E). Thus, for a depreciation of the domestic currency to improve the balance of trade towards surplus, it must increase national income relative to expenditure. We can also view this as a requirement that national saving increase relative to domestic investment in physical capital.

Whether a currency depreciation has these effects depends on the current level of capacity utilization in the economy. When an economy is operating at less than full employment, the currency depreciation makes domestic goods and assets relatively more attractive than foreign goods and assets. The resulting shift in demand away from foreign goods and assets and towards domestic goods and assets will increase both expenditures and income. Because part of the income increase will be saved, national income will increase more than total expenditure, improving the balance of trade.

In a situation where the economy is operating at full employment (capacity), an increase in domestic spending will translate to higher domestic prices, which can reverse the relative price changes of the currency depreciation, resulting in a return to the previous deficit in the balance of trade. A currency depreciation at full capacity does result in a decline in the value of domestic assets. This decline in savers' real wealth will induce an increase in saving to rebuild wealth, initially improving the balance of trade from the currency depreciation. As the real

wealth of savers increases, however, the positive impact on saving will decrease, eventually returning the economy to its previous state and balance of trade.



MODULE QUIZ 15.3

1. The monetary authority of The Stoddard Islands will exchange its currency for U.S. dollars at a one-for-one ratio. As a result, the exchange rate of the Stoddard Islands currency with the U.S. dollar is 1.00, and many businesses in the Islands will accept U.S. dollars in transactions. This exchange rate regime is *best* described as:
 - A. a fixed peg.
 - B. dollarization.
 - C. a currency board.
2. A country that wishes to narrow its trade deficit devalues its currency. If domestic demand for imports is perfectly price-inelastic, whether devaluing the currency will result in a narrower trade deficit is *least likely* to depend on:
 - A. the size of the currency devaluation.
 - B. the country's ratio of imports to exports.
 - C. price elasticity of demand for the country's exports.
3. A devaluation of a country's currency to improve its trade deficit would *most likely* benefit a producer of:
 - A. luxury goods for export.
 - B. export goods that have no close substitutes.
 - C. an export good that represents a relatively small proportion of consumer expenditures.
4. Other things equal, which of the following is *most likely* to decrease a country's trade deficit?
 - A. Increase its capital account surplus.
 - B. Decrease expenditures relative to income.
 - C. Decrease domestic saving relative to domestic investment.

KEY CONCEPTS

LOS 15.a

Currency exchange rates are given as the price of one unit of currency in terms of another. A nominal exchange rate of 1.44 USD/EUR is interpreted as \$1.44 per euro. We refer to the USD as the price currency and the EUR as the base currency.

An increase (decrease) in an exchange rate represents an appreciation (depreciation) of the base currency relative to the price currency.

A spot exchange rate is the rate for immediate delivery. A forward exchange rate is a rate for exchange of currencies at some future date.

A real exchange rate measures changes in relative purchasing power over time.

$$\text{real exchange rate} = \text{nominal exchange rate} \times \left(\frac{\text{CPI}_{\text{base currency}}}{\text{CPI}_{\text{price currency}}} \right)$$

LOS 15.b

For a change in an exchange rate, we can calculate the percentage appreciation (price goes up) or depreciation (price goes down) of the base currency. For example, a decrease in the USD/EUR exchange rate from 1.44 to 1.42 represents a depreciation of the EUR relative to the USD of 1.39% ($1.42 / 1.44 - 1 = -0.0139$) because the price of a euro has fallen 1.39%.

To calculate the appreciation or depreciation of the price currency, we first invert the quote so it is now the base currency and then proceed as above. For example, a decrease in the USD/EUR exchange rate from 1.44 to 1.42 represents an appreciation of the USD relative to the EUR of 1.41%: $(1 / 1.42) / (1 / 1.44) - 1 = \frac{1.44}{1.42} - 1 = 0.0141$

The appreciation is the inverse of the depreciation, $\frac{1}{(1 - 0.0139)} - 1 = 0.0141$.

LOS 15.c

The market for foreign exchange is the largest financial market in terms of the value of daily transactions and has a variety of participants, including large multinational banks (the sell side) and corporations, investment fund managers, hedge fund managers, investors, governments, and central banks (the buy side).

Participants in the foreign exchange markets are referred to as hedgers if they enter into transactions that decrease an existing foreign exchange risk and as speculators if they enter into transactions that increase their foreign exchange risk.

LOS 15.d

Given two exchange rate quotes for three different currencies, we can calculate a currency cross rate. If the MXN/USD quote is 12.1 and the USD/EUR quote is 1.42, we can calculate the cross rate of MXN/EUR as $12.1 \times 1.42 = 17.18$.

LOS 15.e

Points in a foreign currency quotation are in units of the last digit of the quotation. For example, a forward quote of +25.3 when the USD/EUR spot exchange rate is 1.4158 means that the forward exchange rate is $1.4158 + 0.00253 = 1.41833$ USD/EUR.

For a forward exchange rate quote given as a percentage, the percentage (change in the spot rate) is calculated as forward / spot - 1. A forward exchange rate quote of +1.787%, when the spot USD/EUR exchange rate is 1.4158, means that the forward exchange rate is $1.4158 (1 + 0.01787) = 1.4411$ USD/EUR.

LOS 15.f

If a forward exchange rate does not correctly reflect the difference between the interest rates for two currencies, an arbitrage opportunity for a riskless profit exists. In this case, borrowing one currency, converting it to the other currency at the spot rate, investing the proceeds for the period, and converting the end-of-period amount back to the borrowed currency at the forward rate will produce more than enough to pay off the initial loan, with the remainder being a riskless profit on the arbitrage transaction.

LOS 15.g

To calculate a forward premium or forward discount for Currency B using exchange rates quoted as units of Currency A per unit of Currency B, use the following formula:

$$(\text{forward} / \text{spot}) - 1$$

LOS 15.h

The condition that must be met so that there is no arbitrage opportunity available is:

$$\frac{\text{forward}}{\text{spot}} = \frac{(1 + i_{\text{price currency}})}{(1 + i_{\text{base currency}})} \text{ so that forward} = \text{spot} \times \frac{(1 + i_{\text{price currency}})}{(1 + i_{\text{base currency}})}$$

If the spot exchange rate for the euro is 1.25 USD/EUR, the euro interest rate is 4% per year, and the dollar interest rate is 3% per year, the no-arbitrage one-year forward rate can be calculated as:

$$1.25 \times (1.03 / 1.04) = 1.238 \text{ USD/EUR.}$$

LOS 15.i

Exchange rate regimes for countries that do not have their own currency:

- With *formal dollarization*, a country uses the currency of another country.
- In a *monetary union*, several countries use a common currency.

Exchange rate regimes for countries that have their own currency:

- A *currency board arrangement* is an explicit commitment to exchange domestic currency for a specified foreign currency at a fixed exchange rate.
- In a *conventional fixed peg arrangement*, a country pegs its currency within margins of $\pm 1\%$ versus another currency.
- In a system of *pegged exchange rates within horizontal bands* or a *target zone*, the permitted fluctuations in currency value relative to another currency or basket of currencies are wider (e.g., $\pm 2\%$).
- With a *crawling peg*, the exchange rate is adjusted periodically, typically to adjust for higher inflation versus the currency used in the peg.
- With *management of exchange rates within crawling bands*, the width of the bands that identify permissible exchange rates is increased over time.
- With a system of *managed floating exchange rates*, the monetary authority attempts to influence the exchange rate in response to specific indicators, such as the balance of payments, inflation rates, or employment without any specific target exchange rate.
- When a currency is *independently floating*, the exchange rate is market-determined.

LOS 15.j

Elasticities (ϵ) of export and import demand must meet the Marshall-Lerner condition for a depreciation of the domestic currency to reduce an existing trade deficit:

$$W_X \epsilon_X + W_M (\epsilon_M - 1) > 0$$

Under the absorption approach, national income must increase relative to national expenditure in order to decrease a trade deficit. This can also be viewed as a requirement that national saving must increase relative to domestic investment in order to decrease a trade deficit.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 15.1

1. B An increase in the real exchange rate USD/EUR (the number of USD per one EUR) means a euro is worth more in purchasing power (real) terms in the United States. Changes in a real exchange rate depend on the change in the nominal exchange rate relative to the difference in

inflation. By itself, a real exchange rate does not indicate the directions or degrees of change in either the nominal exchange rate or the inflation difference. (LOS 15.a)

2. A Large multinational banks make up the sell side of the foreign exchange market. The buy side includes corporations, real money and leveraged investment accounts, governments and government entities, and retail purchasers of foreign currencies. (LOS 15.c)
3. B $1 / 1.311 = 0.7628$ GBP/USD. (LOS 15.a)
4. C The CAD has appreciated because it is worth a larger number of JPY. The percent appreciation is $(78 - 75) / 75 = 4.0\%$. To calculate the percentage depreciation of the JPY against the CAD, convert the exchange rates to make JPY the base currency: $1 / 75 = 0.0133$ CAD/JPY and $1 / 78 = 0.0128$ CAD/JPY. Percentage depreciation = $(0.0128 - 0.0133) / 0.0133 = -3.8\%$. (LOS 15.b)
5. A Start with one NZD and exchange for $1 / 1.6 = 0.625$ USD. Exchange the USD for $0.625 \times 2,400 = 1,500$ IDR. We get a cross rate of 1,500 IDR/NZD or $1 / 1,500 = 0.00067$ NZD/IDR. (LOS 15.d)
6. A $USD/NZD \ 0.3500 \times NZD/SEK \ 0.3100 = USD/SEK \ 0.1085$.

Notice that the NZD term cancels in the multiplication. (LOS 15.d)

Module Quiz 15.2

1. B The 180-day forward exchange rate is $1.3050 - 0.00425 = CHF/GBP \ 1.30075$. (LOS 15.e)
2. B Interest rates are higher in the United States than in New Zealand. It takes fewer NZD to buy one USD in the forward market than in the spot market. (LOS 15.f)
3. B To calculate a percentage forward premium or discount for the U.S. dollar, we need the dollar to be the base currency. The spot and forward quotes given are U.S. dollars per British pound (USD/GBP), so we must invert them to GBP/USD. The spot GBP/USD price is $1 / 1.533 = 0.6523$ and the forward GBP/USD price is $1 / 1.508 = 0.6631$. Because the forward price is greater than the spot price, we say the dollar is at a forward premium of $0.6631 / 0.6523 - 1 = 1.66\%$. Alternatively, we can calculate this premium with the given quotes as spot/forward - 1 to get $1.533 / 1.508 - 1 = 1.66\%$. (LOS 15.g)
4. B The forward rate in SEK/USD is $9.5238 \left(\frac{1.07}{1.04} \right) = 9.7985$. Since the SEK interest rate is the higher of the two, the SEK must depreciate approximately 3%. (LOS 15.h)
5. A We can solve interest rate parity for the spot rate as follows: With the exchange rates quoted as USD/CHF, the spot is $0.80 \left(\frac{1.04}{1.10} \right) = 0.7564$. Since the interest rate is higher in the United States, it should take fewer USD to buy CHF in the spot market. In other words, the forward USD must be depreciating relative to the spot. (LOS 15.h)

Module Quiz 15.3

1. C This exchange rate regime is a currency board arrangement. The country has not formally dollarized because it continues to issue a domestic currency. A conventional fixed peg allows for a small degree of fluctuation around the target exchange rate. (LOS 15.i)
2. A With perfectly inelastic demand for imports, currency devaluation of any size will increase total expenditures on imports (same quantity at higher prices in the home currency). The trade deficit will narrow only if the increase in export revenues is larger than the increase in import spending. To satisfy the Marshall-Lerner condition when import demand elasticity is zero, export demand elasticity must be larger than the ratio of imports to exports in the country's international trade. (LOS 15.j)
3. A A devaluation of the currency will reduce the price of export goods in foreign currency terms. The greatest benefit would be to producers of goods with more elastic demand. Luxury goods tend to have higher elasticity of demand, while goods that have no close substitutes or represent a small proportion of consumer expenditures tend to have low elasticities of demand. (LOS 15.j)

4. **B** An improvement in a trade deficit requires that domestic savings increase relative to domestic investment, which would decrease a capital account surplus. Decreasing expenditures relative to income means domestic savings increase. Decreasing domestic saving relative to domestic investment is consistent with a larger capital account surplus (an increase in net foreign borrowing) and a greater trade deficit. (LOS 15.j)

TOPIC QUIZ: ECONOMICS

You have now finished the Economics topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Economics. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

FORMULAS

nominal risk-free rate = real risk-free rate + expected inflation rate

required interest rate on a security = nominal risk-free rate
+ default risk premium
+ liquidity premium
+ maturity risk premium

effective annual rate = $(1 + \text{periodic rate})^m - 1$

continuous compounding: $e^r - 1 = \text{EAR}$

$$\text{PV}_{\text{perpetuity}} = \frac{\text{PMT}}{\text{I/Y}}$$

$$\text{FV} = \text{PV}(1 + \text{I/Y})^N$$

$$\text{population mean: } \mu = \frac{\sum_{i=1}^N X_i}{N}$$

$$\text{sample mean: } \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

geometric mean return (R_G): $1 + R_G = \sqrt[n]{(1 + R_1) \times (1 + R_2) \times \dots \times (1 + R_n)}$

$$\text{harmonic mean: } \bar{X}_H = \frac{N}{\frac{1}{\sum_{i=1}^N \frac{1}{X_i}}}$$

$$\text{weighted mean: } \bar{X}_w = \frac{\sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i}$$

$$\text{position of the observation at a given percentile, } y: L_y = (n + 1) \frac{y}{100}$$

range = maximum value – minimum value

excess kurtosis = sample kurtosis – 3

$$\text{MAD} = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

$$\text{population variance} = \sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N},$$

where μ = population mean and N = number of possible outcomes

$$\text{sample variance} = s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1},$$

where \bar{X} = sample mean and n = sample size

$$\text{coefficient of variation: } CV = \frac{s_x}{\bar{X}} = \frac{\text{standard deviation of } x}{\text{average value of } x}$$

$$\text{target downside deviation: } s_{\text{target}} = \sqrt{\frac{\sum_{\text{all } X_i < B} (X_i - B)^2}{n - 1}},$$

joint probability: $P(AB) = P(A \mid B) \times P(B)$

addition rule: $P(A \text{ or } B) = P(A) + P(B) - P(AB)$

multiplication rule: $P(A \text{ and } B) = P(A) \times P(B)$

total probability rule: $P(R) = P(R \mid S_1) \times P(S_1) + P(R \mid S_2) \times P(S_2) + \dots + P(R \mid S_N) \times P(S_N)$

expected value: $E(X) = \sum P(x_i)x_i = P(x_1)x_1 + P(x_2)x_2 + \dots + P(x_n)x_n$

variance from a probability model: $\text{Var}(X) = E\{[X - E(X)]^2\}$

$\text{Cov}(R_i, R_j) = E\{[R_i - E(R_i)][R_j - E(R_j)]\}$

$$\text{Corr}(R_i, R_j) = \frac{\text{Cov}(R_i, R_j)}{\sigma(R_i)\sigma(R_j)}$$

portfolio expected return: $E(R_p) = \sum_{i=1}^N w_i E(R_i) = w_1 E(R_1) + w_2 E(R_2) + \dots + w_n E(R_n)$

portfolio variance: $\text{Var}(R_p) = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \text{Cov}(R_i, R_j)$

where $w_i = \frac{\text{market value of investment in asset } i}{\text{market value of the portfolio}}$

Bayes' formula:

$$\text{updated probability} = \frac{\text{probability of new information for a given event}}{\text{unconditional probability of new information}} \times \text{prior probability of event}$$

$$\text{combination (binomial) formula: } {}_n C_r = \frac{n!}{(n-r)!r!}$$

$$\text{permutation formula: } {}_n P_r = \frac{n!}{(n-r)!}$$

$$\text{binomial probability: } p(x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}$$

for a binomial random variable: $E(X) = np$; variance = $np(1 - p)$

for a normal variable:

90% confidence interval for X is $\bar{X} - 1.65s$ to $\bar{X} + 1.65s$

95% confidence interval for X is $\bar{X} - 1.96s$ to $\bar{X} + 1.96s$

99% confidence interval for X is $\bar{X} - 2.58s$ to $\bar{X} + 2.58s$

$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}} = \frac{x - \mu}{\sigma}$$

$$\text{SFRatio} = \frac{[E(R_p) - R_L]}{\sigma_p}$$

$$\text{continuously compounded rate of return: } r_{cc} = \ln\left(\frac{S_1}{S_0}\right) = \ln(1 + \text{HPR})$$

$$\text{for a uniform distribution: } P(x_1 \leq X \leq x_2) = \frac{(x_2 - x_1)}{(b - a)}$$

sampling error of the mean = sample mean - population mean = $\bar{x} - \mu$

$$\text{standard error of the sample mean, known population variance: } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\text{standard error of the sample mean, unknown population variance: } s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

confidence interval: point estimate \pm (reliability factor \times standard error)

$$\text{confidence interval for the population mean: } \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\text{tests for population mean } = \mu_0: \text{z-statistic} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}, \text{ t-statistic} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

test for equality of variances: $F = \frac{s_1^2}{s_2^2}$, where $s_1^2 > s_2^2$

paired comparisons test: $t\text{-statistic} = \frac{\bar{d} - \mu_{dz}}{s_d}$

test for differences in means:

$$t\text{-statistic} = \frac{(\bar{x}_1 - \bar{x}_2)}{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^{1/2}} \quad (\text{sample variances assumed unequal})$$

$$t\text{-statistic} = \frac{(\bar{x}_1 - \bar{x}_2)}{\left(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}\right)^{1/2}} \quad (\text{sample variances assumed equal})$$

$$\text{test for correlation: } t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

$$\text{regression slope: } \hat{b}_1 = \frac{\text{Cov}_{XY}}{\sigma_X^2}$$

$$\text{coefficient of determination: } R^2 = \frac{\text{regression sum of squares}}{\text{total sum of squares}}$$

$$\text{standard error of estimate} = \sqrt{\text{mean squared error}}$$

$$\text{own-price elasticity} = \frac{\% \text{ change in quantity demanded}}{\% \text{ change in own price}}$$

$$\text{income elasticity} = \frac{\% \text{ change in quantity demanded}}{\% \text{ change in income}}$$

$$\text{cross-price elasticity} = \frac{\% \text{ change in quantity demanded}}{\% \text{ change in price of related good}}$$

breakeven points:

perfect competition: $AR = ATC$

imperfect competition: $TR = TC$

short-run shutdown points:

perfect competition: $AR < AVC$

imperfect competition: $TR < TVC$

$$\begin{aligned}\text{nominal GDP}_t \text{ for year } t &= \sum_{i=1}^N P_{i,t} Q_{i,t} \\ &= \sum_{i=1}^N (\text{price of good } i \text{ in year } t) \\ &\quad \times (\text{quantity of good } i \text{ produced in year } t)\end{aligned}$$

$$\begin{aligned}\text{real GDP for year } t &= \sum_{i=1}^N P_{i,\text{base year}} Q_{i,t} \\ &= \sum_{i=1}^N (\text{price of good } i \text{ in base year}) \\ &\quad \times (\text{quantity of good } i \text{ produced in year } t)\end{aligned}$$

GDP deflator for year t

$$= \frac{\sum_{i=1}^N P_{i,t} Q_{i,t}}{\sum_{i=1}^N P_{i,\text{base year}} Q_{i,t}} \times 100 = \frac{\text{nominal GDP in year } t}{\text{value of year } t \text{ output at base year prices}} \times 100$$

GDP, expenditure approach:

$$\text{GDP} = C + I + G + (X - M)$$

where:

C = consumption spending

I = business investment (capital equipment, inventories)

G = government purchases

X = exports

M = imports

GDP, income approach:

$$\text{GDP} = \text{national income} + \text{capital consumption allowance} + \text{statistical discrepancy}$$

national income = compensation of employees (wages and benefits)

+ corporate and government enterprise profits before taxes

+ interest income

+ unincorporated business net income (business owners' incomes)

+ rent

+ indirect business taxes – subsidies (taxes and subsidies that are included in final prices)

$$\text{growth in potential GDP} = \text{growth in technology} + W_L(\text{growth in labor}) + W_C(\text{growth in capital})$$

where:

W_L = labor's percentage share of national income

W_C = capital's percentage share of national income

growth in per-capita potential GDP = growth in technology + W_C (growth in the capital-to-labor ratio)

where:

W_C = capital's percentage share of national income

consumer price index = $\frac{\text{cost of basket at current prices}}{\text{cost of basket at base period prices}} \times 100$

money multiplier = $\frac{1}{\text{reserve requirement}}$

equation of exchange: money supply \times velocity = price \times real output ($MV = PY$)

Fisher effect: nominal interest rate = real interest rate + expected inflation rate

neutral interest rate = real trend rate of economic growth + inflation target

fiscal multiplier:

$$\frac{1}{1 - MPC(1 - t)}$$

where:

t = tax rate

MPC = marginal propensity to consume

real exchange rate = nominal exchange rate $\times \left(\frac{\text{CPI}_{\text{base currency}}}{\text{CPI}_{\text{price currency}}} \right)$

real exchange rate = $\frac{\text{nominal exchange rate}}{\left(\frac{\text{CPI}_{\text{price currency}}}{\text{CPI}_{\text{base currency}}} \right)}$

forward premium (+) or discount (-) for the base currency:

$$\frac{\text{forward}}{\text{spot}} - 1$$

interest rate parity:

$$\frac{\text{forward}}{\text{spot}} = \frac{(1 + \text{interest rate}_{\text{price currency}})}{(1 + \text{interest rate}_{\text{base currency}})}$$

Marshall-Lerner condition:

$$W_X \epsilon_X + W_M (\epsilon_M - 1) > 0$$

where:

W_M = proportion of trade that is imports

W_X = proportion of trade that is exports

ϵ_M = elasticity of demand for imports

ϵ_X = elasticity of demand for exports

APPENDICES

APPENDIX A: AREAS UNDER THE NORMAL CURVE

Most of the examples in this book have used one version of the z-table to find the area under the normal curve. This table provides the cumulative probabilities (or the area under the entire curve to left of the z-value).

Probability Example

Assume that the annual earnings per share (EPS) for a large sample of firms is normally distributed with a mean of \$5.00 and a standard deviation of \$1.50. What is the approximate probability of an observed EPS value falling between \$3.00 and \$7.25?

If $EPS = x = \$7.25$, then $z = (x - \mu)/\sigma = (\$7.25 - \$5.00)/\$1.50 = +1.50$

If $EPS = x = \$3.00$, then $z = (x - \mu)/\sigma = (\$3.00 - \$5.00)/\$1.50 = -1.33$

Solving Using The Cumulative Z-Table

For z-value of 1.50: Use the row headed 1.5 and the column headed 0 to find the value 0.9332. This represents the area under the curve to the left of the critical value 1.50.

For z-value of -1.33: Use the row headed 1.3 and the column headed 3 to find the value 0.9082. This represents the area under the curve to the left of the critical value +1.33. The area to the left of -1.33 is $1 - 0.9082 = 0.0918$.

The area between these critical values is $0.9332 - 0.0918 = 0.8414$, or 84.14%.

Hypothesis Testing—One-Tailed Test Example

A sample of a stock's returns on 36 nonconsecutive days results in a mean return of 2.0%. Assume the population standard deviation is 20.0%. Can we say with 95% confidence that the mean return is greater than 0%?

$$H_0: \mu \leq 0.0\%, H_a: \mu > 0.0\%. \text{ The test statistic} = z\text{-statistic}$$
$$= \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = (2.0 - 0.0) / (20.0 / \sqrt{36}) = 0.60$$

The significance level = $1.0 - 0.95 = 0.05$, or 5%. Because we are interested in a return greater than 0.0%, this is a one-tailed test.

Using the Cumulative Z-Table

Because this is a one-tailed test with an alpha of 0.05, we need to find the value 0.95 in the cumulative z-table. The closest value is 0.9505, with a corresponding critical z-value of 1.65. Because the test statistic is less than the critical value, we fail to reject H_0 .

Hypothesis Testing—Two-Tailed Test Example

Using the same assumptions as before, suppose that the analyst now wants to determine if he can say with 99% confidence that the stock's return is not equal to 0.0%.

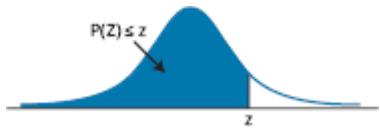
$H_0: \mu = 0.0\%$, $H_a: \mu \neq 0.0\%$. The test statistic (z-value) = $(2.0 - 0.0) / (20.0 / 6) = 0.60$. The significance level = $1.0 - 0.99 = 0.01$, or 1%. Because we are interested in whether or not the stock return is nonzero, this is a two-tailed test.

Using the Cumulative Z-Table

Because this is a two-tailed test with an alpha of 0.01, there is a 0.005 rejection region in both tails. Thus, we need to find the value 0.995 ($1.0 - 0.005$) in the table. The closest value is 0.9951, which corresponds to a critical z-value of 2.58. Because the test statistic is less than the critical value, we fail to reject H_0 and conclude that the stock's return equals 0.0%.

CUMULATIVE Z-TABLE

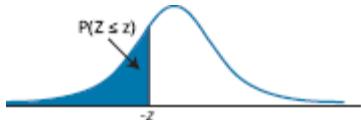
Standard Normal Distribution



$$P(Z \leq z) = N(z) \text{ for } z \geq 0$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Standard Normal Distribution



$$P(Z \leq z) = N(z) \text{ for } z \geq 0$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.7	0.2420	0.2389	0.2358	0.2327	0.2297	0.2266	0.2236	0.2207	0.2177	0.2148
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1057	0.1038	0.1020	0.1003	0.0985
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.4	0.0082	0.0080	0.0078	0.0076	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010

APPENDIX B: STUDENT'S T-DISTRIBUTION

Level of Significance for One-Tailed Test						
df	0.100	0.050	0.025	0.01	0.005	0.0005
Level of Significance for Two-Tailed Test						
df	0.20	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.294
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.768
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

APPENDIX C: *F*-TABLE AT 5% (UPPER TAIL)

***F*-Table, Critical Values, 5% in Upper Tail**

Degrees of freedom for the numerator along top row

Degrees of freedom for the denominator along side row

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	6.01	2.94	2.90	2.86	2.83
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39

APPENDIX D: *F*-TABLE AT 2.5% (UPPER TAIL)

F-Table, Critical Values, 2.5% in Upper Tails

Degrees of freedom for the numerator along top row

Degrees of freedom for the denominator along side row

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40
1	648	799	864	900	922	937	948	957	963	969	977	985	993	997	1001	1006
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.41	4.36	4.31
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48

APPENDIX E: CHI-SQUARED TABLE

Values of χ^2 (Degrees of Freedom, Level of Significance)

Probability in Right Tail

Degrees of Freedom	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0.000157	0.000982	0.003932	0.0158	2.706	3.841	5.024	6.635	7.879
2	0.020100	0.050636	0.102586	0.2107	4.605	5.991	7.378	9.210	10.597
3	0.1148	0.2158	0.3518	0.5844	6.251	7.815	9.348	11.345	12.838
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.086	16.750
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	29.819
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	12.878	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.335
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
50	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
80	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
100	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.170

INDEX

A

absolute advantage, 361
absolute frequency, 32
absorption approach, 390
action lag, 340
addition rule of probability, 73, 74, 75
additivity principle, 19
advertising expenses, 248
aggregate demand curve (AD curve), 276
aggregate supply curve (AS curve), 277
alternative hypothesis, 156
analysis of variance (ANOVA), 202
annuities, 7
annuities due, 7
a priori probability, 72
arithmetic mean, 45, 49
Austrian school, 303
autarky, 353, 359
automatic stabilizers, 335
average cost pricing, 258

B

balance of payments, 369
bar chart, 38
Bayes' formula, 86
Bernoulli random variable, 105
biased estimator, 53
bilateralism, 353
bimodal, 47
bimodal data set, 47
binomial distribution, 105
binomial formula, 88

binomial random variable, 105
black swan risk, 355
bond market vigilantes, 333
bootstrap, 146
box and whisker plot, 51, 52
brand names, 248
breakeven point, 229
broad money, 320
bubble line chart, 42
budget deficit, 319
budget surplus, 319
business cycle, 297
business expectations, 278
buy side, 383

C

capital account, 369
capital consumption allowance (CCA), 273
capital deepening investment, 288
capital restrictions, 367
capital spending, 337
capital transfers, 370
cartel, 251
cash flow additivity principle, 18
categorical data, 30
central limit theorem, 136
chi-square distribution, 178
chi-square distribution (χ^2), 124
closed economy, 359
clustered bar chart, 39
cluster sampling, 135
coefficient of determination, 204
coefficient of variation (CV), 54
combination formula, 88
comparative advantage, 361
comparisons, 44, 65
complements, 222
compounding, 3
compounding frequency, 20

compound interest, 1, 3
concentration measures, 260
conditional expected values, 80
conditional probability, 73
confidence interval, 111, 141
confidence interval for the population mean, 142
confusion matrix, 35
consistent estimator, 139
consumer price index (CPI), 308
contingency table, 35
continuous compounding, 120
continuous data, 29
continuous distribution, 101
continuous random variable, 100
continuous uniform distribution, 104
contraction, 297
contractionary monetary policy, 320, 332
convenience sampling, 135
conventional fixed peg arrangement, 389
cooperative, 351
core inflation, 310
corporations, 383
correlation, 110
correlation coefficient, 61
cost of capital, 6
cost-push inflation, 312
Cournot model, 250
covariance, 60, 82
covariance matrix, 83
crawling peg, 389
credit cycles, 298
critical values, 157
cross-price elasticity of demand, 222
cross rate, 383
cross-sectional data, 30
crowding-out effect, 336
cumulative absolute frequency, 34
cumulative distribution function (cdf), 101
cumulative frequency distribution chart, 38
cumulative relative frequency, 34
currency board arrangement, 389

current account, 369
current spending, 337
cyclically adjusted budget deficit, 341
cyclical unemployment, 306

D

data snooping, 147
data-snooping bias, 147
data table, 32
data types, 29
debt ratio, 336
decile, 51, 66
decision rule, 157, 161
default risk, 4
default risk premium, 4
deflation, 307
degrees of freedom (df), 122, 203, 214, 215
demand for money, 322
demand-pull inflation, 312
dependent variable, 193
desirable properties of an estimator, 138
diminishing marginal productivity, 227, 288
diminishing marginal returns, 227
direct quote, 380
direct taxes, 337
discount factor, 6
discounting, 3, 6
discount rate, 3, 6, 16
discouraged workers, 306
discrete compounding, 120
discrete data, 29
discrete distribution, 100
discretely compounded returns, 120
discrete random variable, 100
discrete uniform random variable, 103
discretionary fiscal policy, 335
diseconomies of scale, 232
disinflation, 307
dispersion, 52

disposable income, 338
distribution function, 101
distributions, 44, 65
dollarization, 389
domestic price, 360
dominant firm model, 252
downside risk, 55
durable goods, 300

E

economic significance, 163
economic tools, 354
effective annual rate (EAR), 23
effective annual yield (EAY), 23
efficient estimator, 139
elasticities approach, 390
elasticity along a linear demand curve, 221
empirical probability, 72
equality of variances, 182
equation of exchange, 322
event, 71
event risk, 354
excess capacity, 247
excess kurtosis, 59
excess reserves, 321
exchange rate regimes, 388
exchange rates, 279, 280, 379
exchange rate targeting, 331
exhaustive events, 71
exogenous risk, 355
expansion, 297
expansionary fiscal policy, 279
expansionary monetary policy, 279, 319, 332
expected inflation, 313
expected value, 77, 78, 106
expenditure approach, 270
exports, 359
export subsidies, 365, 366

F

factorial, 87
factors of production, 226
false positives, 165
F-distribution, 124, 181
financial account, 369
financial tools, 354
fiscal balance, 274
fiscal multiplier, 338
fiscal policy, 319, 335
fiscal policy tools, 337
Fisher effect, 324
Fisher index, 310
foreign direct investment, 354, 360
foreign-owned assets, 370
formal dollarization, 389
forward currency contract, 382
forward discount, 386
forward exchange rate, 381
forward premium, 386
fractional reserve banking, 321
free trade, 359
frequency distribution, 32
frequency polygon, 38
frictional unemployment, 306
full-employment GDP, 278
future value, 1
future value factor, 5
future value (FV), 1, 5, 7, 10, 14

- of an annuity due, 10
- of an ordinary annuity, 7
- of an uneven cash flow series, 14
- of a single sum, 5

future value interest factor, 5

G

GDP deflator, 271
geometric mean, 48, 49

geophysical resource endowment, 352
geopolitical risk, 354
geopolitics, 351
Giffen good, 226
global economic growth, 279
globalization, 352
government entities, 383
government-owned assets abroad, 370
gross domestic income (GDI), 273
gross domestic product (GDP), 269
gross national product, 360
grouped bar chart, 39

H

harmonic mean, 49
headline inflation, 310
heat map, 43
Heckscher-Ohlin model, 363
hedging, 382
hedonic pricing, 310
hegemony, 353
Herfindahl-Hirschman Index (HHI), 260
heteroskedasticity, 199
histogram, 37
homoskedasticity, 199
human capital, 287
hyperinflation, 307
hypothesis, 155
Hypothesis testing, steps, 156

I

impact lag, 340
impact of geopolitical risk, 355
imports, 359
income approach, 270
income effect, 225
income elasticity, 222

income receipts, 370
independent events, 75, 76
independently floating exchange rate, 390
independent variable, 194
indirect quote, 380
indirect taxes, 337
inferior good, 222
inflationary gap, 284
inflation premium, 4
inflation rate, 307
inflation reports, 331
inflation targeting, 331
input prices, 280
institutions, 352
intercept, 196
interest on interest, 1
interest rate effect, 276
interest rate targeting, 331
International Monetary Fund, 371
interquartile range, 51
interval, 32
inventory-sales ratio, 299
investment accounts, 383

J

jackknife, 146
J-curve, 391
joint frequencies, 35
joint probability, 73
judgmental sampling, 136

K

Keynesian school, 302
kinked demand curve model, 249
kurtosis, 59, 60
interpretation, 60

L

labeling, 87
labor force, 287, 306
labor productivity, 280
labor supply, 287
Laspeyres index, 310
leptokurtic distributions, 59
leveraged accounts, 383
likelihood, 73
likelihood of geopolitical risk, 355
line charts, 41
lin-log model, 210
liquidity risk, 4
liquidity risk premium, 4
liquidity trap, 333
location of the mean, median, and mode, 57
log-lin model, 210
log-log model, 211
lognormal distribution, 119
long run, 228
long-run aggregate supply (LRAS) curve, 281
long-run shutdown point, 229
look-ahead bias, 147

M

managed floating exchange rates, 390
management of exchange rates within crawling bands, 389
marginal cost pricing, 258
marginal frequency, 35
marginal probability, 73
markup, 247
Marshall-Lerner condition, 391
maturity risk, 4
maturity risk premium, 4
mean absolute deviation (MAD), 52
mean differences, 172
means of payment, 320
measures of central tendency, 45

measures of location, 51
median, 47
medium of exchange, 320
menu costs, 325
merchandise and services trade, 370
mesokurtic distributions, 59
minimum domestic content requirement, 365
minimum efficient scale, 231
modal interval, 33
mode, 47
Monetarist school, 303
monetary policy, 319
monetary policy tools, 328
monetary transmission mechanism, 328
monetary union, 389
money, 320
money multiplier, 322
money neutrality, 322
monopolistic competition, 238, 246
monopoly, 238
Monte Carlo simulation, 124
multilateralism, 354
multinational corporation, 360
multiplication rule of counting, 89
multiplication rule of probability, 73, 74
multivariate distribution, 110
multivariate normal distribution, 110
mutually exclusive events, 71, 75

N

narrow money, 320
Nash equilibrium, 251
national income, 273
nationalism, 352
national security tools, 354
natural monopoly, 238, 257
natural rate of unemployment (NARU), 312
natural resources, 287
negative skew, 57

Neoclassical school, 302
net exports, 272, 360
New Classical school, 303
New Keynesian school, 303
N-firm concentration ratio, 260
nominal data, 30
nominal exchange rate, 380
nominal GDP, 271
nominal risk-free rate, 4
non-accelerating inflation rate of unemployment (NAIRU), 312
non-cooperative, 351
nondurable goods, 300
nonparametric tests, 183
non-probability sampling, 133
non-state actors, 351
normal distribution, 109
normal good, 222
null hypothesis, 156
numerical data, 29

0

objective of a central bank, 325
objective probabilities, 72
odds, 72
oligopoly, 238
one-dimensional array, 31
one-stage cluster sampling, 135
one-tailed hypothesis test, 158
one-tailed test, 157, 205
opportunity cost, 3, 6
ordinal data, 30
ordinary annuities, 7
ordinary least squares, 196
outcome, 71
outliers, 46, 57, 201

P

Paasche index, 310
paired comparisons test, 174
panel data, 30
parametric tests, 183
participation ratio, 306
peak, 297
pegging, 325
per-capita real GDP, 272
percentile, 51, 66
perfect competition, 238
periodic rates, 23
permutation formula, 88
perpetuity, 12
personal disposable income, 273
personal income, 273
physical capital stock, 287
platykurtic distributions, 59
point estimates, 141
population mean, 45
portfolio expected return, 82
portfolio investment flows, 354
portfolio variance, 83
positive skew, 57
potential GDP, 278
power of a test, 161
predicted value, 195, 208
present value, 1
present value factor, 6
present value interest factor, 6
present value (PV), 1, 6, 8, 11, 12, 14

- of an annuity due, 11
- of an ordinary annuity, 8
- of an uneven cash flow series, 14
- of a perpetuity, 12
- of a single sum, 6

price discrimination, 255
price index, 308
price relative, 120
probability distribution, 99
probability function, 100
probability sampling, 133

probability tree, 80
producer price index (PPI), 310
product innovation, 247
production function, 226, 288
productivity, 307
promissory notes, 321
properties of an estimator, 138
properties of covariance, 82
properties of probability, 71, 72
p-value, 165

Q

qualitative data, 30
quantile, 50
quantitative data, 29
quantitative easing, 333
quantity equation of exchange, 322
quantity theory of money, 322
quartile, 66
quartiles, 50, 51
quintile, 50, 66
quota rents, 366
quotas, 365

R

random sampling, 133
random variable, 71
range, 52
real business cycle theory (RBC), 303
real exchange rate, 380
real exchange rate effect, 276
real GDP, 271
real money accounts, 383
real risk-free rate, 4
recession, 297
recessionary gap, 283
recognition lag, 340

recovery, 298
regionalism, 354
regression line, 195
rejection points, 157
relationships, 44, 65
relative dispersion, 54
relative frequency, 34
rent seeking, 257
required rate of return, 3, 6
residual, 195
retail market, 383
revenue tools, 337
Ricardian equivalence, 339
Ricardian model of trade, 363
risk, types of, 4
roles of central banks, 324
Roy's safety-first criterion, 116

S

sample covariance, 60, 83
sample kurtosis, 60
sample mean, 45
sample selection bias, 147
sample skewness, 58
sample standard deviation, 54
sample variance, 53
sampling and estimation, 133
sampling distribution, 134
sampling error, 134
sanctions, 354
scatter plot, 42
scatter plot matrix, 43
scatter plots, 61
scenario analysis, 355
sell side, 383
services, 300
shoe leather costs, 325
Shortfall risk, 116
short run, 228

short-run aggregate supply (SRAS) curve, 280
short-run shutdown point, 229
shutdown point, 242
significance level, 157
signposts, 355
simple linear regression, 193
simple random sampling, 133, 134
skew, skewness, 57, 59
 interpretation of the skewness measure, 59
slope coefficient, 196
soft power, 352
sources of economic growth, 287
sovereign wealth funds, 383
Spearman rank correlation test, 184
speculative foreign exchange transactions, 382
spending tools, 337
spot exchange rate, 381
spurious correlation, 63
stacked bar chart, 39
Stackelberg model, 250
stagflation, 284
standard deviation, 54
standard error of the sample mean, 137
standardization, 352
standardizing, 113
standard normal distribution, 113
state actors, 351
stated annual interest rates, 23
statistical significance, 163
stratified random sampling, 134
structural budget deficit, 341
structural unemployment, 306
structured data, 31
Student's t-distribution, 121
subjective probability, 72
subsidies, 280
substitutes, 222
substitution effect, 225
sum of squared errors (SSE), 195, 202
sum of squares regression (SSR), 202
sum-of-value-added method, 270

supply of money, 323
survivorship bias, 147
sustainable rate of economic growth, 287
symmetrical distributions, 57
systematic sampling, 134

T

target downside deviation, 55
target semideviation, 55
target zone, 389
tariffs, 365
taxes, 280
t-distribution, 121
technology, 281, 287
terms of trade, 360
test statistic, 160
thematic risk, 355
threshold level, 116
time line, 3
time-period bias, 148
time series, 30
time value of money, 1
tools of geopolitics, 354
total factor productivity, 288
total probability rule, 77
total sum of squares (SST), 202
total variation, 202
trade balance, 274
trade deficit, 360
trade protection, 359
trade surplus, 360
trading blocs, 368
transfer payments, 337
tree map, 40
trimmed mean, 46, 49
trimodal, 47
trimodal data set, 47
trough, 297
t-test, 167, 206

two-dimensional array, 32
two-stage cluster sampling, 135
two-tailed test, 157
Type I error, 160
Type II error, 160

U

unbiased estimator, 139
unconditional probability, 73
underemployed, 306
unemployed, 306
unemployment rate, 306
unilateral transfers, 370
unimodal, 47
unimodal data set, 47
unit labor costs, 312
unit of account, 320
univariate distributions, 110
unstructured data, 31

V

value-of-final-output method, 270
Veblen good, 226
velocity of geopolitical risk, 355
volatility, 79
voluntary export restraint (VER), 365, 366

W

wealth effect, 276
weighted average, 46
weighted mean, 46
wholesale price index (WPI), 310
winsorized mean, 46, 49
word cloud, 40
World Bank, 371

world price, 360

World Trade Organization, 372

Z

z-test, 168



2023 CFA® Exam Prep

SchweserNotes™ Financial Statement Analysis



LEVEL I BOOK 2

KAPLAN SCHWEISER

Book 2: Financial Statement Analysis

SchweserNotes™ 2023

Level I CFA®



SCHWESERNOTES™ 2023 LEVEL I CFA® BOOK 2: FINANCIAL STATEMENT ANALYSIS

©2022 Kaplan, Inc. All rights reserved.

Published in 2021 by Kaplan, Inc.

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

ISBN: 978-1-0788-2586-3

These materials may not be copied without written permission from the author. The unauthorized duplication of these notes is a violation of global copyright laws and the CFA Institute Code of Ethics. Your assistance in pursuing potential violators of this law is greatly appreciated.

Required CFA Institute disclaimer: Kaplan Schweser is a CFA Institute Prep Provider. Only CFA Institute Prep Providers are permitted to make use of CFA Institute copyrighted materials which are the building blocks of the exam. We are also required to update our materials every year and this is validated by CFA Institute.

CFA Institute does not endorse, promote, review or warrant the accuracy or quality of the product and services offered by Kaplan Schweser. CFA Institute®, CFA® and “Chartered Financial Analyst®” are trademarks owned by CFA Institute.

Certain materials contained within this text are the copyrighted property of CFA Institute. The following is the copyright disclosure for these materials: “Copyright, 2022, CFA Institute. Reproduced and republished from 2023 Learning Outcome Statements, Level I, II, and III questions from CFA® Program Materials, CFA Institute Standards of Professional Conduct, and CFA Institute’s Global Investment Performance Standards with permission from CFA Institute. All Rights Reserved.”

Disclaimer: The SchweserNotes should be used in conjunction with the original readings as set forth by CFA Institute in their 2023 Level I CFA Study Guide. The information contained in these Notes covers topics contained in the readings referenced by CFA Institute and is believed to be accurate. However, their accuracy cannot be guaranteed nor is any warranty conveyed as to your ultimate exam success. The authors of the referenced readings have not endorsed or sponsored these Notes.

CONTENTS

Learning Outcome Statements (LOS)

FINANCIAL STATEMENT ANALYSIS

READING 16

Introduction to Financial Statement Analysis

Exam Focus

Module 16.1: Financial Statement Roles

Module 16.2: Footnotes, Audit, and Analysis

Key Concepts

Answer Key for Module Quizzes

READING 17

Financial Reporting Standards

Exam Focus

Module 17.1: Standards Overview

Module 17.2: Financial Reporting Framework

Key Concepts

Answer Key for Module Quizzes

READING 18

Understanding Income Statements

Exam Focus

Module 18.1: Income Statement Overview

Module 18.2: Revenue Recognition

Module 18.3: Expense Recognition

Module 18.4: EPS and Dilutive Securities

Module 18.5: Common-Size Income Statements

Key Concepts

Answer Key for Module Quizzes

READING 19

Understanding Balance Sheets

Exam Focus

Module 19.1: Balance Sheet Introduction

Module 19.2: Assets and Liabilities

Module 19.3: Current Assets and Liabilities

Module 19.4: Noncurrent Assets and Liabilities

Module 19.5: Intangible Assets

Module 19.6: Marketable Securities

Module 19.7: Shareholders' Equity and Ratios

Key Concepts

Answer Key for Module Quizzes

READING 20

Understanding Cash Flow Statements

Exam Focus

Module 20.1: Cash Flow Introduction

Module 20.2: The Direct and Indirect Methods

Module 20.3: Converting Indirect to Direct

Module 20.4: Free Cash Flow and Ratios

Key Concepts

Answer Key for Module Quizzes

READING 21

Financial Analysis Techniques

Exam Focus

Module 21.1: Introduction to Financial Ratios

Module 21.2: Financial Ratios, Part 1

Module 21.3: Financial Ratios, Part 2

Module 21.4: DuPont Analysis

Module 21.5: More Financial Ratios

Key Concepts

Answer Key for Module Quizzes

READING 22

Inventories

Exam Focus

Module 22.1: Cost Flow Methods

Module 22.2: Inventory Systems

Module 22.3: Converting LIFO to FIFO

Module 22.4: Inventory Valuation

Module 22.5: Inventory Analysis

Key Concepts

Answer Key for Module Quizzes

READING 23

Long-Lived Assets

Exam Focus

Module 23.1: Capitalization vs. Expensing

Module 23.2: Depreciation

Module 23.3: Impairment and Revaluation

Module 23.4: Fixed Asset Disclosures

Key Concepts

Answer Key for Module Quizzes

READING 24

Income Taxes

Exam Focus

Module 24.1: Tax Terms

Module 24.2: Deferred Tax Liabilities and Assets

Module 24.3: Change in Tax Rates

Module 24.4: Permanent Differences

Key Concepts

Answer Key for Module Quizzes

READING 25

Non-Current (Long-Term) Liabilities

Exam Focus

Module 25.1: Bond Issuance

Module 25.2: Discount and Premium Bonds

Module 25.3: Issuance Cost, Derecognition, and Disclosures

Module 25.4: Lease and Pension Accounting

Key Concepts

Answer Key for Module Quizzes

READING 26

Financial Reporting Quality

Exam Focus

Module 26.1: Reporting Quality

Module 26.2: Accounting Choices and Estimates

Module 26.3: Warning Signs

Key Concepts

Answer Key for Module Quizzes

READING 27

Applications of Financial Statement Analysis

Exam Focus

Module 27.1: Forecasting

Module 27.2: Credit and Equity Analysis

Key Concepts

Answer Key for Module Quizzes

Topic Quiz: Financial Statement Analysis

Formulas

Index

LEARNING OUTCOME STATEMENTS (LOS)

16. Introduction to Financial Statement Analysis

The candidate should be able to:

- a. describe the roles of financial reporting and financial statement analysis.
- b. describe the roles of the statement of financial position, statement of comprehensive income, statement of changes in equity, and statement of cash flows in evaluating a company's performance and financial position.
- c. describe the importance of financial statement notes and supplementary information—including disclosures of accounting policies, methods, and estimates—and management's commentary.
- d. describe the objective of audits of financial statements, the types of audit reports, and the importance of effective internal controls.
- e. identify and describe information sources that analysts use in financial statement analysis besides annual financial statements and supplementary information.
- f. describe the steps in the financial statement analysis framework.

17. Financial Reporting Standards

The candidate should be able to:

- a. describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation.
- b. describe the roles of financial reporting standard-setting bodies and regulatory authorities in establishing and enforcing reporting standards.
- c. describe the International Accounting Standards Board's conceptual framework, including qualitative characteristics of financial reports, constraints on financial reports, and required reporting elements.
- d. describe general requirements for financial statements under International Financial Reporting Standards (IFRS).
- e. describe implications for financial analysis of alternative financial reporting systems and the importance of monitoring developments in financial reporting standards.

18. Understanding Income Statements

The candidate should be able to:

- a. describe the components of the income statement and alternative presentation formats of that statement.
- b. describe general principles of revenue recognition and accounting standards for revenue recognition.
- c. calculate revenue given information that might influence the choice of revenue recognition method.
- d. describe general principles of expense recognition, specific expense recognition applications, and implications of expense recognition choices for financial analysis.
- e. describe the financial reporting treatment and analysis of non-recurring items (including discontinued operations, unusual or infrequent items) and changes in accounting policies.
- f. contrast operating and non-operating components of the income statement.
- g. describe how earnings per share is calculated and calculate and interpret a company's earnings per share (both basic and diluted earnings per share) for both simple and complex capital structures.

- h. contrast dilutive and antidilutive securities and describe the implications of each for the earnings per share calculation.
- i. formulate income statements into common-size income statements.
- j. evaluate a company's financial performance using common-size income statements and financial ratios based on the income statement.
- k. describe, calculate, and interpret comprehensive income.
- l. describe other comprehensive income and identify major types of items included in it.

19. Understanding Balance Sheets

The candidate should be able to:

- a. describe the elements of the balance sheet: assets, liabilities, and equity.
- b. describe uses and limitations of the balance sheet in financial analysis.
- c. describe alternative formats of balance sheet presentation.
- d. contrast current and non-current assets and current and non-current liabilities.
- e. describe different types of assets and liabilities and the measurement bases of each.
- f. describe the components of shareholders' equity.
- g. demonstrate the conversion of balance sheets to common-size balance sheets and interpret common-size balance sheets.
- h. calculate and interpret liquidity and solvency ratios.

20. Understanding Cash Flow Statements

The candidate should be able to:

- a. compare cash flows from operating, investing, and financing activities and classify cash flow items as relating to one of those three categories given a description of the items.
- b. describe how non-cash investing and financing activities are reported.
(page 81)
- c. contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP).
- d. compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method.
- e. describe how the cash flow statement is linked to the income statement and the balance sheet.
- f. describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data.
- g. demonstrate the conversion of cash flows from the indirect to direct method.
- h. analyze and interpret both reported and common-size cash flow statements.
- i. calculate and interpret free cash flow to the firm, free cash flow to equity, and performance and coverage cash flow ratios.

21. Financial Analysis Techniques

The candidate should be able to:

- a. describe tools and techniques used in financial analysis, including their uses and limitations.
- b. identify, calculate, and interpret activity, liquidity, solvency, profitability, and valuation ratios.
- c. describe relationships among ratios and evaluate a company using ratio analysis.
- d. demonstrate the application of DuPont analysis of return on equity and calculate and interpret effects of changes in its components.

- e. calculate and interpret ratios used in equity analysis and credit analysis.
- f. explain the requirements for segment reporting and calculate and interpret segment ratios.
- g. describe how ratio analysis and other techniques can be used to model and forecast earnings.

22. Inventories

The candidate should be able to:

- a. contrast costs included in inventories and costs recognised as expenses in the period in which they are incurred.
- b. describe different inventory valuation methods (cost formulas).
- c. calculate and compare cost of sales, gross profit, and ending inventory using different inventory valuation methods and using perpetual and periodic inventory systems.
- d. calculate and explain how inflation and deflation of inventory costs affect the financial statements and ratios of companies that use different inventory valuation methods.
- e. explain LIFO reserve and LIFO liquidation and their effects on financial statements and ratios.
- f. demonstrate the conversion of a company's reported financial statements from LIFO to FIFO for purposes of comparison.
- g. describe the measurement of inventory at the lower of cost and net realisable value.
- h. describe implications of valuing inventory at net realisable value for financial statements and ratios.
- i. describe the financial statement presentation of and disclosures relating to inventories.
- j. explain issues that analysts should consider when examining a company's inventory disclosures and other sources of information.
- k. calculate and compare ratios of companies, including companies that use different inventory methods.
- l. analyze and compare the financial statements of companies, including companies that use different inventory methods.

23. Long-Lived Assets

The candidate should be able to:

- a. identify and contrast costs that are capitalised and costs that are expensed in the period in which they are incurred.
- b. compare the financial reporting of the following types of intangible assets: purchased, internally developed, acquired in a business combination.
- c. explain and evaluate how capitalising versus expensing costs in the period in which they are incurred affects financial statements and ratios.
- d. describe the different depreciation methods for property, plant, and equipment and calculate depreciation expense.
- e. describe how the choice of depreciation method and assumptions concerning useful life and residual value affect depreciation expense, financial statements, and ratios.
- f. explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios.
- g. describe the different amortisation methods for intangible assets with finite lives and calculate amortisation expense.
- h. describe how the choice of amortisation method and assumptions concerning useful life and residual value affect amortisation expense, financial statements, and ratios.
- i. describe the revaluation model.

- j. explain the impairment of property, plant, and equipment and intangible assets.
- k. explain the derecognition of property, plant, and equipment and intangible assets.
- l. describe the financial statement presentation of and disclosures relating to property, plant, and equipment and intangible assets.
- m. analyze and interpret financial statement disclosures regarding property, plant, and equipment and intangible assets.
- n. compare the financial reporting of investment property with that of property, plant, and equipment.

24. Income Taxes

The candidate should be able to:

- a. describe the differences between accounting profit and taxable income and define key terms, including deferred tax assets, deferred tax liabilities, valuation allowance, taxes payable, and income tax expense.
- b. explain how deferred tax liabilities and assets are created and the factors that determine how a company's deferred tax liabilities and assets should be treated for the purposes of financial analysis.
- c. calculate income tax expense, income taxes payable, deferred tax assets, and deferred tax liabilities, and calculate and interpret the adjustment to the financial statements related to a change in the income tax rate.
- d. calculate the tax base of a company's assets and liabilities.
- e. evaluate the effect of tax rate changes on a company's financial statements and ratios.
- f. identify and contrast temporary versus permanent differences in pre-tax accounting income and taxable income.
- g. explain recognition and measurement of current and deferred tax items.
- h. describe the valuation allowance for deferred tax assets—when it is required and what effect it has on financial statements.
- i. analyze disclosures relating to deferred tax items and the effective tax rate reconciliation and explain how information included in these disclosures affects a company's financial statements and financial ratios.
- j. identify the key provisions of and differences between income tax accounting under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (GAAP).

25. Non-Current (Long-Term) Liabilities

The candidate should be able to:

- a. determine the initial recognition, initial measurement and subsequent measurement of bonds.
- b. describe the effective interest method and calculate interest expense, amortisation of bond discounts/premiums, and interest payments.
- c. explain the derecognition of debt.
- d. describe the role of debt covenants in protecting creditors.
- e. describe the financial statement presentation of and disclosures relating to debt.
- f. explain motivations for leasing assets instead of purchasing them.
- g. explain the financial reporting of leases from a lessee's perspective.
- h. explain the financial reporting of leases from a lessor's perspective.
- i. compare the presentation and disclosure of defined contribution and defined benefit pension plans.

j. calculate and interpret leverage and coverage ratios.

26. Financial Reporting Quality

The candidate should be able to:

- a. compare and contrast financial reporting quality with the quality of reported results (including quality of earnings, cash flow, and balance sheet items).
- b. describe a spectrum for assessing financial reporting quality.
- c. explain the difference between conservative and aggressive accounting.
- d. describe motivations that might cause management to issue financial reports that are not high quality.
- e. describe conditions that are conducive to issuing low-quality, or even fraudulent, financial reports.
- f. describe mechanisms that discipline financial reporting quality and the potential limitations of those mechanisms.
- g. describe presentation choices, including non-GAAP measures, that could be used to influence an analyst's opinion.
- h. describe accounting methods (choices and estimates) that could be used to manage earnings, cash flow, and balance sheet items.
- i. describe accounting warning signs and methods for detecting manipulation of information in financial reports.

27. Applications of Financial Statement Analysis

The candidate should be able to:

- a. evaluate a company's past financial performance and explain how a company's strategy is reflected in past financial performance.
- b. demonstrate how to forecast a company's future net income and cash flow.
- c. describe the role of financial statement analysis in assessing the credit quality of a potential debt investment.
- d. describe the use of financial statement analysis in screening for potential equity investments.
- e. explain appropriate analyst adjustments to a company's financial statements to facilitate comparison with another company.

Reading 16

INTRODUCTION TO FINANCIAL STATEMENT ANALYSIS

EXAM FOCUS

This introduction may be useful to those who have no previous experience with financial statements. While the income statement, balance sheet, and statement of cash flows are covered in detail in subsequent readings, candidates should pay special attention here to the other sources of information for financial analysis. The nature of the audit report is important, as is the information that is contained in the footnotes to financial statements, proxy statements, Management's Discussion and Analysis, and the supplementary schedules. A useful framework enumerating the steps in financial statement analysis is presented.

MODULE 16.1: FINANCIAL STATEMENT ROLES



LOS 16.a: Describe the roles of financial reporting and financial statement analysis.

Video covering this content is available online.

Financial reporting refers to the way companies show their financial performance to investors, creditors, and other interested parties by preparing and presenting financial statements.

The role of **financial statement analysis** is to use the information in a company's financial statements, along with other relevant information, to make economic decisions. Examples of such decisions include whether to invest in the company's securities or recommend them to investors and whether to extend trade or bank credit to the company. Analysts use financial statement data to evaluate a company's past performance and current financial position in order to form opinions about the company's ability to earn profits and generate cash flow in the future.



PROFESSOR'S NOTE

This reading deals with financial analysis for external users. Management also performs financial analysis in making everyday decisions. However, management may rely on internal financial information that is likely maintained in a different format and unavailable to external users.

LOS 16.b: Describe the roles of the statement of financial position, statement of comprehensive income, statement of changes in equity, and statement of cash flows in evaluating a company's performance and financial position.

The **balance sheet** (also known as the *statement of financial position or statement of financial condition*) reports the firm's financial position at a point in time. The balance sheet consists of three elements:

1. Assets are the resources controlled by the firm.
2. Liabilities are amounts owed to lenders and other creditors.
3. Owners' equity (also *shareholders' equity*, *shareholders' funds*, or *net assets*) is the residual interest in the net assets of an entity that remains after deducting its liabilities from its assets.

Transactions are measured so that the fundamental **accounting equation** holds:

$$\text{assets} = \text{liabilities} + \text{owners' equity}$$

The proportions of liabilities and equity used to finance a company are known as the company's **capital structure**.

The **statement of comprehensive income** reports all changes in equity except for shareholder transactions (e.g., issuing stock, repurchasing stock, and paying dividends). The **income statement** (also known as the *statement of operations* or the *profit and loss statement*) reports on the financial performance of the firm over a period of time. The elements of the income statement include revenues, expenses, and gains and losses.

- *Revenues* are inflows from delivering or producing goods, rendering services, or other activities that constitute the entity's ongoing major or central operations.
- *Expenses* are outflows from delivering or producing goods or services that constitute the entity's ongoing major or central operations.
- *Other income* includes gains that may or may not arise in the ordinary course of business.

The income statement can be combined with "other comprehensive income" and presented as a single statement of comprehensive income. Alternatively, the income statement and the statement of comprehensive income can be presented separately.

The **statement of changes in equity** reports the amounts and sources of changes in equity investors' investment in the firm over a period of time.

The **statement of cash flows** reports the company's cash receipts and payments. These cash flows are classified as follows:

- *Operating cash flows* include the cash effects of transactions that involve the normal business of the firm.
- *Investing cash flows* are those resulting from the acquisition or sale of property, plant, and equipment; of a subsidiary or segment; of securities; and of investments in other firms.
- *Financing cash flows* are those resulting from issuance or retirement of the firm's debt and equity securities and include dividends paid to stockholders.

MODULE 16.2: FOOTNOTES, AUDIT, AND ANALYSIS



Video covering
this content is
available online.

LOS 16.c: Describe the importance of financial statement notes and supplementary information—including disclosures of accounting policies, methods, and estimates—and management's commentary.

Financial statement notes (footnotes) include disclosures that provide further details about the information summarized in the financial statements. Footnotes allow users to improve their assessments of the amount, timing, and uncertainty of the estimates reported in the financial statements. Footnotes:

- Discuss the basis of presentation such as the fiscal period covered by the statements and the inclusion of consolidated entities.
- Provide information about accounting methods, assumptions, and estimates used by management.
- Provide additional information on items such as business acquisitions or disposals, legal actions, employee benefit plans, contingencies and commitments, significant customers, sales to related parties, and segments of the firm.

Management's commentary [also known as management's report, operating and financial review, and **Management's Discussion and Analysis (MD&A)**] is one of the most useful sections of the annual report. In this section, management discusses a variety of issues. IFRS guidance recommends that management commentary address the nature of the business, management's objectives, the company's past performance, the performance measures used, and the company's key relationships, resources, and risks. Analysts must be aware that some parts of management's commentary may be unaudited.

For publicly held firms in the United States, the SEC requires that MD&A discuss trends and identify significant events and uncertainties that affect the firm's liquidity, capital resources, and results of operations. MD&A must also discuss:

- Effects of inflation and changing prices if material.
- Impact of off-balance-sheet obligations and contractual obligations such as purchase commitments.
- Accounting policies that require significant judgment by management.
- Forward-looking expenditures and divestitures.

LOS 16.d: Describe the objective of audits of financial statements, the types of audit reports, and the importance of effective internal controls.

An **audit** is an independent review of an entity's financial statements. Public accountants conduct audits and examine the financial reports and supporting records. The objective of an audit is to enable the auditor to provide an opinion on the fairness and reliability of the financial statements.

The independent certified public accounting firm employed by the Board of Directors is responsible for seeing that the financial statements conform to the applicable accounting standards. The auditor examines the company's accounting and internal control systems, confirms assets and liabilities, and generally tries to determine that there are no material errors in the financial statements. The auditor's report is an important source of information.

The **standard auditor's opinion** contains three parts and states that:

1. Whereas the financial statements are prepared by management and are its responsibility, the auditor has performed an independent review.

2. Generally accepted auditing standards were followed, thus providing *reasonable assurance* that the financial statements contain no material errors.
3. The auditor is satisfied that the statements were prepared in accordance with accepted accounting principles and that the principles chosen and estimates made are reasonable. The auditor's report must also contain additional explanation when accounting methods have not been used consistently between periods.

An *unqualified opinion* (also known as an unmodified or clean opinion) indicates that the auditor believes the statements are free from material omissions and errors. If the statements make any exceptions to the accounting principles, the auditor may issue a *qualified opinion* and explain these exceptions in the audit report. The auditor can issue an *adverse opinion* if the statements are not presented fairly or are materially nonconforming with accounting standards. If the auditor is unable to express an opinion (e.g., in the case of a scope limitation), a *disclaimer of opinion* is issued. Any opinion other than unqualified is sometimes referred to as a *modified opinion*.

The auditor's opinion will also contain an explanatory paragraph when a material loss is probable but the amount cannot be reasonably estimated. These "uncertainties" may relate to the *going concern assumption* (the assumption that the firm will continue to operate for the foreseeable future), the valuation or realization of asset values, or to litigation. This type of disclosure may be a signal of serious problems and may call for close examination by the analyst.

Internal controls are the processes by which the company ensures that it presents accurate financial statements. Internal controls are the responsibility of management. For publicly traded firms in the United States, the auditor must express an opinion on the firm's internal controls. The auditor can provide this opinion separately or as the fourth element of the standard opinion.

An audit report must also contain a section called Key Audit Matters (international reports) or Critical Audit Matters (U.S.), which highlights accounting choices that are of greatest significance to users of financial statements. These would include accounting choices that require significant management judgments and estimates, how significant transactions during a period were accounted for, or choices the auditor finds especially challenging or subjective and which therefore have a significant likelihood of being misstated.

LOS 16.e: Identify and describe information sources that analysts use in financial statement analysis besides annual financial statements and supplementary information.

Besides the annual financial statements, an analyst should examine a company's *quarterly or semiannual reports*. These interim reports typically update the major financial statements and footnotes but are not necessarily audited.

Securities and Exchange Commission (SEC) filings are available from EDGAR (Electronic Data Gathering, Analysis, and Retrieval System, www.sec.gov). These include Form 8-K, which a company must file to report events such as acquisitions and disposals of major assets or changes in its management or corporate governance. Companies' annual and quarterly financial statements are also filed with the SEC (Form 10-K and Form 10-Q, respectively).

Proxy statements are issued to shareholders when there are matters that require a shareholder vote. These statements, which are also filed with the SEC and available from EDGAR, are a good source of information about the election of (and qualifications of) board members, compensation, management qualifications, and the issuance of stock options.

Corporate reports and *press releases* are written by management and are often viewed as public relations or sales materials. Not all of the material is independently reviewed by outside auditors. Such information can often be found on the company's website. Firms often provide **earnings guidance** before the financial statements are released. Once an earnings announcement is made, a conference call may be held whereby senior management is available to answer questions.

An analyst should also review pertinent information on economic conditions and the company's industry and compare the company to its competitors. The necessary information can be acquired from trade journals, statistical reporting services, and government agencies.

LOS 16.f: Describe the steps in the financial statement analysis framework.

The **financial statement analysis framework**¹ consists of six steps:

Step 1: State the objective and context. Determine what questions the analysis seeks to answer, the form in which this information needs to be presented, and what resources and how much time are available to perform the analysis.

Step 2: Gather data. Acquire the company's financial statements and other relevant data on its industry and the economy. Ask questions of the company's management, suppliers, and customers, and visit company sites.

Step 3: Process the data. Make any appropriate adjustments to the financial statements. Calculate ratios. Prepare exhibits such as graphs and common-size balance sheets.

Step 4: Analyze and interpret the data. Use the data to answer the questions stated in the first step. Decide what conclusions or recommendations the information supports.

Step 5: Report the conclusions or recommendations. Prepare a report and communicate it to its intended audience. Be sure the report and its dissemination comply with the Code and Standards that relate to investment analysis and recommendations.

Step 6: Update the analysis. Repeat these steps periodically and change the conclusions or recommendations when necessary.



MODULE QUIZ 16.1, 16.2

1. Which of the following statements *least accurately* describes a role of financial statement analysis?
 - A. Use the information in financial statements to make economic decisions.
 - B. Provide reasonable assurance that the financial statements are free of material errors.
 - C. Evaluate an entity's financial position and past performance to form opinions about its future ability to earn profits and generate cash flow.
2. A firm's financial position at a specific point in time is reported in the:
 - A. balance sheet.
 - B. income statement.
 - C. cash flow statement.

3. Information about accounting estimates, assumptions, and methods chosen for reporting is *most likely* found in:
 - A. the auditor's opinion.
 - B. financial statement notes.
 - C. Management's Discussion and Analysis.
4. If an auditor finds that a company's financial statements have made a specific exception to applicable accounting principles, she is *most likely* to issue a:
 - A. dissenting opinion.
 - B. cautionary note.
 - C. qualified opinion.
5. Information about elections of members to a company's Board of Directors is *most likely* found in:
 - A. a 10-Q filing.
 - B. a proxy statement.
 - C. footnotes to the financial statements.
6. Which of these steps is *least likely* to be a part of the financial statement analysis framework?
 - A. State the purpose and context of the analysis.
 - B. Determine whether the company's securities are suitable for the client.
 - C. Adjust the financial statement data and compare the company to its industry peers.

KEY CONCEPTS

LOS 16.a

The role of financial reporting is to provide a variety of users with useful information about a company's performance and financial position.

The role of financial statement analysis is to use the data from financial statements to support economic decisions.

LOS 16.b

The statement of financial position (balance sheet) shows assets, liabilities, and owners' equity at a point in time.

The statement of comprehensive income shows the results of a firm's business activities over the period. Revenues, the cost of generating those revenues, and the resulting profit or loss are presented on the income statement.

The statement of changes in equity reports the amount and sources of changes in the equity owners' investment in the firm.

The statement of cash flows shows the sources and uses of cash over the period.

LOS 16.c

Important information about accounting methods, estimates, and assumptions is disclosed in the footnotes to the financial statements and supplementary schedules. These disclosures also contain information about segment results, commitments and contingencies, legal proceedings, acquisitions or divestitures, issuance of stock options, and details of employee benefit plans.

Management's commentary (Management's Discussion and Analysis) contains an overview of the company and important information about business trends, future capital needs, liquidity,

significant events, and significant choices of accounting methods requiring management judgment.

LOS 16.d

The objective of audits of financial statements is to provide an opinion on the statements' fairness and reliability.

The auditor's opinion gives evidence of an independent review of the financial statements that verifies that appropriate accounting principles were used, that standard auditing procedures were used to establish reasonable assurance that the statements contain no material errors, and that management's report on the company's internal controls has been reviewed.

An auditor can issue an unqualified (clean) opinion if the statements are free from material omissions and errors, a qualified opinion that notes any exceptions to accounting principles, an adverse opinion if the statements are not presented fairly in the auditor's opinion, or a disclaimer of opinion if the auditor is unable to express an opinion.

A company's management is responsible for maintaining an effective internal control system to ensure the accuracy of its financial statements.

LOS 16.e

Along with the annual financial statements, important information sources for an analyst include a company's quarterly and semiannual reports, proxy statements, press releases, and earnings guidance, as well as information on the industry and peer companies from external sources.

LOS 16.f

The framework for financial analysis has six steps:

1. State the objective of the analysis.
2. Gather data.
3. Process the data.
4. Analyze and interpret the data.
5. Report the conclusions or recommendations.
6. Update the analysis.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 16.1, 16.2

1. **B** This statement describes the role of an auditor, rather than the role of an analyst. The other responses describe the role of financial statement analysis. (Module 16.1, LOS 16.a)
2. **A** The balance sheet reports a company's financial position as of a specific date. The income statement, cash flow statement, and statement of changes in owners' equity show the company's performance during a specific period. (Module 16.1, LOS 16.b)

3. **B** Information about accounting methods and estimates is contained in the footnotes to the financial statements. (Module 16.2, LOS 16.c)
 4. **C** An auditor will issue a qualified opinion if the financial statements make any exceptions to applicable accounting standards and will explain the effect of these exceptions in the auditor's report. (Module 16.2, LOS 16.d)
 5. **B** Proxy statements contain information related to matters that come before shareholders for a vote, such as elections of board members. (Module 16.2, LOS 16.e)
 6. **B** Determining the suitability of an investment for a client is not one of the six steps in the financial statement analysis framework. The analyst would only perform this function if he also had an advisory relationship with the client. Stating the objective and processing the data are two of the six steps in the framework. The others are gathering the data, analyzing the data, updating the analysis, and reporting the conclusions. (Module 16.2, LOS 16.f)
-

1. Hennie van Greuning and Sonja Brajovic Bratanovic, *Analyzing and Managing Banking Risk: Framework for Assessing Corporate Governance and Financial Risk*, International Bank for Reconstruction and Development, April 2003, p. 300.

Reading 17

FINANCIAL REPORTING STANDARDS

EXAM FOCUS

This reading covers accounting standards: why they exist, who issues them, and who enforces them. Know the difference between the roles of standard-setting bodies and government regulatory authorities and be able to name the most important organizations of both kinds. The qualitative characteristics of, required elements for, and constraints on financial reporting presented in the IFRS' conceptual framework are important things to learn from this reading.

MODULE 17.1: STANDARDS OVERVIEW



LOS 17.a: Describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation.

Video covering this content is available online.

According to the IASB *Conceptual Framework for Financial Reporting*, the objective of financial reporting is to provide information about the firm to current and potential investors and creditors that is useful for making their decisions about investing in or lending to the firm.

The conceptual framework is used in the development of accounting standards. Given the variety and complexity of possible transactions and the estimates and assumptions a firm must make when presenting its performance, financial statements could potentially take any form if reporting standards did not exist. Thus, financial reporting standards are needed to provide consistency by narrowing the range of acceptable financial reports.

Reporting standards ensure that transactions are reported by firms similarly. However, standards must remain flexible and allow discretion to management to properly describe the economics of the firm.

Financial reporting is not designed solely for valuation purposes; however, it does provide important inputs for valuation purposes.

LOS 17.b: Describe the roles of financial reporting standard-setting bodies and regulatory authorities in establishing and enforcing reporting standards.

Standard-setting bodies are professional organizations of accountants and auditors that establish financial reporting standards. **Regulatory authorities** are government agencies that have the legal authority to enforce compliance with financial reporting standards.

The two primary standard-setting bodies are the *Financial Accounting Standards Board* (FASB) and the *International Accounting Standards Board* (IASB). In the United States, the FASB sets forth Generally Accepted Accounting Principles (GAAP). Outside the United States, the IASB establishes International Financial Reporting Standards (IFRS). Other national standard-setting

bodies exist as well. Some of the older IASB standards are referred to as International Accounting Standards (IAS).

Regulatory authorities, such as the *Securities and Exchange Commission* (SEC) in the United States and the *Financial Conduct Authority* in the United Kingdom, are established by national governments.

Most national authorities belong to the **International Organization of Securities Commissions** (IOSCO). Together, the members of IOSCO regulate more than 95% of the world's financial markets. IOSCO is not a regulatory body, but its members work together to make national regulations and enforcement more uniform around the world.

The SEC's requirements for financial reporting by U.S. companies are shown in Figure 17.1 as an example of reporting requirements. The SEC has the responsibility of enforcing the Sarbanes-Oxley Act of 2002. The act prohibits a company's external auditor from providing certain additional paid services to the company, to avoid the conflict of interest involved and to promote auditor independence. The act requires a company's executive management to certify that the financial statements are presented fairly and to include a statement about the effectiveness of the company's internal controls of financial reporting. Additionally, the external auditor must provide a statement confirming the effectiveness of the company's internal controls.

Figure 17.1: Securities and Exchange Commission Required Filings

Form S-1. Registration statement filed prior to the sale of new securities to the public. The registration statement includes audited financial statements, risk assessment, underwriter identification, and the estimated amount and use of the offering proceeds.

Form 10-K. Required annual filing that includes information about the business and its management, audited financial statements and disclosures, and disclosures about legal matters involving the firm. Information required in Form 10-K is similar to that which a firm typically provides in its annual report to shareholders. However, a firm's annual report is not a substitute for the required 10-K filing. Equivalent SEC forms for foreign issuers in the U.S. markets are Form 40-F for Canadian companies and Form 20-F for other foreign issuers.

Form 10-Q. U.S. firms are required to file this form quarterly, with updated financial statements (unlike Form 10-K, these statements do not have to be audited) and disclosures about certain events such as significant legal proceedings or changes in accounting policy. Non-U.S. companies are typically required to file the equivalent Form 6-K semiannually.

Form DEF-14A. When a company prepares a proxy statement for its shareholders prior to the annual meeting or other shareholder vote, it also files the statement with the SEC as Form DEF-14A.

Form 8-K. Companies must file this form to disclose material events including significant asset acquisitions and disposals, changes in management or corporate governance, or matters related to its accountants, its financial statements, or the markets in which its securities trade.

Form 144. A company can issue securities to certain qualified buyers without registering the securities with the SEC but must notify the SEC that it intends to do so.

Forms 3, 4, and 5 involve the beneficial ownership of securities by a company's officers and directors. Analysts can use these filings to learn about purchases and sales of company securities by corporate insiders.

In the European Union, each member state has its own securities regulations, but all countries in the EU are required to report using IFRS. The European Commission also has established the European Securities Commission, which advises the European Commission on securities

regulation issues, and the European Securities and Market Authority (ESMA), which coordinates regulation within the EU.



MODULE QUIZ 17.1

1. The objective of financial reporting, according to the IASB framework, is to:
 - A. provide information about the firm to current and potential investors.
 - B. decide the acceptable standards for presenting financial performance.
 - C. minimize management discretion in presenting the financial results of a firm.
2. Standard-setting bodies are responsible for:
 - A. establishing financial reporting standards only.
 - B. establishing and enforcing standards for financial reporting.
 - C. enforcing compliance with financial reporting standards only.
3. Which of the following organizations is *least likely* involved with enforcing compliance with financial reporting standards?
 - A. Financial Conduct Authority.
 - B. Securities and Exchange Commission.
 - C. International Accounting Standards Board.

MODULE 17.2: FINANCIAL REPORTING FRAMEWORK



Video covering
this content is
available online.

LOS 17.c: Describe the International Accounting Standards Board's conceptual framework, including qualitative characteristics of financial reports, constraints on financial reports, and required reporting elements.

The ideas on which the IASB bases its standards are expressed in the “Conceptual Framework for Financial Reporting” that the organization adopted in 2010 and revised in 2018. The IASB framework details the qualitative characteristics of financial statements and specifies the required reporting elements.

At the center of the IASB Conceptual Framework is the objective to provide financial information that is useful in making decisions about providing resources to an entity. The resource providers include investors, lenders, and other creditors. Users of financial statements need information about the firm’s performance, financial position, and cash flow.

Qualitative Characteristics

There are two fundamental characteristics that make financial information useful: relevance and faithful representation.¹

- *Relevance.* Financial statements are relevant if the information in them can influence users’ economic decisions or affect users’ evaluations of past events or forecasts of future events. To be relevant, information should have predictive value, confirmatory value (confirm prior expectations), or both. Materiality is an aspect of relevance.²
- *Faithful representation.* Information that is faithfully representative is complete, neutral (absence of bias), and free from error.

There are four characteristics that enhance relevance and faithful representation: comparability, verifiability, timeliness, and understandability.

- *Comparability*. Financial statement presentation should be consistent among firms and across time periods.
- *Verifiability*. Independent observers, using the same methods, obtain similar results.
- *Timeliness*. Information is available to decision makers before the information is stale.
- *Understandability*. Users with a basic knowledge of business and accounting and who make a reasonable effort to study the financial statements should be able to readily understand the information the statements present. Useful information should not be omitted just because it is complicated.

Required Reporting Elements

The elements of financial statements are the by-now familiar groupings of assets, liabilities, and owners' equity (for measuring financial position) and income and expenses (for measuring performance). The Conceptual Framework describes each of these elements:³

- *Assets*. Resources controlled as a result of past transactions that are expected to provide future economic benefits.
- *Liabilities*. Obligations as a result of past events that are expected to require an outflow of economic resources.
- *Equity*. The owners' residual interest in the assets after deducting the liabilities.
- *Income*. An increase in economic benefits, either increasing assets or decreasing liabilities in a way that increases owners' equity (but not including contributions by owners). Income includes revenues and gains.
- *Expenses*. Decreases in economic benefits, either decreasing assets or increasing liabilities in a way that decreases owners' equity (but not including distributions to owners). Losses are included in expenses.

An item should be *recognized* in its financial statement element if a future economic benefit from the item (flowing to or from the firm) is probable and the item's value or cost can be measured reliably.

The amounts at which items are reported in the financial statement elements depend on their **measurement base**. Measurement bases include *historical cost* (the amount originally paid for the asset), *amortized cost* (historical cost adjusted for depreciation, amortization, depletion, and impairment), *current cost* (the amount the firm would have to pay today for the same asset), *net realizable value* (the estimated selling price of the asset in the normal course of business minus the selling costs), *present value* (the discounted value of the asset's expected future cash flows), and *fair value* (the price at which an asset could be sold, or a liability transferred, in an orderly transaction between willing parties).



PROFESSOR'S NOTE

In upcoming readings, we will discuss these measurement bases in more detail and the situations in which each is appropriate.

Constraints and Assumptions

According to the Conceptual Framework, there is cost-benefit tradeoff of the enhancing characteristics.⁴ Accordingly, the benefit that users gain from the information should be greater than the cost of presenting it. Another constraint, not specifically mentioned in the Conceptual Framework, is the fact that non-quantifiable information about a company (its reputation, brand loyalty, capacity for innovation, etc.) cannot be captured directly in financial statements.

Two important underlying assumptions of financial statements are *accrual accounting* and *going concern*.⁵ Accrual accounting means that financial statements should reflect transactions at the time they actually occur, not necessarily when cash is paid. Going concern assumes the company will continue to exist for the foreseeable future.

LOS 17.d: Describe general requirements for financial statements under International Financial Reporting Standards (IFRS).

International Accounting Standard (IAS) No. 1 defines which financial statements are required and how they must be presented. The **required financial statements** are:

- Balance sheet (statement of financial position).
- Statement of comprehensive income.
- Cash flow statement.
- Statement of changes in owners' equity.
- Explanatory notes, including a summary of accounting policies.

The general **features for preparing financial statements** are stated in IAS No. 1:

- *Fair presentation*, defined as faithfully representing the effects of the entity's transactions and events according to the standards for recognizing assets, liabilities, revenues, and expenses.
- *Going concern basis*, meaning the financial statements are based on the assumption that the firm will continue to exist unless its management intends to (or must) liquidate it.
- *Accrual basis* of accounting is used to prepare the financial statements other than the statement of cash flows.
- *Consistency* between periods in how items are presented and classified, with prior-period amounts disclosed for comparison.
- *Materiality*, meaning the financial statements should be free of misstatements or omissions that could influence the decisions of users of financial statements.
- *Aggregation* of similar items and separation of dissimilar items.
- *No offsetting* of assets against liabilities or income against expenses unless a specific standard permits or requires it.
- *Reporting frequency* must be at least annually.
- *Comparative information* for prior periods should be included unless a specific standard states otherwise.

Also stated in IAS No. 1 are the **structure and content of financial statements**:

- Most entities should present a *classified balance sheet* showing current and noncurrent assets and liabilities.

- *Minimum information* is required on the face of each financial statement and in the notes. For example, the face of the balance sheet must show specific items such as cash and cash equivalents, plant, property and equipment, and inventories. Items listed on the face of the comprehensive income statement must include revenue, profit or loss, tax expense, and finance costs, among others.
- *Comparative information* for prior periods should be included unless a specific standard states otherwise.

LOS 17.e: Describe implications for financial analysis of alternative financial reporting systems and the importance of monitoring developments in financial reporting standards.

As financial reporting standards continue to evolve, analysts need to monitor how these developments will affect the financial statements they use. An analyst should be aware of new products and innovations in the financial markets that generate new types of transactions. These might not fall neatly into the existing financial reporting standards. The analyst can use the financial reporting framework as a guide for evaluating what effect new products or transactions might have on financial statements.

To keep up to date on the evolving standards, an analyst can monitor professional journals and other sources, such as the IASB (www.ifrs.org) and FASB (www.fasb.org) websites. CFA Institute produces position papers on financial reporting issues through the CFA Institute Centre for Financial Market Integrity.

Finally, analysts must monitor company disclosures for significant accounting standards and estimates.



MODULE QUIZ 17.2

1. According to the IASB Conceptual Framework, the fundamental qualitative characteristics that make financial statements useful are:
 - A. verifiability and timeliness.
 - B. relevance and faithful representation.
 - C. understandability and relevance.
2. Which of the following *most accurately* lists a required reporting element that is used to measure a company's financial position and one that is used to measure a company's performance?

<u>Position</u>	<u>Performance</u>
A. Assets	Liabilities
B. Income	Expenses
C. Liabilities	Income
3. International Accounting Standard (IAS) No. 1 *least likely* requires which of the following?
 - A. Neither assets and liabilities, nor income and expenses, may be offset unless required or permitted by a financial reporting standard.
 - B. Audited financial statements and disclosures, along with updated information about the firm and its management, must be filed at least quarterly.
 - C. Fair presentation of financial statements means faithfully representing the firm's events and transactions according to the financial reporting standards.

KEY CONCEPTS

LOS 17.a

The objective of financial statements is to provide economic decision makers with useful information about a firm's financial performance and changes in financial position.

Reporting standards are designed to ensure that different firms' statements are comparable to one another and to narrow the range of reasonable estimates on which financial statements are based. This aids users of the financial statements who rely on them for information about the company's activities, profitability, and creditworthiness.

LOS 17.b

Standard-setting bodies are private sector organizations that establish financial reporting standards. The two primary standard-setting bodies are the International Accounting Standards Board (IASB) and, in the United States, the Financial Accounting Standards Board (FASB).

Regulatory authorities are government agencies that enforce compliance with financial reporting standards. Regulatory authorities include the Securities and Exchange Commission in the United States and the Financial Conduct Authority in the United Kingdom. Many national regulatory authorities belong to the International Organization of Securities Commissions.

LOS 17.c

The IFRS "Conceptual Framework for Financial Reporting" defines the fundamental and enhancing qualitative characteristics of financial statements, specifies the required reporting elements, and notes the constraints and assumptions involved in preparing financial statements.

- The fundamental characteristics of financial statements are relevance and faithful representation. The enhancing characteristics include comparability, verifiability, timeliness, and understandability.
- Elements of financial statements are assets, liabilities, and owners' equity (for measuring financial position) and income and expenses (for measuring performance).
- Constraints on financial statement preparation include cost versus benefit and the difficulty of capturing non-quantifiable information in financial statements.
- The two primary assumptions that underlie the preparation of financial statements are the accrual basis and the going concern assumption.

LOS 17.d

Required financial statements are the balance sheet, comprehensive income statement, cash flow statement, statement of changes in owners' equity, and explanatory notes.

The general features of financial statements according to IAS No. 1 are:

- Fair presentation.
- Going concern.
- Accrual accounting.
- Consistency.
- Materiality.

- Aggregation.
- No offsetting.
- Reporting frequency.
- Comparative information.

Other presentation requirements include a classified balance sheet and specific minimum information that must be reported in the notes and on the face of the financial statements.

LOS 17.e

An analyst should be aware of evolving financial reporting standards and new products and innovations that generate new types of transactions.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 17.1

- A** The IASB Conceptual Framework states that the objective of financial reporting is to provide information about the firm to current and potential investors that is useful for making decisions about investing in or lending to the firm. (LOS 17.a)
- A** Standard-setting bodies are private-sector organizations that establish financial reporting standards. Enforcement is the responsibility of regulatory authorities. (LOS 17.b)
- C** The IASB is a standard-setting body. The Securities and Exchange Commission (in the United States) and the Financial Conduct Authority (in the United Kingdom) are regulatory authorities. (LOS 17.b)

Module Quiz 17.2

- B** The fundamental qualitative characteristics are relevance and faithful representation. Verifiability, timeliness, and understandability are enhancing qualitative characteristics. (LOS 17.c)
- C** Balance sheet reporting elements (assets, liabilities, and owners' equity) measure a company's financial position. Income statement reporting elements (income, expenses) measure its financial performance. (LOS 17.c)
- B** According to IAS No. 1, financial statements must be presented at least *annually*. Fair presentation is one of the IAS No. 1 principles for preparing financial statements. The ban against offsetting is one of the IAS No. 1 principles for presenting financial statements. (LOS 17.d)

1. *Conceptual Framework for Financial Reporting (2010)*. paragraphs QC5–18.
2. Ibid., paragraphs QC19–34.
3. Ibid., paragraphs 4.4–4.23.
4. Ibid., paragraphs QC35–39.

5. Ibid., paragraphs 0B17 and 4.1.

Reading 18

UNDERSTANDING INCOME STATEMENTS

EXAM FOCUS

Now we're getting to the heart of the matter. Since forecasts of future earnings, and therefore estimates of firm value, depend crucially on understanding a firm's income statement, everything in this reading is important. Some of the items requiring calculation include depreciation, COGS, and inventory under different cost flow assumptions, as well as basic and diluted EPS. The separation of items into operating and non-operating categories is important when estimating recurring income as a first step in forecasting future firm earnings. Note that questions regarding the effect on financial ratios of the choice of accounting method and of accounting estimates are one common way to test your understanding of the material on those topics presented here.

MODULE 18.1: INCOME STATEMENT OVERVIEW



LOS 18.a: Describe the components of the income statement and alternative presentation formats of that statement.

Video covering this content is available online.

The income statement reports the revenues and expenses of the firm over a period of time. The income statement is sometimes referred to as the *statement of operations*, the *statement of earnings*, or the *profit and loss statement (P&L)*. The income statement equation is:

$$\text{revenues} - \text{expenses} = \text{net income}$$

Under both U.S. GAAP and IFRS, the income statement and a *statement of other comprehensive income* can be presented separately or presented together as a single *statement of comprehensive income*. Investors examine a firm's income statement for valuation purposes while lenders examine the income statement for information about the firm's ability to make the promised interest and principal payments on its debt.

Revenues are the amounts reported from the sale of goods and services in the normal course of business. Revenue less adjustments for estimated returns and allowances is known as **net revenue**. Details about the presentation of revenue can be found in the footnotes of the financial statements or sometimes in the MD&A.



PROFESSOR'S NOTE

The terms *revenue* and *sales* are sometimes used synonymously. However, sales is just one component of revenue in many firms. In some countries, revenues are referred to as "turnover."

Expenses are the amounts incurred to generate revenue and include cost of goods sold, operating expenses, interest, and taxes. Expenses are grouped together by their nature or

function. Presenting all depreciation expense from manufacturing and administration together in one line of the income statement is an example of grouping by nature of the expense. Combining all costs associated with manufacturing (e.g., raw materials, depreciation, labor, etc.) as cost of goods sold is an example of grouping by function. Grouping expenses by function is sometimes referred to as the cost of sales method.



PROFESSOR'S NOTE

Firms can present columnar data in chronological order from left-to-right or vice versa. Also, some firms present expenses as negative numbers while other firms use parentheses to signify expenses. Still other firms present expenses as positive numbers with the assumption that users know that expenses are subtracted in the income statement. Watch for these different treatments on the exam.

The income statement also includes **gains and losses**, which result in an increase (gains) or decrease (losses) of economic benefits. Gains and losses may or may not result from ordinary business activities. For example, a firm might sell surplus equipment used in its manufacturing operation that is no longer needed. The difference between the sales price and book value is reported as a gain or loss on the income statement. Summarizing, net income is equal to income (revenues + gains) minus expenses (including losses). Thus, the components can be rearranged as follows:

$$\text{net income} = \text{revenues} - \text{ordinary expenses} + \text{other income} - \text{other expense} + \text{gains} - \text{losses}$$

When a firm has a controlling interest in a subsidiary, the statements of the two firms are *consolidated*; the earnings of both firms are included on the income statement. In this case, the share (proportion) of the subsidiary's income not owned by the parent is reported in parent's income statement as the **noncontrolling interest** (also known as **minority interest** or **minority owners' interest**). The noncontrolling interest is subtracted from the consolidated total income to get the net income of the parent company.

Presentation Formats

A firm can present its income statement using a single-step or multi-step format. In a single-step statement, all revenues are grouped together and all expenses are grouped together. A multi-step format includes *gross profit*, revenues minus cost of goods sold.

Figure 18.1 is an example of a multi-step income statement format for the BHG Company.

Figure 18.1: Multi-Step Income Statement

**BHG Company Income Statement
For the year ended December 31, 20X7**

Revenue	\$579,312
Cost of goods sold	<u>(362,520)</u>
Gross profit	216,792
Selling, general, and administrative expense	(109,560)
Depreciation expense	<u>(69,008)</u>
Operating profit	38,224
Interest expense	<u>(2,462)</u>
Income before tax	35,762
Provision for income taxes	<u>(14,305)</u>
Income from continuing operations	21,457
Earnings (losses) from discontinued operations, net of tax	<u>1,106</u>
Net income	<u>\$22,563</u>

Gross profit is the amount that remains after the direct costs of producing a product or service are subtracted from revenue. Subtracting operating expenses, such as selling, general, and administrative expenses, from gross profit results in another subtotal known as **operating profit** or operating income. For nonfinancial firms, operating profit is profit before financing costs, income taxes, and non-operating items are considered. Subtracting interest expense and income taxes from operating profit results in the firm's net income, sometimes referred to as "earnings" or the "bottom line."



PROFESSOR'S NOTE

Interest expense is usually considered an operating expense for financial firms. Although there may be some differences between operating income and earnings before interest and taxes (EBIT), we often use EBIT as a proxy for operating income in analysis.

MODULE 18.2: REVENUE RECOGNITION



LOS 18.b: Describe general principles of revenue recognition and accounting standards for revenue recognition.

Video covering this content is available online.

LOS 18.c: Calculate revenue given information that might influence the choice of revenue recognition method.

In a sale of goods where the goods are exchanged for cash and returns are not allowed, the recognition of revenue is straightforward: it is recognized at the time of the exchange. The recognition of revenue is not, however, dependent on receiving cash payment. If a sale of goods is made on credit, revenue can be recognized at the time of sale, and an asset, **accounts receivable**, is created on the balance sheet.

If payment for the goods is received prior to the transfer of the goods, a liability, **unearned revenue**, is created when the cash is received (offsetting the increase in the asset *cash*.) Revenue

is recognized as the goods are transferred to the buyer. As an example, consider a magazine subscription; when the subscription is purchased, an unearned revenue liability is created, and as magazine issues are delivered, revenue is recorded and the liability is decreased.

Converged standards under IFRS and U.S. GAAP take a principles-based approach to revenue recognition issues. The central principle is that a firm should recognize revenue when it has transferred a good or service to a customer. This is consistent with the familiar accrual accounting principle that revenue should be recognized when earned.

The converged standards identify a five-step process¹ for recognizing revenue:

1. Identify the contract(s) with a customer.
2. Identify the separate or distinct performance obligations in the contract.
3. Determine the transaction price.
4. Allocate the transaction price to the performance obligations in the contract.
5. Recognize revenue when (or as) the entity satisfies a performance obligation.

The standard defines a **contract** as an agreement between two or more parties that specifies their obligations and rights. Collectability must be probable for a contract to exist, but “probable” is defined differently under IFRS and U.S. GAAP so an identical activity could still be accounted for differently by IFRS and U.S. GAAP reporting firms.

A **performance obligation** is a promise to deliver a distinct good or service. A “distinct” good or service is one that meets the following criteria:

- The customer can benefit from the good or service on its own or combined with other resources that are readily available.
- The promise to transfer the good or service can be identified separately from any other promises.

A **transaction price** is the amount a firm expects to receive from a customer in exchange for transferring a good or service to the customer. A transaction price is usually a fixed amount but can also be variable, for example, if it includes a bonus for early delivery.

A firm should recognize revenue only when it is highly probable they will not have to reverse it. For example, a firm may need to recognize a liability for a refund obligation (and an offsetting asset for the right to returned goods) if revenue from a sale cannot be estimated reliably.

For long-term contracts, revenue is recognized based on a firm’s progress toward completing a performance obligation. Progress toward completion can be measured from the input side (e.g., using the percentage of completion costs incurred as of the statement date). Progress can also be measured from the output side, using engineering milestones or percentage of the total output delivered to date.

The following summarizes some examples from IFRS 15 of appropriate revenue recognition under various circumstances.

EXAMPLE: Revenue recognition

1. Performance obligation and progress towards completion

A contractor agrees to build a warehouse for a price of \$10 million and estimates the total costs of construction at \$8 million. Although there are several *identifiable components* of the building (site preparation, foundation, electrical components, roof, etc.), these components are not *separate deliverables*, and the performance obligation is the completed building.

During the first year of construction, the builder incurs \$4 million of costs, 50% of the estimated total costs of completion. Based on this expenditure and a belief that the percentage of costs incurred represents an appropriate measure of progress towards completing the performance obligation, the builder recognizes \$5 million (50% of the transaction price of \$10 million) as revenue for the year. This treatment is consistent with the percentage-of-completion method previously in use, although the new standards do not call it that.

2. Variable consideration—performance bonus

Consider this construction contract with the addition of a promised bonus payment of \$1 million if the building is completed in three years. At the end of the first year, the contractor has some uncertainty about whether he can complete building by the end of the third year because of environmental concerns. Because revenue should be recognized only when it is *highly probable that it will not be reversed*, the builder does not consider the possible bonus as part of the transaction price. In this case, year 1 revenue is still \$5 million, calculated just as we did previously.

During the second year of construction, the contractor incurred an additional \$2 million in costs and the environmental concerns have been resolved. The contractor has no doubt that the building will be finished in time to receive the bonus payment.

The percentage of total costs incurred over the first two years is now $(\$4 \text{ million} + \$2 \text{ million}) / \$8 \text{ million} = 75\%$. The total revenue to be recognized to date, with the bonus payment included in transaction value, is $0.75 \times \$11 \text{ million} = \8.25 million . Because \$5 million of revenue had been recognized in year 1, \$3.25 million ($= \$8.25 \text{ million} - \5 million) of revenue will be recognized in year 2.

3. Contract revisions

Contracts are often changed over the construction period. The issue for revenue recognition is whether to treat a contract modification as an extension of the existing contract or as a new contract. Returning to our example, a contract revision requires installation of refrigeration to provide cold storage in part of the warehouse. In this case, the contract revision should be considered an extension of the existing contract because the goods and services to be provided are not *distinct from those already transferred*.

The contractor agrees to the revisions during the second year of construction and believes they will increase his costs by \$2 million, to \$10 million. The transaction value is increased by \$3 million, to \$14 million, including the bonus, which he believes is still the appropriate treatment.

As before, the contractor has incurred \$6 million in costs through the end of the second year. Now he calculates the percentage of the contract obligations completed to be $\$6 \text{ million} / \$10 \text{ million} = 60\%$. The total revenue to be recognized to date is $60\% \times \$14 \text{ million} = \8.4 million . He will report \$3.4 million ($= \$8.4 \text{ million} - \5 million) of revenue for the second year.

4. Acting as an agent

Consider a travel agent who arranges a first-class ticket for a customer flying to Singapore. The ticket price is \$10,000, made by nonrefundable payment at purchase, and the travel agent receives a \$1,000 commission on the sale. Because the travel agent is not responsible for providing the flight and bears no inventory or credit risk, she is *acting as an agent*. Because she is an agent, rather than a *principal*, she should report revenue equal to her commission of \$1,000, the net amount of the sale. If she were a principal in the transaction, she would report revenue of \$10,000, the gross amount of the sale, and an expense of \$9,000 for the ticket.

The costs to secure a long-term contract, such as sales commissions, must be capitalized; that is, the expense for these costs is spread over the life of the contract.

There are a significant number of required disclosures under the converged standards. They include:

- Contracts with customers by category.

- Assets and liabilities related to contracts, including balances and changes.
- Outstanding performance obligations and the transaction prices allocated to them.
- Management judgments used to determine the amount and timing of revenue recognition, including any changes to those judgments.



MODULE QUIZ 18.1, 18.2

1. For a nonfinancial firm, are depreciation expense and interest expense included or excluded from operating expenses in the income statement?

<u>Depreciation expense</u>	<u>Interest expense</u>
A. Included	Included
B. Included	Excluded
C. Excluded	Included

2. Which of the following expense items is *best* described as being classified by function rather than classified in nature?
 - A. Cost of goods sold.
 - B. Depreciation.
 - C. Wage expense.
3. The first step in the revenue recognition process is to:
 - A. determine the price.
 - B. identify the contract.
 - C. identify the obligations.

MODULE 18.3: EXPENSE RECOGNITION



LOS 18.d: Describe general principles of expense recognition, specific expense recognition applications, and implications of expense recognition choices for financial analysis.

Video covering
this content is
available online.

Expenses are subtracted from revenue to calculate net income. According to the IASB, expenses are decreases in economic benefits during the accounting period in the form of outflows or depletions of assets or incurrence of liabilities that result in decreases in equity other than those relating to distributions to equity participants.²

If the financial statements were prepared on a cash basis, neither revenue recognition nor expense recognition would be an issue. The firm would simply recognize cash received as revenue and cash payments as expense.

Under the accrual method of accounting, expense recognition is based on the **matching principle** whereby expenses to generate revenue are recognized in the same period as the revenue. Inventory provides a good example. Assume inventory is purchased during the fourth quarter of one year and sold during the first quarter of the following year. Using the matching principle, both the revenue and the expense (cost of goods sold) are recognized in the first quarter, when the inventory is sold, not the period in which the inventory was purchased.

Not all expenses can be directly tied to revenue generation. These costs are known as **period costs**. Period costs, such as administrative costs, are expensed in the period incurred.

Inventory Expense Recognition

If a firm can identify exactly which items were sold and which items remain in inventory, it can use the **specific identification** method. For example, an auto dealer records each vehicle sold or in inventory by its identification number.

Under the **first-in, first-out** (FIFO) method, the first item purchased is assumed to be the first item sold. The cost of inventory acquired first (beginning inventory and early purchases) is used to calculate the cost of goods sold for the period. The cost of the most recent purchases is used to calculate ending inventory. FIFO is appropriate for inventory that has a limited shelf life. For example, a food products company will sell its oldest inventory first to keep the inventory on hand fresh.

Under the **last-in, first-out** (LIFO) method, the last item purchased is assumed to be the first item sold. The cost of inventory most recently purchased is assigned to the cost of goods sold for the period. The costs of beginning inventory and earlier purchases are assigned to ending inventory. LIFO is appropriate for inventory that does not deteriorate with age. For example, a coal distributor will sell coal off the top of the pile.

In the United States, LIFO is popular because of its income tax benefits. In an inflationary environment, LIFO results in higher cost of goods sold. Higher cost of goods sold results in lower taxable income and, therefore, lower income taxes.

The **weighted average cost** method makes no assumption about the physical flow of the inventory. It is popular because of its ease of use. The cost per unit is calculated by dividing cost of available goods by total units available, and this average cost is used to determine both cost of goods sold and ending inventory. Average cost results in cost of goods sold and ending inventory values between those of LIFO and FIFO.

FIFO and average cost are permitted under both U.S. GAAP and IFRS. LIFO is allowed under U.S. GAAP but is prohibited under IFRS.

Figure 18.2 summarizes the effects of the inventory methods.

Figure 18.2: Inventory Method Comparison

Method	Assumption	Cost of Goods Sold Consists of...	Ending Inventory Consists of...
FIFO (U.S. and IFRS)	The items first purchased are the first to be sold.	first purchased	most recent purchases
LIFO (U.S. only)	The items last purchased are the first to be sold.	last purchased	earliest purchases
Weighted average cost (U.S. and IFRS)	Items sold are a mix of purchases.	average cost of all items	average cost of all items



PROFESSOR'S NOTE

We will illustrate how to calculate inventory and cost of goods sold using each of these three cost flow assumptions in our reading on Inventories.

Depreciation Expense Recognition

The cost of long-lived assets must also be matched with revenues. Long-lived assets are expected to provide economic benefits beyond one accounting period. The allocation of cost over an asset's life is known as depreciation (tangible assets), depletion (natural resources), or amortization (intangible assets). Most firms use the **straight-line depreciation** method for financial reporting purposes. The straight-line method recognizes an equal amount of depreciation expense each period. However, most assets generate more benefits in the early years of their economic life and fewer benefits in the later years. In this case, an *accelerated depreciation method* is more appropriate for matching the expenses to revenues.

In the early years of an asset's life, the straight-line method will result in lower depreciation expense as compared to an accelerated method. Lower expense results in higher net income. In the later years of the asset's life, the effect is reversed, and straight-line depreciation results in higher expense and lower net income compared to accelerated methods.

Straight-line depreciation (SL) allocates an equal amount of depreciation each year over the asset's useful life as follows:

$$\text{SL depreciation expense} = \frac{\text{cost} - \text{residual value}}{\text{useful life}}$$

Accelerated depreciation speeds up the recognition of depreciation expense in a systematic way to recognize more depreciation expense in the early years of the asset's life and less depreciation expense in the later years of its life. Total depreciation expense over the life of the asset will be the same as it would be if straight-line depreciation were used.

The **declining balance method** (DB) applies a constant rate of depreciation to an asset's (declining) book value each year.



PROFESSOR'S NOTE

The declining balance method is also known as the diminishing balance method.

The most common declining balance method is *double-declining balance* (DDB), which applies two times the straight-line rate to the declining balance. If an asset's life is ten years, the straight-line rate is 1/10 or 10%, and the DDB rate would be 2/10 or 20%.

$$\text{DDB depreciation} = \left(\frac{2}{\text{useful life}} \right) (\text{cost} - \text{accumulated depreciation})$$

DB does not explicitly use the asset's residual value in the calculations, but depreciation ends once the estimated residual value has been reached. If the asset is expected to have no residual value, the DB method will never fully depreciate it, so the DB method is typically changed to straight-line at some point in the asset's life.



PROFESSOR'S NOTE

We will illustrate how to calculate depreciation expense in our reading on Long-Lived Assets.

Amortization Expense Recognition

Amortization is the allocation of the cost of an intangible asset (such as a franchise agreement) over its useful life. Amortization expense should match the proportion of the asset's economic benefits used during the period. Most firms use the straight-line method to calculate annual amortization expense for financial reporting. Straight-line amortization is calculated exactly like straight-line depreciation.

Intangible assets with indefinite lives (e.g., goodwill) are not amortized. However, they must be tested for impairment at least annually. If the asset value is impaired, an expense equal to the impairment amount is recognized on the income statement.

Bad Debt Expense and Warranty Expense Recognition

If a firm sells goods or services on credit or provides a warranty to the customer, the matching principle requires the firm to estimate bad debt expense and/or warranty expense. By doing so, the firm is recognizing the expense in the period of the sale, rather than a later period.

Implications for Financial Analysis

Like revenue recognition, expense recognition requires a number of estimates. Since estimates are involved, it is possible for firms to delay or accelerate the recognition of expenses. Delayed expense recognition increases current net income and is therefore more aggressive.

Analysts must consider the underlying reasons for a change in an expense estimate. If a firm's bad debt expense has recently decreased, did the firm lower its expense estimate because its collection experience improved, or was the expense decreased to manipulate net income?

Analysts should also compare a firm's estimates to those of other firms within the firm's industry. If a firm's warranty expense is significantly less than that of a peer firm, is the lower warranty expense a result of higher quality products, or is the firm's expense recognition more aggressive than that of the peer firm?

Firms disclose their accounting policies and significant estimates in the financial statement footnotes and in the management discussion and analysis (MD&A) section of the annual report.

LOS 18.e: Describe the financial reporting treatment and analysis of non-recurring items (including discontinued operations, unusual or infrequent items) and changes in accounting policies.

Non-Recurring Items

Discontinued operations. A *discontinued operation* is one that management has decided to dispose of, but either has not yet done so, or has disposed of in the current year after the operation had generated income or losses. To be accounted for as a discontinued operation, the business—in terms of assets, operations, and investing and financing activities—must be physically and operationally distinct from the rest of the firm.

The date when the company develops a formal plan for disposing of an operation is referred to as the *measurement date*, and the time between the measurement period and the actual disposal date is referred to as the *phaseout period*. Any income or loss from discontinued operations is reported separately in the income statement, net of tax, after income from continuing operations. Any past income statements presented must be restated, separating the income or

loss from the discontinued operations. On the measurement date, the company will accrue any estimated loss during the phaseout period and any estimated loss on the sale of the business. Any expected gain on the disposal cannot be reported until after the sale is completed.

Analytical implications: The analysis is straightforward. Discontinued operations do not affect net income from continuing operations. For this reason, analysts may exclude discontinued operations when forecasting future earnings. The actual event of discontinuing a business segment or selling assets may provide information about the future cash flows of the firm, however.

Unusual or infrequent items. The definition of these items is obvious—these events are either unusual in nature or infrequent in occurrence. Examples of items that could be considered unusual or infrequent include:

- Gains or losses from the sale of assets or part of a business, if these activities are not a firm's ordinary operations.
- Impairments, write-offs, write-downs, and restructuring costs.

Unusual or infrequent items are included in income from continuing operations and are reported before tax.

Analytical implications: Even though unusual or infrequent items affect net income from continuing operations, an analyst may want to review them to determine whether they truly should be included when forecasting future firm earnings. Some companies appear to be accident-prone and have “unusual or infrequent” losses every year or every few years.

Changes in Accounting Policies and Estimates

Accounting changes include changes in accounting policies, changes in accounting estimates, and prior-period adjustments. Such changes may require either **retrospective application** or **prospective application**. With retrospective application, any prior-period financial statements presented in a firm's current financial statements must be restated, applying the new policy to those statements as well as future statements. Retrospective application enhances the comparability of the financial statements over time. With prospective application, prior statements are not restated, and the new policies are applied only to future financial statements.

Standard setting bodies, at times, issue a **change in accounting policy**. Sometimes a firm may change which accounting policy it applies, for example, by changing its inventory costing method or capitalizing rather than expensing specific purchases. Unless it is impractical, changes in accounting policies require retrospective application.

In the recent change to revenue recognition standards, firms were given the option of *modified retrospective application*. This application does not require restatement of prior-period statements; however, beginning values of affected accounts are adjusted for the cumulative effects of the change.

Generally, a **change in accounting estimate** is the result of a change in management's judgment, usually due to new information. For example, management may change the estimated useful life of an asset because new information indicates the asset has a longer or shorter life than originally expected. Changes in accounting estimates are applied prospectively and do not require the restatement of prior financial statements.

Analytical implications: Accounting estimate changes typically do not affect cash flow. An analyst should review changes in accounting estimates to determine their impact on future operating results.

Sometimes a change from an incorrect accounting method to one that is acceptable under GAAP or IFRS is required. A correction of an accounting error made in previous financial statements is reported as a **prior-period adjustment** and requires retrospective application. Prior-period results are restated. Disclosure of the nature of any significant prior-period adjustment and its effect on net income is also required.

Analytical implications: Prior-period adjustments usually involve errors or new accounting standards and do not typically affect cash flow. Analysts should review adjustments carefully because errors may indicate weaknesses in the firm's internal controls.

LOS 18.f: Contrast the operating and non-operating components of the income statement.

Operating and nonoperating transactions are usually reported separately in the income statement. For a nonfinancial firm, nonoperating transactions may result from investment income and financing expenses. For example, a nonfinancial firm may receive dividends and interest from investments in other firms. The investment income and any gains and losses from the sale of these securities are not a part of the firm's normal business operations. Interest expense is based on the firm's capital structure, which is also independent of the firm's operations. Conversely, for a financial firm, investment income and financing expenses are usually considered operating activities.



MODULE QUIZ 18.3

- When accounting for inventory, are the first-in, first-out (FIFO) and last-in, first-out (LIFO) cost flow assumptions permitted under U.S. GAAP?

<u>FIFO</u>	<u>LIFO</u>
A. Yes	Yes
B. Yes	No
C. No	Yes

- Which of the following *best* describes the impact of depreciating equipment with a useful life of 6 years and no salvage value using the declining balance method as compared to the straight-line method?
 - Total depreciation expense will be higher over the life of the equipment.
 - Depreciation expense will be higher in the first year.
 - Scraping the equipment after five years will result in a larger loss.

- CC Corporation reported the following inventory transactions (in chronological order) for the year:

Purchase	Sales
40 units at \$30	13 units at \$35
20 units at \$40	35 units at \$45
90 units at \$50	60 units at \$60

Assuming inventory at the beginning of the year was zero, calculate the year-end inventory using FIFO and LIFO.

<u>FIFO</u>	<u>LIFO</u>
A. \$5,220	\$1,040
B. \$2,100	\$1,280
C. \$2,100	\$1,040

4. At the beginning of the year, Triple W Corporation purchased a new piece of equipment to be used in its manufacturing operation. The cost of the equipment was \$25,000. The equipment is expected to be used for 4 years and then sold for \$4,000. Depreciation expense to be reported for the second year using the double-declining-balance method is *closest* to:
- A. \$5,250.
 - B. \$6,250.
 - C. \$7,000.
5. Changing an accounting estimate:
- A. is reported prospectively.
 - B. requires restatement of all prior-period statements presented in the current financial statements.
 - C. is reported by adjusting the beginning balance of retained earnings for the cumulative effect of the change.
6. Which of the following transactions would *most likely* be reported below income from continuing operations, net of tax?
- A. Gain or loss from the sale of equipment used in a firm's manufacturing operation.
 - B. A change from the accelerated method of depreciation to the straight-line method.
 - C. The operating income of a physically and operationally distinct division that is currently for sale, but not yet sold.
7. Which of the following statements about nonrecurring items is *least accurate*?
- A. Discontinued operations are reported net of taxes at the bottom of the income statement before net income.
 - B. Unusual or infrequent items are reported before taxes above net income from continuing operations.
 - C. A change in accounting principle is reported in the income statement net of taxes after extraordinary items and before net income.
8. Which of the following is *least likely* considered a nonoperating transaction from the perspective of a manufacturing firm?
- A. Dividends received from available-for-sale securities.
 - B. Interest expense on subordinated debentures.
 - C. Accruing bad debt expense for goods sold on credit.

MODULE 18.4: EPS AND DILUTIVE SECURITIES



LOS 18.g: Describe how earnings per share is calculated and calculate and interpret a company's earnings per share (both basic and diluted earnings per share) for both simple and complex capital structures.

Video covering this content is available online.

LOS 18.h: Contrast dilutive and antidilutive securities and describe the implications of each for the earnings per share calculation.

Earnings per share (EPS) is one of the most commonly used corporate profitability performance measures for publicly-traded firms (nonpublic companies are not required to

report EPS data). EPS is reported only for shares of common stock (also known as ordinary stock).

A company may have either a simple or complex capital structure:

- A **simple capital structure** is one that contains *no* potentially dilutive securities. A simple capital structure contains only common stock, nonconvertible debt, and nonconvertible preferred stock.
- A **complex capital structure** contains *potentially dilutive securities* such as options, warrants, or convertible securities.

All firms with complex capital structures must report both *basic* and *diluted* EPS. Firms with simple capital structures report only basic EPS.

Basic EPS

The basic EPS calculation does not consider the effects of any dilutive securities in the computation of EPS.

$$\text{basic EPS} = \frac{\text{net income} - \text{preferred dividends}}{\text{weighted average number of common shares outstanding}}$$

The current year's preferred dividends are subtracted from net income because EPS refers to the per-share earnings *available to common shareholders*. Net income minus preferred dividends is the income available to common stockholders. Common stock dividends are *not* subtracted from net income because they are a part of the net income available to common shareholders.

The **weighted average number of common shares** is the number of shares outstanding during the year, weighted by the portion of the year they were outstanding.

Effect of Stock Dividends and Stock Splits

A **stock dividend** is the distribution of additional shares to each shareholder in an amount proportional to their current number of shares. If a 10% stock dividend is paid, the holder of 100 shares of stock would receive 10 additional shares.

A **stock split** refers to the division of each "old" share into a specific number of "new" (post-split) shares. The holder of 100 shares will have 200 shares after a 2-for-1 split or 150 shares after a 3-for-2 split.

The important thing to remember is that each shareholder's proportional ownership in the company is unchanged by either of these events. Each shareholder has more shares but the same percentage of the total shares outstanding.



PROFESSOR'S NOTE

For our purposes here, a stock dividend and a stock split are two ways of doing the same thing. For example, a 50% stock dividend and a 3-for-2 stock split both result in three "new" shares for every two "old" shares. Stock dividends and stock splits are explained further in the Equity Investments topic area.

EXAMPLE: Weighted average shares outstanding

Johnson Company has 10,000 shares outstanding at the beginning of the year. On April 1, Johnson issues 4,000 new shares. On July 1, Johnson distributes a 10% stock dividend. On September 1, Johnson repurchases 3,000 shares. Calculate Johnson's weighted average number of shares outstanding for the year, for its reporting of basic earnings per share.

Answer:

Shares outstanding are weighted by the portion of the year the shares were outstanding. Any shares that were outstanding before the 10% stock dividend must be adjusted for it. Transactions that occur after the stock dividend do not need to be adjusted.



PROFESSOR'S NOTE

Think of the shares before the stock dividend as "old" shares and shares after the stock dividend as "new" shares that each represent ownership of a smaller portion of the company, in this example 10/11ths of that of an old (pre-stock dividend) share. The weighted average number of shares for the year will be in new shares.

Shares outstanding on January 1: $10,000 \times 1.10 \times 12/12$ of the year	= 11,000
Shares issued April 1: $4,000 \times 1.10 \times 9/12$ of the year	= 3,300
Shares repurchased September 1: $-3,000 \times 4/12$ of the year	= <u>-1,000</u>
Weighted average shares outstanding	= 13,300

EXAMPLE: Basic earnings per share

Johnson Company has net income of \$10,000, paid \$1,000 cash dividends to its preferred shareholders, and paid \$1,750 cash dividends to its common shareholders. Calculate Johnson's basic EPS using the weighted average number of shares from the previous example.

Answer:

$$\text{basic EPS} = \frac{\text{net income} - \text{pref. div.}}{\text{wt. avg. shares of common}} = \frac{\$10,000 - \$1,000}{13,300} = \$0.68$$



PROFESSOR'S NOTE

Remember, the payment of a cash dividend on common shares is not considered in the calculation of EPS.

Things to know about the weighted average shares outstanding calculation:

- The weighting system is days outstanding divided by the number of days in a year, but on the exam, the monthly approximation method will probably be used.
- Shares issued enter into the computation from the date of issuance.
- Reacquired shares are excluded from the computation from the date of reacquisition.
- Shares sold or issued in a purchase of assets are included from the date of issuance.
- A stock split or stock dividend is applied to all shares outstanding prior to the split or dividend and to the beginning-of-period weighted average shares. A stock split or stock dividend adjustment is not applied to any shares issued or repurchased after the split or dividend date.

Diluted EPS

Before calculating diluted EPS, it is necessary to understand the following terms:

- **Dilutive securities** are stock options, warrants, convertible debt, or convertible preferred stock that would *decrease* EPS if exercised or converted to common stock.
- **Antidilutive securities** are stock options, warrants, convertible debt, or convertible preferred stock that would *increase* EPS if exercised or converted to common stock.

The numerator of the basic EPS equation contains income available to common shareholders (net income less preferred dividends). In the case of diluted EPS, if there are dilutive securities, then the numerator must be adjusted as follows:

- If convertible preferred stock is dilutive (meaning EPS will fall if it is converted to common stock), the convertible preferred dividends must be added to earnings available to common shareholders.
- If convertible bonds are dilutive, then the bonds' after-tax interest expense is not considered an interest expense for diluted EPS. Hence, interest expense multiplied by (1 - the tax rate) must be added back to the numerator.



PROFESSOR'S NOTE

Interest paid on bonds is typically tax deductible for the firm. If convertible bonds are converted to stock, the firm saves the interest cost but loses the tax deduction. Thus, only the after-tax interest savings are added back to income available to common shareholders.

The basic EPS denominator is the weighted average number of shares. When the firm has dilutive securities outstanding, the denominator is the basic EPS denominator adjusted for the equivalent number of common shares that would be created by the conversion of all dilutive securities outstanding (convertible bonds, convertible preferred shares, warrants, and options), with each one considered separately to determine if it is dilutive.

If a dilutive security was issued during the year, the increase in the weighted average number of shares for diluted EPS is based on only the portion of the year the dilutive security was outstanding.

Dilutive stock options or warrants increase the number of common shares outstanding in the denominator for diluted EPS. There is no adjustment to the numerator.

The **diluted EPS equation** is:

diluted EPS =

$$\frac{\text{adjusted income available for common shares}}{\text{weighted-average common and potential common shares outstanding}}$$

where *adjusted income available for common shares* is:

- net income – preferred dividends
- + dividends on convertible preferred stock
- + after-tax interest on convertible debt

Therefore, diluted EPS is:

$$\text{diluted EPS} = \frac{\left[\text{net income} - \frac{\text{preferred}}{\text{dividends}} \right] + \left[\begin{array}{c} \text{convertible} \\ \text{preferred} \\ \text{dividends} \end{array} \right] + \left(\begin{array}{c} \text{convertible} \\ \text{debt} \\ \text{interest} \end{array} \right) (1 - t)}{\left(\begin{array}{c} \text{weighted} \\ \text{average} \\ \text{shares} \end{array} \right) + \left(\begin{array}{c} \text{shares from} \\ \text{conversion of} \\ \text{conv. pfd. shares} \end{array} \right) + \left(\begin{array}{c} \text{shares from} \\ \text{conversion of} \\ \text{conv. debt} \end{array} \right) + \left(\begin{array}{c} \text{shares} \\ \text{issuable from} \\ \text{stock options} \end{array} \right)}$$

The effect of conversion to common shares is included in the calculation of diluted EPS for a given security only if it is, in fact, dilutive. If a firm has more than one potentially dilutive security outstanding, each potentially dilutive security must be examined separately to determine if it is actually dilutive (i.e., would reduce EPS if converted to common stock).

Example: EPS with convertible preferred stock

During 20X6, ZZZ reported net income of \$4,350,000 and had 2,000,000 shares of common stock outstanding for the entire year. ZZZ's 7%, \$5,000,000 par value preferred stock is convertible into common stock at a conversion rate of 1.1 shares for every \$10 of par value. Compute basic and diluted EPS.

Answer:

$$\text{basic EPS} = \frac{\$4,350,000 - (0.07)(\$5,000,000)}{2,000,000} = \$2.00$$

Step 2: Calculate diluted EPS:

- Compute the increase in common stock outstanding if the preferred stock is converted to common stock at the beginning of 20X6: $(\$5,000,000 / \$10) \times 1.1 = 550,000$ shares.
- If the convertible preferred shares were converted to common stock, there would be no preferred dividends paid. Therefore, you should add back the convertible preferred dividends that had previously been subtracted from net income in the numerator.

Compute diluted EPS as if the convertible preferred stock were converted into common stock:

$$\text{diluted EPS} = \frac{\text{net. inc.} - \text{pref. div.} + \text{convert. pref. dividends}}{\text{wt. avg. shares} + \text{convert. pref. common shares}}$$

$$\text{diluted EPS} = \frac{\$4,350,000}{2,000,000 + 550,000} = \$1.71$$

- Check to see if diluted EPS is less than basic EPS ($\$1.71 < \2.00). If the answer is yes, the preferred stock is dilutive and must be included in diluted EPS as computed above. If the answer is no, the preferred stock is antidilutive and conversion effects are not included in diluted EPS.

A quick way to check whether convertible preferred stock is dilutive is to divide the preferred dividend by the number of shares that will be created if the preferred stock is converted. For ZZZ: $\frac{\$5,000,000 \times 0.07}{550,000} = \0.64 . Since this is less than basic EPS, the convertible preferred is dilutive.

Example: EPS with convertible debt

During 20X6, YYY Corp. had earnings available to common shareholders of \$2,500,000 and had 1,000,000 shares of common stock outstanding for the entire year, for basic EPS of \$2.50. During 20X5, YYY issued 2,000, \$1,000 par, 5% bonds for \$2,000,000 (issued at par). Each of these bonds is convertible to 120 shares of common stock. The tax rate is 30%. Compute the 20X6 diluted EPS.

Answer:

Compute the increase in common stock outstanding if the convertible debt is converted to common stock at the beginning of 20X6:

$$\text{shares issuable for debt conversion} = (2,000)(120) = 240,000 \text{ shares}$$

If the convertible debt is considered converted to common stock at the beginning of 20X6, then there would be no interest expense related to the convertible debt. Therefore, it is necessary to increase YYY's after-tax net income for the after-tax effect of the decrease in interest expense:

$$\text{increase in income} = [(2,000)(\$1,000)(0.05)] (1 - 0.30) = \$70,000$$

Compute diluted EPS as if the convertible debt were common stock:

$$\text{diluted EPS} = \frac{\text{net. inc.} - \text{pref. div.} + \text{convert. int.} (1 - t)}{\text{wt. avg. shares} + \text{convertible debt shares}}$$

$$\text{diluted EPS} = \frac{\$2,500,000 + \$70,000}{1,000,000 + 240,000} = \$2.07$$

Check to make sure that *diluted EPS* is less than *basic EPS* ($\$2.07 < \2.50). If diluted EPS is more than the basic EPS, the convertible bonds are *antidilutive* and should not be treated as common stock in computing diluted EPS.

A quick way to determine whether the convertible debt is dilutive is to calculate its per share impact by:

$$\frac{\text{convertible debt interest} (1 - t)}{\text{convertible debt shares}}$$

If this per share amount is greater than basic EPS, the convertible debt is antidilutive, and the effects of conversion should not be included when calculating diluted EPS.

If this per share amount is less than basic EPS, the convertible debt is dilutive, and the effects of conversion should be included in the calculation of diluted EPS.

For YYY:

$$\frac{\$70,000}{240,000} = \$0.29$$

The company's basic EPS is \$2.50, so the convertible debt is dilutive, and the effects of conversion should be included in the calculation of diluted EPS.

Stock options and warrants are dilutive only when their exercise prices are less than the average market price of the stock over the year. If the options or warrants are dilutive, use the **treasury stock method** to calculate the number of shares used in the denominator.

- The treasury stock method assumes that the funds received by the company from the exercise of the options would be used to hypothetically purchase shares of the company's common stock in the market at the average market price.
- The net increase in the number of shares outstanding (the adjustment to the denominator) is the number of shares created by exercising the options less the number of shares hypothetically repurchased with the proceeds of exercise.

Example: EPS with stock options

During 20X6, XXX reported earnings available to common shareholders of \$1,200,000 and had 500,000 shares of common stock outstanding for the entire year, for basic EPS of \$2.40. XXX has 100,000 stock options (or warrants) outstanding the entire year. Each option allows its holder to purchase one share of common stock at \$15 per share. The average market price of XXX's common stock during 20X6 is \$20 per share. Compute diluted EPS.

Answer:

Number of common shares created if the options are exercised: 100,000 shares

Cash inflow if the options are exercised (\$15/share)(100,000): \$1,500,000

Number of shares that can be purchased with these funds is:
\$1,500,000 / \$20 75,000 shares

Net increase in common shares outstanding from the exercise of the stock options (100,000 – 75,000) 25,000 shares

$$\text{diluted EPS} = \frac{\$1,200,000}{500,000 + 25,000} = \$2.29$$

A quick way to calculate the net increase in common shares from the potential exercise of stock options or warrants when the exercise price is less than the average market price is:

$$\left[\frac{\text{AMP} - \text{EP}}{\text{AMP}} \right] \times N$$

where:

AMP = average market price over the year

EP = exercise price of the options or warrants

N = number of common shares that the options and warrants can be converted into

For XXX:

$$\frac{\$20 - \$15}{\$20} \times 100,000 \text{ shares} = 25,000 \text{ shares}$$



MODULE QUIZ 18.4

1. The Hall Corporation had 100,000 shares of common stock outstanding at the beginning of the year. Hall issued 30,000 shares of common stock on May 1. On July 1, the company issued a 10% stock dividend. On September 1, Hall issued 1,000, 10% bonds, each convertible into 21 shares of common stock. What is the weighted average number of shares to be used in computing basic and diluted EPS, assuming the convertible bonds are dilutive?

Average shares, basic

- A. 132,000
- B. 132,000
- C. 139,000

Average shares, dilutive

- 139,000
- 146,000
- 146,000

2. Given the following information, how many shares should be used in computing diluted EPS?

- 300,000 shares outstanding.
- 100,000 warrants exercisable at \$50 per share.
- Average share price is \$55.
- Year-end share price is \$60.
 - A. 9,091.
 - B. 90,909.
 - C. 309,091.

3. An analyst gathered the following information about a company:

- 100,000 common shares outstanding from the beginning of the year.
- Earnings of \$125,000.
- 1,000, 7%, \$1,000 par bonds convertible into 25 shares each, outstanding as of the beginning of the year.
- The tax rate is 40%.

The company's diluted EPS is *closest* to:

- A. \$1.22.
- B. \$1.25.
- C. \$1.34.

4. An analyst has gathered the following information about a company:

- 50,000 common shares outstanding from the beginning of the year.
- Warrants outstanding all year on 50,000 shares, exercisable at \$20 per share.
- Stock is selling at year-end for \$25.
- The average price of the company's stock for the year was \$15.

How many shares should be used in calculating the company's diluted EPS?

- A. 16,667.
- B. 50,000.
- C. 66,667.

MODULE 18.5: COMMON-SIZE INCOME STATEMENTS



Video covering this content is available online.

LOS 18.i: Formulate income statements into common-size income statements.

A vertical **common-size income statement** expresses each category of the income statement as a percentage of revenue. The common-size format standardizes the income statement by eliminating the effects of size. This allows for comparison of income statement items over time (time-series analysis) and across firms (cross-sectional analysis). For example, the following are year-end income statements of industry competitors North Company and South Company:

	North Co.	South Co.
Revenue	\$75,000,000	\$3,500,000
Cost of goods sold	<u>52,500,000</u>	<u>700,000</u>
Gross profit	\$22,500,000	\$2,800,000
Administrative expense	11,250,000	525,000
Research expense	<u>3,750,000</u>	<u>700,000</u>
Operating profit	\$7,500,000	\$1,575,000

Notice that North is significantly larger and more profitable than South when measured in absolute dollars. North's gross profit is \$22,500,000, as compared to South's gross profit of \$2,800,000. Similarly, North's operating profit of \$7,500,000 is significantly greater than South's operating profit of \$1,575,000.

Once we convert the income statements to common-size format, we can see that South is the more profitable firm on a relative basis. South's gross profit of 80% and operating profit of 45% are significantly greater than North's gross profit of 30% and operating profit of 10%.

	North Co.	South Co.
Revenue	100%	100%
Cost of goods sold	<u>70%</u>	<u>20%</u>
Gross profit	30%	80%
Administrative expense	15%	15%
Research expense	<u>5%</u>	<u>20%</u>
Operating profit	10%	45%

Common-size analysis can also be used to examine a firm's strategy. South's higher gross profit margin may be the result of technologically superior products. Notice that South spends more on research than North on a relative basis. This may allow South to charge a higher price for its products.

In most cases, expressing expenses as a percentage of revenue is appropriate. One exception is income tax expense. Tax expense is more meaningful when expressed as a percentage of pretax income. The result is known as the **effective tax rate**.

LOS 18.j: Evaluate a company's financial performance using common-size income statements and financial ratios based on the income statement.

Margin ratios can be used to measure a firm's profitability quickly. **Gross profit margin** is the ratio of gross profit (revenue minus cost of goods sold) to revenue (sales).

$$\text{gross profit margin} = \frac{\text{gross profit}}{\text{revenue}}$$

Gross profit margin can be increased by raising prices or reducing production costs. A firm might be able to increase prices if its products can be differentiated from other firms' products as a result of factors such as brand names, quality, technology, or patent protection. This was illustrated in the previous example whereby South's gross profit margin was higher than North's.

Another popular margin ratio is **net profit margin**. Net profit margin is the ratio of net income to revenue.

$$\text{net profit margin} = \frac{\text{net income}}{\text{revenue}}$$

Net profit margin measures the profit generated after considering all expenses. Like gross profit margin, net profit margin should be compared over time and with the firm's industry peers.

Any subtotal found in the income statement can be expressed as a percentage of revenue. For example, operating profit divided by revenue is known as **operating profit margin**. Pretax accounting profit divided by revenue is known as **pretax margin**.

LOS 18.k: Describe, calculate, and interpret comprehensive income.

LOS 18.l: Describe other comprehensive income and identify major types of items included in it.

At the end of each accounting period, the net income of the firm (less any dividends declared) is added to stockholders' equity through an account known as **retained earnings**. Therefore, any transaction that affects the income statement (net income) will also affect stockholders' equity.

Recall that net income is equal to revenue minus expenses. **Comprehensive income** is a more inclusive measure that includes all changes in equity except for owner contributions and distributions. That is, comprehensive income is the sum of net income and **other comprehensive income (OCI)**. Under both U.S. GAAP and IFRS, other comprehensive income includes transactions that are *not* included in net income, such as:

1. Foreign currency translation gains and losses.
2. Adjustments for minimum pension liability.
3. Unrealized gains and losses from cash flow hedging derivatives.
4. Unrealized gains and losses from available-for-sale securities.

Gains or losses in the value of securities that a firm owns and has not yet sold are referred to as *unrealized* gains and losses. Whether unrealized gains and losses on investment securities are reported in the income statement, reported in other comprehensive income, or not reported in either, depends on how a firm has classified the securities. Interest or dividends received from securities owned by the firm are reported on the income statement.

U.S. GAAP

Debt securities that a firm owns, but intends to sell, are classified as **trading securities**, and any unrealized gains and losses during the period are reported on the income statement.

Debt securities the firm does not intend to sell prior to maturity are classified as **held to maturity**. Securities classified as held to maturity are reported at amortized cost on the balance sheet (not fair value). Therefore, unrealized gains and losses are not reported on either the income statement or as other comprehensive income.

Debt securities that are not expected to be held to maturity or sold in the near term are classified as **available-for-sale securities**. Unrealized gains and losses on available-for-sale securities are reported as other comprehensive income, not on the income statement.

EXAMPLE: Calculating comprehensive income

Calculate comprehensive income for Triple C Corporation, which reports under U.S. GAAP, using the selected financial statement data found in the following table.

Triple C Corporation – Selected Financial Statement Data

Net income	\$1,000
Dividends received from available-for-sale securities	60
Unrealized loss from foreign currency translation	(15)
Dividends paid	(110)
Reacquire common stock	(400)
Unrealized gain from cash flow hedge	30
Unrealized loss from available-for-sale securities	(10)
Realized gain on sale of land	65

Answer:

Net income	\$1,000
Unrealized loss from foreign currency translation	(15)
Unrealized gain from cash flow hedge	30
Unrealized loss from available-for-sale securities	<u>(10)</u>
Comprehensive income	\$1,005

The dividends received for available-for-sale securities and the realized gain on the sale of land are already included in net income. Dividends paid and the reacquisition of common stock are transactions with shareholders, so they are not included in comprehensive income. The three items reported as other comprehensive income are added to net income to calculate comprehensive income.

IFRS

Under IFRS, unrealized gains and losses can also be reported on the income statement, included in other comprehensive income, or not reported (for securities carried at amortized cost rather than fair value). While these alternatives are consistent with those under U.S. GAAP, the terms used to classify securities are different. The three categories for securities under IFRS reflect how they are valued and how unrealized gains and losses are reported in the financial statements. The IFRS categories are:

- *Securities measured at fair value through profit and loss* (corresponds to trading securities under U.S. GAAP).
- *Securities measured at amortized cost* (corresponds to held-to-maturity under U.S. GAAP).
- *Securities measured at fair value through other comprehensive income* (corresponds to available-for-sale under U.S. GAAP).

While both IFRS and U.S. GAAP have the same three available treatments of unrealized gains and losses, specific securities may be classified differently under IFRS or U.S. GAAP. An analyst must account for these differences when comparing ratios that involve net income, such as net profit margin and price-to-earnings. The potential impact of such accounting differences on specific ratios can be examined by comparing the ratios calculated with net income to those calculated using comprehensive income.



PROFESSOR'S NOTE

We describe the differences between IFRS and U.S. GAAP classifications of securities in our review of Understanding Balance Sheets.



MODULE QUIZ 18.5

1. A vertical common-size income statement expresses each category of the income statement as a percentage of:
 - A. assets.
 - B. gross profit.
 - C. revenue.
2. Which of the following would *most likely* result in higher gross profit margin, assuming no fixed costs?
 - A. A 10% increase in the number of units sold.
 - B. A 5% decrease in production cost per unit.
 - C. A 7% decrease in administrative expenses.
3. Which of the following transactions affects owners' equity but does not affect net income?
 - A. Foreign currency translation gain.
 - B. Repaying the face amount on a bond issued at par.
 - C. Dividends received from available-for-sale securities.
4. Which of the following is *least likely* to be included when calculating comprehensive income?
 - A. Unrealized loss from cash flow hedging derivatives.
 - B. Unrealized gain from available-for-sale securities.
 - C. Dividends paid to common shareholders.

KEY CONCEPTS

LOS 18.a

The income statement shows an entity's revenues, expenses, gains and losses during a reporting period.

A multi-step income statement provides a subtotal for gross profit and a single step income statement does not. Expenses on the income statement can be grouped by the nature of the expense items or by their function, such as with expenses grouped into cost of goods sold.

LOS 18.b

Revenue is recognized when earned and expenses are recognized when incurred.

Accounting standards identify a five-step process for recognizing revenue:

1. Identify the contract(s) with a customer.
2. Identify the performance obligations in the contract.
3. Determine the transaction price.
4. Allocate the transaction price to the performance obligations in the contract.
5. Recognize revenue when (or as) the entity satisfies a performance obligation.

LOS 18.c

Information that can influence the choice of revenue recognition method includes progress toward completion of a performance obligation, variable considerations and their likelihood of being earned, revisions to contracts, and whether the firm is acting as a principal or an agent in a transaction.

LOS 18.d

The matching principle requires that firms match revenues recognized in a period with the expenses required to generate them. One application of the matching principle is seen in accounting for inventory, with cost of goods sold as the cost of units sold from inventory that are included in current-period revenue. Other costs, such as depreciation of fixed assets or administrative overhead, are period costs and are taken without regard to revenues generated during the period.

Depreciation methods:

- Straight-line: Equal amount of depreciation expense in each year of the asset's useful life.
- Declining balance: Apply a constant rate of depreciation to the declining book value until book value equals residual value.

Inventory valuation methods:

- FIFO: Inventory reflects cost of most recent purchases, COGS reflects cost of oldest purchases.
- LIFO: COGS reflects cost of most recent purchases, inventory reflects cost of oldest purchases.
- Average cost: Unit cost equals cost of goods available for sale divided by total units available and is used for both COGS and inventory.
- Specific identification: Each item in inventory is identified and its historical cost is used for calculating COGS when the item is sold.

Intangible assets with limited lives should be amortized using a method that reflects the flow over time of their economic benefits. Intangible assets with indefinite lives (e.g., goodwill) are not amortized.

Users of financial data should analyze the reasons for any changes in estimates of expenses and compare these estimates with those of peer companies.

LOS 18.e

Results of discontinued operations are reported below income from continuing operations, net of tax, from the date the decision to dispose of the operations is made. These results are segregated because they likely are non-recurring and do not affect future net income.

Unusual or infrequent items are reported before tax and above income from continuing operations. An analyst should determine how "unusual" or "infrequent" these items really are for the company when estimating future earnings or firm value.

Changes in accounting standards, changes in accounting methods applied, and corrections of accounting errors require retrospective restatement of all prior-period financial statements included in the current statement. A change in an accounting estimate, however, is applied prospectively (to subsequent periods) with no restatement of prior-period results.

LOS 18.f

Operating income is generated from the firm's normal business operations. For a nonfinancial firm, income that results from investing or financing transactions is classified as non-operating

income, while it is operating income for a financial firm since its business operations include investing in and financing securities.

LOS 18.g

$$\text{basic EPS} = \frac{\text{net income} - \text{preferred dividends}}{\text{weighted average number of common shares outstanding}}$$

When a company has potentially dilutive securities, it must report diluted EPS.

For any convertible preferred stock, convertible debt, warrants, or stock options that are dilutive, the calculation of diluted EPS is:

$$\text{diluted EPS} =$$

$$\frac{\left[\text{net income} - \text{preferred dividends} \right] + \left[\begin{array}{c} \text{convertible} \\ \text{preferred} \\ \text{dividends} \end{array} \right] + \left[\begin{array}{c} \text{convertible} \\ \text{debt} \\ \text{interest} \end{array} \right] (1 - t)}{\left(\begin{array}{c} \text{weighted} \\ \text{average} \\ \text{shares} \end{array} \right) + \left(\begin{array}{c} \text{shares from} \\ \text{conversion of} \\ \text{conv. pfd. shares} \end{array} \right) + \left(\begin{array}{c} \text{shares from} \\ \text{conversion of} \\ \text{conv. debt} \end{array} \right) + \left(\begin{array}{c} \text{shares} \\ \text{issuable from} \\ \text{stock options} \end{array} \right)}$$

LOS 18.h

A dilutive security is one that, if converted to its common stock equivalent, would decrease EPS. An antidilutive security is one that would not reduce EPS if converted to its common stock equivalent.

LOS 18.i

A vertical common-size income statement expresses each item as a percentage of revenue. The common-size format standardizes the income statement by eliminating the effects of size. Common-size income statements are useful for trend analysis and for comparisons with peer firms.

LOS 18.j

Common-size income statements are useful in examining a firm's business strategies.

Two popular profitability ratios are gross profit margin (gross profit / revenue) and net profit margin (net income / revenue). A firm can often achieve higher profit margins by differentiating its products from the competition.

LOS 18.k

Comprehensive income is the sum of net income and other comprehensive income. It measures all changes to equity other than those from transactions with shareholders.

LOS 18.l

Transactions with shareholders, such as dividends paid and shares issued or repurchased, are not reported on the income statement.

Other comprehensive income includes other transactions that affect equity but do not affect net income, including:

- Gains and losses from foreign currency translation.
- Pension obligation adjustments.

- Unrealized gains and losses from cash flow hedging derivatives.
- Unrealized gains and losses on available-for-sale securities.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 18.1, 18.2

- B** Depreciation is included in the computation of operating expenses. Interest expense is a financing cost. Thus, it is excluded from operating expenses. (Module 18.1, LOS 18.a)
 - A** Cost of goods sold includes a number of expenses related to the same function, the production of inventory. Depreciation and wages are examples of expenses classified by nature. (Module 18.1, LOS 18.a)
 - B** The five steps in revenue recognition are:
 1. Identify the contract or contracts with the customer.
 2. Identify the performance obligations in the contract(s).
 3. Determine a transaction price.
 4. Allocate the transaction price to the performance obligations.
 5. Recognize revenue when/as the performance obligations have been satisfied.
- (Module 18.2, LOS 18.b)

Module Quiz 18.3

- A** LIFO and FIFO are both permitted under U.S. GAAP. LIFO is prohibited under IFRS. (LOS 18.d)
- B** Accelerated depreciation will result in higher depreciation in the early years and lower depreciation in the later years compared to the straight-line method. Total depreciation expense will be the same under both methods. The book value would be higher in the later years using straight-line depreciation, so the loss from scrapping the equipment under an accelerated method is less compared to the straight-line method. (LOS 18.d)
- B** 108 units were sold ($13 + 35 + 60$) and 150 units were available for sale (beginning inventory of 0 plus purchases of $40 + 20 + 90$), so there are $150 - 108 = 42$ units in ending inventory. Under FIFO, units from the last batch purchased would remain in inventory: $42 \times \$50 = \$2,100$. Under LIFO, the first 42 units purchased would be in inventory: $(40 \times \$30) + (2 \times \$40) = \$1,280$. (LOS 18.d)
- B** Year 1: $(2 / 4) \times 25,000 = \$12,500$. Year 2: $(2 / 4) \times (25,000 - 12,500) = \$6,250$. (LOS 18.d)
- A** A change in an accounting estimate is reported prospectively. No restatement of prior period statements is necessary. (LOS 18.e)
- C** A physically and operationally distinct division that is currently for sale is treated as a discontinued operation. The income from the division is reported net of tax below income from continuing operations. Gains and losses on sales of operating assets, as well as depreciation expense, are reported pretax, above income from continuing operations. (LOS 18.e)

7. **C** A change in accounting principle requires retrospective application; that is, all prior period financial statements currently presented are restated to reflect the change. (LOS 18.e)

8. **C** Bad debt expense is an operating expense. The other choices are nonoperating items from the perspective of a manufacturing firm. (LOS 18.f)

Module Quiz 18.4

1. **A** The new stock is weighted by 8 / 12. The bonds are weighted by 4 / 12 and are not affected by the stock dividend.

$$\begin{aligned}\text{Basic shares} &= \{[100,000 \times (12 / 12)] + [30,000 \times (8 / 12)]\} \times 1.10 \\ &= 132,000\end{aligned}$$

$$\text{Diluted shares} = 132,000 + [21,000 \times (4 / 12)] = 139,000$$

(LOS 18.g, 18.h)

2. **C** Since the exercise price of the warrants is less than the average share price, the warrants are dilutive. Using the treasury stock method to determine the denominator impact:

$$\frac{\$55 - \$50}{\$55} \times 100,000 \text{ shares} = 9,091 \text{ shares}$$

Thus, the denominator will increase by 9,091 shares to 309,091 shares. The question asks for the total, not just the impact of the warrants. (LOS 18.g, 18.h)

3. **B** First, calculate basic EPS = $\frac{\$125,000}{100,000} = \1.25

Next, check if the convertible bonds are dilutive:

$$\text{numerator impact} = (1,000 \times 1,000 \times 0.07) \times (1 - 0.4) = \$42,000$$

$$\text{denominator impact} = (1,000 \times 25) = 25,000 \text{ shares}$$

$$\text{per share impact} = \frac{\$42,000}{25,000 \text{ shares}} = \$1.68$$

Since \$1.68 is greater than the basic EPS of \$1.25, the bonds are antidilutive. Thus, diluted EPS = basic EPS = \$1.25.

(LOS 18.g, 18.h)

Module Quiz 18.5

1. **C** Each category of the income statement is expressed as a percentage of revenue (sales). (LOS 18.i)

2. **B** A 5% decrease in per unit production cost will increase gross profit by reducing cost of goods sold. Assuming no fixed costs, gross profit margin will remain the same if sale quantities increase. Administrative expenses are not included in gross profit margin. (LOS 18.j)

3. **A** A foreign currency translation gain is not included in net income but the gain increases owners' equity. Dividends received are reported in the income statement. The repayment of principal does not affect owners' equity. (LOS 18.k, LOS 18.l)

4. C Comprehensive income includes all changes in equity except transactions with shareholders. Therefore, dividends paid to common shareholders do not affect current period comprehensive income. (LOS 18.k, 18.l)

1. IFRS 15 *Revenue From Contracts With Customers*.

2. IASB *Framework for the Preparation and Presentation of Financial Statements*, paragraph 4.25(b).

Reading 19

UNDERSTANDING BALANCE SHEETS

EXAM FOCUS

While the income statement presents a picture of a firm's economic activities over a period of time, its balance sheet is a snapshot of its financial and physical assets and its liabilities at a point in time. Just as with the income statement, understanding balance sheet accounts, how they are valued, and what they represent, is also crucial to the financial analysis of a firm. Again, different choices of accounting methods and different accounting estimates will affect a firm's financial ratios, and an analyst must be careful to make the necessary adjustments in order to compare two or more firms. Special attention should be paid to the method by which each balance sheet item is calculated and how changes in balance sheet values relate to the income statement and to shareholders' equity. Later readings in this topic area include more detailed information on several balance sheet accounts, including inventories, long-lived assets, deferred taxes, and long-term liabilities.

MODULE 19.1: BALANCE SHEET INTRODUCTION



LOS 19.a: Describe the elements of the balance sheet: assets, liabilities, and equity.

Video covering this content is available online.

The **balance sheet** (also known as the statement of financial position or statement of financial condition) reports the firm's financial position at a point in time. The balance sheet consists of assets, liabilities, and equity.

Assets: Resources controlled as a result of past transactions that are expected to provide future economic benefits.

Liabilities: Obligations as a result of past events that are expected to require an outflow of economic resources.

Equity: The owners' residual interest in the assets after deducting the liabilities. Equity is also referred to as stockholders' equity, shareholders' equity, or owners' equity. Analysts sometimes refer to equity as "net assets."

A financial statement item should be recognized if a future economic benefit from the item (flowing to or from the firm) is *probable* and the item's value or cost can be measured reliably.

MODULE 19.2: ASSETS AND LIABILITIES



LOS 19.b: Describe uses and limitations of the balance sheet in financial analysis.

Video covering this content is available online.

The balance sheet can be used to assess a firm's liquidity, solvency, and ability to make distributions to shareholders. From the firm's perspective, **liquidity** is the ability to meet short-term obligations and **solvency** is the ability to meet long-term obligations.

The balance sheet elements (assets, liabilities, and equity) should not be interpreted as market value or intrinsic value. For most firms, the balance sheet consists of a mixture of values. For example, some assets are reported at historical cost, some are reported at amortized cost, and others may be reported at fair value. There are numerous valuation bases. Even if the balance sheet was reported at fair value, the value may have changed since the balance sheet date. Also, there are a number of assets and liabilities that do not appear on the balance sheet but certainly have value. For example, the value of a firm's employees and reputation is not reported on the balance sheet.

LOS 19.c: Describe alternative formats of balance sheet presentation.

Both IFRS and U.S. GAAP require firms to separately report their current assets and noncurrent assets and current and noncurrent liabilities. The current/noncurrent format is known as a **classified balance sheet** and is useful in evaluating liquidity.

Under IFRS, firms can choose to use a **liquidity-based format** if the presentation is more relevant and reliable. Liquidity-based presentations, which are often used in the banking industry, present assets and liabilities in the order of liquidity.



MODULE QUIZ 19.1, 19.2

1. Which of the following is *most likely* an essential characteristic of an asset?
 - A. An asset is tangible.
 - B. An asset is obtained at a cost.
 - C. An asset provides future benefits.
2. Which of the following statements about analyzing the balance sheet is *most accurate*?
 - A. The value of the firm's reputation is reported on the balance sheet at amortized cost.
 - B. Shareholders' equity is equal to the intrinsic value of the firm.
 - C. The balance sheet can be used to measure the firm's capital structure.
3. Century Company's balance sheet follows:

Century Company Balance Sheet (in millions)		
	20X7	20X6
Current assets	\$340	\$280
Noncurrent assets	<u>660</u>	<u>630</u>
Total assets	\$1,000	\$910
Current liabilities	\$170	\$110
Noncurrent liabilities	<u>50</u>	<u>50</u>
Total liabilities	\$220	\$160
Equity	<u>\$780</u>	<u>\$750</u>
Total liabilities and equity	<u>\$1,000</u>	<u>\$910</u>

Century's balance sheet presentation is known as:

- A. a classified balance sheet.
- B. a liquidity-based balance sheet.
- C. an account form balance sheet.

MODULE 19.3: CURRENT ASSETS AND LIABILITIES



Video covering this content is available online.

LOS 19.d: Contrast current and non-current assets and current and non-current liabilities.

Current assets include cash and other assets that will likely be converted into cash or used up within one year or one operating cycle, whichever is greater. The **operating cycle** is the time it takes to produce or purchase inventory, sell the product, and collect the cash. Current assets are usually presented in the order of their liquidity, with cash being the most liquid. Current assets reveal information about the operating activities of the firm.

Current liabilities are obligations that will be satisfied within one year or one operating cycle, whichever is greater. More specifically, a liability that meets any of the following criteria is considered current:

- Settlement is expected during the normal operating cycle.
- Settlement is expected within one year.
- Held primarily for trading purposes.
- There is not an unconditional right to defer settlement for more than one year.

Current assets minus current liabilities equals **working capital**. Not enough working capital may indicate liquidity problems. Too much working capital may be an indication of inefficient use of assets.

Noncurrent assets do not meet the definition of current assets because they will not be converted into cash or used up within one year or operating cycle. Noncurrent assets provide information about the firm's investing activities, which form the foundation upon which the firm operates.

Noncurrent liabilities do not meet the criteria of current liabilities. Noncurrent liabilities provide information about the firm's long-term financing activities.

LOS 19.e: Describe different types of assets and liabilities and the measurement bases of each.

Current Assets

Current assets include cash and other assets that will be converted into cash or used up within one year or operating cycle, whichever is greater.

Cash and cash equivalents. Cash equivalents are short-term, highly liquid investments that are readily convertible to cash and near enough to maturity that interest rate risk is insignificant. Examples of cash equivalents include Treasury bills, commercial paper, and money market funds. Cash and equivalents are considered financial assets. Generally, financial assets are reported on the balance sheet at amortized cost or fair value. For cash equivalents, either measurement base should result in about the same value.

Marketable securities. Marketable securities are financial assets that are traded in a public market and whose value can be readily determined. Examples include Treasury bills, notes, bonds, and equity securities. Details of the investment are disclosed in the financial footnotes. Measurement bases for marketable securities will be discussed later in this reading.

Accounts receivable. Accounts receivable (also known as trade receivables) are financial assets that represent amounts owed to the firm by customers for goods or services sold on credit. Accounts receivable are reported at **net realizable value**, which is based on estimated **bad debt expense**. Bad debt expense increases the **allowance for doubtful accounts**, a contra-asset account. A **contra account** is used to reduce the value of its controlling account. Thus, gross receivables less the allowance for doubtful accounts is equal to accounts receivable at net realizable value, the amount the firm expects to collect. When receivables are "written off" (removed from the balance sheet because they are uncollectible), both gross receivables and the allowance account are reduced.

Firms are required to disclose significant concentrations of credit risk, including customer, geographic, and industry concentrations.

Analyzing receivables relative to sales can reveal collection problems. The allowance for doubtful accounts should also be considered relative to the level and growth rate of sales. Firms can underestimate bad debt expense, thereby increasing reported earnings.

Inventories. Inventories are goods held for sale to customers or used in manufacture of goods to be sold. Manufacturing firms separately report inventories of raw materials, work-in-process, and finished goods. The costs included in inventory include purchase cost, conversion costs, and other costs necessary to bring the inventory to its present location and condition.

Standard costing and the retail method are used by some firms to measure inventory costs. **Standard costing**, often used by manufacturing firms, involves assigning predetermined amounts of materials, labor, and overhead to goods produced. Firms that use the **retail method** measure inventory at retail prices and then subtract gross profit in order to determine cost.

Using different cost flow assumptions (also known as cost flow methods), firms assign inventory costs to the income statement (cost of goods sold). As discussed in our reading on Understanding Income Statements, FIFO and average cost are permitted under both IFRS and U.S. GAAP. LIFO is permitted under U.S. GAAP but is prohibited under IFRS.

Inventories are reported at the lower of cost or net realizable value under IFRS, and under U.S. GAAP for companies that use inventory cost methods other than LIFO or retail. Net realizable value is equal to the selling price less any completion costs and disposal (selling) costs. Under U.S. GAAP, companies using LIFO or the retail method report inventories at the lower of cost or market. Market is usually equal to replacement cost; however, market cannot be greater than net realizable value or less than net realizable value less a normal profit margin. If net realizable value (IFRS) or market (U.S. GAAP) is less than the inventory's carrying value, the inventory is written down and a loss is recognized in the income statement. If there is a subsequent recovery in value, the inventory can be written back up under IFRS. No write-up is allowed under U.S. GAAP; the firm simply reports greater profit when the inventory is sold.



PROFESSOR'S NOTE

Inventories are described in more detail in a later reading.

Other current assets. Other current assets include amounts that may not be material if shown separately; thus, the items are combined into a single amount. An example is **prepaid expenses**, which are operating costs that have been paid in advance. As the costs are actually incurred, an expense is recognized in the income statement and prepaid expenses (an asset) decrease. For example, if a firm makes an annual rent payment of \$400,000 at the beginning of the year, an asset (cash) decreases and another asset (prepaid rent) increases by the amount of the payment. At the end of three months, one-quarter of the prepaid rent has been used. At this point, the firm will recognize \$100,000 of rent expense in its income statement and reduce assets (prepaid rent) by \$100,000.

Current Liabilities

Current liabilities are obligations that will be satisfied within one year or operating cycle, whichever is greater.

Accounts payable. Accounts payable (also known as trade payables) are amounts the firm owes to suppliers for goods or services purchased on credit. Analyzing payables relative to purchases can signal credit problems with suppliers.

Notes payable and current portion of long-term debt. Notes payable are obligations in the form of promissory notes owed to creditors and lenders. Notes payable can also be reported as noncurrent liabilities if their maturities are greater than one year. The current portion of long-term debt is the principal portion of debt due within one year or operating cycle, whichever is greater.

Accrued liabilities. Accrued liabilities (accrued expenses) are expenses that have been recognized in the income statement but are not yet contractually due. Accrued liabilities result from the accrual method of accounting, under which expenses are recognized as incurred. For example, consider a firm that is required to make annual year-end interest payments of \$100,000 on an outstanding bank loan. At the end of March, the firm would recognize one-quarter (\$25,000) of the total interest expense in its income statement and an accrued liability would be increased by the same amount, even though the liability is not actually due until the end of the year.

Some firms include income tax payable as an accrued liability. **Taxes payable** are current taxes that have been recognized in the income statement but have not yet been paid. Other examples of accrued liabilities include interest payable, wages payable, and accrued warranty expense.

Unearned revenue. Unearned revenue (also known as unearned income, deferred revenue, or deferred income) is cash collected in advance of providing goods and services. For example, a magazine publisher receives subscription payments in advance of delivery. When payment is received, assets (cash) and liabilities (unearned revenue) increase by the same amount. As the magazines are delivered, the publisher recognizes revenue in the income statement and reduces the liability.

When analyzing liquidity, keep in mind that unearned revenue does not require a future outflow of cash like accounts payable. Also, unearned revenue may be an indication of future growth as the revenue will ultimately be recognized in the income statement.

MODULE 19.4: NONCURRENT ASSETS AND LIABILITIES



Video covering this content is available online.

Property, plant, and equipment. Property, plant, and equipment (PP&E) are tangible assets used in the production of goods and services. PP&E includes land and buildings, machinery and equipment, furniture, and natural resources. Under IFRS, PP&E can be reported using the **cost model** or the **revaluation model**. Under U.S. GAAP, only the cost model is allowed.

Under the cost model, PP&E other than land is reported at amortized cost (historical cost minus accumulated depreciation, amortization, depletion, and impairment losses). Land is not depreciated because it has an indefinite life. **Historical cost** includes the purchase price plus any cost necessary to get the asset ready for use, such as delivery and installation costs. As discussed in the reading on Understanding Income Statements, there are several depreciation methods (e.g., straight-line and declining balance methods) used to allocate the cost to the income statement over time. Thus, the balance sheet and income statement are affected by the depreciation method and related estimates (i.e., salvage value and useful life of assets).

Also under the cost model, PP&E must be tested for **impairment**. An asset is impaired if its carrying value exceeds the **recoverable amount**. Under IFRS, the recoverable amount of an asset is the greater of fair value less any selling costs, or the asset's **value in use**. Value in use is the present value of the asset's future cash flow stream. If impaired, the asset is written down to its recoverable amount and a loss is recognized in the income statement. Loss recoveries are allowed under IFRS but not under U.S. GAAP.

Under the revaluation model, PP&E is reported at fair value less any accumulated depreciation. Changes in fair value are reflected in shareholders' equity and may be recognized in the income statement in certain circumstances.

Investment property. Under IFRS, investment property includes assets that generate rental income or capital appreciation. U.S. GAAP does not have a specific definition of investment property. Under IFRS, investment property can either be reported at amortized cost (just like PP&E) or fair value. Under the **fair value model**, any change in fair value is recognized in the income statement.

Deferred tax assets. As we will discuss in our reading on Income Taxes, deferred taxes are the result of temporary differences between financial reporting income and tax reporting income. **Deferred tax assets** are created when the amount of taxes payable exceeds the amount of income tax expense recognized in the income statement. This can occur when expenses or losses are recognized in the income statement before they are tax deductible, or when revenues or gains are taxable before they are recognized in the income statement. Eventually, the deferred tax asset will reverse when the expense is deducted for tax purposes or the revenue is recognized in the income statement. Deferred tax assets can also be created by unused tax losses from prior periods, which have value because they can be used to reduce taxes in subsequent periods.



MODULE QUIZ 19.3, 19.4

1. Which of the following would *most likely* result in a current liability?
 - A. Possible warranty claims.
 - B. Recognizing impairment of PP&E.
 - C. Estimated income taxes for the current year.
2. How should the proceeds received from the advance sale of tickets to a sporting event be treated by the seller, assuming the tickets are nonrefundable?
 - A. Unearned revenue is recognized to the extent that costs have been incurred.
 - B. Revenue is recognized to the extent that costs have been incurred.
 - C. Revenue is deferred until the sporting event is held.
3. Which of the following inventory valuation methods is required by the accounting standard-setting bodies?
 - A. Lower of cost or net realizable value.
 - B. Weighted average cost.
 - C. First-in, first-out.
4. Under IFRS, a firm may report the value of property, plant, and equipment using:
 - A. only the cost model.
 - B. the cost model or the fair value model.
 - C. the cost model or the revaluation model.

MODULE 19.5: INTANGIBLE ASSETS



Intangible assets are non-monetary assets that lack physical substance. Securities are not considered intangible assets. Intangible assets are either identifiable or unidentifiable. **Identifiable intangible assets** can be acquired separately or are the result of rights or privileges conveyed to their owner. Examples of identifiable intangibles are patents, trademarks, and copyrights. **Unidentifiable intangible**

Video covering
this content is
available online.

assets cannot be acquired separately and may have an unlimited life. The best example of an unidentifiable intangible asset is goodwill.

Under IFRS, identifiable intangibles that are *purchased* can be reported on the balance sheet using the cost model or the revaluation model, although the revaluation model can only be used if an active market for the intangible asset exists. Both models are basically the same as the measurement models used for PP&E. Under U.S. GAAP, only the cost model is allowed.

Except for certain legal costs, intangible assets that are *created internally*, such as research and development costs, are expensed as incurred under U.S. GAAP. Under IFRS, a firm must identify the research stage (discovery of new scientific or technical knowledge) and the development stage (using research results to plan or design products). Under IFRS, the firm must expense costs incurred during the research stage but can capitalize costs incurred during the development stage.

Finite-lived intangible assets are amortized over their useful lives and tested for impairment in the same way as PP&E. The amortization method and useful life estimates are reviewed at least annually. Intangible assets with indefinite lives are not amortized, but are tested for impairment at least annually.

Under IFRS and U.S. GAAP, all of the following should be expensed as incurred:

- Start-up and training costs.
- Administrative overhead.
- Advertising and promotion costs.
- Relocation and reorganization costs.
- Termination costs.

Some analysts choose to eliminate intangible assets for analytical purposes. However, analysts should consider the value to the firm of each intangible asset before making any adjustments.

Goodwill. Goodwill is the excess of purchase price over the fair value of the identifiable net assets (assets minus liabilities) acquired in a business acquisition. Acquirers are often willing to pay more than the fair value of the target's identifiable net assets because the target may have assets that are not reported on its balance sheet. For example, the target's reputation and customer loyalty certainly have value; however, the value is not quantifiable. Also, the target may have research and development assets that remain off-balance-sheet because of current accounting standards. Finally, part of the acquisition price may reflect perceived synergies from the business combination. For example, the acquirer may be able to eliminate duplicate facilities and reduce payroll as a result of the acquisition.



PROFESSOR'S NOTE

Occasionally, the purchase price of an acquisition is less than fair value of the identifiable net assets. In this case, the difference is immediately recognized as a gain in the acquirer's income statement.

Goodwill is only created in a purchase acquisition. Internally generated goodwill is expensed as incurred.

Goodwill is not amortized but must be tested for impairment at least annually. If impaired, goodwill is reduced and a loss is recognized in the income statement. The impairment loss does

not affect cash flow. As long as goodwill is not impaired, it can remain on the balance sheet indefinitely.

Since goodwill is not amortized, firms can manipulate net income upward by allocating more of the acquisition price to goodwill and less to the identifiable assets. The result is less depreciation and amortization expense, resulting in higher net income.

Accounting goodwill should not be confused with economic goodwill. Economic goodwill derives from the expected future performance of the firm, while accounting goodwill is the result of past acquisitions.

When computing ratios, analysts should eliminate goodwill from the balance sheet and goodwill impairment charges from the income statement for comparability. Also, analysts should evaluate future acquisitions in terms of the price paid relative to the earning power of the acquired assets.

MODULE 19.6: MARKETABLE SECURITIES



Financial instruments are contracts that give rise to both a financial asset of one entity and a financial liability or equity instrument of another entity.¹ Financial instruments can be found on the asset side and the liability side of the balance sheet. Financial assets include investment securities (stocks and bonds), derivatives, loans, and receivables.

Video covering
this content is
available online.

Financial instruments are measured at historical cost, amortized cost, or fair value. Financial assets measured at cost include unquoted equity investments (for which fair value cannot be reliably measured) and loans to and notes receivable from other entities.

Under U.S. GAAP, debt securities acquired with the intent to hold them until they mature are classified as **held-to-maturity securities** and measured at amortized cost. Amortized cost is equal to the original issue price minus any principal payments, plus any amortized discount or minus any amortized premium, minus any impairment losses. Subsequent changes in market value are ignored.



PROFESSOR'S NOTE

Amortization of bond discounts and premiums (for bonds issued by a firm, rather than owned by a firm) is addressed in our review of Non-Current (Long-Term) Liabilities.

Financial assets measured at fair value, also known as **mark-to-market** accounting, include trading securities, available-for-sale securities, and derivatives.

Trading securities (also known as held-for-trading securities) are debt securities acquired with the intent to sell them over the near term. Trading securities are reported on the balance sheet at fair value, and the unrealized gains and losses (changes in market value before the securities are sold) are recognized in the income statement. All equity securities holdings (except those that give a company significant influence over a firm) are treated in the same manner, with unrealized gains and losses reported on the income statement. Unrealized gains and losses are also known as holding period gains and losses. **Derivative instruments** are treated the same as trading securities.



PROFESSOR'S NOTE

Accounting for intercorporate investments when the share owner has significant influence or control over a firm is addressed at Level II.

Available-for-sale securities are debt securities that are not expected to be held to maturity or traded in the near term. Like trading securities, available-for-sale securities are reported on the balance sheet at fair value. However, any unrealized gains and losses are not recognized in the income statement, but are reported in other comprehensive income as a part of shareholders' equity.

For all financial securities, dividend and interest income and realized gains and losses (actual gains or losses when the securities are sold) are recognized in the income statement.

Figure 19.1 summarizes the different classifications and measurement bases of financial assets under U.S. GAAP.

Figure 19.1: Financial Asset Measurement Bases—U.S. GAAP

Historical Cost	Amortized Cost	Fair Value
Unlisted equity investments	Held-to-maturity securities	Trading securities
Loans and notes receivable		Available-for-sale securities Derivatives

EXAMPLE: Classification of investment securities

Triple D Corporation, a U.S. GAAP reporting firm, purchased a 6% bond, at par, for \$1,000,000 at the beginning of the year. Interest rates have recently increased and the market value of the bond declined \$20,000. Determine the bond's effect on Triple D's financial statements under each classification of securities.

Answer:

If the bond is classified as a *held-to-maturity* security, the bond is reported on the balance sheet at \$1,000,000. Interest income of \$60,000 [$\$1,000,000 \times 6\%$] is reported in the income statement.

If the bond is classified as a *trading* security, the bond is reported on the balance sheet at \$980,000. The \$20,000 unrealized loss and \$60,000 of interest income are both recognized in the income statement.

If the bond is classified as an *available-for-sale* security, the bond is reported on the balance sheet at \$980,000. Interest income of \$60,000 is recognized in the income statement. The \$20,000 unrealized loss is reported as part of other comprehensive income.

IFRS Treatment of Marketable Securities

Recall from our discussion of accounting for marketable securities in the review of Understanding Income Statements that the three classifications of investment securities under IFRS are as follows:

- Securities measured at amortized cost (corresponds to the treatment of held-to-maturity securities under U.S. GAAP).
- Securities measured at fair value through other comprehensive income (corresponds to the treatment of available-for-sale securities under U.S. GAAP).

- Securities measured at fair value through profit and loss (corresponds to the treatment of trading securities under U.S. GAAP).

While the three different treatments are essentially the same as those used under U.S. GAAP, there are significant differences in how securities are classified under IFRS and under U.S. GAAP. Similarities and differences are as follows:

- Under both IFRS and U.S. GAAP, loans, notes receivable, debt securities a firm intends to hold until maturity, and unlisted securities for which fair value cannot be reliably determined, are all *measured at (amortized) historical cost*.
- Under IFRS, debt securities for which a firm intends to collect the interest payments but also to sell the securities are *measured at fair value through other comprehensive income*. This is similar to the treatment of available-for-sale securities under U.S. GAAP.
- Under IFRS, firms may make an irrevocable choice at the time of purchase to account for equity securities as *measured at fair value through other comprehensive income*. Equity securities cannot be classified as available for sale under U.S. GAAP.
- Under IFRS, financial assets that do not fit either of the other two classifications are *measured at fair value through profit and loss* (unrealized gains and losses reported on the income statement).
- Under IFRS, firms can make an irrevocable choice to carry any financial asset at *fair value through profit and loss*. This choice is not available under U.S. GAAP.

Figure 19.2 summarizes the different classifications of financial assets under IFRS.

Figure 19.2: Financial Asset Classifications—IFRS

Measured at amortized cost	Measured at fair value through other comprehensive income	Measured at fair value through profit and loss
Debt securities acquired with the intent to hold them to maturity	Debt securities acquired with intent to collect interest payments but sell before maturity	Debt securities acquired with intent to sell in near term
Loans receivable		Equity securities (unless fair value through OCI is chosen at time of purchase)
Notes receivable	Equity securities only if this treatment is chosen at time of purchase	Derivatives
Unlisted equity securities if fair value cannot be determined reliably		Any security not assigned to the other two categories
		Any security for which this treatment is chosen at time of purchase

Non-Current Liabilities

Long-term financial liabilities. Financial liabilities include bank loans, notes payable, bonds payable, and derivatives. If the financial liabilities are not issued at face amount, the liabilities are usually reported on the balance sheet at amortized cost. Amortized cost is equal to the issue price minus any principal payments, plus any amortized discount or minus any amortized premium.

In some cases, financial liabilities are reported at fair value. Examples include held-for-trading liabilities such as a short position in a stock (which may be classified as a short-term liability), derivative liabilities, and non-derivative liabilities with exposures hedged by derivatives.

Deferred tax liabilities. Deferred tax liabilities are amounts of income taxes payable in future periods as a result of taxable temporary differences. Deferred tax liabilities are created when the amount of income tax expense recognized in the income statement is greater than taxes payable. This can occur when expenses or losses are tax deductible before they are recognized in the income statement. A good example is when a firm uses an accelerated depreciation method for tax purposes and the straight-line method for financial reporting. Deferred tax liabilities are also created when revenues or gains are recognized in the income statement before they are taxable. For example, a firm often recognizes the earnings of a subsidiary before any distributions (dividends) are made. Eventually, deferred tax liabilities will reverse when the taxes are paid.

MODULE QUIZ 19.5, 19.6

1. SF Corporation has created employee goodwill by reorganizing its retirement benefit package. An independent management consultant estimated the value of the goodwill at \$2 million. In addition, SF recently purchased a patent that was developed by a competitor. The patent has an estimated useful life of five years. Should SF report the goodwill and patent on its balance sheet?

<u>Goodwill</u>	<u>Patent</u>
A. Yes	No
B. No	Yes
C. No	No

2. At the beginning of the year, Parent Company purchased all 500,000 shares of Sub Incorporated for \$15 per share. Just before the acquisition date, Sub's balance sheet reported net assets of \$6 million. Parent determined the fair value of Sub's property and equipment was \$1 million higher than reported by Sub. What amount of goodwill should Parent report as a result of its acquisition of Sub?
- A. \$0.
B. \$500,000.
C. \$1,500,000.
3. At the beginning of the year, Company P purchased \$80,000 face value of Company S corporate bonds for \$77,000. Company P intends to hold these bonds for several years but sell them before they mature. At the end of the year, the market value of the bonds was \$75,000. What amount should Company P report on its balance sheet at year-end for the investment in Company S bonds?
- A. \$75,000
B. \$77,000
C. \$80,000

MODULE 19.7: SHAREHOLDERS' EQUITY AND RATIOS



Video covering this content is available online.

LOS 19.f: Describe the components of shareholders' equity.

Owners' equity is the residual interest in assets that remains after subtracting an entity's liabilities. Owners' equity includes contributed capital, preferred stock, treasury stock, retained earnings, non-controlling interest, and accumulated other comprehensive income.

Contributed capital (also known as issued capital) is the amount contributed by equity shareholders.

The **par value** of common stock is a stated or legal value. Par value has no relationship to fair value. Some common shares are even issued without a par value. When par value exists, it is reported separately in stockholders' equity. In that case, the total proceeds from issuing an equity security are the par value of the issued shares plus "additional paid-in capital."

Also disclosed is the number of common shares that are authorized, issued, and outstanding.

Authorized shares are the number of shares that may be sold under the firm's articles of incorporation. **Issued shares** are the number of shares that have actually been sold to shareholders. The number of **outstanding shares** is equal to the issued shares less shares that have been reacquired by the firm (i.e., treasury stock).

Preferred stock has certain rights and privileges not conferred by common stock. For example, preferred shareholders are paid dividends at a specified rate, usually expressed as a percentage of par value, and have priority over the claims of the common shareholders in the event of liquidation.

Preferred stock can be classified as debt or equity, depending on the terms. For example, perpetual preferred stock that is non-redeemable is considered equity. However, preferred stock that calls for *mandatory redemption* in fixed amounts is considered a financial liability.

Noncontrolling interest (minority interest) is the minority shareholders' pro-rata share of the net assets (equity) of a subsidiary that is not wholly owned by the parent.

Retained earnings are the undistributed earnings (net income) of the firm since inception, the cumulative earnings that have not been paid out to shareholders as dividends.

Treasury stock is stock that has been reacquired by the issuing firm but not yet retired. Treasury stock reduces stockholders' equity. It does not represent an investment in the firm. Treasury stock has no voting rights and does not receive dividends.

Accumulated other comprehensive income includes all changes in stockholders' equity except for transactions recognized in the income statement (net income) and transactions with shareholders, such as issuing stock, reacquiring stock, and paying dividends.

As discussed in the reading on Understanding Income Statements, comprehensive income aggregates net income and certain special transactions that are not reported in the income statement but that affect stockholders' equity. These special transactions comprise what is known as "other comprehensive income." Comprehensive income is equal to net income plus other comprehensive income.



PROFESSOR'S NOTE

It is easy to confuse the two terms "comprehensive income" and "accumulated other comprehensive income." Comprehensive income is an income measure over a period of time. It includes net income and other comprehensive income for the period. Accumulated other comprehensive income does not include net income but is a component of stockholders' equity at a point in time.

The **statement of changes in stockholders' equity** summarizes all transactions that increase or decrease the equity accounts for the period. The statement includes transactions with

shareholders and reconciles the beginning and ending balance of each equity account, including capital stock, additional paid-in-capital, retained earnings, and accumulated other comprehensive income. In addition, the components of accumulated other comprehensive income are disclosed (i.e., unrealized gains and losses from available-for-sale securities, cash flow hedging derivatives, foreign currency translation, and adjustments for minimum pension liability).

A statement of changes in stockholders' equity is illustrated in Figure 19.3.

Figure 19.3: Sample Statement of Changes in Stockholders' Equity

	Common Stock	Retained Earnings (in thousands)	Accumulated Other Comprehensive Income (loss)	Total
Beginning balance	\$49,234	\$26,664	(\$406)	\$75,492
Net income		6,994		6,994
Net unrealized loss on available-for-sale securities			(40)	(40)
Net unrealized loss on cash flow hedges			(56)	(56)
Minimum pension liability			(26)	(26)
Cumulative translation adjustment			42	42
Comprehensive income				6,914
Issuance of common stock	1,282			1,282
Repurchases of common stock	(6,200)			(6,200)
Dividends		(2,360)		(2,360)
Ending balance	<u>\$44,316</u>	<u>\$31,298</u>	<u>(\$486)</u>	<u>\$75,128</u>

LOS 19.g: Demonstrate the conversion of balance sheets to common-size balance sheets and interpret common-size balance sheets.

A vertical **common-size balance sheet** expresses each item of the balance sheet as a percentage of total assets. The common-size format standardizes the balance sheet by eliminating the effects of size. This allows for comparison over time (time-series analysis) and across firms (cross-sectional analysis). For example, following are the balance sheets of industry competitors East Company and West Company.

	East	West
Cash	\$2,300	\$1,500
Accounts receivable	3,700	1,100
Inventory	<u>5,500</u>	<u>900</u>
Current assets	11,500	3,500
Plant and equipment	32,500	11,750
Goodwill	<u>1,750</u>	<u>0</u>
Total assets	\$45,750	\$15,250
Current liabilities	\$10,100	\$1,000
Long-term debt	<u>26,500</u>	<u>5,100</u>
Total liabilities	36,600	6,100
Equity	<u>9,150</u>	<u>9,150</u>
Total liabilities & equity	\$45,750	\$15,250

East is obviously the larger company. By converting the balance sheets to common-size format, we can eliminate the size effect.

	East	West
Cash	5%	10%
Accounts receivable	8%	7%
Inventory	<u>12%</u>	<u>6%</u>
Current assets	25%	23%
Plant and equipment	71%	77%
Goodwill	<u>4%</u>	<u>0%</u>
Total assets	100%	100%
Current liabilities	22%	7%
Long-term debt	<u>58%</u>	<u>33%</u>
Total liabilities	80%	40%
Equity	<u>20%</u>	<u>60%</u>
Total liabilities & equity	100%	100%

East's investment in current assets of 25% of total assets is slightly higher than West's current assets of 23%. However, East's current liabilities of 22% of total assets are significantly higher than West's current liabilities of 7%. Thus, East is less liquid and may have more difficulty paying its current obligations when due. However, West's superior working capital position may not be an efficient use of resources. The investment returns on working capital are usually lower than the returns on long-term assets.

A closer look at current assets reveals that East reports less cash as a percentage of assets than West. In fact, East does not have enough cash to satisfy its current liabilities without selling more inventory and collecting receivables. East's inventories of 12% of total assets are higher

than West's inventories of 6%. Carrying higher inventories may be an indication of inventory obsolescence. Further analysis of inventory is necessary.

Not only are East's current liabilities higher than West's, but East's long-term debt of 58% of total assets is much greater than West's long-term debt of 33%. Thus, East may have trouble satisfying its long-term obligations since its capital structure consists of more debt.

Common-size analysis can also be used to examine a firm's strategies. East appears to be growing through acquisitions since it is reporting goodwill. West is growing internally since no goodwill is reported. It could be that East is financing the acquisitions with debt.

LOS 19.h: Calculate and interpret liquidity and solvency ratios.

Balance sheet ratios compare balance sheet items only. Balance sheet ratios, along with common-size analysis, can be used to evaluate a firm's liquidity and solvency. The results should be compared over time (time-series analysis) and across firms (cross-sectional analysis).

Liquidity ratios measure the firm's ability to satisfy its short-term obligations as they come due. Liquidity ratios include the current ratio, the quick ratio, and the cash ratio.

$$\text{current ratio} = \frac{\text{current assets}}{\text{current liabilities}}$$

$$\text{quick ratio} = \frac{\text{cash} + \text{marketable securities} + \text{receivables}}{\text{current liabilities}}$$

$$\text{cash ratio} = \frac{\text{cash} + \text{marketable securities}}{\text{current liabilities}}$$

Although all three ratios measure the firm's ability to pay current liabilities, they should be considered collectively. For example, assume Firm A has a higher current ratio but a lower quick ratio as compared to Firm B. This is the result of higher inventory as compared to Firm B. The quick ratio (also known as the acid-test ratio) is calculated by excluding inventory from current assets. Similar analysis can be performed by comparing the quick ratio and the cash ratio. The cash ratio is calculated by excluding inventory and receivables.

Solvency ratios measure the firm's ability to satisfy its long-term obligations. Solvency ratios include the long-term debt-to-equity ratio, the total debt-to-equity ratio, the debt ratio, and the financial leverage ratio.

$$\text{long-term debt-to-equity} = \frac{\text{long-term debt}}{\text{total equity}}$$

$$\text{total debt-to-equity} = \frac{\text{total debt}}{\text{total equity}}$$

$$\text{debt ratio} = \frac{\text{total debt}}{\text{total assets}}$$

$$\text{financial leverage} = \frac{\text{total assets}}{\text{total equity}}$$

All four ratios measure solvency but they should be considered collectively. For example, Firm A might have a higher long-term debt-to-equity ratio but a lower total debt-to-equity ratio as

compared to Firm B. This is an indication that Firm B is utilizing more short-term debt to finance itself.

When calculating solvency ratios, debt is considered to be any interest-bearing obligation. On the other hand, the financial leverage ratio captures the impact of all obligations, both interest bearing and non-interest bearing.

Analysts must understand the limitations of balance sheet ratio analysis:

- Comparisons with peer firms are limited by differences in accounting standards and estimates.
- Lack of homogeneity as many firms operate in different industries.
- Interpretation of ratios requires significant judgment.
- Balance sheet data are only measured at a single point in time.



MODULE QUIZ 19.7

1. Miller Corporation has 160,000 shares of common stock authorized. There are 92,000 shares issued and 84,000 shares outstanding. How many shares of treasury stock does Miller own?

- A. 8,000.
- B. 68,000.
- C. 76,000.

2. Selected data from Alpha Company's balance sheet at the end of the year follows:

Investment in Beta Company, at fair value	\$150,000
Deferred taxes	\$86,000
Common stock, \$1 par value	\$550,000
Preferred stock, \$100 par value	\$175,000
Retained earnings	\$893,000
Accumulated other comprehensive income	\$46,000

The investment in Beta Company had an original cost of \$120,000. Assuming the investment in Beta is classified as available-for-sale, Alpha's total owners' equity at year-end is *closest* to:

- A. \$1,618,000.
- B. \$1,664,000.
- C. \$1,714,000.

3. A vertical common-size balance sheet expresses each category of the balance sheet as a percentage of:

- A. assets.
- B. equity.
- C. revenue.

4. Which of the following ratios are used to measure a firm's liquidity and solvency?

<u>Liquidity</u>	<u>Solvency</u>
A. Current ratio	Quick ratio
B. Debt-to-equity ratio	Financial leverage ratio
C. Cash ratio	Total debt ratio

KEY CONCEPTS

LOS 19.a

Assets are resources controlled as result of past transactions that are expected to provide future economic benefits. Liabilities are obligations as a result of past events that are expected to require an outflow of economic resources. Equity is the owners' residual interest in the assets after deducting the liabilities.

A financial statement item should be recognized if a future economic benefit to or from the firm is probable and the item's value or cost can be measured reliably.

LOS 19.b

The balance sheet can be used to assess a firm's liquidity, solvency, and ability to pay dividends to shareholders.

Balance sheet assets, liabilities, and equity should not be interpreted as market value or intrinsic value. For most firms, the balance sheet consists of a mixture of values including historical cost, amortized cost, and fair value.

Some assets and liabilities are difficult to quantify and are not reported on the balance sheet.

LOS 19.c

A classified balance sheet separately reports current and noncurrent assets and current and noncurrent liabilities. Alternatively, liquidity-based presentations, often used in the banking industry, present assets and liabilities in order of liquidity.

LOS 19.d

Current (noncurrent) assets are those expected to be used up or converted to cash in less than (more than) one year or the firm's operating cycle, whichever is greater.

Current (noncurrent) liabilities are those the firm expects to satisfy in less than (more than) one year or the firm's operating cycle, whichever is greater.

LOS 19.e

Cash equivalents are short-term, highly liquid financial assets that are readily convertible to cash. Their balance sheet values are generally close to identical using either amortized cost or fair value.

Accounts receivable are reported at net realizable value by estimating bad debt expense.

Inventories are reported at the lower of cost or net realizable value (IFRS) or the lower of cost or market (U.S. GAAP). Cost can be measured using standard costing or the retail method. Different cost flow assumptions can affect inventory values.

Property, plant, and equipment (PP&E) can be reported using the cost model or the revaluation model under IFRS. Under U.S. GAAP, only the cost model is allowed. PP&E is impaired if its carrying value exceeds the recoverable amount. Recoveries of impairment losses are allowed under IFRS but not U.S. GAAP.

Intangible assets created internally are expensed as incurred. Purchased intangibles are reported similar to PP&E. Under IFRS, research costs are expensed as incurred and development costs are capitalized. Both research and development costs are expensed under U.S. GAAP.

Goodwill is the excess of purchase price over the fair value of the identifiable net assets (assets minus liabilities) acquired in a business acquisition. Goodwill is not amortized but must be tested for impairment at least annually.

Under IFRS, debt securities acquired with intent hold them to maturity are measured at amortized cost. Debt securities acquired with the intent to collect interest payments but sell before maturity are measured at fair value through other comprehensive income. Debt securities acquired with the intent to sell them in the near term, as well as equity securities and derivatives, are measured at fair value through profit and loss.

IFRS permits firms to elect, irrevocably at the time of purchase, to measure equity securities at fair value through other comprehensive income, or any security at fair value through profit and loss.

Under U.S. GAAP, held-to-maturity securities are reported at amortized cost. Trading securities, available-for-sale securities, and derivatives are reported at fair value. For trading securities and derivatives, unrealized gains and losses are recognized in the income statement. Unrealized gains and losses for available-for-sale securities are reported in equity (other comprehensive income).

Accounts payable are amounts owed to suppliers for goods or services purchased on credit. Accrued liabilities are expenses that have been recognized in the income statement but are not yet contractually due. Unearned revenue is cash collected in advance of providing goods and services.

Financial liabilities not issued at face value, like bonds payable, are reported at amortized cost. Held-for-trading liabilities and derivative liabilities are reported at fair value.

LOS 19.f

Owners' equity includes:

- Contributed capital—the amount paid in by common shareholders.
- Preferred stock—capital stock that has certain rights and privileges not possessed by the common shareholders. Classified as debt if mandatorily redeemable.
- Treasury stock—issued common stock that has been repurchased by the firm.
- Retained earnings—the cumulative undistributed earnings of the firm since inception.
- Noncontrolling (minority) interest—the portion of a subsidiary that is not owned by the parent.
- Accumulated other comprehensive income—includes all changes to equity from sources other than net income and transactions with shareholders.

The statement of changes in stockholders' equity summarizes the transactions during a period that increase or decrease equity, including transactions with shareholders.

LOS 19.g

A vertical common-size balance sheet expresses each item of the balance sheet as a percentage of total assets. The common-size format standardizes the balance sheet by eliminating the effects of size. This allows for comparison over time (time-series analysis) and across firms (cross-sectional analysis).

LOS 19.h

Balance sheet ratios, along with common-size analysis, can be used to evaluate a firm's liquidity and solvency. Liquidity ratios measure the firm's ability to satisfy its short-term obligations as they come due. Liquidity ratios include the current ratio, the quick ratio, and the cash ratio.

Solvency ratios measure the firm's ability to satisfy its long-term obligations. Solvency ratios include the long-term debt-to-equity ratio, the total debt-to-equity ratio, the debt ratio, and the financial leverage ratio.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 19.1, 19.2

1. **C** An asset is a future economic benefit obtained or controlled as a result of past transactions. Some assets are intangible (e.g., goodwill), and others may be donated. (Module 19.1, LOS 19.a)
2. **C** The balance sheet lists the firm's assets, liabilities, and equity. The capital structure is measured by the mix of debt and equity used to finance the business. (Module 19.2, LOS 19.b)
3. **A** A classified balance sheet groups together similar items (e.g., current and noncurrent assets and liabilities) to arrive at significant subtotals. (Module 19.2, LOS 19.c)

Module Quiz 19.3, 19.4

1. **C** Estimated income taxes for the current year are likely reported as a current liability. To recognize the warranty expense, it must be probable, not just possible. Recognizing impairment of PP&E does not create a liability. (Module 19.3, LOS 19.d)
2. **C** The ticket revenue should not be recognized until it is earned. Even though the tickets are nonrefundable, the seller is still obligated to hold the event. (Module 19.3, LOS 19.e)
3. **A** Inventories are required to be valued at the lower of cost or net realizable value (or "market" under U.S. GAAP). FIFO and average cost are two of the inventory cost flow assumptions among which a firm has a choice. (Module 19.3, LOS 19.e)
4. **C** IFRS permits either the cost model or the revaluation model for property, plant, and equipment. (Module 19.4, LOS 19.d, 19.e)

Module Quiz 19.5, 19.6

1. **B** Goodwill developed internally is expensed as incurred. The purchased patent is reported on the balance sheet. (Module 19.5, LOS 19.e)
2. **B** Purchase price of \$7,500,000 [$\$15 \text{ per share} \times 500,000 \text{ shares}$] – fair value of net assets of \$7,000,000 [$\$6,000,000 \text{ book value} + \$1,000,000 \text{ increase in property and equipment}$] = goodwill of \$500,000. (Module 19.5, LOS 19.e)

- 3. A** Debt securities acquired with the intent to sell before maturity are reported on the balance sheet at their fair values. (Module 19.6, LOS 19.e)

Module Quiz 19.7

- 1. A** The difference between the issued shares and the outstanding shares is the treasury shares. (LOS 19.f)
- 2. B** Total stockholders' equity consists of common stock of \$550,000, preferred stock of \$175,000, retained earnings of \$893,000, and accumulated other comprehensive income of \$46,000, for a total of \$1,664,000. The \$30,000 unrealized gain from the investment in Beta is already included in accumulated other comprehensive income. (LOS 19.f)
- 3. A** Each category of the balance sheet is expressed as a percentage of total assets. (LOS 19.g)
- 4. C** The current ratio, quick ratio, and cash ratio measure liquidity. Debt-to-equity, the total debt ratio, and the financial leverage ratio measure solvency. (LOS 19.h)

1. IAS 32, *Financial Instruments: Presentation*, 32.11.

Reading 20

UNDERSTANDING CASH FLOW STATEMENTS

EXAM FOCUS

This reading covers the third important required financial statement: the statement of cash flows. Since the income statement is based on the accrual method, net income may not represent cash generated from operations. A company may be generating positive and growing net income but may be headed for insolvency because insufficient cash is being generated from operating activities. Constructing a statement of cash flows, by either the direct or indirect method, is therefore very important in an analysis of a firm's activities and prospects. Make sure you understand the preparation of a statement of cash flows by either method, the classification of various cash flows as operating, financing, or investing cash flows, and the key differences in these classifications between U.S. GAAP and international accounting standards.

MODULE 20.1: CASH FLOW INTRODUCTION



The **cash flow statement** provides information beyond that available from the income statement, which is based on accrual, rather than cash, accounting. The cash flow statement provides the following:

Video covering
this content is
available online.

- Information about a company's cash receipts and cash payments during an accounting period.
- Information about a company's operating, investing, and financing activities.
- An understanding of the impact of accrual accounting events on cash flows.

The cash flow statement provides information to assess the firm's liquidity, solvency, and financial flexibility. An analyst can use the statement of cash flows to determine whether:

- Regular operations generate enough cash to sustain the business.
- Enough cash is generated to pay off existing debts as they mature.
- The firm is likely to need additional financing.
- Unexpected obligations can be met.
- The firm can take advantage of new business opportunities as they arise.

LOS 20.a: Compare cash flows from operating, investing, and financing activities and classify cash flow items as relating to one of those three categories given a description of the items.

Items on the cash flow statement come from two sources: (1) income statement items and (2) changes in balance sheet accounts. A firm's cash receipts and payments are classified on the

cash flow statement as either operating, investing, or financing activities.

Cash flow from operating activities (CFO), sometimes referred to as “cash flow from operations” or “operating cash flow,” consists of the inflows and outflows of cash resulting from transactions that affect a firm’s net income.

Cash flow from investing activities (CFI) consists of the inflows and outflows of cash resulting from the acquisition or disposal of long-term assets and certain investments.

Cash flow from financing activities (CFF) consists of the inflows and outflows of cash resulting from transactions affecting a firm’s capital structure.

Examples of each cash flow classification, in accordance with U.S. GAAP, are presented in Figure 20.1.

Note that the acquisition of debt and equity investments (other than trading securities) and loans made to others are reported as investing activities; however, the income from these investments (interest and dividends received) is reported as an operating activity. Also, note that principal amounts borrowed from others are reported as financing activities; however, the interest paid is reported as an operating activity. Finally, note that dividends paid to the firm’s shareholders are financing activities.



PROFESSOR'S NOTE

Don’t confuse dividends received and dividends paid. Under U.S. GAAP, dividends received are operating cash flows and dividends paid are financing cash flows.

Figure 20.1: U.S. GAAP Cash Flow Classifications

Operating Activities	
Inflows	Outflows
Cash collected from customers	Cash paid to employees and suppliers
Interest and dividends received	Cash paid for other expenses
Sale proceeds from trading securities	Acquisition of trading securities
	Interest paid on debt or leases
	Taxes paid
Investing Activities	
Inflows	Outflows
Sale proceeds from fixed assets	Acquisition of fixed assets
Sale proceeds from debt and equity investments	Acquisition of debt and equity investments
Principal received from loans made to others	Loans made to others
Financing Activities	
Inflows	Outflows
Principal amounts of debt issued	Principal paid on debt or leases
Proceeds from issuing stock	Payments to reacquire stock
	Dividends paid to shareholders

LOS 20.b: Describe how non-cash investing and financing activities are reported.

Noncash investing and financing activities are not reported in the cash flow statement since they do not result in inflows or outflows of cash.

For example, if a firm acquires real estate with financing provided by the seller, the firm has made an investing and financing decision. This transaction is the equivalent of borrowing the purchase price. However, since no cash is involved in the transaction, it is not reported as an investing and financing activity in the cash flow statement.

Another example of a noncash transaction is an exchange of debt for equity. Such an exchange results in a reduction of debt and an increase in equity. However, since no cash is involved in the transaction, it is not reported as a financing activity in the cash flow statement.

Noncash transactions must be disclosed in either a footnote or supplemental schedule to the cash flow statement. Analysts should be aware of the firm's noncash transactions, incorporate them into analysis of past and current performance, and include their effects in estimating future cash flows.

LOS 20.c: Contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP).

Recall from Figure 20.1 that under U.S. GAAP, dividends paid to the firm's shareholders are reported as financing activities while interest paid is reported in operating activities. Interest received and dividends received from investments are also reported as operating activities.

International Financial Reporting Standards (IFRS) allow more flexibility in the classification of cash flows. Under IFRS, interest and dividends received may be classified as either operating *or* investing activities. Dividends paid to the company's shareholders and interest paid on the company's debt may be classified as either operating *or* financing activities.

Another important difference relates to income taxes paid. Under U.S. GAAP, all taxes paid are reported as operating activities, even taxes related to investing and financing transactions. Under IFRS, income taxes are also reported as operating activities unless the expense is associated with an investing or financing transaction.

For example, consider a company that sells land that was held for investment for \$1 million. Income taxes on the sale total \$160,000. Under U.S. GAAP, the firm reports an inflow of cash from investing activities of \$1 million and an outflow of cash from operating activities of \$160,000. Under IFRS, the firm can report a net inflow of \$840,000 from investing activities.

LOS 20.d: Compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method.

There are two methods of presenting the cash flow statement: the direct method and the indirect method. Both methods are permitted under U.S. GAAP and IFRS. The use of the direct method, however, is encouraged by both standard setters. Regrettably, most firms use the indirect method. The difference between the two methods relates to the presentation of cash flow from operating activities. The presentation of cash flows from investing activities and financing activities is exactly the same under both methods.

Direct Method

Under the **direct method**, each line item of the accrual-based income statement is converted into cash receipts or cash payments. Recall that under the accrual method of accounting, the timing of revenue and expense recognition may differ from the timing of the related cash flows. Under cash-basis accounting, revenue and expense recognition occur when cash is received or paid. Simply stated, the direct method converts an accrual-basis income statement into a cash-basis income statement.

Figure 20.2 contains an example of a presentation of operating cash flow for Seagraves Supply Company using the direct method.

Figure 20.2: Direct Method of Presenting Operating Cash Flow

Seagraves Supply Company	
Operating Cash Flow – Direct Method	
For the year ended December 31, 20X7	
Cash collections from customers	\$429,980
Cash paid to suppliers	(265,866)
Cash paid for operating expenses	(124,784)
Cash paid for interest	(4,326)
Cash paid for taxes	<u>(14,956)</u>
Operating cash flow	\$20,048

Notice the similarities of the direct method cash flow presentation and an income statement. The direct method begins with cash inflows from customers and then deducts cash outflows for purchases, operating expenses, interest, and taxes.

Indirect Method

Under the **indirect method**, net income is converted to operating cash flow by making adjustments for transactions that affect net income but are not cash transactions. These adjustments include eliminating noncash expenses (e.g., depreciation and amortization), nonoperating items (e.g., gains and losses), and changes in balance sheet accounts resulting from accrual accounting events.

Figure 20.3 contains an example of a presentation of operating cash flow for Seagraves Supply Company under the indirect method.

Figure 20.3: Indirect Method of Presenting Operating Cash Flow

Seagraves Supply Company
Operating Cash Flow – Indirect Method
For the year ended December 31, 20X7

Net income	\$18,788
Adjustments to reconcile net income to cash flow provided by operating activities:	
Depreciation and amortization	7,996
Deferred income taxes	416
Increase in accounts receivable	(1,220)
Increase in inventory	(20,544)
Decrease in prepaid expenses	494
Increase in accounts payable	13,406
Increase in accrued liabilities	<u>712</u>
Operating cash flow	<u>\$20,048</u>

Notice that under the indirect method, the starting point is net income, the “bottom line” of the income statement. Under the direct method, the starting point is the top of the income statement, revenues, adjusted to show cash received from customers. Total cash flow from operating activities is exactly the same under both methods, only the presentation methods differ.

Arguments in Favor of Each Method

The primary advantage of the direct method is that it presents the firm’s operating cash receipts and payments, while the indirect method only presents the net result of these receipts and payments. Therefore, the direct method provides more information than the indirect method. This knowledge of past receipts and payments is useful in estimating future operating cash flows.

The main advantage of the indirect method is that it focuses on the difference between net income and operating cash flow. This provides a useful link to the income statement when forecasting future operating cash flow. Analysts forecast net income and then derive operating cash flow by adjusting net income for the differences between accrual accounting and the cash basis of accounting.

Disclosure Requirements

Under U.S. GAAP, a direct method presentation must also disclose the adjustments necessary to reconcile net income to cash flow from operating activities. This disclosure is the same information that is presented in an indirect method cash flow statement. This reconciliation is not required under IFRS.

Under IFRS, payments for interest and taxes must be disclosed separately in the cash flow statement under either method (direct or indirect). Under U.S. GAAP, payments for interest and taxes can be reported in the cash flow statement or disclosed in the footnotes.

LOS 20.e: Describe how the cash flow statement is linked to the income statement and the balance sheet.

The cash flow statement reconciles the beginning and ending balances of cash over an accounting period. The change in cash is a result of the firm's operating, investing, and financing activities as follows:

$$\begin{aligned} & \text{Operating cash flow} \\ + & \text{Investing cash flow} \\ + & \underline{\text{Financing cash flow}} \\ = & \text{Change in cash balance} \\ + & \underline{\text{Beginning cash balance}} \\ = & \text{Ending cash balance} \end{aligned}$$

With a few exceptions, operating activities relate to the firm's current assets and current liabilities. Investing activities typically relate to the firm's noncurrent assets, and financing activities typically relate to the firm's noncurrent liabilities and equity.

Transactions for which the timing of revenue or expense recognition differs from the receipt or payment of cash are reflected in changes in balance sheet accounts. For example, when revenues (sales) exceed cash collections, the firm has sold items on credit and accounts receivable (an asset) increase. The opposite occurs when customers repay more on their outstanding accounts than the firm extends in new credit: cash collections exceed revenues and accounts receivable decrease. When purchases from suppliers exceed cash payments, accounts payable (a liability) increase. When cash payments exceed purchases, payables decrease.

It is helpful to understand how transactions affect each balance sheet account. For example, accounts receivable are increased by sales and decreased by cash collections. We can summarize this relationship as follows:

$$\begin{aligned} & \underline{\text{Beginning accounts receivable}} \\ + & \text{Sales} \\ - & \underline{\text{Cash collections}} \\ = & \text{Ending accounts receivable} \end{aligned}$$

Knowing three of the four variables, we can solve for the fourth. For example, if beginning accounts receivable are €10,000, ending accounts receivable are €15,000, and sales are €68,000, then cash collections must equal €63,000.

Understanding these interrelationships is not only useful in preparing the cash flow statement, but is also helpful in uncovering accounting shenanigans.



MODULE QUIZ 20.1

1. Which of the following items is *least likely* considered a cash flow from financing activity under U.S. GAAP?
 - A. Receipt of cash from the sale of bonds.
 - B. Payment of cash for dividends.
 - C. Payment of interest on debt.

2. Which of the following would be *least likely* to cause a change in investing cash flow?
 - A. The sale of a division of the company.
 - B. The purchase of new machinery.
 - C. An increase in depreciation expense.
3. Which of the following is *least likely* a change in cash flow from operations under U.S. GAAP?
 - A. A decrease in notes payable.
 - B. An increase in interest expense.
 - C. An increase in accounts payable.
4. Sales of inventory would be classified as:
 - A. operating cash flow.
 - B. investing cash flow.
 - C. financing cash flow.
5. Issuing bonds would be classified as:
 - A. investing cash flow.
 - B. financing cash flow.
 - C. no cash flow impact.
6. Sale of land would be classified as:
 - A. operating cash flow.
 - B. investing cash flow.
 - C. financing cash flow.
7. The write-off of obsolete equipment would be classified as:
 - A. operating cash flow.
 - B. investing cash flow.
 - C. no cash flow impact.
8. Under IFRS, interest expense would be classified as:
 - A. either operating cash flow or financing cash flow.
 - B. operating cash flow only.
 - C. financing cash flow only.
9. Under U.S. GAAP, dividends received from investments would be classified as:
 - A. operating cash flow.
 - B. investing cash flow.
 - C. financing cash flow.
10. Torval, Inc., retires debt securities by issuing equity securities. This is considered a:
 - A. cash flow from investing.
 - B. cash flow from financing.
 - C. noncash transaction.
11. Where are dividends paid to shareholders reported in the cash flow statement under U.S. GAAP and IFRS?

<u>U.S. GAAP</u>	<u>IFRS</u>
A. Operating or financing activities	Operating or financing activities
B. Financing activities	Operating or financing activities
C. Operating activities	Financing activities

12. From an analyst's perspective, an advantage of the indirect method for presenting operating cash flow is that the indirect method:
 - A. shows operating cash received and paid.
 - B. provides more information than the direct method.
 - C. shows the difference between net income and operating cash flow.
13. Which balance sheet items are *most likely* to be linked to cash flows from financing?

- A. Long-lived assets.
- B. Current assets and liabilities.
- C. Long-term liabilities and equity.

MODULE 20.2: THE DIRECT AND INDIRECT METHODS



Video covering this content is available online.

LOS 20.f: Describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data.



PROFESSOR'S NOTE

Throughout the discussion of the direct and indirect methods, remember the following points:

- CFO is calculated differently, but the result is the same under both methods.
- The calculation of CFI and CFF is identical under both methods.
- There is an inverse relationship between changes in assets and changes in cash flows. In other words, an increase in an asset account is a use of cash, and a decrease in an asset account is a source of cash.
- There is a direct relationship between changes in liabilities and changes in cash flow. In other words, an increase in a liability account is a source of cash, and a decrease in a liability is a use of cash.
- Sources of cash are positive numbers (cash inflows) and uses of cash are negative numbers (cash outflows).

Direct Method

The direct method of presenting a firm's statement of cash flows shows only cash payments and cash receipts over the period. The sum of these inflows and outflows is the company's CFO. The direct method gives the analyst more information than the indirect method. The analyst can see the actual amounts that went to each use of cash and that were received from each source of cash. This information can help the analyst to better understand the firm's performance over time and to forecast future cash flows.

The following are common components of cash flow that appear on a statement of cash flow presented under the direct method:

- Cash collected from customers, typically the main component of CFO.
- Cash used in the production of goods and services (cash inputs).
- Cash operating expenses.
- Cash paid for interest.
- Cash paid for taxes.



PROFESSOR'S NOTE

A common "trick" in direct method questions is to provide information on depreciation expense along with other operating cash flow components. When using the direct method, ignore depreciation expense—it's a noncash charge. We'll see later that we do consider depreciation expense in indirect method computations, but we do this solely because depreciation expense and

other noncash expenses have been subtracted in calculating net income (our starting point) and need to be added back to get cash flow.

Investing cash flows (CFI) are calculated by examining the change in the gross asset accounts that result from investing activities, such as property, plant, and equipment, intangible assets, and investment securities. Related accumulated depreciation or amortization accounts are ignored since they do not represent cash expenses.

PROFESSOR'S NOTE

In this context, "gross" simply means an amount that is presented on the balance sheet before deducting any accumulated depreciation or amortization.

When calculating cash paid for a new asset, it is necessary to determine whether old assets were sold. If assets were sold during the period, you must use the following formula:

$$\text{cash paid for new asset} = \text{ending gross assets} + \text{gross cost of old assets sold} - \text{beginning gross assets}$$

PROFESSOR'S NOTE

It may be easier to think in terms of the account reconciliation format discussed earlier. That is, beginning gross assets + cash paid for new assets - gross cost of assets sold = ending gross assets. Given three of the variables, simply solve for the fourth.

When calculating the cash flow from an asset that has been sold, it is necessary to consider any gain or loss from the sale using the following formula:

$$\text{cash from asset sold} = \text{book value of the asset} + \text{gain (or - loss) on sale}$$

Financing cash flows (CFF) are determined by measuring the cash flows occurring between the firm and its suppliers of capital. Cash flows between the firm and its creditors result from new borrowings (positive CFF) and debt principal repayments (negative CFF). Note that interest paid is technically a cash flow to creditors, but it is included in CFO under U.S. GAAP. Cash flows between the firm and its shareholders occur when equity is issued, shares are repurchased, or dividends are paid. CFF is the sum of these two measures:

$$\text{net cash flows from creditors} = \text{new borrowings} - \text{principal amounts repaid}$$

$$\text{net cash flows from shareholders} = \text{new equity issued} - \text{share repurchases} - \text{cash dividends paid}$$

Cash dividends paid can be calculated from dividends declared and any changes in dividends payable.

Finally, total cash flow is equal to the sum of CFO, CFI, and CFF. If calculated correctly, the total cash flow will equal the change in cash from one balance sheet to the next.

Indirect Method

Cash flow from operations is presented differently under the indirect method, but the amount of CFO is the same under either method. Cash flow from financing and cash flow from investing are presented in the same way on cash flow statements prepared under both the direct and indirect methods of presenting the statement of cash flows.

Under the indirect method of presenting CFO, we begin with net income and adjust it for differences between accounting items and actual cash receipts and cash disbursements. Depreciation, for example, is deducted in calculating net income, but requires no cash outlay in the current period. Therefore, we must add depreciation (and amortization) to net income for the period in calculating CFO.

Another adjustment to net income on an indirect statement of cash flows is to subtract gains on the disposal of assets. Proceeds from the sale of fixed assets are an investing cash flow. Since gains are a portion of such proceeds, we need to subtract them from net income in calculating CFO under the indirect method. Conversely, a loss would be added back to net income in calculating CFO under the indirect method.

Under the indirect method, we also need to adjust net income for change in balance sheet accounts. If, for example, accounts receivable went up during the period, we know that sales during the period were greater than the cash collected from customers. Since sales were used to calculate net income under the accrual method, we need to reduce net income to reflect the fact that credit sales, rather than cash collected were used in calculating net income.

A change in accounts payable indicates a difference between purchases and the amount paid to suppliers. An increase in accounts payable, for example, results when purchases are greater than cash paid to suppliers. Since purchases were subtracted in calculating net income, we need to add any increase in accounts payable to net income so that CFO reflects the actual cash disbursements for purchases (rather than total purchases).

The steps in calculating CFO under the indirect method can be summarized as follows:

Step 1: Begin with net income.

Step 2: Add or subtract changes to balance sheet operating accounts as follows:

- Increases in the operating asset accounts (uses of cash) are subtracted, while decreases (sources of cash) are added.
- Increases in the operating liability accounts (sources of cash) are added, while decreases (uses of cash) are subtracted.

Step 3: Add back all noncash charges to income (such as depreciation and amortization) and subtract all noncash components of revenue.

Step 4: Subtract gains or add losses that resulted from financing or investing cash flows (such as gains from sale of land).

EXAMPLE: Statement of cash flows using the indirect method

Use the following balance sheet and income statement to prepare a statement of cash flows under the indirect method.

Income Statement for 20X7

Sales	\$100,000
Expense	
Cost of goods sold	40,000
Wages	5,000
Depreciation	7,000
Interest	<u>1,000</u>
Total expenses	<u>\$53,000</u>
Income from continuing operations	\$47,000
Gain from sale of land	<u>10,000</u>
Pretax income	57,000
Provision for taxes	<u>20,000</u>
Net income	<u>\$37,000</u>
Common dividends declared	<u>\$8,500</u>

Balance Sheets for 20X7 and 20X6

	20X7	20X6
Assets		
Current assets		
Cash	\$33,000	\$9,500
Accounts receivable	10,000	9,000
Inventory	5,000	7,000
Noncurrent assets		
Land	\$35,000	\$40,000
Gross plant and equipment	85,000	60,000
Less: Accumulated depreciation	<u>(16,000)</u>	<u>(9,000)</u>
Net plant and equipment	<u>\$69,000</u>	<u>\$51,000</u>
Goodwill	<u>10,000</u>	<u>10,000</u>
Total assets	<u>\$162,000</u>	<u>\$126,500</u>

Liabilities		
Current liabilities		
Accounts payable	\$9,000	\$5,000
Wages payable	4,500	8,000
Interest payable	3,500	3,000
Taxes payable	5,000	4,000
Dividends payable	<u>6,000</u>	<u>1,000</u>
Total current liabilities	28,000	21,000
Noncurrent liabilities		
Bonds	\$15,000	\$10,000
Deferred tax liability	20,000	15,000
Total liabilities	\$63,000	\$46,000
Stockholders' equity		
Common stock	\$40,000	\$50,000
Retained earnings	<u>59,000</u>	<u>30,500</u>
Total equity	\$99,000	\$80,000
Total liabilities and stockholders' equity	\$162,000	\$126,500

Any discrepancies between the changes in accounts reported on the balance sheet and those reported in the statement of cash flows are typically due to business combinations and changes in exchange rates.

Answer:

Operating cash flow:

Step 1: Start with net income of \$37,000.

Step 2: Subtract gain from sale of land of \$10,000.

Step 3: Add back noncash charges of depreciation of \$7,000.

Step 4: Subtract increases in receivables and inventories and add increases of payables and deferred taxes.

Net income	\$37,000
Gain from sale of land	(10,000)
Depreciation	<u>7,000</u>
Subtotal	<u>\$34,000</u>
Changes in operating accounts	
Increase in receivables	(\$1,000)
Decrease in inventories	2,000
Increase in accounts payable	4,000
Decrease in wages payable	(3,500)
Increase in interest payable	500
Increase in taxes payable	1,000
Increase in deferred taxes	<u>5,000</u>
Cash flow from operations	\$42,000

Investing cash flow:

In this example, we have two components of investing cash flow: the sale of land and the change in gross plant and equipment (P&E).

cash from sale of land = decrease in asset + gain on sale = \$5,000 + \$10,000 = \$15,000 (source)

beginning land + land purchased - gross cost of land sold = ending land = \$40,000 + \$0 - \$5,000 = \$35,000

Note: If the land had been sold at a loss, we would have subtracted the loss amount from the decrease in land.

P&E purchased = ending gross P&E + gross cost of P&E sold - beginning gross P&E = \$85,000 + \$0 - \$60,000 = \$25,000 (use)

beginning gross P&E + P&E purchased - gross cost of P&E sold = ending P&E = \$60,000 + \$25,000 - \$0 = \$85,000

Cash from sale of land	\$15,000
Purchase of plant and equipment	(25,000)
Cash flow from investments	<u>(\$10,000)</u>

Financing cash flow:

cash from bond issue = ending bonds payable + bonds repaid - beginning bonds payable = \$15,000 + \$0 - \$10,000 = \$5,000 (source)

beginning bonds payable + bonds issued - bonds repaid = ending bonds payable = \$10,000 + \$5,000 - \$0 = \$15,000

cash to reacquire stock = beginning common stock + stock issued - ending common stock = \$50,000 + \$0 - \$40,000 = \$10,000 (use, or a net share repurchase of \$10,000)

beginning common stock + stock issued - stock reacquired = ending common stock = \$50,000 + \$0 - \$10,000 = \$40,000

cash dividends = -dividend declared + increase in dividends payable = -\$8,500* + \$5,000 = -\$3,500 (use)

beginning dividends payable + dividends declared - dividends paid = ending dividends payable = \$1,000 + \$8,500 - \$3,500 = \$6,000

*Note: If the dividend declared amount is not provided, you can calculate the amount as follows: dividends declared = beginning retained earnings + net income - ending retained earnings. Here, \$30,500 + \$37,000 - \$59,000 = \$8,500.

Sale of bonds	\$5,000
Repurchase of stock	(10,000)
Cash dividends	<u>(3,500)</u>
Cash flow from financing	<u>(\$8,500)</u>

Total cash flow:

Cash flow from operations	\$42,000
Cash flow from investments	(10,000)
Cash flow from financing	<u>(8,500)</u>
Total cash flow	<u>\$23,500</u>

The total cash flow of \$23,500 is equal to the increase in the cash account. The difference between beginning cash and ending cash should be used as a check figure to ensure that the total cash flow calculation is correct.

Both IFRS and U.S. GAAP encourage the use of a statement of cash flows in the direct format. Under U.S. GAAP, a statement of cash flows under the direct method must include footnote disclosure of the indirect method. Most companies however, report cash flows using the indirect method, which requires no additional disclosure. The next LOS illustrates the method an analyst will use to create a statement of cash flows in the direct method format when the company reports using the indirect method.



MODULE QUIZ 20.2

1. Using the following information, what is the firm's cash flow from operations?

Net income	\$120
Decrease in accounts receivable	20
Depreciation	25
Increase in inventory	10
Increase in accounts payable	7
Decrease in wages payable	5
Increase in deferred tax liabilities	15
Profit from the sale of land	2

- A. \$158.
- B. \$170.
- C. \$174.

Assuming U.S. GAAP, use the following data to answer Questions 2 through 4.

Net income	\$45
Depreciation	75
Taxes paid	25
Interest paid	5
Dividends paid	10
Cash received from sale of company building	40
Issuance of preferred stock	35
Repurchase of common stock	30
Purchase of machinery	20
Issuance of bonds	50
Debt retired through issuance of common stock	45
Paid off long-term bank borrowings	15
Profit on sale of building	20

2. Cash flow from operations is:

- A. \$70.
- B. \$100.
- C. \$120.

3. Cash flow from investing activities is:

- A. -\$30.
 B. \$20.
 C. \$50.
4. Cash flow from financing activities is:
 A. \$30.
 B. \$55.
 C. \$75.
5. Given the following:

Sales	\$1,500
Increase in inventory	100
Depreciation	150
Increase in accounts receivable	50
Decrease in accounts payable	70
After-tax profit margin	25%
Gain on sale of machinery	\$30

Cash flow from operations is:

- A. \$115.
 B. \$275.
 C. \$375.
6. Net income for Monique, Inc., for the year ended December 31, 20X7 was \$78,000. Its accounts receivable balance at December 31, 20X7 was \$121,000, and this balance was \$69,000 at December 31, 20X6. The accounts payable balance at December 31, 20X7 was \$72,000 and was \$43,000 at December 31, 20X6. Depreciation for 20X7 was \$12,000, and there was an unrealized gain of \$15,000 included in 20X7 income from the change in value of trading securities. Which of the following amounts represents Monique's cash flow from operations for 20X7?
 A. \$52,000.
 B. \$67,000.
 C. \$82,000.

7. Martin, Inc., had the following transactions during 20X7:

- Purchased new fixed assets for \$75,000.
- Converted \$70,000 worth of preferred shares to common shares.
- Received cash dividends of \$12,000. Paid cash dividends of \$21,000.
- Repaid mortgage principal of \$17,000.

Assuming Martin follows U.S. GAAP, which of the following amounts represents Martin's cash flows from investing and cash flows from financing in 20X7, respectively?

<u>Cash flows from investing</u>	<u>Cash flows from financing</u>
A. (\$5,000)	(\$21,000)
B. (\$75,000)	(\$21,000)
C. (\$75,000)	(\$38,000)

MODULE 20.3: CONVERTING INDIRECT TO DIRECT



Video covering
this content is

LOS 20.g: Demonstrate the conversion of cash flows from the indirect to direct method.

available online.

The only difference between the indirect and direct methods of presentation is in the cash flow from operations (CFO) section. CFO under the direct method can be computed using a combination of the income statement and a statement of cash flows prepared under the indirect method.

There are two major sections in CFO under the direct method: cash inflows (receipts) and cash outflows (payments). We will illustrate the conversion process using some frequently used accounts. Please note that the following list is for illustrative purposes only and is far from all-inclusive of what may be encountered in practice. The general principle here is to adjust each income statement item for its corresponding balance sheet accounts and to eliminate noncash and nonoperating transactions.

Cash collections from customers:

1. Begin with net sales from the income statement.
2. Subtract (add) any increase (decrease) in the accounts receivable balance as reported in the indirect method. If the company has sold more on credit than has been collected from customers, accounts receivable will increase and cash collections will be less than net sales.
3. Add (subtract) an increase (decrease) in unearned revenue. Unearned revenue includes cash advances from customers. Cash received from customers when the goods or services have yet to be delivered is not included in net sales, so the advances must be added to net sales in order to calculate cash collections.

Cash payments to suppliers:

1. Begin with cost of goods sold (COGS) as reported in the income statement.
2. If depreciation and/or amortization have been included in COGS (they increase COGS), these noncash expenses must be added back when computing the cash paid to suppliers.
3. Reduce (increase) COGS by any increase (decrease) in the accounts payable balance as reported in the indirect method. If payables have increased, then more was spent on credit purchases during the period than was paid on existing payables, so cash payments are reduced by the amount of the increase in payables.
4. Add (subtract) any increase (decrease) in the inventory balance as disclosed in the indirect method. Increases in inventory are not included in COGS for the period but still represent the purchase of inputs, so they increase cash paid to suppliers.
5. Subtract an inventory write-off that occurred during the period. An inventory write-off, as a result of applying the lower of cost or market rule, will reduce ending inventory and increase COGS for the period. However, no cash flow is associated with the write-off.

Other items in a direct method cash flow statement follow the same principles. Cash taxes paid, for example, can be derived by starting with income tax expense on the income statement. Adjustment must be made for changes in related balance sheet accounts (deferred tax assets and liabilities, and income taxes payable).

Cash operating expense is equal to selling, general, and administrative expense (SG&A) from the income statement, increased (decreased) for any increase (decrease) in prepaid expenses. Any increase in prepaid expenses is a cash outflow that is not included in SG&A for the current period.

EXAMPLE: Direct method for computing CFO

Prepare a cash flow statement using the direct method, based on the indirect statement of cash flows, balance sheet, and income statement from the previous example.

Answer:



PROFESSOR'S NOTE

There are many ways to think about these calculations and lots of sources and uses and pluses and minuses to keep track of. It's easier if you use a "+" sign for net sales and a "-" sign for cost of goods sold and other cash expenses used as the starting points. Doing so will allow you to consistently follow the rule that an increase in assets or decrease in liabilities is a use of cash and a decrease in assets or an increase in liabilities is a source. We'll use this approach in the answer to the example. Remember, sources are always + and uses are always -.

The calculations that follow include a reconciliation of each account, analyzing the transactions that increase and decrease the account for the period. As previously discussed, this reconciliation is useful in understanding the interrelationships between the balance sheet, income statement, and cash flow statement.

Cash from operations:

Keep track of the balance sheet items used to calculate CFO by marking them off the balance sheet. They will not be needed again when determining CFI and CFF.

$$\text{cash collections} = \text{sales} - \text{increase in accounts receivable} = \$100,000 - \$1,000 = \$99,000$$

$$\text{beginning receivables} + \text{sales} - \text{cash collections} = \text{ending receivables} = \$9,000 + \$100,000 - \$99,000 = \$10,000$$

$$\text{cash paid to suppliers} = -\text{COGS} + \text{decrease in inventory} + \text{increase in accounts payable} = -\$40,000 + \$2,000 + \$4,000 = -\$34,000$$

$$\text{beginning inventory} + \text{purchases} - \text{COGS} = \text{ending inventory} = \$7,000 + \$38,000 \text{ (not provided)} - \$40,000 = \$5,000$$

$$\text{beginning accounts payable} + \text{purchases} - \text{cash paid to suppliers} = \text{ending accounts payable} = \$5,000 + \$38,000 \text{ (not provided)} - \$34,000 = \$9,000$$

$$\text{cash wages} = -\text{wages} - \text{decrease in wages payable} = -\$5,000 - \$3,500 = -\$8,500$$

$$\text{beginning wages payable} + \text{wages expense} - \text{wages paid} = \text{ending wages payable} = \$8,000 + \$5,000 - \$8,500 = \$4,500$$

$$\text{cash interest} = -\text{interest expense} + \text{increase in interest payable} = -\$1,000 + \$500 = -\$500$$

$$\text{beginning interest payable} + \text{interest expense} - \text{interest paid} = \text{ending interest payable} = \$3,000 + \$1,000 - \$500 = \$3,500$$

$$\text{cash taxes} = -\text{tax expense} + \text{increase in taxes payable} + \text{increase in deferred tax liability} = -\$20,000 + \$1,000 + \$5,000 = -\$14,000$$

$$\text{beginning taxes payable} + \text{beginning deferred tax liability} + \text{tax expense} - \text{taxes paid} = \text{ending taxes payable} + \text{ending deferred tax liability} = \$4,000 + \$15,000 + \$20,000 - \$14,000 = \$5,000 + \$20,000$$

Cash collections	\$99,000
Cash to suppliers	(34,000)
Cash wages	(8,500)
Cash interest	(500)
Cash taxes	<u>(14,000)</u>
Cash flow from operations	<u>\$42,000</u>

LOS 20.h: Analyze and interpret both reported and common-size cash flow statements.

Major Sources and Uses of Cash

Cash flow analysis begins with an evaluation of the firm's sources and uses of cash from operating, investing, and financing activities. Sources and uses of cash change as the firm moves through its life cycle. For example, when a firm is in the early stages of growth, it may experience negative operating cash flow as it uses cash to finance increases in inventory and receivables. This negative operating cash flow is usually financed externally by issuing debt or equity securities. These sources of financing are not sustainable. Eventually, the firm must begin generating positive operating cash flow or the sources of external capital may no longer be available. Over the long term, successful firms must be able to generate operating cash flows that exceed capital expenditures and provide a return to debt and equity holders.

Operating Cash Flow

An analyst should identify the major determinants of operating cash flow. Positive operating cash flow can be generated by the firm's earnings-related activities. However, positive operating cash flow can also be generated by decreasing noncash working capital, such as liquidating inventory and receivables or increasing payables. Decreasing noncash working capital is not sustainable, since inventories and receivables cannot fall below zero and creditors will not extend credit indefinitely unless payments are made when due.

Operating cash flow also provides a check of the quality of a firm's earnings. A stable relationship of operating cash flow and net income is an indication of quality earnings. (This relationship can also be affected by the business cycle and the firm's life cycle.) Earnings that significantly exceed operating cash flow may be an indication of aggressive (or even improper) accounting choices such as recognizing revenues too soon or delaying the recognition of expenses. The variability of net income and operating cash flow should also be considered.

Investing Cash Flow

The sources and uses of cash from investing activities should be examined. Increasing capital expenditures, a use of cash, is usually an indication of growth. Conversely, a firm may reduce capital expenditures or even sell capital assets in order to save or generate cash. This may result in higher cash outflows in the future as older assets are replaced or growth resumes. As mentioned previously, generating operating cash flow that exceeds capital expenditures is a desirable trait.

Financing Cash Flow

The financing activities section of the cash flow statement reveals information about whether the firm is generating cash flow by issuing debt or equity. It also provides information about whether the firm is using cash to repay debt, reacquire stock, or pay dividends. For example, an analyst would certainly want to know if a firm issued debt and used the proceeds to reacquire stock or pay dividends to shareholders.

Common-Size Cash Flow Statement

Like the income statement and balance sheet, common-size analysis can be used to analyze the cash flow statement.

The cash flow statement can be converted to common-size format by expressing each line item as a percentage of revenue. Alternatively, each inflow of cash can be expressed as a percentage of total cash inflows, and each outflow of cash can be expressed as a percentage of total cash outflows.

A revenue based common-size cash flow statement is useful in identifying trends and forecasting future cash flow. Since each line item of the cash flow statement is stated in terms of revenue, once future revenue is forecast, cash flows can be estimated for those items that are tied to revenue.

EXAMPLE: Common-size cash flow statement analysis

Triple Y Corporation's common-size cash flow statement is shown in the table below. Explain the decrease in Triple Y's total cash flow as a percentage of revenues.

Triple Y Corporation

Year	Cash Flow Statement (Percent of Revenues)		
	20X9	20X8	20X7
Net income	13.4%	13.4%	13.5%
Depreciation	4.0%	3.9%	3.9%
Accounts receivable	-0.6%	-0.6%	-0.5%
Inventory	-10.3%	-9.2%	-8.8%
Prepaid expenses	0.2%	-0.2%	0.1%
Accrued liabilities	5.5%	5.5%	5.6%
Operating cash flow	12.2%	12.8%	13.8%
Cash from sale of fixed assets	0.7%	0.7%	0.7%
Purchase of plant and equipment	-12.3%	-12.0%	-11.7%
Investing cash flow	-11.6%	-11.3%	-11.0%
Sale of bonds	2.6%	2.5%	2.6%
Cash dividends	-2.1%	-2.1%	-2.1%
Financing cash flow	0.5%	0.4%	0.5%
Total cash flow	1.1%	1.9%	3.3%

Answer:

Operating cash flow has decreased as a percentage of revenues. This appears to be due largely to accumulating inventories. Investing activities, specifically purchases of plant and equipment, have also required an increasing percentage of the firm's cash flow.

MODULE 20.4: FREE CASH FLOW AND RATIOS



LOS 20.i: Calculate and interpret free cash flow to the firm, free cash flow to equity, and performance and coverage cash flow ratios.

Video covering this content is available online.

Free cash flow is a measure of cash that is available for discretionary purposes. This is the cash flow that is available once the firm has covered its capital expenditures. This is a fundamental cash flow measure and is often used for valuation. There are several measures of free cash flow. Two of the more common measures are free cash flow to the firm and free cash flow to equity.

Free Cash Flow to the Firm

Free cash flow to the firm (FCFF) is the cash available to all investors, both equity owners and debt holders. FCFF can be calculated by starting with either net income or operating cash flow.

FCFF is calculated from net income as:

$$\text{FCFF} = \text{NI} + \text{NCC} + [\text{Int} \times (1 - \text{tax rate})] - \text{FCInv} - \text{WCInv}$$

where:

NI = net income

NCC = noncash charges (depreciation and amortization)

Int = cash interest paid

FCInv = fixed capital investment (net capital expenditures)

WCInv = working capital investment



PROFESSOR'S NOTE

Fixed capital investment is cash spent on fixed assets minus cash received from selling fixed assets. It is not the same as CFI, which includes cash flows from fixed investments, investments in securities, and repaid principal from loans made.

Note that cash interest paid, net of tax, is added back to net income. This is because FCFF is the cash flow available to stockholders and debt holders. Since interest is paid to (and therefore "available to") the debt holders, it must be included in FCFF.

FCFF can also be calculated from operating cash flow as:

$$\text{FCFF} = \text{CFO} + [\text{Int} \times (1 - \text{tax rate})] - \text{FCInv}$$

where:

CFO = cash flow from operations

Int = cash interest paid

FCInv = fixed capital investment (net capital expenditures)

It is not necessary to adjust for noncash charges and changes in working capital when starting with CFO, since they are already reflected in the calculation of CFO. For firms that follow IFRS, it is not necessary to adjust for interest that is included as a part of financing activities.

Additionally, firms that follow IFRS can report dividends paid as operating activities. In this case, the dividends paid would be added back to CFO. Again, the goal is to calculate the cash flow that is available to the shareholders and debt holders. It is not necessary to adjust dividends for taxes since dividends paid are not tax deductible.

Free Cash Flow to Equity

Free cash flow to equity (FCFE) is the cash flow that would be available for distribution to common shareholders. FCFE can be calculated as follows:

$$\text{FCFE} = \text{CFO} - \text{FCInv} + \text{net borrowing}$$

where:

CFO = cash flow from operations

FCInv = fixed capital investment (net capital expenditures)

net borrowing = debt issued - debt repaid



PROFESSOR'S NOTE

If net borrowing is negative (debt repaid exceeds debt issued), we would subtract net borrowing in calculating FCFE.

If firms that follow IFRS have subtracted dividends paid in calculating CFO, dividends must be added back when calculating FCFE.

EXAMPLE: Free cash flow

Using the financial statements from the examples presented earlier, calculate the company's free cash flow to the firm and free cash flow to equity. Assume a tax rate of 40%.

Answer:

Free cash flow to the firm = CFO + [interest paid \times (1 - tax rate)] - fixed capital investment = \$42,000 + \$500(1 - 0.4) - \$10,000 = \$32,300

Free cash flow to equity = CFO - fixed capital investment + net borrowing = \$42,000 - \$10,000 + \$5,000 = \$37,000

Alternatively, free cash flow to equity = FCFF - [interest paid \times (1 - tax rate)] + net borrowing = \$32,300 - \$500(1 - 0.4) + \$5,000 = \$37,000

Other Cash Flow Ratios

Just as with the income statement and balance sheet, the cash flow statement can be analyzed by comparing the cash flows either over time or to those of other firms. Cash flow ratios can be categorized as performance ratios and coverage ratios.

Performance Ratios

The **cash flow-to-revenue ratio** measures the amount of operating cash flow generated for each dollar of revenue.

$$\text{cash flow-to-revenue} = \frac{\text{CFO}}{\text{net revenue}}$$

The **cash return-on-assets ratio** measures the return of operating cash flow attributed to all providers of capital.

$$\text{cash return-on-assets} = \frac{\text{CFO}}{\text{average total assets}}$$

The **cash return-on-equity ratio** measures the return of operating cash flow attributed to shareholders.

$$\text{cash return-on-equity} = \frac{\text{CFO}}{\text{average total equity}}$$

The **cash-to-income ratio** measures the ability to generate cash from firm operations.

$$\text{cash-to-income} = \frac{\text{CFO}}{\text{operating income}}$$

Cash flow per share is a variation of basic earnings per share measured by using CFO instead of net income.

$$\text{cash flow per share} = \frac{\text{CFO} - \text{preferred dividends}}{\text{weighted average number of common shares}}$$

Note: If common dividends were classified as operating activities under IFRS, they should be added back to CFO for purposes of calculating cash flow per share.

Coverage Ratios

The **debt coverage ratio** measures financial risk and leverage.

$$\text{debt coverage} = \frac{\text{CFO}}{\text{total debt}}$$

The **interest coverage ratio** measures the firm's ability to meet its interest obligations.

$$\text{interest coverage} = \frac{\text{CFO} + \text{interest paid} + \text{taxes paid}}{\text{interest paid}}$$

Note: If interest paid was classified as a financing activity under IFRS, no interest adjustment is necessary.

The **reinvestment ratio** measures the firm's ability to acquire long-term assets with operating cash flow.

$$\text{reinvestment} = \frac{\text{CFO}}{\text{cash paid for long-term assets}}$$

The **debt payment ratio** measures the firm's ability to satisfy long-term debt with operating cash flow.

$$\text{debt payment} = \frac{\text{CFO}}{\text{cash long-term debt repayment}}$$

The **dividend payment ratio** measures the firm's ability to make dividend payments from operating cash flow.

$$\text{dividend payment} = \frac{\text{CFO}}{\text{dividends paid}}$$

The **investing and financing ratio** measures the firm's ability to purchase assets, satisfy debts, and pay dividends.

$$\text{investing and financing} = \frac{\text{CFO}}{\text{cash outflows from investing and financing activities}}$$

MODULE QUIZ 20.3, 20.4

1. Continental Corporation reported sales revenue of \$150,000 for the current year. If accounts receivable decreased \$10,000 during the year and accounts payable increased \$4,000 during the year, cash collections were:
 - A. \$154,000.
 - B. \$160,000.
 - C. \$164,000.
2. In preparing a common-size cash flow statement, each cash flow is expressed as a percentage of:
 - A. total assets.
 - B. total revenues.
 - C. the change in cash.
3. To calculate free cash flow to the firm based on operating cash flow, an analyst should add interest expense net of tax and subtract:
 - A. noncash charges.
 - B. fixed capital investment.
 - C. working capital investment.
4. The reinvestment ratio measures a firm's ability to use its operating cash flow to:
 - A. pay dividends.
 - B. invest in working capital.
 - C. acquire long-lived assets.

KEY CONCEPTS

LOS 20.a

Cash flow from operating activities (CFO) consists of the inflows and outflows of cash resulting from transactions that affect a firm's net income.

Cash flow from investing activities (CFI) consists of the inflows and outflows of cash resulting from the acquisition or disposal of long-term assets and certain investments.

Cash flow from financing activities (CFF) consists of the inflows and outflows of cash resulting from transactions affecting a firm's capital structure, such as issuing or repaying debt and issuing or repurchasing stock.

LOS 20.b

Noncash investing and financing activities, such as taking on debt to the seller of a purchased asset, are not reported in the cash flow statement but must be disclosed in the footnotes or a supplemental schedule.

LOS 20.c

Under U.S. GAAP, dividends paid are financing cash flows. Interest paid, interest received, and dividends received are operating cash flows. All taxes paid are operating cash flows.

Under IFRS, dividends paid and interest paid can be reported as either operating or financing cash flows. Interest received and dividends received can be reported as either operating or investing cash flows. Taxes paid are operating cash flows unless they arise from an investing or financing transaction.

LOS 20.d

Under the direct method of presenting CFO, each line item of the accrual-based income statement is adjusted to get cash receipts or cash payments. The main advantage of the direct method is that it presents clearly the firm's operating cash receipts and payments.

Under the indirect method of presenting CFO, net income is adjusted for transactions that affect net income but do not affect operating cash flow, such as depreciation and gains or losses on asset sales, and for changes in balance sheet items. The main advantage of the indirect method is that it focuses on the differences between net income and operating cash flow. This provides a useful link to the income statement when forecasting future operating cash flow.

LOS 20.e

Operating activities typically relate to the firm's current assets and current liabilities. Investing activities typically relate to noncurrent assets. Financing activities typically relate to noncurrent liabilities and equity.

Timing of revenue or expense recognition that differs from the receipt or payment of cash is reflected in changes in balance sheet accounts.

LOS 20.f

The direct method of calculating CFO is to sum cash inflows and cash outflows for operating activities.

- Cash collections from customers—sales adjusted for changes in receivables and unearned revenue.
- Cash paid for inputs—COGS adjusted for changes in inventory and accounts payable.
- Cash operating expenses—SG&A adjusted for changes in related accrued liabilities or prepaid expenses.
- Cash interest paid—interest expense adjusted for the change in interest payable.
- Cash taxes paid—income tax expense adjusted for changes in taxes payable and changes in deferred tax assets and liabilities.

The indirect method of calculating CFO begins with net income and adjusts it for gains or losses related to investing or financing cash flows, noncash charges to income, and changes in balance sheet operating items.

CFI is calculated by determining the changes in asset accounts that result from investing activities. The cash flow from selling an asset is its book value plus any gain on the sale (or minus any loss on the sale).

CFF is the sum of net cash flows from creditors (new borrowings minus principal repaid) and net cash flows from shareholders (new equity issued minus share repurchases minus cash dividends paid).

LOS 20.g

An indirect cash flow statement can be converted to a direct cash flow statement by adjusting each income statement account for changes in associated balance sheet accounts and by eliminating noncash and non-operating items.

LOS 20.h

An analyst should determine whether a company is generating positive operating cash flow over time that is greater than its capital spending needs and whether the company's accounting policies are causing reported earnings to diverge from operating cash flow.

A common-size cash flow statement shows each item as a percentage of revenue or shows each cash inflow as a percentage of total inflows and each outflow as a percentage of total outflows.

LOS 20.i

Free cash flow to the firm (FCFF) is the cash available to all investors, both equity owners and debt holders.

- $\text{FCFF} = \text{net income} + \text{noncash charges} + [\text{cash interest paid} \times (1 - \text{tax rate})] - \text{fixed capital investment} - \text{working capital investment.}$
- $\text{FCFF} = \text{CFO} + [\text{cash interest paid} \times (1 - \text{tax rate})] - \text{fixed capital investment.}$

Free cash flow to equity (FCFE) is the cash flow that is available for distribution to the common shareholders after all obligations have been paid.

$$\text{FCFE} = \text{CFO} - \text{fixed capital investment} + \text{net borrowing}$$

Cash flow performance ratios, such as cash return on equity or on assets, and cash coverage ratios, such as debt coverage or cash interest coverage, provide information about the firm's operating performance and financial strength.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 20.1

1. **C** The payment of interest on debt is an *operating* cash flow under U.S. GAAP. (LOS 20.a)
2. **C** Depreciation does not represent a cash flow. To the extent that it affects the firm's taxes, an increase in depreciation changes operating cash flows, but not investing cash flows. (LOS 20.a)
3. **A** A change in notes payable is a financing cash flow. (LOS 20.a)
4. **A** Sales of inventory would be classified as operating cash flow. (LOS 20.a)
5. **B** Issuing bonds would be classified as financing cash flow. (LOS 20.a)
6. **B** Sale of land would be classified as investing cash flow. (LOS 20.a)
7. **C** Write-off of obsolete equipment has no cash flow impact. (LOS 20.a)
8. **A** Under IFRS, interest expense can be classified as either an operating cash flow or financing cash flow. (LOS 20.a)

9. **A** Dividends received from investments would be classified as operating cash flow under U.S. GAAP. (LOS 20.a)
10. **C** The exchange of debt securities for equity securities is a noncash transaction. (LOS 20.b)
11. **B** Under U.S. GAAP, dividends paid are reported as financing activities. Under IFRS, dividends paid can be reported as either operating or financing activities. (LOS 20.c)
12. **C** The indirect method reconciles the difference between net income and CFO. The direct method shows operating cash received and paid and, therefore, provides more information on its face than the indirect method. (LOS 20.d)
13. **C** Financing cash flows are linked primarily to changes in long-term liabilities and equity. Changes in current assets and liabilities tend to be linked to operating cash flows. Changes in long-lived assets are typically linked to investing cash flows. (LOS 20.e)

Module Quiz 20.2

1. **B** Net income – profits from sale of land + depreciation + decrease in receivables – increase in inventories + increase in accounts payable – decrease in wages payable + increase in deferred tax liabilities = $120 - 2 + 25 + 20 - 10 + 7 - 5 + 15 = \170 . Note that the profit on the sale of land should be subtracted from net income because this transaction is classified as investing, not operating. (LOS 20.a, 20.f)
2. **B** Net income – profit on sale of building + depreciation = $45 - 20 + 75 = \$100$. Note that taxes and interest are already deducted in calculating net income, and that the profit on the sale of the building should be subtracted from net income. (LOS 20.a, 20.f)
3. **B** Cash from sale of building – purchase of machinery = $40 - 20 = \$20$. (LOS 20.a, 20.f)
4. **A** Sale of preferred stock + issuance of bonds – principal payments on bank borrowings – repurchase of common stock – dividends paid = $35 + 50 - 15 - 30 - 10 = \30 . Note that we did not include \$45 of debt retired through issuance of common stock since this was a noncash transaction. Knowing how to handle noncash transactions is important. (LOS 20.a, 20.f)
5. **B** Net income = $\$1,500 \times 0.25 = \375 , and cash flow from operations = net income – gain on sale of machinery + depreciation – increase in accounts receivable – increase in inventory – decrease in accounts payable = $375 - 30 + 150 - 50 - 100 - 70 = \275 . (LOS 20.a, 20.f)
6. **A**

Net income	\$78,000
Depreciation	12,000
Unrealized gain	(15,000)
Increase in accounts receivable	(52,000)
Increase in accounts payable	<u>29,000</u>
Cash flow from operations	\$52,000

(LOS 20.f)

(LOS 20.f)

7. C Purchased new fixed assets for \$75,000 – cash outflow from investing

Converted \$70,000 of preferred shares to common shares – noncash transaction

Received dividends of \$12,000 – cash inflow from operations

Paid dividends of \$21,000 – cash outflow from financing

Mortgage repayment of \$17,000 – cash outflow from financing

CFI = -75,000

CFF = -21,000 - 17,000 = -\$38,000

(LOS 20.a, 20.f)

Module Quiz 20.3, 20.4

1. B \$150,000 sales + \$10,000 decrease in accounts receivable = \$160,000 cash collections.
The change in accounts payable does not affect cash collections. Accounts payable result from a firm's purchases from its suppliers. (Module 20.3, LOS 20.f, 20.g)
2. B The cash flow statement can be converted to common-size format by expressing each line item as a percentage of revenue. (Module 20.3, LOS 20.h)
3. B FCFF can be calculated from CFO by adding interest expense net of tax and subtracting fixed capital investment. (Module 20.4, LOS 20.i)
4. C The reinvestment ratio is CFO / cash paid for long-term assets. (Module 20.4, LOS 20.i)

Reading 21

FINANCIAL ANALYSIS TECHNIQUES

EXAM FOCUS

This reading presents a “tool box” for an analyst. It would be nice if you could calculate all these ratios, but it is imperative that you understand what firm characteristic each one is measuring, and even more important, that you know whether a higher or lower ratio is better in each instance. Different analysts calculate some ratios differently. It would be helpful if analysts were always careful to distinguish between total liabilities, total interest-bearing debt, long-term debt, and creditor and trade debt, but they do not. Some analysts routinely add deferred tax liabilities to debt or exclude goodwill when calculating assets and equity; others do not. Statistical reporting services almost always disclose how each of the ratios they present was calculated. So do not get too tied up in the details of each ratio, but understand what each one represents and what factors would likely lead to significant changes in a particular ratio. The DuPont formulas have been with us a long time and were in the curriculum when I took the exams back in the 1980s. Decomposing ROE into its components is an important analytic technique and it should definitely be in your tool box.

MODULE 21.1: INTRODUCTION TO FINANCIAL RATIOS



Video covering this content is available online.

LOS 21.a: Describe tools and techniques used in financial analysis, including their uses and limitations.

Various tools and techniques are used to convert financial statement data into formats that facilitate analysis. These include ratio analysis, common-size analysis, graphical analysis, and regression analysis.

Ratio Analysis

Ratios are useful tools for expressing relationships among data that can be used for internal comparisons and comparisons across firms. They are often most useful in identifying questions that need to be answered, rather than answering questions directly. Specifically, ratios can be used to do the following:

- Project future earnings and cash flow.
- Evaluate a firm’s flexibility (the ability to grow and meet obligations even when unexpected circumstances arise).
- Assess management’s performance.
- Evaluate changes in the firm and industry over time.
- Compare the firm with industry competitors.

Analysts must also be aware of the limitations of ratios, including the following:

- Financial ratios are not useful when viewed in isolation. They are only informative when compared to those of other firms or to the company's historical performance.
- Comparisons with other companies are made more difficult by different accounting treatments. This is particularly important when comparing U.S. firms to non-U.S. firms.
- It is difficult to find comparable industry ratios when analyzing companies that operate in multiple industries.
- Conclusions cannot be made by calculating a single ratio. All ratios must be viewed relative to one another.
- Determining the target or comparison value for a ratio is difficult, requiring some range of acceptable values.

It is important to understand that the definitions of ratios can vary widely among the analytical community. For example, some analysts use all liabilities when measuring leverage, while other analysts only use interest-bearing obligations. Consistency is paramount. Analysts must also understand that reasonable values of ratios can differ among industries.

Common-Size Analysis

Common-size statements normalize balance sheets and income statements and allow the analyst to more easily compare performance across firms and for a single firm over time.

- A vertical common-size balance sheet expresses all balance sheet accounts as a percentage of total assets.
- A vertical common-size income statement expresses all income statement items as a percentage of sales.

In addition to comparisons of financial data across firms and time, common-size analysis is appropriate for quickly viewing certain financial ratios. For example, the gross profit margin, operating profit margin, and net profit margin are all clearly indicated within a common-size income statement. Vertical common-size income statement ratios are especially useful for studying trends in costs and profit margins.

$$\text{vertical common-size income statement ratios} = \frac{\text{income statement account}}{\text{sales}}$$

Balance sheet accounts can also be converted to common-size ratios by dividing each balance sheet item by total assets.

$$\text{vertical common-size balance-sheet ratios} = \frac{\text{balance sheet account}}{\text{total assets}}$$

EXAMPLE: Constructing common-size statements

The common-size statements in Figure 21.1 show balance sheet items as percentages of assets, and income statement items as percentages of sales.

- You can convert all asset and liability amounts to their actual values by multiplying the percentages listed below by their total assets of \$57,100; \$55,798; and \$52,071, respectively for 20X6, 20X5, and 20X4 (data is USD millions).
- Also, all income statement items can be converted to their actual values by multiplying the given percentages by total sales, which were \$29,723; \$29,234; and \$22,922, respectively, for 20X6, 20X5, and

20X4.

Figure 21.1: Vertical Common-Size Balance Sheet and Income Statement

Balance Sheet, fiscal year-end	20X6	20X5	20X4
Assets			
Cash & cash equivalents	0.38%	0.29%	0.37%
Accounts receivable	5.46%	5.61%	6.20%
Inventories	5.92%	5.42%	5.84%
Deferred income taxes	0.89%	0.84%	0.97%
Other current assets	0.41%	0.40%	0.36%
Total current assets	13.06%	12.56%	13.74%
Gross fixed assets	25.31%	23.79%	25.05%
Accumulated depreciation	8.57%	7.46%	6.98%
Net gross fixed assets	16.74%	16.32%	18.06%
Other long-term assets	70.20%	71.12%	68.20%
Total assets	100.00%	100.00%	100.00%
Liabilities			
Accounts payable	3.40%	3.40%	3.79%
Short-term debt	1.00%	2.19%	1.65%
Other current liabilities	8.16%	10.32%	9.14%
Total current liabilities	12.56%	15.91%	14.58%
Long-term debt	18.24%	14.58%	5.18%
Other long-term liabilities	23.96%	27.44%	53.27%
Total liabilities	54.76%	57.92%	73.02%
Preferred equity	0.00%	0.00%	0.00%
Common equity	45.24%	42.08%	26.98%
Total liabilities & equity	100.00%	100.00%	100.00%
Income Statement, fiscal year	20X6	20X5	20X4
Revenues	100.00%	100.00%	100.00%
Cost of goods sold	59.62%	60.09%	60.90%
Gross profit	40.38%	39.91%	39.10%
Selling, general & administrative	16.82%	17.34%	17.84%
Depreciation	2.39%	2.33%	2.18%
Amortization	0.02%	3.29%	2.33%
Other operating expenses	0.58%	0.25%	-0.75%
Operating income	20.57%	16.71%	17.50%
Interest and other debt expense	2.85%	4.92%	2.60%
Income before taxes	17.72%	11.79%	14.90%
Provision for income taxes	6.30%	5.35%	6.17%
Net income	11.42%	6.44%	8.73%

Even a cursory inspection of the income statement in Figure 21.1 can be quite instructive. Beginning at the bottom, we can see that the profitability of the company has increased nicely in 20X6 after falling slightly in 20X5. We can examine the 20X6 income statement values to find the source of this greatly improved profitability. Cost of goods sold seems to be stable, with an improvement (decrease) in 20X6 of only 0.48%. SG&A was down approximately one-half percent as well.

These improvements from (relative) cost reduction, however, only begin to explain the 5% increase in the net profit margin for 20X6. Improvements in two items, “amortization” and “interest and other debt expense,” appear to be the most significant factors in the firm’s improved profitability in 20X6. Clearly the analyst must investigate further in both areas to learn whether these improvements represent permanent improvements or whether these items can be expected to return to previous percentage-of-sales levels in the future.

We can also note that interest expense as a percentage of sales was approximately the same in 20X4 and 20X6. We must investigate the reasons for the higher interest costs in 20X5 to determine whether the current level of 2.85% can be expected to continue into the next period. In addition, more than 3% of the 5% increase in net profit margin in 20X6 is due to a decrease in amortization expense. Since this is a noncash expense, the decrease may have no implications for cash flows looking forward.

This discussion should make clear that common-size analysis doesn’t tell an analyst the whole story about this company, but can certainly point the analyst in the right direction to find out the circumstances that led to the increase in the net profit margin and to determine the effects, if any, on firm cash flow going forward.

Another way to present financial statement data that is quite useful when analyzing trends over time is a **horizontal common-size balance sheet or income statement**. The divisor here is the first-year values, so they are all standardized to 1.0 by construction. Figure 21.2 illustrates this approach.

Figure 21.2: Horizontal Common-Size Balance Sheet Data

	20X4	20X5	20X6
Inventory	1.0	1.1	1.4
Cash and marketable securities	1.0	1.3	1.2
Long-term debt	1.0	1.6	1.8
PP&E (net of depreciation)	1.0	0.9	0.8

Trends in the values of these items, as well as the relative growth in these items, are readily apparent from a horizontal common-size balance sheet.



PROFESSOR’S NOTE

We have presented data in Figure 21.1 with information for the most recent period on the left, and in Figure 21.2 we have presented the historical values from left to right. Both presentation methods are common, and on the exam, you should pay special attention to which method is used in the data presented for any question.

We can view the values in the common-size financial statements as ratios. Net income is shown on the common-size income statement as net income/revenues, which is the net profit margin, and tells the analyst the percentage of each dollar of sales that remains for shareholders after all expenses related to the generation of those sales are deducted. One measure of financial leverage, long-term debt to total assets, can be read directly from the vertical common-size financial statements. Specific ratios commonly used in financial analysis and interpretation of their values are covered in detail in this review.

Graphical Analysis

Graphs can be used to visually present performance comparisons and composition of financial statement elements over time.

A **stacked column graph** (also called a *stacked bar graph*) shows the changes in items from year to year in graphical form. Figure 21.3 presents such data for a hypothetical corporation.

Another alternative for graphic presentation of data is a **line graph**. Figure 21.4 presents the same data as Figure 21.3, but as a line graph. The increase in trade payables and the decrease in cash are evident in either format and would alert the analyst to potential liquidity problems that require further investigation and analysis.

Regression Analysis

Regression analysis can be used to identify relationships between variables. The results are often used for forecasting. For example, an analyst might use the relationship between GDP and sales to prepare a sales forecast.

Figure 21.3: Stacked Column (Stacked Bar) Graph

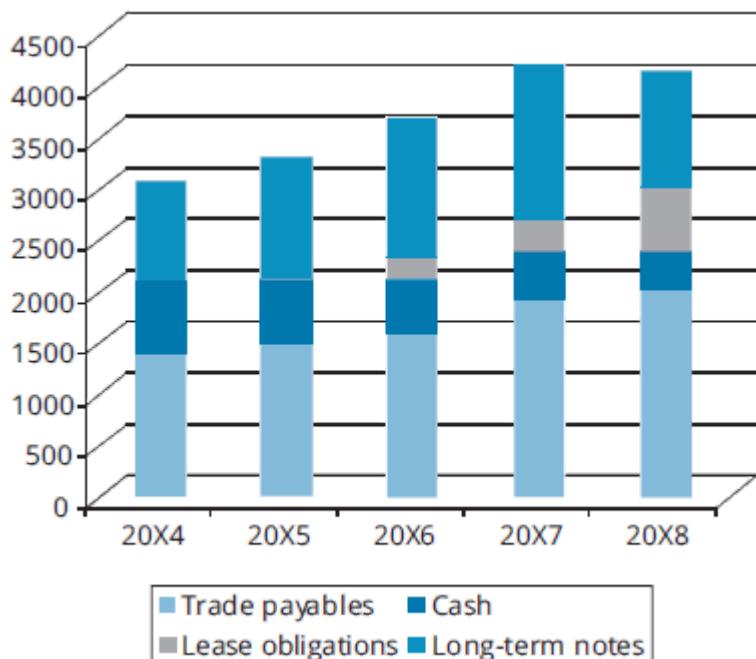
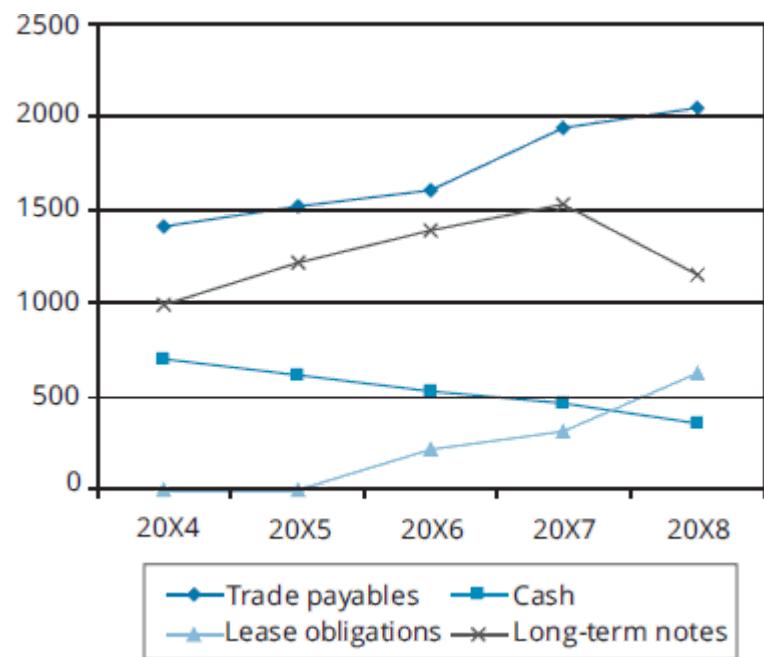


Figure 21.4: Line Graph



Video covering
this content is
available online.

MODULE 21.2: FINANCIAL RATIOS, PART 1

LOS 21.b: Identify, calculate, and interpret activity, liquidity, solvency, profitability, and valuation ratios.

Financial ratios can be segregated into different classifications by the type of information about the company they provide. One such classification scheme is:

- **Activity ratios.** This category includes several ratios also referred to asset utilization or turnover ratios (e.g., inventory turnover, receivables turnover, and total assets turnover). They often give indications of how well a firm utilizes various assets such as inventory and fixed assets.
- **Liquidity ratios.** Liquidity here refers to the ability to pay short-term obligations as they come due.
- **Solvency ratios.** Solvency ratios give the analyst information on the firm's financial leverage and ability to meet its longer-term obligations.
- **Profitability ratios.** Profitability ratios provide information on how well the company generates operating profits and net profits from its sales.
- **Valuation ratios.** Sales per share, earnings per share, and price to cash flow per share are examples of ratios used in comparing the relative valuation of companies.



PROFESSOR'S NOTE

We examine valuation ratios in another LOS concerning equity analysis later in this review, and in the Equity Investments topic area.

It should be noted that these categories are not mutually exclusive. An activity ratio such as payables turnover may also provide information about the liquidity of a company, for example. There is no one standard set of ratios for financial analysis. Different analysts use different ratios and different calculation methods for similar ratios. Some ratios are so commonly used that there is very little variation in how they are defined and calculated. We will note some alternative treatments and alternative terms for single ratios as we detail the commonly used ratios in each category.

Activity Ratios

Activity ratios (also known as asset utilization ratios or operating efficiency ratios) measure how efficiently the firm is managing its assets.

- A measure of accounts receivable turnover is *receivables turnover*:

$$\text{receivables turnover} = \frac{\text{annual sales}}{\text{average receivables}}$$



PROFESSOR'S NOTE

In most cases when a ratio compares a balance sheet account (such as receivables) with an income or cash flow item (such as sales), the balance sheet item will be the average of the account instead of simply the end-of-year balance. Averages are calculated by adding the beginning-of-year account value to the end-of-year account value, then dividing the sum by two.

A high receivables turnover ratio could be the result of doing an excellent job of managing credit terms and collections. On the other hand, it might indicate that a company has stringent credit terms, offers a large discount for early payment, or charges high penalties for late payment. A company that has excessively stringent credit terms will lose sales as a result. Insight into why a company has a high receivables turnover rate can be gained from looking at the company's revenue growth compared to peers. Slower growth could indicate that credit terms may be too stringent, while a high receivables turnover together with revenue growth at or above the peer group average could indicate superior credit terms and collections management.

- The inverse of the receivables turnover times 365 is the *average collection period*, or *days of sales outstanding*, which is the average number of days it takes for the company's customers to pay their bills:

$$\text{days of sales outstanding} = \frac{365}{\text{receivables turnover}}$$

- A measure of a firm's efficiency with respect to its processing and inventory management is *inventory turnover*:

$$\text{inventory turnover} = \frac{\text{cost of goods sold}}{\text{average inventory}}$$



PROFESSOR'S NOTE

Pay careful attention to the numerator in the turnover ratios. For inventory turnover, be sure to use cost of goods sold, not sales.

Inventory turnover that is high may indicate effective management of inventory, but could also result from holding inventory levels too low so that sales are lost when orders cannot be filled immediately. A low inventory turnover ratio relative to peers could indicate that some inventory is obsolete and slow-selling. In either case, examining revenue growth relative to peers can provide more insight into whether inventory is well or poorly managed.

- The inverse of the inventory turnover times 365 is the *average inventory processing period*, *number of days of inventory*, or *days of inventory on hand*:

$$\text{days of inventory on hand} = \frac{365}{\text{inventory turnover}}$$

- A measure of the use of trade credit by the firm is the *payables turnover* ratio:

$$\text{payables turnover} = \frac{\text{purchases}}{\text{average trade payables}}$$



PROFESSOR'S NOTE

You can use the inventory equation to calculate purchases from the financial statements. Purchases = ending inventory – beginning inventory + cost of goods sold.

- The inverse of the payables turnover ratio multiplied by 365 is the *payables payment period* or *number of days of payables*, which is the average amount of time it takes the company to pay its bills:

$$\text{number of days of payables} = \frac{365}{\text{payables turnover ratio}}$$



PROFESSOR'S NOTE

We have shown days calculations for payables, receivables, and inventory based on annual turnover and a 365-day year. If turnover ratios are for a quarter rather than a year, the number of days in the quarter should be divided by the quarterly turnover ratios in order to get the “days” form of these ratios.

A high payables turnover ratio relative to peers may indicate that a company is not fully taking advantage of supplier credit terms, or that the company is paying suppliers early to take advantage of discounts. A payables turnover rate that is low relative to that of peer companies may indicate that a company is having problems with short-term cash flows or, alternatively, that a company is simply taking advantage of lenient terms negotiated with suppliers. As with inventory turnover, examining other ratios (in this case liquidity ratios) can provide insight into which interpretation of a relatively high or low payables turnover ratio is more likely.

- The effectiveness of the firm's use of its total assets to create revenue is measured by its *total asset turnover*:

$$\text{total asset turnover} = \frac{\text{revenue}}{\text{average total assets}}$$

Different types of industries might have considerably different turnover ratios. Manufacturing businesses that are capital-intensive might have asset turnover ratios near one, while retail businesses might have turnover ratios near 10. As was the case with the current asset turnover ratios discussed previously, it is desirable for the total asset turnover ratio to be close to the industry norm. Low asset turnover ratios might mean that the company has too much capital tied up in its asset base. A turnover ratio that is too high might imply that the firm has too few assets for potential sales, or that the asset base is outdated.

- The utilization of fixed assets is measured by the *fixed asset turnover* ratio:

$$\text{fixed asset turnover} = \frac{\text{revenue}}{\text{average net fixed assets}}$$

Low fixed asset turnover might mean that the company has too much capital tied up in its asset base or is using the assets it has inefficiently. A turnover ratio that is too high might imply that the firm has obsolete equipment, or at a minimum, that the firm will probably have to incur capital expenditures in the near future to increase capacity to support growing revenues. Since “net” here refers to net of accumulated depreciation, firms with more recently acquired assets will typically have lower fixed asset turnover ratios.

- How effectively a company is using its working capital is measured by the *working capital turnover* ratio:

$$\text{working capital turnover} = \frac{\text{revenue}}{\text{average working capital}}$$

Working capital (sometimes called *net* working capital) is current assets minus current liabilities. The working capital turnover ratio gives us information about the utilization of working capital in terms of dollars of sales per dollar of working capital. Some firms may have very low working capital if outstanding payables equal or exceed inventory and receivables. In this case the working capital turnover ratio will be very large, may vary significantly from period to period, and is less informative about changes in the firm's operating efficiency.

Liquidity Ratios

Liquidity ratios are employed by analysts to determine the firm's ability to pay its short-term liabilities.

- The *current ratio* is the best-known measure of liquidity:

$$\text{current ratio} = \frac{\text{current assets}}{\text{current liabilities}}$$

The higher the current ratio, the more likely it is that the company will be able to pay its short-term bills. A current ratio of less than one means that the company has negative working capital and is probably facing a liquidity crisis. Working capital equals current assets minus current liabilities.

- The *quick ratio* is a more stringent measure of liquidity because it does not include inventories and other assets that might not be very liquid:

$$\text{quick ratio} = \frac{\text{cash} + \text{marketable securities} + \text{receivables}}{\text{current liabilities}}$$

The higher the quick ratio, the more likely it is that the company will be able to pay its short-term bills. Marketable securities are short-term debt instruments, typically liquid and of good credit quality.

- The most conservative liquidity measure is the *cash ratio*:

$$\text{cash ratio} = \frac{\text{cash} + \text{marketable securities}}{\text{current liabilities}}$$

The higher the cash ratio, the more likely it is that the company will be able to pay its short-term bills.

The current, quick, and cash ratios differ only in the assumed liquidity of the current assets that the analyst projects will be used to pay off current liabilities.

- The *defensive interval ratio* is another measure of liquidity that indicates the number of days of average cash expenditures the firm could pay with its current liquid assets:

$$\text{defensive interval} = \frac{\text{cash} + \text{marketable securities} + \text{receivables}}{\text{average daily expenditures}}$$

Expenditures here include cash expenses for costs of goods, SG&A, and research and development. If these items are taken from the income statement, noncash charges such as depreciation should be added back just as in the preparation of a statement of cash flows by the indirect method.

- The *cash conversion cycle* is the length of time it takes to turn the firm's cash investment in inventory back into cash, in the form of collections from the sales of that inventory. The cash conversion cycle is computed from days sales outstanding, days of inventory on hand, and number of days of payables:

$$\text{cash conversion cycle} = \text{days sales outstanding} + \text{days of inventory on hand} - \text{number of days of payables}$$

High cash conversion cycles are considered undesirable. A conversion cycle that is too high implies that the company has an excessive amount of capital investment in the sales process.



MODULE 21.3: FINANCIAL RATIOS, PART 2

Solvency Ratios

Solvency ratios measure a firm's financial leverage and ability to meet its long-term obligations. Solvency ratios include various **debt ratios** that are based on the balance sheet and **coverage ratios** that are based on the income statement.

- A measure of the firm's use of fixed-cost financing sources is the *debt-to-equity ratio*:

$$\text{debt-to-equity} = \frac{\text{total debt}}{\text{total shareholders' equity}}$$

Increases and decreases in this ratio suggest a greater or lesser reliance on debt as a source of financing.

Total debt is calculated differently by different analysts and different providers of financial information. Here, we will define it as long-term debt plus interest-bearing short-term debt.

Some analysts include the present value of lease obligations and/or non-interest-bearing current liabilities, such as trade payables.

- Another way of looking at the usage of debt is the *debt-to-capital ratio*:

$$\text{debt-to-capital} = \frac{\text{total debt}}{\text{total debt} + \text{total shareholders' equity}}$$

Capital equals all short-term and long-term debt plus preferred stock and equity. Increases and decreases in this ratio suggest a greater or lesser reliance on debt as a source of financing.

- A slightly different way of analyzing debt utilization is the *debt-to-assets ratio*:

$$\text{debt-to-assets} = \frac{\text{total debt}}{\text{total assets}}$$

Increases and decreases in this ratio suggest a greater or lesser reliance on debt as a source of financing.

- Another measure that is used as an indicator of a company's use of debt financing is the *financial leverage ratio* (or leverage ratio):

$$\text{financial leverage} = \frac{\text{average total assets}}{\text{average total equity}}$$

Average here means the average of the values at the beginning and at the end of the period. Greater use of debt financing increases financial leverage and, typically, risk to equity holders and bondholders alike.

- The remaining risk ratios help determine the firm's ability to repay its debt obligations. The first of these is the *interest coverage ratio*:

$$\text{interest coverage} = \frac{\text{earnings before interest and taxes}}{\text{interest payments}}$$

The lower this ratio, the more likely it is that the firm will have difficulty meeting its debt payments.

Because depreciation and amortization are not cash expenses, another ratio that reflects a firm's ability to meet its debt obligations is the *debt-to-EBITDA* ratio:

$$\text{debt-to-EBITDA} = \frac{\text{total debt}}{\text{EBITDA}}$$

- Another indicator of a company's ability to meet its obligations is the *fixed charge coverage* ratio:

$$\text{fixed charge coverage} = \frac{\text{earnings before interest and taxes} + \text{lease payments}}{\text{interest payments} + \text{lease payments}}$$

Here, lease payments are added back to operating earnings in the numerator and also added to interest payments in the denominator. Significant lease obligations will reduce this ratio significantly compared to the interest coverage ratio. Fixed charge coverage is the more meaningful measure for companies that lease a large portion of their assets, such as some airlines.



PROFESSOR'S NOTE

With all solvency ratios, the analyst must consider the variability of a firm's cash flows when determining the reasonableness of the ratios. Firms with stable cash flows are usually able to carry more debt.

Profitability Ratios

Profitability ratios measure the overall performance of the firm relative to revenues, assets, equity, and capital.

- The *net profit margin* is the ratio of net income to revenue:

$$\text{net profit margin} = \frac{\text{net income}}{\text{revenue}}$$

Analysts should be concerned if this ratio is too low. The net profit margin should be based on net income from continuing operations, because analysts should be primarily concerned about future expectations, and below-the-line items such as discontinued operations will not affect the company in the future.

Operating profitability ratios look at how good management is at turning their efforts into profits. Operating ratios compare the top of the income statement (sales) to profits. The different ratios are designed to isolate specific costs.

- The *gross profit margin* is the ratio of gross profit (sales less cost of goods sold) to sales:

$$\text{gross profit margin} = \frac{\text{gross profit}}{\text{revenue}}$$

An analyst should be concerned if this ratio is too low. Gross profit can be increased by raising prices or reducing costs. However, the ability to raise prices may be limited by competition.

- The *operating profit margin* is the ratio of operating profit (gross profit less selling, general, and administrative expenses) to sales. Operating profit is also referred to as earnings before interest and taxes (EBIT):

$$\text{operating profit margin} = \frac{\text{operating income}}{\text{revenue}} \text{ or } \frac{\text{EBIT}}{\text{revenue}}$$

Strictly speaking, EBIT includes some nonoperating items, such as gains on investment. The analyst, as with other ratios with various formulations, must be consistent in his calculation method and know how published ratios are calculated. Analysts should be concerned if this ratio is too low. Some analysts prefer to calculate the operating profit margin by adding back depreciation and any amortization expense to arrive at earnings before interest, taxes, depreciation, and amortization (EBITDA).

- Sometimes profitability is measured using earnings before tax (EBT), which can be calculated by subtracting interest from EBIT or from operating earnings. The *pretax margin* is calculated as:

$$\text{pretax margin} = \frac{\text{EBT}}{\text{revenue}}$$

- Another set of profitability ratios measures profitability relative to funds invested in the company by common stockholders, preferred stockholders, and suppliers of debt financing. The first of these measures is the *return on assets* (ROA). Typically, ROA is calculated using net income:

$$\text{return on assets (ROA)} = \frac{\text{net income}}{\text{average total assets}}$$

This measure is a bit misleading, however, because interest is excluded from net income but total assets include debt as well as equity. Adding interest adjusted for tax back to net income puts the returns to both equity and debt holders in the numerator. The interest expense that should be added back is gross interest expense, not net interest expense (which is gross interest expense less interest income). This results in an alternative calculation for ROA:

$$\text{return on assets (ROA)} = \frac{\text{net income} + \text{interest expense} (1 - \text{tax rate})}{\text{average total assets}}$$

- A measure of return on assets that includes both taxes and interest in the numerator is the *operating return on assets*:

$$\text{operating return on assets} = \frac{\text{operating income}}{\text{average total assets}} \text{ or } \frac{\text{EBIT}}{\text{average total assets}}$$

- The *return on total capital* (ROTC) is the ratio of net income before interest and taxes to total capital:

$$\text{return on total capital} = \frac{\text{EBIT}}{\text{average total capital}}$$

Total capital includes short- and long-term debt, preferred equity, and common equity. Analysts should be concerned if this ratio is too low.

- The *return on equity* (ROE) is the ratio of net income to average total equity (including preferred stock):

$$\text{return on equity} = \frac{\text{net income}}{\text{average total equity}}$$

Analysts should be concerned if this ratio is too low. It is sometimes called return on total equity.

- A similar ratio to the return on equity is the *return on common equity*:

$$\begin{aligned}\text{return on common equity} &= \frac{\text{net income} - \text{preferred dividends}}{\text{average common equity}} \\ &= \frac{\text{net income available to common}}{\text{average common equity}}\end{aligned}$$

This ratio differs from the return on total equity in that it only measures the accounting profits available to, and the capital invested by, common stockholders, instead of common and preferred stockholders. That is why preferred dividends are deducted from net income in the numerator. Analysts should be concerned if this ratio is too low.

The return on common equity is often more thoroughly analyzed using the DuPont decomposition, which is described later in this reading.

EXAMPLE: Calculating ratios

A balance sheet and income statement for Sedgwick Company are shown in the following tables for this year and the previous year.

Using the company information provided, calculate the following ratios for the current year: current ratio, total asset turnover, net profit margin, return on common equity, and total debt to equity.

Sedgwick Company Balance Sheet		
Year	Current Year	Previous Year
Assets		
Cash and marketable securities	\$105	\$95
Receivables	205	195
Inventories	310	290
Total current assets	620	580
Gross property, plant, and equipment	1,800	\$1,700
Accumulated depreciation	360	340
Net property, plant, and equipment	1,440	1,360
Total assets	\$2,060	\$1,940
Liabilities		
Payables	\$110	\$90
Short-term debt	160	140
Current portion of long-term debt	55	45
Current liabilities	325	\$275
Long-term debt	610	\$690
Deferred taxes	105	95
Common stock at par	300	300
Additional paid in capital	400	400
Retained earnings	320	180
Common shareholders' equity	1,020	880
Total liabilities and equity	\$2,060	\$1,940

Sedgwick Company Income Statement	
Year	Current Year
Sales	\$4,000
Cost of goods sold	3,000
Gross profit	1,000
Operating expenses	650
Operating profit	350
Interest expense	50
Earnings before taxes	300
Taxes	100
Net income	200
Common dividends	60

Answer:

$$\text{current ratio} = \frac{\text{current assets}}{\text{current liabilities}}$$

$$\text{current ratio} = \frac{620}{325} = 1.9$$

$$\text{total asset turnover} = \frac{\text{revenue}}{\text{average assets}}$$

$$\text{total asset turnover} = \frac{4,000}{(2,060 + 1,940) / 2} = 2.0$$

$$\text{net profit margin} = \frac{\text{net income}}{\text{revenue}}$$

$$\text{net profit margin} = \frac{200}{4,000} = 5.0\%$$

$$\text{return on common equity} = \frac{\text{net income} - \text{preferred dividends}}{\text{average common equity}}$$

$$\text{return on common equity} = \frac{200}{(1,020 + 880) / 2} = 21.1\%$$

$$\text{debt-to-equity ratio} = \frac{\text{total debt}}{\text{total equity}}$$

$$\text{debt-to-equity ratio} = \frac{610 + 160 + 55}{1,020} = 80.9\%$$

Note that preferred equity would be included in the denominator if there were any, and that we have included short-term debt and the current portion of long-term debt in calculating total (interest-bearing) debt.

LOS 21.c: Describe relationships among ratios and evaluate a company using ratio analysis.

EXAMPLE: Relationships among ratios

An analyst calculates the following activity and liquidity ratios for a company over the last three years:

	20X8	20X7	20X6
Current ratio	2.0	1.5	1.2
Quick ratio	0.5	0.8	1.0
Days of inventory	60	50	30
Days' sales outstanding	20	30	40

Determine what the analyst should infer from these ratios taken together.

Answer:

The current ratio has increased over this period, while the quick ratio has decreased. This could result from inventories increasing or from current assets other than inventories decreasing. The increase in days of inventory on hand suggests increasing inventories explain the opposing trends in the current and quick ratios.

The decrease in days' sales outstanding indicates that the company has been collecting cash from customers sooner than it had been in the past. Taken together, these ratios suggest the company may be accelerating its collections to make up for a drain on cash from poor inventory management.

EXAMPLE: Using ratios to evaluate a company

An analyst is comparing the values of several ratios for the current year to their prior-year values, the current industry average values, and industry average ratios for Sedgwick Company. These selected ratio values are shown in the following table.

	Current Year	Previous Year	Industry Average
Current ratio	1.9	2.1	1.5
Total asset turnover	2.0	2.3	2.4
Net profit margin	5.0%	5.8%	6.5%
Return on common equity	21.1%	24.1%	19.8%
Debt-to-equity	80.9%	99.4%	35.7%

Discuss how these ratios compare with the company's performance last year and with the industry performance.

Answer:

Although the firm's liquidity, as measured by the current ratio, has decreased over the past year, it remains above the industry average.

Total asset turnover has declined over the past year and now appears to be significantly lower than the industry average.

Net profit margin is lower than last year and much lower than the industry average.

Return on equity is lower than last year but still higher than the industry average. Given the decline in net profit margin and total asset turnover, it is likely that higher-than-average financial leverage is the reason for this ROE outperformance.

Our supposition about the firm's financial leverage is confirmed by its debt-to-equity ratios. While the current year's ratio is lower than last year's, it is still more than twice the industry-average ratio. The significant decrease in the company's debt-to-equity ratio over the past year suggests the company is trying to get its debt level more in line with the industry average.

MODULE QUIZ 21.1, 21.2, 21.3



1. To study trends in a firm's cost of goods sold (COGS), the analyst should standardize the cost of goods sold numbers to a common-sized basis by dividing COGS by:
 - A. assets.
 - B. sales.
 - C. net income.
2. Which of the following is *least likely* a limitation of financial ratios?
 - A. Data on comparable firms are difficult to acquire.
 - B. Determining the target or comparison value for a ratio requires judgment.
 - C. Different accounting treatments require the analyst to adjust the data before comparing ratios.

3. RGB, Inc.'s purchases during the year were \$100,000. The balance sheet shows an average accounts payable balance of \$12,000. RGB's payables payment period is *closest* to:
- 37 days.
 - 44 days.
 - 52 days.
4. RGB, Inc., has a gross profit of \$45,000 on sales of \$150,000. The balance sheet shows average total assets of \$75,000 with an average inventory balance of \$15,000. RGB's total asset turnover and inventory turnover are *closest* to:
- | <u>Asset turnover</u> | <u>Inventory turnover</u> |
|-----------------------|---------------------------|
| A. 7.00 times | 2.00 times |
| B. 2.00 times | 7.00 times |
| C. 0.50 times | 0.33 times |
5. If RGB, Inc., has annual sales of \$100,000, average accounts payable of \$30,000, and average accounts receivable of \$25,000, RGB's receivables turnover and average collection period are *closest* to:
- | <u>Receivables turnover</u> | <u>Average collection period</u> |
|-----------------------------|----------------------------------|
| A. 2.1 times | 174 days |
| B. 3.3 times | 111 days |
| C. 4.0 times | 91 days |
6. A company's current ratio is 1.9. If some of the accounts payable are paid off from the cash account, the:
- numerator would decrease by a greater percentage than the denominator, resulting in a lower current ratio.
 - denominator would decrease by a greater percentage than the numerator, resulting in a higher current ratio.
 - numerator and denominator would decrease proportionally, leaving the current ratio unchanged.
7. A company's quick ratio is 1.2. If inventory were purchased for cash, the:
- numerator would decrease more than the denominator, resulting in a lower quick ratio.
 - denominator would decrease more than the numerator, resulting in a higher current ratio.
 - numerator and denominator would decrease proportionally, leaving the current ratio unchanged.
8. All other things held constant, which of the following transactions will increase a firm's current ratio if the ratio is greater than one?
- Accounts receivable are collected and the funds received are deposited in the firm's cash account.
 - Fixed assets are purchased from the cash account.
 - Accounts payable are paid with funds from the cash account.
9. RGB, Inc.'s receivable turnover is ten times, the inventory turnover is five times, and the payables turnover is nine times. RGB's cash conversion cycle is *closest* to:
- 69 days.
 - 104 days.
 - 150 days.
10. An analyst who is interested in a company's long-term solvency would *most likely* examine the:
- return on total capital.
 - defensive interval ratio.
 - fixed charge coverage ratio.

11. RGB, Inc.'s income statement shows sales of \$1,000, cost of goods sold of \$400, pre-interest operating expense of \$300, and interest expense of \$100. RGB's interest coverage ratio is closest to:
- 2 times.
 - 3 times.
 - 4 times.

MODULE 21.4: DUPONT ANALYSIS



Video covering this content is available online.

The **DuPont system of analysis** is an approach that can be used to analyze return on equity (ROE). It uses basic algebra to break down ROE into a function of different ratios, so an analyst can see the impact of leverage, profit margins, and turnover on shareholder returns. There are two variants of the DuPont system: The original three-part approach and the extended five-part system.

For the **original approach**, start with ROE defined as:

$$\text{return on equity} = \left(\frac{\text{net income}}{\text{average equity}} \right)$$

Average or year-end values for equity can be used. Multiplying ROE by (revenue/revenue) and rearranging terms produces:

$$\text{return on equity} = \left(\frac{\text{net income}}{\text{revenue}} \right) \left(\frac{\text{revenue}}{\text{average equity}} \right)$$

The first term is the profit margin, and the second term is the equity turnover:

$$\text{return on equity} = \left(\frac{\text{net profit}}{\text{margin}} \right) \left(\frac{\text{equity}}{\text{turnover}} \right)$$

We can expand this further by multiplying these terms by (assets/assets), and rearranging terms:

$$\text{return on equity} = \left(\frac{\text{net income}}{\text{revenue}} \right) \left(\frac{\text{revenue}}{\text{average total assets}} \right) \left(\frac{\text{average total assets}}{\text{equity}} \right)$$

The first term is still the profit margin, the second term is now asset turnover, and the third term is a financial leverage ratio that will increase as the use of debt financing increases:

$$\text{return on equity} = \text{net profit margin} \times \text{asset turnover} \times \text{leverage ratio}$$



PROFESSOR'S NOTE

The leverage ratio is sometimes called the "equity multiplier."

This is the original DuPont equation. It is arguably the most important equation in ratio analysis, since it breaks down a very important ratio (ROE) into three key components. If ROE is relatively low, it must be that at least one of the following is true: The company has a poor profit margin, the company has poor asset turnover, or the firm has too little leverage.

PROFESSOR'S NOTE



Often candidates get confused and think the DuPont method is a way to calculate ROE. While you can calculate ROE given the components of either the original or extended DuPont equations, this isn't necessary if you have the financial statements. If you have net income and equity, you can calculate ROE. The DuPont method is a way to decompose ROE, to better see what changes are driving the changes in ROE.

EXAMPLE: Decomposition of ROE with original DuPont

Staret, Inc., has maintained a stable and relatively high ROE of approximately 18% over the last three years. Use traditional DuPont analysis to decompose this ROE into its three components and comment on trends in company performance.

Staret, Inc., Selected Balance Sheet and Income Statement Items (Millions)			
Year	20X3	20X4	20X5
Net income	21.5	22.3	21.9
Revenue	305	350	410
Average equity	119	124	126
Average assets	230	290	350

Answer:

ROE

$$20X3: 21.5 / 119 = 18.1\%$$

$$20X4: 22.3 / 124 = 18.0\%$$

$$20X5: 21.9 / 126 = 17.4\%$$

DuPont

$$20X3: 7.0\% \times 1.33 \times 1.93$$

$$20X4: 6.4\% \times 1.21 \times 2.34$$

$$20X5: 5.3\% \times 1.17 \times 2.78$$

(some rounding in values)

While ROE has dropped only slightly, both total asset turnover and net profit margin have declined. The effects of declining net margins and turnover on ROE have been offset by a significant increase in leverage. The analyst should be concerned about the decrease in net margin and determine the combination of pricing pressure and/or increasing expenses that caused this. Note that the company has become more risky due to increased debt financing.

EXAMPLE: Computing ROE using original DuPont

A company has a net profit margin of 4%, asset turnover of 2.0, and a debt-to-assets ratio of 60%. What is the ROE?

Answer:

Debt-to-assets = 60%, which means equity to assets is 40%; this implies assets to equity (the leverage ratio) is $1 / 0.4 = 2.5$

$$\text{ROE} = (\text{net profit margin})(\text{total asset turnover})(\text{average assets} / \text{average equity}) = (0.04)(2.00)(2.50) = 0.20, \text{ or } 20\%$$

The **extended (5-way) DuPont equation** takes the net profit margin and breaks it down further.

$$ROE = \left(\frac{\text{net income}}{\text{EBT}} \right) \left(\frac{\text{EBT}}{\text{EBIT}} \right) \left(\frac{\text{EBIT}}{\text{revenue}} \right) \left(\frac{\text{revenue}}{\text{average assets}} \right) \left(\frac{\text{average assets}}{\text{average equity}} \right)$$

Note that the first term in the 3-part DuPont equation, net profit margin, has been decomposed into three terms:

$\frac{\text{net income}}{\text{EBT}}$ is called the *tax burden* and is equal to $(1 - \text{tax rate})$.

$\frac{\text{EBT}}{\text{EBIT}}$ is called the *interest burden*.

$\frac{\text{EBIT}}{\text{revenue}}$ is called the *EBIT margin*.

We then have:

$$ROE = \left(\frac{\text{tax burden}}{\text{burden}} \right) \left(\frac{\text{interest burden}}{\text{burden}} \right) \left(\frac{\text{EBIT margin}}{\text{margin}} \right) \left(\frac{\text{asset turnover}}{\text{turnover}} \right) \left(\frac{\text{financial leverage}}{\text{leverage}} \right)$$

An increase in interest expense as proportion of EBIT will increase the interest burden (i.e., decrease the interest burden ratio). Increases in either the tax burden or the interest burden (i.e., decreases in the ratios) will tend to decrease ROE.

EBIT in the second two expressions can be replaced by operating earnings. In this case, we have the operating margin rather than the EBIT margin. The interest burden term would then show the effects of nonoperating income as well as the effect of interest expense.

Note that in general, high profit margins, leverage, and asset turnover will lead to high levels of ROE. However, this version of the formula shows that more leverage *does not always* lead to higher ROE. As leverage rises, so does the interest burden. Hence, the positive effects of leverage can be offset by the higher interest payments that accompany more debt. Note that higher taxes will always lead to lower levels of ROE.

EXAMPLE: Extended DuPont analysis

An analyst has gathered data from two companies in the same industry. Calculate the ROE for both companies and use the extended DuPont analysis to explain the critical factors that account for the differences in the two companies' ROEs.

Selected Income and Balance Sheet Data		
	Company A	Company B
Revenues	\$500	\$900
EBIT	35	100
Interest expense	5	0
EBT	30	100
Taxes	10	40
Net income	20	60
Average assets	250	300
Total debt	100	50
Average equity	\$150	\$250

Answer:

EBIT margin = EBIT / revenue

Company A: EBIT margin = $35 / 500 = 7.0\%$

Company B: EBIT margin = $100 / 900 = 11.1\%$

asset turnover = revenue / average assets

Company A: asset turnover = $500 / 250 = 2.0$

Company B: asset turnover = $900 / 300 = 3.0$

interest burden = EBT / EBIT

Company A: interest burden = $30 / 35 = 85.7\%$

Company B: interest burden = $100 / 100 = 1$

financial leverage = average assets / average equity

Company A: financial leverage = $250 / 150 = 1.67$

Company B: financial leverage = $300 / 250 = 1.2$

tax burden = net income / EBT

Company A: tax burden = $20 / 30 = 66.7\%$

Company B: tax burden = $60 / 100 = 60.0\%$

Company A: ROE = $0.667 \times 0.857 \times 0.07 \times 2.0 \times 1.67 = 13.4\%$

Company B: ROE = $0.608 \times 1.0 \times 0.111 \times 3.0 \times 1.2 = 24\%$

Company B has a higher tax burden but a lower interest burden (a lower ratio indicates a higher burden). Company B has better EBIT margins and better asset utilization (perhaps management of inventory, receivables, or payables, or a lower cost basis in its fixed assets due to their age), and less leverage. Its higher EBIT margins and asset turnover are the main factors leading to its significantly higher ROE, which it achieves with less leverage than Company A.

MODULE QUIZ 21.4



1. Return on equity using the traditional DuPont formula equals:
 - (net profit margin) (interest component) (solvency ratio).
 - (net profit margin) (total asset turnover) (tax retention rate).
 - (net profit margin) (total asset turnover) (financial leverage multiplier).
2. RGB, Inc., has a net profit margin of 12%, a total asset turnover of 1.2 times, and a financial leverage multiplier of 1.2 times. RGB's return on equity is *closest* to:
 - 12.0%.
 - 14.2%.
 - 17.3%.
3. Use the following information for RGB, Inc.:
 - EBIT / revenue = 10%
 - Tax retention rate = 60%
 - Revenue / assets = 1.8 times
 - Current ratio = 2 times
 - EBT / EBIT = 0.9 times
 - Assets / equity = 1.9 timesRGB, Inc.'s return on equity is *closest* to:
 - 10.5%.
 - 14.0%.
 - 18.5%.
4. Which of the following equations *least accurately* represents return on equity?
 - (net profit margin)(equity turnover).
 - (net profit margin)(total asset turnover)(assets / equity).
 - (ROA)(interest burden)(tax retention rate).

5. Paragon Co. has an operating profit margin (EBIT / revenue) of 11%; an asset turnover ratio of 1.2; a financial leverage multiplier of 1.5 times; an average tax rate of 35%; and an interest burden of 0.7. Paragon's return on equity is *closest* to:
- A. 9%.
 - B. 10%.
 - C. 11%.

MODULE 21.5: MORE FINANCIAL RATIOS



Video covering
this content is
available online.

LOS 21.e: Calculate and interpret ratios used in equity analysis and credit analysis.

Valuation ratios are used in analysis for investment in common equity. The most widely used valuation ratio is the *price-to-earnings* (P/E) ratio, the ratio of the current market price of a share of stock divided by the company's earnings per share. Related measures based on price per share are the *price-to-cash flow*, the *price-to-sales*, and the *price-to-book value* ratios.



PROFESSOR'S NOTE

The use of the previous valuation ratios is covered in Equity Investments.

Per-share valuation measures include *earnings per share* (EPS). *Basic EPS* is net income available to common divided by the weighted average number of common shares outstanding.

Diluted EPS is a "what if" value. It is calculated to be the lowest possible EPS that could have been reported if all firm securities that can be converted into common stock, and that would decrease basic EPS if they had been, were converted. That is, if all dilutive securities had been converted. Potentially dilutive securities include convertible debt and convertible preferred stock, as well as options and warrants issued by the company. The numerator of diluted EPS is increased by the after-tax interest savings on any dilutive debt securities and by the dividends on any dilutive convertible preferred stock. The denominator is increased by the common shares that would result from conversion or exchange of dilutive securities into common shares.



PROFESSOR'S NOTE

Refer back to our reading on Understanding Income Statements for details and examples of how to calculate basic and diluted EPS.

Other per-share measures include *cash flow per share*, *EBIT per share*, and *EBITDA per share*. Per share measures are not comparable because the number of outstanding shares differ among firms. For example, assume Firm A and Firm B both report net income of \$100. If Firm A has 100 shares outstanding, its EPS is \$1 per share. If Firm B has 20 shares outstanding, its EPS is \$5 per share.

Dividends

Dividends are declared on a per-common-share basis. Total dividends on a firm-wide basis are referred to as *dividends declared*. Neither EPS nor net income is reduced by the payment of common stock dividends. Net income minus dividends declared is retained earnings, the earnings that are used to grow the corporation rather than being distributed to equity holders.

The proportion of a firm's net income that is retained to fund growth is an important determinant of the firm's *sustainable growth rate*.

To estimate the sustainable growth rate for a firm, the rate of return on resources is measured as the return on equity capital, or the ROE. The proportion of earnings reinvested is known as the *retention rate* (RR).

- The formula for the sustainable growth rate, which is how fast the firm can grow without additional external equity issues while holding leverage constant, is:

$$g = RR \times ROE$$

- The calculation of the retention rate is:

$$\begin{aligned} \text{retention rate} &= \frac{\text{net income available to common} - \text{dividends declared}}{\text{net income available to common}} \\ &= 1 - \text{dividend payout ratio} \end{aligned}$$

where:

$$\text{dividend payout ratio} = \frac{\text{dividends declared}}{\text{net income available to common}}$$

EXAMPLE: Calculating sustainable growth

The following figure provides data for three companies.

Growth Analysis Data			
Company	A	B	C
Earnings per share	\$3.00	\$4.00	\$5.00
Dividends per share	1.50	1.00	2.00
Return on equity	14%	12%	10%

Calculate the sustainable growth rate for each company.

Answer:

$$RR = 1 - (\text{dividends} / \text{earnings})$$

$$\text{Company A: } RR = 1 - (1.50 / 3.00) = 0.500$$

$$\text{Company B: } RR = 1 - (1.00 / 4.00) = 0.750$$

$$\text{Company C: } RR = 1 - (2.00 / 5.00) = 0.600$$

$$g = RR \times ROE$$

$$\text{Company A: } g = 0.500 \times 14\% = 7.0\%$$

$$\text{Company B: } g = 0.750 \times 12\% = 9.0\%$$

$$\text{Company C: } g = 0.600 \times 10\% = 6.0\%$$

Some ratios have specific applications in certain industries.

Net income per employee and *sales per employee* are used in the analysis and valuation of service and consulting companies.

Growth in same-store sales is used in the restaurant and retail industries to indicate growth without the effects of new locations that have been opened. It is a measure of how well the firm is doing at attracting and keeping existing customers and, in the case of locations with overlapping markets, may indicate that new locations are taking customers from existing ones.

Sales per square foot is another metric commonly used in the retail industry.

Business Risk

The standard deviation of revenue, standard deviation of operating income, and the standard deviation of net income are all indicators of the variation in and the uncertainty about a firm's performance. Since they all depend on the size of the firm to a great extent, analysts employ a size-adjusted measure of variation. The **coefficient of variation** for a variable is its standard deviation divided by its expected value.



PROFESSOR'S NOTE

We saw this before as a measure of portfolio risk in Quantitative Methods.

Certainly, different industries have different levels of uncertainty about revenues, expenses, taxes, and nonoperating items. Comparing coefficients of variation for a firm across time, or among a firm and its peers, can aid the analyst in assessing both the relative and absolute degree of risk a firm faces in generating income for its investors.

$$CV_{sales} = \frac{\text{standard deviation of sales}}{\text{mean sales}}$$

$$CV_{operating\ income} = \frac{\text{standard deviation of operating income}}{\text{mean operating income}}$$

$$CV_{net\ income} = \frac{\text{standard deviation of net income}}{\text{mean net income}}$$

Banks, insurance companies, and other financial firms carry their own challenges for analysts. Part of the challenge is to understand the commonly used terms and the ratios they represent.

Capital adequacy typically refers to the ratio of some dollar measure of the risk, both operational and financial, of the firm to its equity capital. Other measures of capital are also used. A common measure of capital risk is *value-at-risk*, which is an estimate of the dollar size of the loss that a firm will exceed only some specific percent of the time, over a specific period of time.

Banks are subject to minimum *reserve requirements*. Their ratios of various liabilities to their central bank reserves must be above the minimums. The ratio of a bank's liquid assets to certain liabilities is called the *liquid asset requirement*.

The performance of financial companies that lend funds is often summarized as the *net interest margin*, which is simply interest income divided by the firm's interest-earning assets.

Credit Analysis

Credit analysis is based on many of the ratios that we have already covered in this review. In assessing a company's ability to service and repay its debt, analysts use interest coverage ratios (calculated with EBIT or EBITDA), return on capital, and debt-to-assets ratios. Other ratios focus on various measures of cash flow to total debt.

Ratios have been used to analyze and predict firm bankruptcies. Altman (2000)¹ developed a Z-score that is useful in predicting firm bankruptcies (a low score indicates high probability of

failure). The predictive model was based on a firm's working capital to assets, retained earnings to assets, EBIT to assets, market to book value of a share of stock, and revenues to assets.



PROFESSOR'S NOTE

Credit analysis is covered in more detail in Fixed Income.

LOS 21.f: Explain the requirements for segment reporting and calculate and interpret segment ratios.

A **business segment** is a portion of a larger company that accounts for more than 10% of the company's revenues, assets, or income and is distinguishable from the company's other lines of business in terms of the risk and return characteristics of the segment. **Geographic segments** are also identified when they meet the size criterion given previously and the geographic unit has a business environment that is different from that of other segments or the remainder of the company's business.

For example, in its 2016 annual report, Boeing described its business segments as follows (see www.Boeing.com):

We are organized based on the products and services we offer. We operate in five principal segments:

- Commercial Airplanes;
- Our Defense, Space & Security (BDS) business comprises three segments:
 - Boeing Military Aircraft (BMA)
 - Network & Space Systems (N&SS)
 - Global Services & Support (GS&S)
- Boeing Capital (BCC).

Both U.S. GAAP and IFRS require companies to report segment data, but the required disclosure items are only a subset of the required disclosures for the company as a whole. Nonetheless, an analyst can prepare a more detailed analysis and forecast by examining the performance of business or geographic segments separately. Segment profit margins, asset utilization (turnover), and return on assets can be very useful in gaining a clear picture of a firm's overall operations. For forecasting, growth rates of segment revenues and profits can be used to estimate future sales and profits and to determine the changes in company characteristics over time.

LOS 21.g: Describe how ratio analysis and other techniques can be used to model and forecast earnings.

Ratio analysis can be used in preparing pro forma financial statements that provide estimates of financial statement items for one or more future periods. The preparation of pro forma financial statements and related forecasts is covered in some detail in the Corporate Issuers topic area. Here, some examples will suffice.

A forecast of financial results that begins with an estimate of a firm's next-period revenues might use the most recent COGS, or an average of COGS, from a common-size income statement. On a common-size income statement, COGS is calculated as a percentage of revenue. If the analyst has no reason to believe that COGS in relation to sales will change for the next period,

the COGS percentage from a common-size income statement can be used in constructing a pro forma income statement for the next period based on the estimate of sales.

Similarly, the analyst may believe that certain ratios will remain the same or change in one direction or the other for the next period. In the absence of any information indicating a change, an analyst may choose to incorporate the operating profit margin from the prior period into a pro forma income statement for the next period. Beginning with an estimate of next-period sales, the estimated operating profit margin can be used to forecast operating profits for the next period.

Rather than point estimates of sales and net and operating margins, the analyst may examine possible changes in order to create a range of possible values for key financial variables.

Three methods of examining the variability of financial outcomes around point estimates are: *sensitivity analysis*, *scenario analysis*, and *simulation*. Sensitivity analysis is based on “what if” questions such as: What will be the effect on net income if sales increase by 3% rather than the estimated 5%? Scenario analysis is based on specific scenarios (a specific set of outcomes for key variables) and will also yield a range of values for financial statement items. Simulation is a technique in which probability distributions for key variables are selected and a computer is used to generate a distribution of values for outcomes based on repeated random selection of values for the key variables.

MODULE QUIZ 21.5

1. A firm has a dividend payout ratio of 40%, a net profit margin of 10%, an asset turnover of 0.9 times, and a financial leverage multiplier of 1.2 times. The firm’s sustainable growth rate is *closest* to:
 - A. 4.3%.
 - B. 6.5%.
 - C. 8.0%.
2. Accounting standards require segment reporting for a distinguishable part of a firm that comprises at least:
 - A. 10% of assets.
 - B. 5% of revenues.
 - C. 20% of earnings.
3. An analyst who needs to model and forecast a company’s earnings for the next three years would be *least likely* to:
 - A. assume that key financial ratios will remain unchanged for the forecast period.
 - B. use common-size financial statements to estimate expenses as a percentage of net income.
 - C. examine the variability of the predicted outcomes by performing a sensitivity or scenario analysis.



Video covering
this content is
available online.

KEY CONCEPTS

LOS 21.a

Ratios can be used to project earnings and future cash flow, evaluate a firm’s flexibility, assess management’s performance, evaluate changes in the firm and industry over time, and compare the firm with industry competitors.

Vertical common-size data are stated as a percentage of sales for income statements or as a percentage of total assets for balance sheets. Horizontal common-size data present each item as a percentage of its value in a base year.

Ratio analysis has limitations. Ratios are not useful when viewed in isolation and require adjustments when different companies use different accounting treatments. Comparable ratios may be hard to find for companies that operate in multiple industries. Ratios must be analyzed relative to one another, and determining the range of acceptable values for a ratio can be difficult.

LOS 21.b

Activity ratios indicate how well a firm uses its assets. They include receivables turnover, days of sales outstanding, inventory turnover, days of inventory on hand, payables turnover, payables payment period, and turnover ratios for total assets, fixed assets, and working capital.

Liquidity ratios indicate a firm's ability to meet its short-term obligations. They include the current, quick, and cash ratios, the defensive interval, and the cash conversion cycle.

Solvency ratios indicate a firm's ability to meet its long-term obligations. They include the debt-to-equity, debt-to-capital, debt-to-assets, financial leverage, interest coverage, and fixed charge coverage ratios.

Profitability ratios indicate how well a firm generates operating income and net income. They include net, gross, and operating profit margins, pretax margin, return on assets, operating return on assets, return on total capital, return on total equity, and return on common equity.

Valuation ratios are used to compare the relative values of stocks. They include earnings per share and price-to-earnings, price-to-sales, price-to-book value, and price-to-cash-flow ratios.

LOS 21.c

An analyst should use an appropriate combination of different ratios to evaluate a company over time and relative to comparable companies. The interpretation of an increase in ROE, for example, may be quite different for a firm that has significantly increased its financial leverage compared to one that has maintained or decreased its financial leverage.

LOS 21.d

Basic DuPont equation:

$$\text{ROE} = \left(\frac{\text{net income}}{\text{sales}} \right) \left(\frac{\text{sales}}{\text{assets}} \right) \left(\frac{\text{assets}}{\text{equity}} \right)$$

Extended DuPont equation:

$$\text{ROE} = \left(\frac{\text{net income}}{\text{EBT}} \right) \left(\frac{\text{EBT}}{\text{EBIT}} \right) \left(\frac{\text{EBIT}}{\text{revenue}} \right) \left(\frac{\text{revenue}}{\text{total assets}} \right) \left(\frac{\text{total assets}}{\text{total equity}} \right)$$

LOS 21.e

Ratios used in equity analysis include price-to-earnings, price-to-cash flow, price-to-sales, and price-to-book value ratios, and basic and diluted earnings per share. Other ratios are relevant to specific industries such as retail and financial services.

Credit analysis emphasizes interest coverage ratios, return on capital, debt-to-assets ratios, and cash flow to total debt.

LOS 21.f

A business or geographic segment is a portion of a firm that has risk and return characteristics distinguishable from the rest of the firm and accounts for more than 10% of the firm's sales or assets.

Firms are required to report some items for significant business and geographic segments. Profitability, leverage, and turnover ratios by segment can give the analyst a better understanding of the performance of the overall business.

LOS 21.g

Ratio analysis in conjunction with other techniques can be used to construct pro forma financial statements based on a forecast of sales growth and assumptions about the relation of changes in key income statement and balance sheet items to growth of sales.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 21.1, 21.2, 21.3

1. **B** With a vertical common-size income statement, all income statement accounts are divided by sales. (Module 21.1, LOS 21.a)
2. **A** Company and industry data are widely available from numerous private and public sources. The other statements describe limitations of financial ratios. (Module 21.1, LOS 21.a)
3. **B** payables turnover = $(\text{purchases} / \text{avg. AP}) = 100 / 12 = 8.33$
payables payment period = $365 / 8.33 = 43.8$ days
(Module 21.2, LOS 21.b)
4. **B** total asset turnover = $(\text{sales} / \text{total assets}) = 150 / 75 = 2$ times
inventory turnover = $(\text{COGS} / \text{avg. inventory}) = (150 - 45) / 15 = 7$ times
(Module 21.2, LOS 21.b)
5. **C** receivables turnover = $(S / \text{avg. AR}) = 100 / 25 = 4$
average collection period = $365 / 4 = 91.25$ days
(Module 21.2, LOS 21.b)
6. **B** Current ratio = current assets / current liabilities. If cash (a current asset) and AP (a current liability) decrease by the same amount and the current ratio is greater than 1, then the numerator decreases less in percentage terms than the denominator, and the current ratio increases. (Module 21.2, LOS 21.b)
7. **A** Quick ratio = $(\text{cash} + \text{marketable securities} + \text{AR}) / \text{current liabilities}$. If cash decreases, the quick ratio will also decrease. The denominator is unchanged. (Module 21.2, LOS 21.b)
8. **C** Current ratio = current assets / current liabilities. If CR is > 1, then if CA and CL both fall, the overall ratio will increase. (Module 21.2, LOS 21.b)

9. A $(365 / 10 + 365 / 5 - 365 / 9) = 69$ days

(Module 21.2, LOS 21.b)

10. C Fixed charge coverage is a solvency ratio. Return on total capital is a measure of profitability and the defensive interval ratio is a liquidity measure. (Module 21.3, LOS 21.b)

11. B Interest coverage ratio = $EBIT / I = (1,000 - 400 - 300) / 100 = 3$ times

(Module 21.3, LOS 21.b)

Module Quiz 21.4

1. C This is the correct formula for the three-ratio DuPont model for ROE. (LOS 21.d)

2. C

$$\text{return on equity} = \left(\frac{\text{net income}}{\text{sales}} \right) \left(\frac{\text{sales}}{\text{assets}} \right) \left(\frac{\text{assets}}{\text{equity}} \right) = (0.12)(1.2)(1.2) =$$

$$0.1728 = 17.28\%$$

(LOS 21.d)

3. C Tax burden = $(1 - \text{tax rate}) = \text{tax retention rate} = 0.6$.

$$\text{ROE} = 0.6 \times 0.9 \times 0.1 \times 1.8 \times 1.9 = 0.1847 = 18.47\%$$

(LOS 21.d)

4. C $(\text{ROA})(\text{interest burden})(\text{tax retention rate})$ is not one of the DuPont models for calculating ROE. (LOS 21.d)

5. A Tax burden = $1 - 0.35 = 0.65$.

$$\text{ROE} = 0.65 \times 0.7 \times 0.11 \times 1.2 \times 1.5 = 0.0901.$$

(LOS 21.d)

Module Quiz 21.5

1. B $g = (\text{retention rate})(\text{ROE})$

$$\text{ROE} = \text{net profit margin} \times \text{asset turnover} \times \text{equity multiplier} = (0.1)(0.9)(1.2) = 0.108$$

$$g = (1 - 0.4)(0.108) = 6.5\%$$

(LOS 21.e)

2. A For segment reporting, a business segment is a distinguishable portion of the overall company that produces more than 10% of its revenues or accounts for more than 10% of its assets. (LOS 21.f)

3. B An earnings forecast model would typically estimate expenses as a percentage of sales. (LOS 21.g)

1. Edward I. Altman, "Predicting Financial Distress of Companies: Revisiting the Z-Score and Zeta[®] Models," July 2000.

Reading 22

INVENTORIES

EXAM FOCUS

This reading discusses the different inventory cost flow methods: FIFO, LIFO, and weighted average cost. You must understand how to calculate COGS, ending inventory, and gross profit under each of these methods. Also, you must understand the effects of each method on a firm's liquidity, profitability, activity, and solvency ratios. Be able to apply the appropriate inventory valuation method under IFRS (lower of cost or net realizable value) and U.S. GAAP (lower of cost or market), and calculate inventory losses and loss reversals, if allowed. Finally, be able to evaluate a firm's effectiveness in managing its inventory.

MODULE 22.1: COST FLOW METHODS



Merchandising firms, such as wholesalers and retailers, purchase inventory that is ready for sale. In this case, inventory is reported in one account on the balance sheet. Manufacturing firms normally report inventory using three separate accounts: raw materials, work-in-process, and finished goods.

Video covering this content is available online.

Cost of goods sold (COGS), also referred to as cost of sales (COS) under IFRS, is related to the beginning balance of inventory, purchases, and the ending balance of inventory. The relationship is summarized in the following equation:

$$\text{COGS} = \text{beginning inventory} + \text{purchases} - \text{ending inventory}$$

This equation can be rearranged to solve for any of the four variables:

$$\text{purchases} = \text{ending inventory} - \text{beginning inventory} + \text{COGS}$$

$$\text{beginning inventory} = \text{COGS} - \text{purchases} + \text{ending inventory}$$

$$\text{ending inventory} = \text{beginning inventory} + \text{purchases} - \text{COGS}$$



PROFESSOR'S NOTE

Many candidates find the inventory equation easiest to remember in this last form. If you start with beginning inventory, add the goods that came in (purchases), and subtract the goods that went out (COGS), the result must be ending inventory.

LOS 22.a: Contrast costs included in inventories and costs recognised as expenses in the period in which they are incurred.

Cost is the basis for most inventory valuation. The main issue involves determining the amounts that should be included in cost.

The costs included in inventory are similar under IFRS and U.S. GAAP. These costs, known as **product costs**, are capitalized in the Inventories account on the balance sheet and include:

- Purchase cost less trade discounts and rebates.
- Conversion (manufacturing) costs including labor and overhead.
- Other costs necessary to bring the inventory to its present location and condition.

By capitalizing inventory cost as an asset, expense recognition is delayed until the inventory is sold and revenue is recognized.

Not all inventory costs are capitalized; some costs are expensed in the period incurred. These costs, known as **period costs**, include:

- Abnormal waste of materials, labor, or overhead.
- Storage costs (unless required as part of production).
- Administrative overhead.
- Selling costs.

EXAMPLE: Costs included in inventory

Vindaloo Company manufactures a single product. The following information was taken from the company's production and cost records last year:

Units produced	5,000
Raw materials	\$15,000
Conversion cost for finished goods	\$20,000
Freight-in to plant	\$800
Storage cost for finished goods	\$500
Abnormal waste	\$100
Freight-out customers	\$1,100

Assuming no abnormal waste is included in conversion cost, calculate the capitalized cost of one unit.

Answer:

Capitalized inventory cost includes the raw materials cost, conversion cost, and freight-in to plant, as follows:

Raw materials	\$15,000
Conversion cost	\$20,000
Freight-in to plant	\$800
Total capitalized cost	\$35,800
Units produced	5,000
Capitalized cost per unit	\$7.16 (\$35,800 / 5,000 units)

The storage cost, abnormal waste, and the freight-out to customers are expensed as incurred.

LOS 22.b: Describe different inventory valuation methods (cost formulas).

If the cost of inventory remains constant over time, determining the firm's COGS and ending inventory is simple. To compute COGS, simply multiply the number of units sold by the cost per unit. Similarly, to compute ending inventory, multiply the number of units remaining by the cost per unit.

However, it is likely that the cost of purchasing or producing inventory will change over time. As a result, firms must select a cost flow method (known as the *cost flow assumption* under U.S. GAAP and *cost flow formula* under IFRS) to allocate the inventory cost to the income statement (COGS) and the balance sheet (ending inventory).

Under IFRS, the permissible methods are:

- Specific identification.
- First-in, first-out.
- Weighted average cost.

U.S. GAAP permits these same cost flow methods, as well as the last-in, first-out (LIFO) method. LIFO is not allowed under IFRS.

A firm can use one or more of the inventory cost flow methods. However, the firm must employ the same cost flow method for inventories of similar nature and use.

Under the **specific identification** method, each unit sold is matched with the unit's actual cost. Specific identification is appropriate when inventory items are not interchangeable and is commonly used by firms with a small number of costly and easily distinguishable items such as jewelry. Specific identification is also appropriate for special orders or projects outside a firm's normal course of business.

Under the **first-in, first-out** (FIFO) method, the first item purchased is assumed to be the first item sold. The advantage of FIFO is that ending inventory is valued based on the most recent purchases, arguably the best approximation of current cost. Conversely, FIFO COGS is based on the earliest purchase costs. In an inflationary environment, COGS will be understated compared to current cost. As a result, earnings will be overstated.

Under the **last-in, first-out** (LIFO) method, the item purchased most recently is assumed to be the first item sold. In an inflationary environment, LIFO COGS will be higher than FIFO COGS, and earnings will be lower. Lower earnings translate into lower income taxes, which increase cash flow. Under LIFO, ending inventory on the balance sheet is valued using the earliest costs. Therefore, in an inflationary environment, LIFO ending inventory is less than current cost.



PROFESSOR'S NOTE

The income tax advantages of using LIFO explain its popularity among U.S. firms. The tax savings result in the peculiar situation where lower reported earnings are associated with higher cash flow from operations.

Weighted average cost is a simple and objective method. The average cost per unit of inventory is computed by dividing the total cost of goods available for sale (beginning inventory + purchases) by the total quantity available for sale. To compute COGS, the average cost per unit is multiplied by the number of units sold. Similarly, to compute ending inventory, the average cost per unit is multiplied by the number of units that remain.

During inflationary or deflationary periods, the weighted average cost method will produce an inventory value between those produced by FIFO and LIFO.

Figure 22.1: Inventory Cost Flow Comparison

Method	Assumption	Cost of Goods Sold Consists of...	Ending Inventory Consists of...
FIFO (U.S. and IFRS)	The items first purchased are the first to be sold.	first purchased	most recent purchases
LIFO (U.S. only)	The items last purchased are the first to be sold.	last purchased	earliest purchases
Weighted average cost (U.S. and IFRS)	Items sold are a mix of purchases.	average cost of all items	average cost of all items

LOS 22.c: Calculate and compare cost of sales, gross profit, and ending inventory using different inventory valuation methods and using perpetual and periodic inventory systems.

The following example demonstrates how to calculate COGS and ending inventory using the FIFO, LIFO, and weighted average cost flow methods.

EXAMPLE: Inventory cost flow methods

Use the inventory data in the following figure to calculate the cost of goods sold and ending inventory under the FIFO, LIFO, and weighted average cost methods.

Inventory Data		
January 1 (beginning inventory)	2 units @ \$2 per unit =	\$4
January 7 purchase	3 units @ \$3 per unit =	\$9
January 19 purchase	5 units @ \$5 per unit =	\$25
Cost of goods available	10 units	\$38
Units sold during January	7 units	

Answer:

FIFO cost of goods sold. Value the seven units sold at the unit cost of the first units purchased. Start with the earliest units purchased and work down, as illustrated in the following figure.

FIFO COGS Calculation		
From beginning inventory	2 units @ \$2 per unit =	\$4
From first purchase	3 units @ \$3 per unit =	\$9
From second purchase	2 units @ \$5 per unit =	\$10
FIFO cost of goods sold	7 units	\$23
Ending inventory	3 units @ \$5 =	\$15

LIFO cost of goods sold. Value the seven units sold at the unit cost of the last units purchased. Start with the most recently purchased units and work up, as illustrated in the following figure.

LIFO COGS Calculation		
From second purchase	5 units @ \$5 per unit =	\$25
From first purchase	2 units @ \$3 per unit =	\$6
LIFO cost of goods sold	7 units	\$31
Ending inventory	2 units @ \$2 + 1 unit @ \$3 =	\$7

Average cost of goods sold. Value the seven units sold at the average unit cost of goods available.

Weighted Average COGS Calculation		
Average unit cost	\$38 / 10 =	\$3.80 per unit
Weighted average cost of goods sold	7 units @ \$3.80 per unit =	\$26.60
Ending inventory	3 units @ \$3.80 per unit =	\$11.40

Summary		
Inventory system	COGS	Ending Inventory
FIFO	\$23.00	\$15.00
LIFO	\$31.00	\$7.00
Average cost	\$26.60	\$11.40

Note that prices and inventory levels were rising over the period and that purchases during the period were the same for all cost flow methods.

MODULE QUIZ 22.1

-  1. Which of the following is *most likely* included in a firm's ending inventory?
- A. Storage costs of finished goods.
 - B. Variable production overhead.
 - C. Selling and administrative costs.
2. Under which inventory cost flow assumption does inventory on the balance sheet *best* approximate its current cost?
- A. First-in, first-out.
 - B. Weighted average cost.
 - C. Last-in, first-out.
3. During the year, a firm's inventory purchases were as follows:

Quarter	Units Purchased	Cost per Unit	Total
1	400	\$3.30	\$1,320
2	100	3.60	360
3	200	3.90	780
4	50	4.20	210
	750		\$2,670

- The firm uses a periodic inventory system and calculates inventory and COGS at the end of the year.
- Beginning inventory was 200 units at \$3 per unit = \$600.
- Sales for the year were 600 units.

Compute COGS for the year under FIFO and LIFO.

<u>FIFO</u>	<u>LIFO</u>
A. \$1,920	\$2,175
B. \$1,920	\$1,850
C. \$2,070	\$2,175

4. During May, a firm's inventory account included the following transactions:

May 1	Inventory	25 units @ \$4.00
May 12	Purchased	60 units @ \$4.20
May 16	Sold	40 units @ \$6.00
May 27	Purchased	30 units @ \$4.25
May 29	Sold	40 units @ \$6.10

Assuming periodic FIFO inventory costing, gross profit for May was:

- A. \$132.
 - B. \$147.
 - C. \$153.
5. In periods of rising prices and stable inventory quantities, which of the following *best* describes the effect on gross profit of using LIFO as compared to using FIFO?
- A. Lower.
 - B. Higher.
 - C. The same.

MODULE 22.2: INVENTORY SYSTEMS



Firms account for changes in inventory using either a periodic or perpetual system. In a **periodic inventory system**, inventory values and COGS are determined at the end of the accounting period. No detailed records of inventory are maintained; rather, inventory acquired during the period is reported in a Purchases account. At the end of the period, purchases are added to beginning inventory to arrive at cost of goods available for sale. To calculate COGS, ending inventory is subtracted from goods available for sale.

Video covering
this content is
available online.

In a **perpetual inventory system**, inventory values and COGS are updated continuously. Inventory purchased and sold is recorded directly in inventory when the transactions occur. Thus, a Purchases account is not necessary.

For the FIFO and specific identification methods, ending inventory values and COGS are the same whether a periodic or perpetual system is used. However, periodic and perpetual inventory systems can produce different values for inventory and COGS under the LIFO and weighted average cost methods.

The following example illustrates the differences.

EXAMPLE: Periodic vs. perpetual inventory system

Our earlier cost flow illustration was actually an example of a periodic system. Accordingly, we waited until the end of January to calculate COGS and ending inventory. Now assume the purchases and sales occurred as follows:

January 1 (beginning inventory)	2 units @ \$2 per unit
January 7 purchase	3 units @ \$3 per unit
January 12 sale	4 units
January 19 purchase	5 units @ \$5 per unit
January 29 sale	3 units

Recalculate COGS and ending inventory under the FIFO and LIFO cost flow methods using a perpetual inventory system.

Answer:

In the case of FIFO, ending inventory and COGS will be the same as with the periodic system illustrated in the earlier example.

FIFO Perpetual System

The January 12 sale of 4 units consists of:

Units	From	Cost
2	Jan 1 beginning inventory	$2 \text{ units} \times \$2 = \4
2	Jan 7 purchase	$2 \text{ units} \times \$3 = \6
		<u>\$10</u>

The January 29 sale of 3 units consists of:

Units	From	Cost
1	Jan 7 purchase	$1 \text{ unit} \times \$3 = \3
2	Jan 19 purchase	$2 \text{ units} \times \$5 = \10
		<u>\$13</u>

Total FIFO COGS for January \$23

January ending inventory consists of:

Units	From	Cost
3	Jan 19 purchase	$3 \text{ units} \times \$5 = \15

FIFO COGS and ending inventory are the same whether a perpetual or periodic system is used because the first-in (and therefore the first-out) values are the same regardless of subsequent purchases.

In the case of LIFO, COGS and ending inventory under a periodic system will be different from those calculated under a perpetual system. In our earlier example, LIFO COGS and ending inventory for January were \$31 and \$7, respectively, using a periodic system. Using a perpetual system, LIFO COGS and ending inventory are \$26 and \$12.

LIFO Perpetual System

The January 12 sale of 4 units consists of:

Units	From	Cost
3	Jan 7 purchase	$3 \text{ units} \times \$3 = \9
1	Jan 1 purchase	$1 \text{ units} \times \$2 = \2
		<u>\$11</u>

The January 29 sale of 3 units consists of:

Units	From	Cost
3	Jan 19 purchase	$3 \text{ units} \times \$5 = \15

Total LIFO COGS for January \$26

January ending inventory consists of:

Units	From	Cost
1	Jan 1 beginning inventory	$1 \text{ units} \times \$2 = \2
2	Jan 19 purchase	$2 \text{ units} \times \$5 = \10
		<u>\$12</u>

A periodic system matches the total purchases for the month with the total withdrawals of inventory units for the month. Conversely, a perpetual system matches each unit withdrawn with the immediately preceding purchases.

Summary

Inventory System	FIFO COGS	LIFO COGS	FIFO Inventory	LIFO Inventory
Periodic	\$23	\$31	\$15	\$7
Perpetual	\$23	\$26	\$15	\$12

Notice the relationship of higher COGS under LIFO and lower ending inventory under LIFO (assuming inflation) still holds whether the firm uses a periodic or perpetual inventory system. The point of this example is that under a perpetual system, LIFO COGS and ending inventory will differ from those calculated under a periodic system.

LOS 22.d: Calculate and explain how inflation and deflation of inventory costs affect the financial statements and ratios of companies that use different inventory valuation methods.

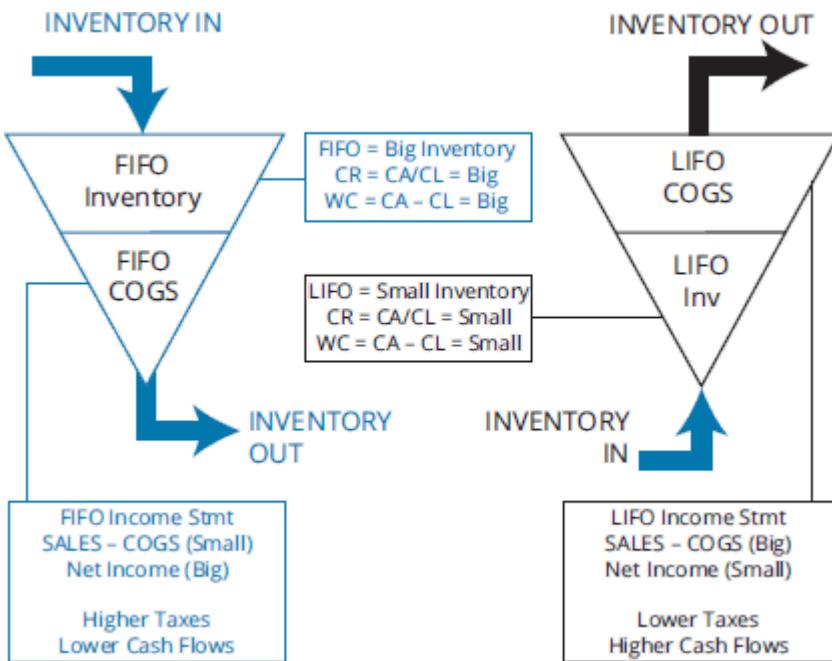
During inflationary periods and with stable or increasing inventory quantities, LIFO COGS is higher than FIFO COGS. This is because the last units purchased have a higher cost than the first units purchased. Under LIFO, the more costly last units purchased are assumed to be the first units sold (to COGS). Of course, higher COGS under LIFO will result in lower gross profit and net income compared to FIFO.

Using similar logic, we can see that LIFO ending inventory is lower than FIFO ending inventory because under LIFO, ending inventory is valued using older, lower costs.

During deflationary periods and stable or increasing inventory quantities, the cost flow effects of using LIFO and FIFO will be reversed; that is, LIFO COGS will be lower and LIFO ending inventory will be higher. This makes sense because the most recent lower-cost purchases are assumed to be sold first under LIFO, and the units in ending inventory are assumed to be the earliest purchases with higher costs.

Consider the diagram in Figure 22.2 to help visualize the FIFO-LIFO difference during periods of rising prices and growing inventory levels.

Figure 22.2: LIFO and FIFO Diagram—Rising Prices and Growing Inventory Balances



Remember, it's not the older or newer physical inventory units that are reported in the income statement and balance sheet; rather, it is the *costs* that are assigned to the units sold and to the units remaining in inventory.

PROFESSOR'S NOTE

Be able to describe the effects of LIFO and FIFO, assuming inflation, in your sleep. When prices are falling, the effects are simply reversed. When you are finished with this review, take the time to look at these graphs and relationships again to solidify the concepts in your mind.

During periods of stable prices, all three cost flow methods will yield the same results for inventory, COGS, and gross profit. During periods of trending prices (up or down), different cost flow methods may result in significant differences in these items.

PROFESSOR'S NOTE

The presumption in this section is that inventory quantities are stable or increasing.

Ending inventory. When prices are rising or falling, FIFO provides the most useful measure of ending inventory. This is a critical point. Recall that FIFO inventory is made up of the most recent purchases. These purchase costs can be viewed as a better approximation of current cost, and thus a better approximation of economic value. LIFO inventory, by contrast, is based on older costs that may differ significantly from current economic value.

Cost of goods sold. Changing prices can also produce significant differences between COGS under LIFO and FIFO. Recall that LIFO COGS is based on the most recent purchases. As a result, when prices are rising, LIFO COGS will be higher than FIFO COGS. When prices are falling, LIFO COGS will be lower than FIFO COGS. Because LIFO COGS is based on the most recent purchases, LIFO produces a better approximation of current cost in the income statement.

When prices are changing, the weighted average cost method will produce values of COGS and ending inventory between those of FIFO and LIFO.

Gross profit. Because COGS is subtracted from revenue in calculating gross profit, gross profit is also affected by the choice of cost flow method. Assuming inflation, higher COGS under LIFO will result in lower gross profit. In fact, all profitability measures (gross profit, operating profit, income before taxes, and net income) will be affected by the choice of cost flow method.

Figure 22.3: Effects of Inventory Valuation Methods

	FIFO	LIFO
Cost of sales	Lower	Higher
Ending inventory	Higher	Lower
Gross profit	Higher	Lower

Note: Assumes increasing prices and stable or increasing inventory levels.

MODULE QUIZ 22.2

1. A firm's purchases and sales during a period occur in the following order:

Beginning inventory	3 units @ \$390 per unit
Purchase	7 units @ \$385 per unit
Sale	5 units
Purchase	4 units @ \$380 per unit
Sale	8 units
Purchase	5 units @ \$370 per unit

Using LIFO and a perpetual inventory system, the firm's cost of sales for the period is:

- \$4,605.
 - \$4,995.
 - \$5,145.
2. In a period of falling prices, a firm reporting under LIFO, compared to reporting under FIFO, will have a higher:
 - cost of sales.
 - gross profit margin.
 - inventory turnover ratio.
 3. Compared to reporting under FIFO for both tax and financial statements, a firm that chooses to report under LIFO during a period of falling prices would be *most likely* to report a lower:
 - inventory.
 - gross profit.
 - cash balance.

MODULE 22.3: CONVERTING LIFO TO FIFO

LOS 22.e: Explain LIFO reserve and LIFO liquidation and their effects on financial statements and ratios.

Video covering this content is available online.

LOS 22.f: Demonstrate the conversion of a company's reported financial statements from LIFO to FIFO for purposes of comparison.

When prices are rising, firms that report inventory under LIFO will report lower inventory values and higher cost of goods sold than firms that report under FIFO. Because these

differences affect key ratios used to evaluate and compare companies, analysts may adjust the financial statements of LIFO firms so they can be compared to those of FIFO firms.

These four relations hold when prices have been rising over the relevant period:

1. LIFO inventory < FIFO inventory.
2. LIFO COGS > FIFO COGS.
3. LIFO net income < FIFO net income.
4. LIFO tax < FIFO tax.

Firms that report under LIFO must also report a **LIFO reserve**, the amount by which LIFO inventory is less than FIFO inventory. To make financial statements prepared under LIFO comparable to those of FIFO firms, an analyst must:

1. add the LIFO reserve to LIFO inventory on the balance sheet.
2. increase the retained earnings component of shareholders' equity by the LIFO reserve.

When prices are increasing, a LIFO firm will pay less in taxes than it would pay under FIFO. For this reason, analysts often decrease a LIFO firm's cash by the tax rate times the LIFO reserve and increase its retained earnings by the LIFO reserve times $(1 - \text{tax rate})$ instead of the full LIFO reserve.

For example, consider a firm with a LIFO reserve of \$150 that faces a tax rate of 40%. To convert the balance sheet to FIFO, increase inventory by \$150, decrease cash by \$60 ($\$150 \times 40\%$), and increase stockholders' equity (retained earnings) by \$90 [$\$150 \times (1 - 40\%)$]. This will bring the accounting equation back into balance. The net effect of the adjustments is an increase in assets and shareholders' equity of \$90, which is equal to the LIFO reserve net of tax.

For comparison purposes, it is also necessary to convert the LIFO firm's COGS to FIFO COGS. The difference between LIFO COGS and FIFO COGS is equal to the *change* in the LIFO reserve for the period. To convert COGS from LIFO to FIFO, simply subtract the change in the LIFO reserve:

$$\text{FIFO COGS} = \text{LIFO COGS} - (\text{ending LIFO reserve} - \text{beginning LIFO reserve})$$

Assuming inflation, FIFO COGS is lower than LIFO COGS, so subtracting the change in the LIFO reserve from LIFO COGS makes intuitive sense. When prices are falling, we still subtract the change in the LIFO reserve to convert from LIFO COGS to FIFO COGS. In this case, however, the change in the LIFO reserve is negative and subtracting it will result in higher COGS. This again makes sense. When prices are falling, FIFO COGS is greater than LIFO COGS.

EXAMPLE: Converting financial statements from LIFO to FIFO

Witz Company reported that at the end of last year, LIFO inventory was \$14,000, cost of goods sold was \$40,000, and net income was \$2,400. Witz's LIFO reserve was \$8,000 at the beginning of the year and \$10,000 at year-end. The company's tax rate was 40%. Determine the effects on the balance sheet and income statement had the FIFO inventory costing method been used, including an adjustment for taxes.

Answer:

Adjustments to the balance sheet

Ending inventory under FIFO would be higher by the amount of the LIFO reserve.

$$\text{FIFO ending inventory} = \text{LIFO ending inventory} + \text{LIFO reserve} = \$14,000 + \$10,000 = \$24,000$$

If FIFO had been used, cash would be lower (because taxes would have been cumulatively higher) by the LIFO reserve times the tax rate, or $\$10,000 \times 40\% = \$4,000$.

With inventories \$10,000 higher and cash \$4,000 lower, total assets would be \$6,000 higher if FIFO had been used. This would be balanced by a \$6,000 increase in shareholders' equity. This adjustment is equal to the LIFO reserve times (1 – tax rate).

Adjustments to the income statement

Cost of goods sold under FIFO would have been lower by the change in the LIFO reserve.

$$\text{FIFO cost of sales} = \text{LIFO cost of sales} - (\text{ending LIFO reserve} - \text{beginning LIFO reserve}) = \$40,000 - (\$10,000 - \$8,000) = \$38,000$$

Had COGS been \$2,000 less, pretax income would have been higher by \$2,000, so taxes would have been higher by $\$2,000 \times 40\% = \800 .

As a result, net income would have been higher by the change in the LIFO reserve net of tax.

$$\text{FIFO net income} = \text{LIFO net income} + (1 - \text{tax rate})(\text{ending LIFO reserve} - \text{beginning LIFO reserve}) = \$2,400 + (\$2,000)(1 - 0.40) = \$3,600$$



PROFESSOR'S NOTE

Later in this reading, we present a more comprehensive example of converting financial statements from LIFO to FIFO for comparison with other companies.

Effects on Ratios

Again assuming increasing prices, we can trace the ratio effects of the adjustments of LIFO values to FIFO values.

Profitability. As compared to FIFO, LIFO produces higher COGS in the income statement and results in lower earnings. Any profitability measure that includes COGS will be higher under FIFO. For example, reducing COGS will result in higher gross, operating, and net profit margins as compared to LIFO.

Liquidity. Compared to FIFO, LIFO results in a lower inventory value on the balance sheet. Because inventory (a current asset) is higher under FIFO, the current ratio, a popular measure of liquidity, is also higher under FIFO. Working capital is higher under FIFO as well, because current assets are higher.

Activity. Inventory turnover (COGS / average inventory) is higher for firms that use LIFO compared to firms that use FIFO. Under LIFO, COGS is valued at more recent, higher costs (higher numerator), while inventory is valued at older, lower costs (lower denominator). Adjusting to FIFO values will result in lower turnover and higher days of inventory on hand (365 / inventory turnover).

Solvency. Adjusting to FIFO results in higher total assets because inventory is higher. Higher total assets under FIFO result in higher stockholders' equity (assets – liabilities). Because total assets and stockholders' equity are higher under FIFO, the debt ratio and the debt-to-equity ratio are lower under FIFO compared to LIFO.

LIFO Liquidation

Recall that the LIFO reserve is equal to the difference between LIFO inventory and FIFO inventory. The LIFO reserve will increase when prices are rising and inventory quantities are

stable or increasing. If a firm is liquidating its inventory, or if prices are falling, the LIFO reserve will decline.

A **LIFO liquidation** occurs when a LIFO firm's inventory quantities decline. Older, lower costs are included in COGS compared to a situation in which inventory quantities are not declining. LIFO liquidation results in higher profit margins and higher income taxes compared to what they would be if inventory quantities were not declining. The extra profit reported with a LIFO liquidation inflates operating margins by recognizing historical inflationary gains from increasing inventory prices as income in the current period. Increases in profit margins from LIFO liquidation are not sustainable, however, because a firm cannot continue forever to sell existing inventory without replenishment.

Management could use a LIFO liquidation (draw down inventory) to artificially inflate current period earnings. Inventory declines can also be caused by events outside management's control, such as strikes or materials shortages at a key supplier that make inventory reduction involuntary, or a decline in expected customer orders that results in a voluntary reduction in inventory to suit market conditions.

Analysts must look to the LIFO reserve disclosures in the footnotes to see if the LIFO reserve has decreased over the period, which would indicate the possibility of a LIFO liquidation that requires adjustment of profit margins if its impact has been significant.

Regardless of the underlying reason for a LIFO liquidation, the resulting decrease in COGS will increase gross profits, pretax income, and net income. Decreased cash expenses (from not producing inventory) will increase operating cash flow, although higher income taxes on higher earnings will partially offset this increase in cash flows.

The following example illustrates the calculation of the extra profits earned as a result of inventory liquidation for a firm that uses LIFO for inventory valuation.

EXAMPLE: LIFO liquidation

At the beginning of 20X7, Big 4 Manufacturing Company had 400 units of inventory as follows:

Year Purchased	Number of Units	Cost Per Unit	Total Cost
20X4	120	\$10	\$1,200
20X5	140	11	1,540
20X6	<u>140</u>	12	<u>1,680</u>
	400		\$4,420

Big 4 reports inventory under LIFO. Due to a strike, no units were produced during 20X7. During 20X7, Big 4 sold 280 units. In the absence of the strike, Big 4 would have had a cost of \$14 for each unit produced. Compute the extra profit that resulted from the inventory liquidation.

Answer:

Because of the LIFO liquidation, actual COGS was \$3,220 as follows:

	Units	Cost
Beginning inventory	400	\$4,420
+ Purchases	-0-	-0-
- Ending inventory	<u>120</u>	<u>1,200</u> (\$10 × 120 units)
= COGS (actual)	280	\$3,220

Had Big 4 produced 280 units during 20X7, COGS would have been $280 \times \$14 = \$3,920$.

Due to the LIFO liquidation, COGS was lower by \$700 ($\$3,920 - \$3,220$); thus, pretax profit was higher by \$700. The higher profit is unsustainable because Big 4 will need to produce units at the new higher cost in future periods.



MODULE QUIZ 22.3

1. In an inflationary environment, a LIFO liquidation will *most likely* result in an increase in:
 - A. inventory.
 - B. accounts payable.
 - C. operating profit margin.
2. Bangor Company discloses that its LIFO reserve was \$625,000 at the end of the previous year and \$675,000 at the end of the current year. For the current year, beginning inventory was \$2,350,000 and ending inventory was \$2,525,000. The firm's tax rate is 30%. What would Bangor's ending inventory have been using FIFO?
 - A. \$2,575,000.
 - B. \$2,997,500.
 - C. \$3,200,000.
3. A firm that uses LIFO for inventory accounting reported COGS of \$300,000 and ending inventory of \$200,000 for the current period, and a LIFO reserve that decreased from \$40,000 to \$35,000 over the period. If the firm had reported using FIFO, its gross profit would have been:
 - A. the same.
 - B. \$5,000 higher.
 - C. \$5,000 lower.

MODULE 22.4: INVENTORY VALUATION



LOS 22.g: Describe the measurement of inventory at the lower of cost and net realisable value.

Video covering this content is available online.

Under IFRS, inventory is reported on the balance sheet at the lower of cost or net realizable value. **Net realizable value (NRV)** is equal to the expected sales price less the estimated selling costs and completion costs. If net realizable value is less than the balance sheet value of inventory, the inventory is “written down” to net realizable value and the loss is recognized in the income statement. If there is a subsequent recovery in value, the inventory can be “written up” and the gain is recognized in the income statement by reducing COGS by the amount of the recovery. Because inventory is valued at the lower of cost or net realizable value, inventory cannot be written up by more than it was previously written down.



PROFESSOR'S NOTE

The write-down, or subsequent write-up, of inventory is usually accomplished through the use of a valuation allowance account. A valuation allowance account is a contra-asset account, similar to accumulated depreciation. By using a valuation allowance account, the firm is able to separate the original cost of inventory from the carrying value of the inventory.

Under U.S. GAAP, companies that use inventory cost methods other than LIFO or the retail method report inventories at the lower of cost or NRV. For companies using LIFO or the retail method, inventory is reported on the balance sheet at the **lower of cost or market**. Market is usually equal to replacement cost, but cannot be greater than NRV or less than NRV minus a

normal profit margin. If replacement cost exceeds NRV, then market is NRV. If replacement cost is less than NRV minus a normal profit margin, then market is NRV minus a normal profit margin.



PROFESSOR'S NOTE

Think of lower of cost or market, where "market" cannot be outside a range of values. The range is from net realizable value minus a normal profit margin, to net realizable value. So the size of the range is the normal profit margin. "Net" means sales price less selling and completion costs.

If cost exceeds market, the inventory is written down to market on the balance sheet. The decrease in value is recognized in the income statement by increasing COGS for relatively small changes in value or by recording the loss from the inventory write-down separately for a relatively large change in value. The market value becomes the new cost basis.

If there is a subsequent recovery in value, no write-up is allowed under U.S. GAAP. This applies to companies using lower of cost or NRV as well as those using lower of cost or market.

EXAMPLE: Inventory write-down

Zoom, Inc., sells digital cameras. Per-unit cost information pertaining to Zoom's inventory is as follows:

Original cost	\$210
Estimated selling price	\$225
Estimated selling costs	\$22
Net realizable value	\$203
Replacement cost	\$197
Normal profit margin	\$12

What are the per-unit carrying values of Zoom's inventory using *lower of cost or NRV* and *lower of cost or market*?

Answer:

Using the lower of cost or net realizable value, because original cost of \$210 exceeds net realizable value ($\$225 - \$22 = \$203$), the inventory is written down to the net realizable value of \$203 and the \$7 decrease in value ($\$203 \text{ net realizable value} - \$210 \text{ original cost}$) is reported in the income statement.

Using the lower of cost or market, market is equal to replacement cost of \$197, since net realizable value of \$203 is greater than replacement cost, and net realizable value minus a normal profit margin ($\$203 - \$12 = \$191$) is less than replacement cost. Since original cost exceeds market (replacement cost), the inventory is written down to \$197 and a \$13 loss ($\$197 \text{ replacement cost} - \$210 \text{ original cost}$) is reported in the income statement.

EXAMPLE: Inventory write-up

Assume that in the year after the write-down in the previous example, net realizable value and replacement cost both increase by \$10. What is the impact of the recovery under IFRS, and under U.S. GAAP if lower of cost or market is used?

Answer:

Under IFRS, Zoom will write up inventory to \$210 per unit and recognize a \$7 gain in its income statement. The write-up (gain) is limited to the original write-down of \$7. The carrying value cannot exceed original cost.

Under U.S. GAAP, no write-up is allowed. The per-unit carrying value will remain at \$197. Zoom will simply recognize higher profit when the inventory is sold.

Recall that LIFO ending inventory is based on older, lower costs (assuming inflation) than under FIFO. Because cost is the basis for determining whether an impairment has occurred, LIFO firms are less likely to recognize inventory write-downs than firms using FIFO or weighted average cost.

Analysts must understand how an inventory write-down or write-up affects a firm's ratios. For example, a write-down may significantly affect inventory turnover in current and future periods. Thus, comparability of ratios across periods may be an issue.

In certain industries, reporting inventory above historical cost is permitted under IFRS and U.S. GAAP. This exception applies primarily to producers and dealers of commodity-like products, such as agricultural and forest products, mineral ores, and precious metals. Under this exception, inventory is reported at net realizable value and any unrealized gains and losses from changing market prices are recognized in the income statement. If an active market exists for the commodity, the quoted market price is used to value the inventory. Otherwise, recent market transactions are used.

LOS 22.h: Describe implications of valuing inventory at net realisable value for financial statements and ratios.

A write-down of inventory to net realizable value affects the financial statements and ratios in several ways. Assuming the write-down is reported as part of the cost of sales, these effects in the period of the write-down include:

- As inventory is part of current assets, an inventory write-down decreases both current and total assets.
- Current ratio (CA/CL) decreases. However, the quick ratio is unaffected because inventories are not included in the numerator of the quick ratio.
- Inventory turnover (COGS/average inventory) is increased, which decreases days' inventory on hand and the cash conversion cycle.
- The decrease in total assets increases total asset turnover and increases the debt-to-assets ratio.
- Equity is decreased, increasing the debt-to-equity ratio.
- The increase in COGS reduces gross margin, operating margin, and net margin.
- The percentage decrease in net income can be expected to be greater than the percentage decrease assets or equity. As a result, both ROA and ROE are decreased.

For periods subsequent to a write-down of inventory to net realizable value, COGS may be decreased by lower inventory carrying values, which will increase profitability. Together with the decreases in assets and equity from the prior-period write-down, an increase in net income from decreased COGS will increase reported ROA and ROE in subsequent periods.

LOS 22.i: Describe the financial statement presentation of and disclosures relating to inventories.

Inventory disclosures, usually found in the financial statement footnotes, are useful in evaluating the firm's inventory management. The disclosures are also useful in making adjustments to facilitate comparisons with other firms in the industry.

Required inventory disclosures are similar under U.S. GAAP and IFRS and include:

- The cost flow method (LIFO, FIFO, etc.) used.
- Total carrying value of inventory, with carrying value by classification (raw materials, work-in-process, and finished goods) if appropriate.
- Carrying value of inventories reported at fair value less selling costs.
- The cost of inventory recognized as an expense (COGS) during the period.
- Amount of inventory write-downs during the period.
- Reversals of inventory write-downs during the period, including a discussion of the circumstances of reversal (IFRS only because U.S. GAAP does not allow reversals).
- Carrying value of inventories pledged as collateral.

Inventory Changes

Although rare, a firm can change inventory cost flow methods. In most cases, the change is made retrospectively; that is, the prior years' financial statements are recast based on the new cost flow method. The cumulative effect of the change is reported as an adjustment to the beginning retained earnings of the earliest year presented.

Under IFRS, the firm must demonstrate that the change will provide reliable and more relevant information. Under U.S. GAAP, the firm must explain why the change in cost flow method is preferable.

An exception to retrospective application applies when a firm changes *to LIFO* from another cost flow method. In this case, the change is applied prospectively; no adjustments are made to the prior periods. With prospective application, the carrying value of inventory under the old method simply becomes the first layer of inventory under LIFO in the period of the change.

LOS 22.j: Explain issues that analysts should consider when examining a company's inventory disclosures and other sources of information.

Merchandising firms, such as wholesalers and retailers, purchase inventory that is ready for sale. In this case, inventory is reported in one account on the balance sheet. On the other hand, manufacturing firms normally report inventory using three separate accounts: raw materials, work-in-process, and finished goods. Analysts can use these disclosures, along with other sources of information such as Management's Discussion and Analysis, economic data specific to the industry, industry trade publications, and other sections of the firm's financial reports, as a signal of a firm's future revenues and earnings.

For example, an increase in raw materials and/or work-in-process inventory may be an indication of an expected increase in demand. Higher demand should result in higher revenues and earnings. Conversely, an increase in finished goods inventory, while raw materials and work-in-process are decreasing, may be an indication of decreasing demand and potential inventory write-downs in the future.

Analysts should also examine the relationship between sales and finished goods. Finished goods inventory growing faster than sales may indicate declining demand and excessive or potentially obsolete inventory. Obsolete inventory will result in lower earnings in the future when the inventory is written down. In addition, too much inventory is costly as the firm may incur

storage costs, insurance premiums, and inventory taxes. Too much inventory uses cash that might be more efficiently used somewhere else.

The inventory turnover ratio measures how quickly a firm is selling its inventory. Generally, high inventory turnover (low days of inventory on hand) is desirable. However, inventory turnover can be too high. A firm with an inventory turnover ratio that is too high may not be carrying enough inventory to satisfy customers' needs, which can cause the firm to lose sales. High inventory turnover may also indicate that inventory write-downs have occurred. Write-downs are usually the result of poor inventory management.

To further assess the explanation for high inventory turnover, we can look at inventory turnover relative to sales growth within the firm and industry. High turnover with slower growth may be an indication of inadequate inventory quantities. Alternatively, sales growth at or above the industry average supports the conclusion that high inventory turnover reflects greater efficiency.

MODULE 22.5: INVENTORY ANALYSIS



LOS 22.k: Calculate and compare ratios of companies, including companies that use different inventory methods.

Video covering this content is available online.

LOS 22.l: Analyze and compare the financial statements of companies, including companies that use different inventory methods.

In the following example, we calculate key ratios for a corporation that reports under LIFO and compare them to those that would be reported if the firm had used the FIFO cost flow method.

EXAMPLE: Comparing financial ratios under LIFO and FIFO

Viper Corp. is a high-performance bicycle manufacturer. Viper's balance sheets for 20X5 and 20X6 and an income statement for 20X6 are as shown. The balance sheets and income statement were prepared using LIFO. Calculate the current ratio, inventory turnover, long-term debt-to-equity ratio, gross profit margin, net profit margin, and return on assets ratio for 20X6 for both LIFO and FIFO inventory cost flow methods.

Viper Balance Sheet

(Prepared using LIFO)	20X6	20X5
Assets		
Cash	\$115	\$95
Receivables	205	195
Inventories	310	290
Total current assets	\$630	\$580
Gross plant and equipment	\$1,800	\$1,700
Accumulated depreciation	360	340
Net plant and equipment	\$1,440	\$1,360
Total assets	\$2,070	\$1,940
Liabilities and equity		
Payables	\$110	\$90
Short-term debt	215	185
Current liabilities	\$325	\$275
Long-term debt	715	785
Common stock	300	300
Additional paid-in-capital	400	400
Retained earnings	330	180
Total liabilities and equity	\$2,070	\$1,940

Viper Income Statement

(Prepared using LIFO)	20X6
Revenue	\$4,000
Cost of goods sold	3,000
Gross profit	\$1,000
Operating expenses	650
Operating profit	350
Interest expense	50
Earnings before tax	300
Taxes	60
Net income	240
Common dividends	\$90

Inventory footnote: The company uses the LIFO inventory cost flow method. Had FIFO been used, inventories would have been \$100 higher in 20X6 and \$90 higher in 20X5.

Income tax footnote: The tax rate for 20X5 and 20X6 was 20%.

Answer:

The analyst would recast the financial statements assuming FIFO for comparison purposes as follows:

Viper Balance Sheet

(Adjusted from LIFO to FIFO)	20X6	20X5
Assets		
Cash ¹	\$95	\$77
Receivables	205	195
Inventories ²	410	380
Total current assets	\$710	\$652
Gross plant and equipment	\$1,800	\$1,700
Accumulated depreciation	360	340
Net plant and equipment	\$1,440	\$1,360
Total assets	\$2,150	\$2,012
Liabilities and equity		
Payables	\$110	\$90
Short-term debt	215	185
Current liabilities	\$325	\$275
Long-term debt	715	785
Common stock	300	300
Additional paid-in-capital	400	400
Retained earnings ³	410	252
Total liabilities and equity	\$2,150	\$2,012

¹ Subtract taxes on LIFO reserve of \$20 and \$18 for 20X6 and 20X5, respectively.

² Add LIFO reserve of \$100 and \$90 for 20X6 and 20X5, respectively.

³ Add LIFO reserve (net of tax) of \$80 and \$72 for 20X6 and 20X5, respectively.

Viper Income Statement

(Adjusted from LIFO to FIFO)	20X6
Revenue	\$4,000
Cost of goods sold ⁴	2,990
Gross profit	\$1,010
Operating expenses	650
Operating profit	360
Interest expense	50
Earnings before tax	310
Taxes ⁵	62
Net income	248
Common dividends	\$90

⁴ Subtract \$10 change in reserve for 20X6.

⁵ Add \$2 taxes on change in the reserve for 20X6.

Current ratio

The current ratio (current assets / current liabilities) under LIFO is \$630 / \$325 = 1.9.

To convert to FIFO, the 20X6 LIFO reserve of \$100 is added to current assets (inventory) and income taxes on the LIFO reserve of \$20 are subtracted from cash.

Thus, under FIFO, the current ratio is (\$630 + \$100 LIFO reserve - \$20 taxes) / \$325 = 2.2. The current ratio is higher under FIFO as ending inventory now approximates current cost.

Inventory turnover

The inventory turnover ratio (COGS / average inventory) under LIFO is $\$3,000 / \$300 = 10.0$.

To convert to FIFO COGS, it is necessary to subtract the change in the LIFO reserve from LIFO COGS. The change in the LIFO reserve is $\$100$ ending reserve – $\$90$ beginning reserve = $\$10$.

To convert LIFO average inventory to FIFO, the average LIFO reserve is added to average LIFO inventory: $(\$90 \text{ beginning reserve} + \$100 \text{ ending reserve}) / 2 = \95 . Alternatively, we can calculate average FIFO inventory by averaging the beginning and ending FIFO inventory: $(\$290 \text{ beginning LIFO inventory} + \$90 \text{ beginning LIFO reserve} + \$310 \text{ ending LIFO inventory} + \$100 \text{ ending LIFO reserve}) / 2 = \395 .

Thus, under FIFO, inventory turnover is $(\$3,000 - 10 \text{ change in LIFO reserve}) / (\$300 + \$95 \text{ average LIFO reserve}) = 7.6$. Inventory turnover is lower under FIFO due to higher average inventory in the denominator and lower COGS in the numerator (assuming inflation).

Long-term debt-to-equity ratio

The long-term debt-to-equity ratio (long-term debt / stockholders' equity) under LIFO is $\$715 / \$1,030 = 0.6942$.

To convert to FIFO, the 20X6 LIFO reserve, net of tax, is added to stockholders' equity. The adjustment to stockholders' equity is necessary to make the accounting equation balance. The 20X6 LIFO reserve of $\$100$ was added to inventory and $\$20$ of income taxes was subtracted from cash, so the difference of $\$80$ is added to stockholders' equity.

Thus, under FIFO, long-term debt-to-equity is $\$715 / (\$1,030 + \$80 \text{ ending LIFO reserve, net of tax}) = 0.6441$. Long-term debt-to-equity is lower under FIFO (assuming inflation) because stockholders' equity is higher, since it reflects the effects of bringing the LIFO reserve onto the balance sheet.

Gross profit margin

The gross profit margin (gross profit / revenue) under LIFO is $\$1,000 / \$4,000 = 25.0\%$.

To convert to FIFO gross profit margin, the $\$10$ change in the LIFO reserve is subtracted from LIFO COGS. Thus, under FIFO, gross profit margin is $(\$1,000 + \$10 \text{ change in LIFO reserve}) / \$4,000 = 25.3\%$. Gross profit margin is higher under FIFO because COGS is lower under FIFO.

Net profit margin

The net profit margin (net income / revenue) under LIFO is $\$240 / \$4,000 = 6.0\%$.

To convert to FIFO net profit margin, subtract the $\$10$ change in the LIFO reserve from LIFO COGS to get FIFO COGS and increase income taxes $\$2$ ($\$10$ increase in reserve \times 20% tax rate). The increase in income taxes is the result of applying the 20X6 tax rate to the increase in taxable profit (lower COGS).

Thus, under FIFO, net profit margin is $(\$240 + \$10 \text{ change in LIFO reserve} - \$2 \text{ taxes}) / \$4,000 = 6.2\%$. The net profit margin is greater under FIFO because COGS is less under FIFO (assuming inflation).



PROFESSOR'S NOTE

We did not recognize the entire tax effect of the 20X6 LIFO reserve in the 20X6 income statement. The change from LIFO to FIFO is handled retrospectively. In other words, had we been using FIFO all along, the resulting higher taxes would have already been recognized in the previous years' income statements.

Return on assets

Return on assets (net income/average assets) under LIFO is $\$240 / \$2,005 = 11.97\%$.

To convert to FIFO return on assets, LIFO net income is increased by the change in the LIFO reserve, net of tax. Thus, FIFO net income is equal to $\$240 + \$10 \text{ change in reserve} - \$2 \text{ taxes} = \$248$.

To convert LIFO average assets, add the beginning and ending LIFO reserves, net of tax, to total assets. Thus, FIFO average assets is equal to $(\$2,070 \text{ 20X6 assets} + \$80 \text{ 20X6 reserve, net of tax} + \$1,940 \text{ 20X5 assets} + \$72 \text{ 20X5 reserve, net of tax}) / 2 = \$2,081$.

Thus, the FIFO return on assets is $\$248 / \$2,081 = 11.92\%$. In this example, the increase in FIFO net income is roughly proportionate to the increase in FIFO average assets. This is not always the case.

For comparison purposes, the following table summarizes our findings. The results of Viper's peer group have been included for analytical purposes.

Figure 22.4: Ratio Comparison

Year Ended 20X6	Viper		Peer Group
	LIFO	FIFO	FIFO
Current ratio	1.9	2.2	1.7
Inventory turnover	10.0	7.6	9.8
Long-term debt-to-equity	0.7	0.6	0.6
Gross profit margin	25.0%	25.3%	32.1%
Net profit margin	6.0%	6.2%	6.8%
Return on assets	11.97%	11.92%	12.5%

Analysis of Viper's Ratios Relative to a Group of Industry Peers

Liquidity: The after-tax LIFO adjustment resulted in an increase in Viper's current ratio. The adjusted ratio exceeds the peer group, indicating greater liquidity. Because inventory is the largest component of Viper's current assets, additional analysis is needed.

Activity: Viper's adjusted inventory turnover declined as expected due to the decrease in COGS and the increase in average inventory. Adjusted inventory turnover is less than the peer group, which indicates that it takes Viper longer to sell its goods. In terms of inventory days (365 / inventory turnover), Viper has 48.0 days of inventory on hand while the peer group has 37.2 days on hand. Too much inventory is costly and can also be an indication of obsolescence.

Solvency: Viper's adjusted long-term debt-to-equity ratio of 0.6 is in line with the peer group.

Profitability: As expected, Viper's adjusted gross profit and net profit margin ratios increased because COGS is lower under FIFO. However, the adjusted margin ratios are significantly less than the peer group's ratios. Coupled with lower adjusted inventory turnover, Viper's lower gross profit margin may be an indication that Viper is reducing sales prices to move its inventory. This is another indication that some of Viper's inventory may be obsolete. As previously discussed, obsolete (impaired) inventory must be written down.



MODULE QUIZ 22.4, 22.5

- Kamp, Inc., sells specialized bicycle shoes. At year-end, due to a sudden increase in manufacturing costs, the replacement cost per pair of shoes is \$55. The original cost is \$43, and the current selling price is \$50. The normal profit margin is 10% of the selling price, and the selling costs are \$3 per pair. Using the lower of cost or market method under U.S. GAAP, which of the following amounts should each pair of shoes be reported on Kamp's year-end balance sheet?
 - \$42.
 - \$43.
 - \$47.
- Poulter Products reports under IFRS and wrote its inventory value down from cost of \$400,000 to net realizable value of \$380,000. The *most likely* financial statement effect of this change is:
 - an increase in cost of sales.
 - a decrease in depreciation charges.
 - a loss reported as other comprehensive income.
- Which of the following inventory disclosures would *least likely* be found in the footnotes of a firm following IFRS?
 - The amount of loss reversals, from previously written-down inventory, recognized during the period.

- B. The carrying value of inventories that collateralize a short-term loan.
 - C. The separate carrying values of raw materials, work-in-process, and finished goods computed under the LIFO cost flow method.
4. Paul Neimer calculates the following horizontal common-size inventory data for Redpine Manufacturing, Inc.:

	Year 1	Year 2	Year 3	Year 4
Sales	1.00	1.10	1.18	1.25
Inventories:				
Raw materials	1.00	1.09	1.07	1.04
Work in process	1.00	1.11	1.15	1.17
Finished goods	1.00	1.10	1.21	1.33

Based on these data, Neimer should *most likely* conclude that Redpine:

- A. has an increasing inventory turnover ratio.
 - B. anticipates declining demand for its products.
 - C. might be losing sales due to inadequate inventory.
5. Which of the following is *most likely* for a firm with high inventory turnover and lower sales growth than the industry average? The firm:
- A. is managing its inventory effectively.
 - B. may have obsolete inventory that requires a write-down.
 - C. may be losing sales by not carrying enough inventory.
6. During a period of increasing prices, compared to reporting under LIFO, a firm that reports using average cost for inventory will have a:
- A. lower gross margin.
 - B. higher current ratio.
 - C. higher asset turnover.

KEY CONCEPTS

LOS 22.a

Costs included in inventory on the balance sheet include purchase cost, conversion costs, and other costs necessary to bring the inventory to its present location and condition. All of these costs for inventory acquired or produced in the current period are added to beginning inventory value and then allocated either to cost of goods sold for the period or to ending inventory.

Period costs, such as abnormal waste, most storage costs, administrative costs, and selling costs, are expensed as incurred.

LOS 22.b

Inventory cost flow methods:

- FIFO: The cost of the first item purchased is the cost of the first item sold. Ending inventory is based on the cost of the most recent purchases, thereby approximating current cost.
- LIFO: The cost of the last item purchased is the cost of the first item sold. Ending inventory is based on the cost of the earliest items purchased. LIFO is prohibited under IFRS.

- Weighted average cost: COGS and inventory values are between their FIFO and LIFO values.
- Specific identification: Each unit sold is matched with the unit's actual cost.

LOS 22.c

Under LIFO, cost of sales reflects the most recent purchase or production costs, and balance sheet inventory values reflect older outdated costs.

Under FIFO, cost of sales reflects the oldest purchase or production costs for inventory, and balance sheet inventory values reflect the most recent costs.

Under the weighted average cost method, cost of sales and balance sheet inventory values are between those of LIFO and FIFO.

When purchase or production costs are rising, LIFO cost of sales is higher than FIFO cost of sales, and LIFO gross profit is lower than FIFO gross profit as a result. LIFO inventory is lower than FIFO inventory.

When purchase or production costs are falling, LIFO cost of sales is lower than FIFO cost of sales, and LIFO gross profit is higher than FIFO gross profit as a result. LIFO inventory is higher than FIFO inventory.

In either case, LIFO cost of sales and FIFO inventory values better represent economic reality (replacement costs).

In a periodic system, inventory values and COGS are determined at the end of the accounting period. In a perpetual system, inventory values and COGS are updated continuously.

In the case of FIFO and specific identification, ending inventory values and COGS are the same whether a periodic or perpetual system is used. LIFO and weighted average cost, however, can produce different inventory values and COGS depending on whether a periodic or perpetual system is used.

LOS 22.d

When prices are *increasing* and inventory quantities are stable or increasing:

LIFO results in:	FIFO results in:
higher COGS	lower COGS
lower gross profit	higher gross profit
lower inventory balances	higher inventory balances
higher inventory turnover	lower inventory turnover

When prices are *decreasing* and inventory quantities are stable or increasing:

LIFO results in:	FIFO results in:
lower COGS	higher COGS
higher gross profit	lower gross profit
higher inventory balances	lower inventory balances
lower inventory turnover	higher inventory turnover

The weighted average cost method results in values between those of LIFO and FIFO if prices are increasing or decreasing.

LOS 22.e

A firm that reports under LIFO must disclose a LIFO reserve, which is the difference between LIFO inventory reported and inventory had the firm used the FIFO method. LIFO reserve will be positive during periods of rising inventory costs and negative during periods of falling inventory costs.

A LIFO liquidation occurs when a firm using LIFO sells more inventory during a period than it produces. During periods of rising prices, this drawdown in inventory reduces cost of goods sold because the lower cost of previously produced inventory is used, resulting in an unsustainable increase in gross profit margin.

LOS 22.f

To convert a firm's financial statements from LIFO to what they would have been under FIFO:

1. Add the LIFO reserve to LIFO inventory.
2. Subtract the change in the LIFO reserve for the period from COGS.
3. Decrease cash by LIFO reserve \times tax rate.
4. Increase retained earnings (equity) by LIFO reserve \times (1 – tax rate).

LOS 22.g

Under IFRS, inventories are valued at the lower of cost or net realizable value. Inventory write-ups are allowed, but only to the extent that a previous write-down to net realizable value was recorded.

Under U.S. GAAP, inventories are valued at the lower of cost or net realizable value for companies using cost methods other than LIFO or the retail method. For companies using LIFO or the retail method, inventories are valued at the lower of cost or market. Market is usually equal to replacement cost but cannot exceed net realizable value or be less than net realizable value minus a normal profit margin. No subsequent write-up is allowed for any company reporting under U.S. GAAP.

LOS 22.h

A write-down of inventory value from cost to net realizable value will:

- Decrease inventory, assets, and equity.
- Increase asset turnover, the debt-to-equity ratio and the debt-to-assets ratio.
- Result in a loss on the income statement, which will decrease net income and the net profit margin, as well as ROA and ROE for a typical firm.

LOS 22.i

Required inventory disclosures:

- The cost flow method (LIFO, FIFO, etc.) used.
- Total carrying value of inventory and carrying value by classification (raw materials, work-in-process, and finished goods) if appropriate.

- Carrying value of inventories reported at fair value less selling costs.
- The cost of inventory recognized as an expense (COGS) during the period.
- Amount of inventory write-downs during the period.
- Reversals of inventory write-downs during the period (IFRS only because U.S. GAAP does not allow reversals).
- Carrying value of inventories pledged as collateral.

LOS 22.j

An analyst should examine inventory disclosures to determine whether:

- The finished goods category is growing while raw materials and goods in process are declining, which may indicate decreasing demand and potential future inventory write-downs.
- Raw materials and goods in process are increasing, which may indicate increasing future demand and higher earnings.
- Increases in finished goods are greater than increases in sales, which may indicate decreasing demand or inventory obsolescence and potential future inventory write-downs.

LOS 22.k

Inventory turnover, days of inventory on hand, and gross profit margin can be used to evaluate the quality of a firm's inventory management.

- Inventory turnover that is too low (high days of inventory on hand) may be an indication of slow-moving or obsolete inventory.
- High inventory turnover together with low sales growth relative to the industry may indicate inadequate inventory levels and lost sales because customer orders could not be fulfilled.
- High inventory turnover together with high sales growth relative to the industry average suggests that high inventory turnover reflects greater efficiency rather than inadequate inventory.

LOS 22.l

Comparison of company financial statements may require statements to be adjusted to reflect the same inventory costing methods for both firms, or for the subject firm and any industry or peer group of firms used for comparison.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 22.1

- B** Variable production overhead is capitalized as a part of inventory. Storage costs not related to the production process, and selling and administrative costs are expensed as incurred. (LOS 22.a)
- A** Under FIFO, ending inventory is made up of the most recent purchases, thereby providing a closer approximation of current cost. (LOS 22.b)

3. A

FIFO COGS

$$\begin{aligned} 200 \text{ units from beginning inventory} \times \$3.00 &= \$600 \\ 400 \text{ units from 1st quarter} \times \$3.30 &= \underline{\$1,320} \\ &\quad \$1,920 \end{aligned}$$

LIFO COGS

$$\begin{aligned} 50 \text{ units from 4th quarter} \times \$4.20 &= \$210 \\ 200 \text{ units from 3rd quarter} \times \$3.90 &= 780 \\ 100 \text{ units from 2nd quarter} \times \$3.60 &= 360 \\ 250 \text{ units from 1st quarter} \times \$3.30 &= \underline{\$825} \\ &\quad \$2,175 \end{aligned}$$

Note the shortcut. Once FIFO COGS of \$1,920 is calculated, look at the LIFO column. We know that during inflation and stable or increasing inventory quantities, LIFO COGS is higher than FIFO. Only LIFO COGS of \$2,175 meets this condition. (LOS 22.c)

4. C Under FIFO, the first units purchased are the first units sold. FIFO COGS is the same under a periodic system and a perpetual system.

Revenue	\$484	(40 units × \$6.00) + (40 units × \$6.10)
COGS	<u>\$331</u>	(25 units × \$4.00) + (55 units × \$4.20)
Gross profit	\$153	

(LOS 22.c)

(LOS 22.c)

5. A Compared to FIFO, COGS calculated under LIFO will be higher because the most recent, higher cost units are assumed to be the first units sold. Higher COGS under LIFO will result in lower gross profit (revenue – COGS). (LOS 22.c)

Module Quiz 22.2

1. B The cost of the first five units sold is \$385 per unit. Of the next eight units sold, the cost of four of them is \$380 per unit, two have a cost of \$385 per unit, and two have a cost of \$390 per unit. Cost of sales = $5 \times \$385 + 4 \times \$380 + 2 \times \$385 + 2 \times \$390 = \$4,995$. (LOS 22.c)
2. B With falling prices, LIFO COGS will include the cost of lower priced inventory and COGS will be less when compared to FIFO COGS. Because of this, the firm would report a higher gross profit margin under LIFO than under FIFO, while LIFO inventory will be higher and inventory turnover lower. (LOS 22.d)
3. C When prices are falling, LIFO would result in lower COGS (higher gross profit) and higher ending inventory than FIFO. Higher gross profit with LIFO would result in higher taxes payable which would reduce cash balances (as long they pay their taxes). (LOS 22.d)

Module Quiz 22.3

1. **C** In a LIFO liquidation, older and lower costs are included in cost of sales. Thus, cost of sales per unit decreases and profit margins increase. (LOS 22.e)
2. **C** FIFO inventory = LIFO inventory + LIFO reserve = \$2,525,000 + \$675,000 = \$3,200,000. (LOS 22.f)
3. **C** FIFO COGS = LIFO COGS - (ending LIFO reserve - beginning LIFO reserve)

Ending LIFO reserve - beginning LIFO reserve = \$35,000 - \$40,000 = -\$5,000

With FIFO COGS \$5,000 greater than LIFO COGS, gross profit under FIFO would be \$5,000 lower than under LIFO. (LOS 22.f)

Module Quiz 22.4, 22.5

1. **B** Market is equal to the replacement cost as long as replacement cost is within a specific range. The upper bound is net realizable value (NRV) which is equal to the selling price (\$50) less selling costs (\$3) for a NRV of \$47. The lower bound is NRV (\$47) less normal profit margin (10% of selling price = \$5) for a net amount of \$42. Because replacement cost is greater than NRV (\$47), market equals NRV (\$47). Additionally, we have to use the lower of cost (\$43) or market (\$47) principle, so the shoes should be recorded at a cost of \$43. (Module 22.4, LOS 22.g)
2. **A** The write-down in inventory value from cost to net realizable value is reported on the income statement either as an addition to cost of sales or as a separate line item, not as other comprehensive income. Depreciation will not be affected as inventory is not depreciated. (Module 22.4, LOS 22.h)
3. **C** While the separate carrying values of raw materials, work-in-process, and finished goods are required disclosure for some firms, LIFO is not permitted under IFRS. (Module 22.4, LOS 22.i)
4. **B** Redpine's finished goods inventory is growing faster than sales, while work-in-process inventory is growing more slowly than sales and raw materials inventory is decreasing. These data are consistent with Redpine reducing production in response to decreasing demand. Inventory turnover ratios cannot be calculated directly from the common-size data given, but finished goods inventory increasing faster than sales suggests inventory turnover is likely decreasing. (Module 22.4, LOS 22.j)
5. **C** High inventory turnover coupled with low sales growth relative to the industry may be an indication of inadequate inventory levels. In this case, the firm may be losing sales by not carrying enough inventory. (Module 22.5, LOS 22.k)
6. **B** Compared to using LIFO, using average cost would produce lower COGS, higher gross operating income, and higher ending inventory, so current assets and the current ratio would be higher. Consequently, gross margin would be higher and asset turnover would be lower under the average cost inventory method. (Module 22.5, LOS 22.k, 22.l)

Reading 23

LONG-LIVED ASSETS

EXAM FOCUS

Long-lived assets include tangible assets, intangible assets, and financial assets. Firms make many estimates and choices in accounting for long-lived assets that affect the firms' profits, ratios, and cash flow classifications. You must understand the effects of and issues concerning capitalization versus immediate expensing of various costs, including construction interest, research and development, and software costs. For capitalized costs, you must be familiar with the effects of the different depreciation and amortization methods and be able to determine if an asset is impaired. Finally, you must be familiar with the revaluation (fair value) model under IFRS and the disclosure requirements for financial reporting.

MODULE 23.1: CAPITALIZATION VS. EXPENSING



LOS 23.a: Identify and contrast costs that are capitalised and costs that are expensed in the period in which they are incurred.

Video covering this content is available online.

When a firm makes an expenditure, it can either **capitalize** the cost as an asset on the balance sheet or **expense** the cost in the income statement in the period incurred. As a general rule, an expenditure that is expected to provide a future economic benefit over multiple accounting periods is capitalized; however, if the future economic benefit is unlikely or highly uncertain, the expenditure is expensed in the period incurred.

An expenditure that is capitalized is initially recorded as an asset on the balance sheet at cost, typically its fair value at acquisition plus any costs necessary to prepare the asset for use. Except for land and intangible assets with indefinite lives (such as acquisition goodwill), the cost is then allocated to the income statement over the life of the asset as **depreciation** expense (for tangible assets) or **amortization** expense (for intangible assets with finite lives).

Alternatively, if an expenditure is immediately expensed, current period pretax income is reduced by the amount of the expenditure.

Once an asset is capitalized, subsequent related expenditures that provide more future economic benefits (e.g., rebuilding the asset) are also capitalized. Subsequent expenditures that merely sustain the usefulness of the asset (e.g., regular maintenance) are expensed when incurred.

EXAMPLE: Capitalizing versus expensing

Northwood Corp. purchased new equipment to be used in its manufacturing plant. The cost of the equipment was \$250,000 including \$5,000 freight and \$12,000 of taxes. In addition to the equipment cost, Northwood paid \$10,000 to install the equipment and \$7,500 to train its employees to use the equipment. Over the asset's

life, Northwood paid \$35,000 for repair and maintenance. At the end of five years, Northwood extended the life of the asset by rebuilding the equipment's motors at a cost of \$85,000.

What amounts should be capitalized on Northwood's balance sheet and what amounts should be expensed in the period incurred?

Answer:

Northwood should capitalize all costs that provide future economic benefits, including the costs that are necessary to get the asset ready for use. Rebuilding the equipment's motors extended its life and thus increased its future benefits.

Capitalized Costs

Purchase price	\$250,000 (including freight & taxes)
Installation costs	10,000
Rebuilt motors	<u>85,000</u>
	\$345,000

Costs that do not provide future economic benefits are expensed in the period incurred. The initial training costs are not necessary to get the asset ready for use. Rather, the training costs are necessary to get the employees ready to use the asset. Thus, the training costs are immediately expensed. Repair and maintenance costs are operating expenditures that do not extend the life of the equipment.

Costs Expensed When Incurred

Initial training costs	\$7,500
Repair and maintenance	<u>35,000</u>
	\$42,500

Capitalized Interest

When a firm constructs an asset for its own use or, in limited circumstances, for resale, the interest that accrues during the construction period is capitalized as a part of the asset's cost. The reasons for capitalizing interest are to accurately measure the cost of the asset and to better match the cost with the revenues generated by the constructed asset. The treatment of construction interest is similar under U.S. GAAP and IFRS.

Capitalized interest is not reported in the income statement as interest expense. Once construction interest is capitalized, the interest cost is allocated to the income statement through depreciation expense (if the asset is held for use), or COGS (if the asset is held for sale).

Generally, capitalized interest is reported in the cash flow statement as an outflow from investing activities, while interest expense is reported as an outflow from operating activities under U.S. GAAP. Note, however, that interest expense can be an operating, financing, or investing cash flow under IFRS.

For an analyst, both capitalized and expensed interest should be used when calculating interest coverage ratios. Any depreciation of capitalized interest on the income statement should be added back when calculating income measures.

LOS 23.b: Compare the financial reporting of the following types of intangible assets: purchased, internally developed, acquired in a business combination.

Intangible assets are long-term assets that lack physical substance, such as patents, brand names, copyrights, and franchises. Some intangible assets have finite lives while others have indefinite lives.

The cost of a finite-lived intangible asset is amortized over its useful life. Indefinite-lived intangible assets are not amortized, but are tested for impairment at least annually. If impaired, the reduction in value is recognized in the income statement as a loss in the period in which the impairment is recognized.

Intangible assets are also considered either identifiable or unidentifiable. Under IFRS, an **identifiable intangible asset** must be:

- Capable of being separated from the firm or arise from a contractual or legal right.
- Controlled by the firm.
- Expected to provide future economic benefits.

In addition, the future economic benefits must be probable and the asset's cost must be reliably measurable.

An **unidentifiable intangible asset** is one that cannot be purchased separately and may have an indefinite life. The most common example of an unidentifiable intangible asset is goodwill. Goodwill is the excess of purchase price over the fair value of the identifiable assets (net of liabilities) acquired in a business combination.

Not all intangible assets are reported on the balance sheet. Accounting for an intangible asset depends on whether the asset was created internally, purchased externally, or obtained as part of a business combination.

Intangible Assets Created Internally

With some exceptions, costs to create intangible assets are expensed as incurred. Important exceptions are research and development costs (under IFRS) and software development costs.

Research and development costs. Under IFRS, **research costs**, which are costs aimed at the discovery of new scientific or technical knowledge and understanding, are expensed as incurred. However, **development costs** may be capitalized. Development costs are incurred to translate research findings into a plan or design of a new product or process. To recognize an intangible asset in development, a firm must show that it can complete the asset and intends to use or sell the completed asset, among other criteria.

Under U.S. GAAP, both research and development costs are generally expensed as incurred. However, the costs of creating software for sale to others are treated in a manner similar to the treatment of research and development costs under IFRS. Costs incurred to develop software for sale to others are expensed as incurred until the product's technological feasibility has been established, after which the costs of developing a salable product are capitalized.

Purchased Intangible Assets

Like tangible assets, an intangible asset purchased from another party is initially recorded on the balance sheet at cost, typically its fair value at acquisition.

If the intangible asset is purchased as part of a group, the total purchase price is allocated to each asset on the basis of its fair value. For analytical purposes, an analyst is usually more interested in the type of asset acquired rather than the value assigned on the balance sheet. For

example, recently acquired franchise rights may provide insight into the firm's future operating performance. In this case, the allocation of cost is not as important.

The financial statement effects of capitalizing intangible assets are the same as the effects of capitalizing other expenditures. Capitalizing results in higher net income in the first year and lower net income in the subsequent years. Similarly, assets, equity, and operating cash flow are all higher when expenditures are capitalized.

Intangible Assets Obtained in a Business Combination

The **acquisition method** is used to account for business combinations. Under the acquisition method, the purchase price is allocated to the identifiable assets and liabilities of the acquired firm on the basis of fair value. Any remaining amount of the purchase price is recorded as **goodwill**. Goodwill is said to be an unidentifiable asset that cannot be separated from the business itself.

Only goodwill created in a business combination is capitalized on the balance sheet. The costs of any internally generated "goodwill" are expensed in the period incurred.

LOS 23.c: Explain and evaluate how capitalising versus expensing costs in the period in which they are incurred affects financial statements and ratios.

Although it may make no operational difference, the choice between capitalizing costs and expensing them will affect net income, shareholders' equity, total assets, cash flow from operations, cash flow from investing, and numerous financial ratios.

Net Income

Capitalizing an expenditure delays the recognition of an expense in the income statement. Thus, in the period that an expenditure is capitalized, the firm will report higher net income compared to immediately expensing. In subsequent periods, the firm will report lower net income compared to expensing, as the capitalized expenditure is allocated to the income statement through depreciation expense. This allocation process reduces the variability of net income by spreading the expense over multiple periods.



PROFESSOR'S NOTE

For growing firms, capitalizing expenditures may result in earnings that are higher over many periods compared to an otherwise identical expensing firm. This is because the amount of depreciation from previously capitalized expenditures is less than the amount of additional costs that are being newly capitalized each period.

Conversely, if a firm expenses an expenditure in the current period, net income is reduced by the after-tax amount of the expenditure. In subsequent periods, no allocation of cost is necessary. Thus, net income in future periods is higher than if the expenditure had been capitalized.

Over the life of an asset, *total* net income is identical whether the asset's cost is capitalized or expensed. Timing of the expense recognition in the income statement is the only difference.

Shareholders' Equity

Because capitalization results in higher net income in the period of the expenditure compared to expensing, it also results in higher shareholders' equity because retained earnings are greater. Total assets are greater with capitalization and liabilities are unaffected, so the accounting equation ($A = L + E$) remains balanced. As the cost is allocated to the income statement in subsequent periods, net income, retained earnings, and shareholders' equity will be reduced.

If the expenditure is immediately expensed, retained earnings and shareholders' equity will reflect the entire reduction in net income in the period of the expenditure.

Cash Flow From Operations

A capitalized expenditure is usually reported in the cash flow statement as an outflow from investing activities. If immediately expensed, the expenditure is reported as an outflow from operating activities. Thus, capitalizing an expenditure will result in higher operating cash flow and lower investing cash flow compared to expensing. Assuming no differences in tax treatment, *total* cash flow will be the same. The classification of the cash flow is the only difference.

Recall that when an expenditure is capitalized, depreciation expense is recognized in subsequent periods. Depreciation is a noncash expense and, aside from any tax effects, does not affect operating cash flow.



PROFESSOR'S NOTE

If the tax treatment is changed to match the financial reporting treatment of the expenditure, expensing will result in higher operating cash flow in the first year because of the tax savings. However, if the tax treatment is independent of the financial reporting treatment, taxes, and therefore cash flows, are unaffected by the choice.

Financial Ratios

Capitalizing an expenditure initially results in higher assets and higher equity compared to expensing. Thus, both the debt-to-assets ratio and the debt-to-equity ratio are lower (they have larger denominators) with capitalization.

Capitalizing an expenditure will *initially* result in higher return on assets (ROA) and higher return on equity (ROE). This is the result of higher net income in the first year. In subsequent years, ROA and ROE will be lower for a capitalizing firm because net income is reduced by the depreciation expense.

Because an expensing firm recognizes the entire expense in the first year, ROA and ROE will be lower in the first year and higher in the subsequent years. After the first year, net income (numerator) is higher, and assets and equity (denominators) are lower, than they would be if the firm had capitalized the expenditure. Analysts must be careful when comparing firms because immediately expensing an expenditure gives the appearance of growth after the first year.

The interest coverage ratio (EBIT / interest expense) measures a firm's ability to make required interest payments on its debt. In the year of the expenditure, capitalizing interest results in lower interest expense compared to expensing. The result is a higher interest coverage ratio (smaller denominator) when interest is capitalized.

Many analysts calculate the interest coverage ratio based on total interest expense, including capitalized interest. Because the interest is a required payment, this may be a better measure of

the firm's solvency. Treating the capitalized interest as interest expense for analytical purposes reduces the interest coverage ratio. Bond rating agencies often make this adjustment.

The financial effects of capitalizing versus expensing are summarized in Figure 23.1.

Figure 23.1: Financial Statement Effects of Capitalizing vs. Expensing

Method	Assumption	Cost of Goods Sold Consists of...	Ending Inventory Consists of...
FIFO (U.S. and IFRS)	The items first purchased are the first to be sold.	first purchased	most recent purchases
LIFO (U.S. only)	The items last purchased are the first to be sold.	last purchased	earliest purchases
Weighted average cost (U.S. and IFRS)	Items sold are a mix of purchases.	average cost of all items	average cost of all items

MODULE QUIZ 23.1

-  If a company purchases an asset with future economic benefits that are highly uncertain, the company should:
 - expense the purchase.
 - use straight-line depreciation.
 - use an accelerated depreciation method.
- The cost of an intangible asset is *most likely* to be amortized if the asset has:
 - a finite life and was purchased.
 - a finite life and was created internally.
 - an indefinite life and was acquired in a business combination.
- Red Company immediately expenses its development costs while Black Company capitalizes its development costs. All else equal, Red Company will:
 - show smoother reported earnings than Black Company.
 - report higher operating cash flow than Black Company.
 - report higher asset turnover than Black Company.

MODULE 23.2: DEPRECIATION



LOS 23.d: Describe the different depreciation methods for property, plant, and equipment and calculate depreciation expense.

Video covering this content is available online.

Depreciation is the systematic allocation of an asset's cost over time. Two important terms are:

- **Carrying (book) value.** The net value of an asset or liability on the balance sheet. For property, plant, and equipment, carrying value equals historical cost minus accumulated depreciation.
- **Historical cost.** The original purchase price of the asset including installation and transportation costs. Historical cost is also known as *gross investment in the asset*.

Depreciation is a real and significant operating expense. Even though depreciation doesn't require current cash expenditures (the cash outflow was made in the past when the asset was purchased), it is an expense nonetheless and cannot be ignored.

The analyst must decide whether the reported depreciation expense is more or less than *economic depreciation*, which is the actual decline in the value of the asset over the period. One chain of video rental stores was found to be overstating income by depreciating its stock of movies by equal amounts each year. In fact, a greater portion of the decrease in the value of newly released movies occurs in the first year. Depreciating the rental assets by a greater amount during the first year would have better approximated economic depreciation and reduced reported income.

Depreciation Methods

Depreciation of a capitalized cost (asset) may be reported using straight-line, accelerated, or units-of-production methods.

Straight-line depreciation is the predominant method of computing depreciation for financial reporting. Depreciation is the same amount each year over the asset's estimated life:

$$\text{depreciation expense} = \frac{\text{original cost} - \text{salvage value}}{\text{depreciable life}}$$

EXAMPLE: Calculating straight-line depreciation expense

Littlefield Company recently purchased a machine at a cost of \$12,000. The machine is expected to have a residual value of \$2,000 at the end of its useful life in five years. Calculate depreciation expense using the straight-line method.

Answer:

The annual depreciation expense each year will be:

$$\frac{\text{cost} - \text{residual}}{\text{useful life}} = \frac{(\$12,000 - \$2,000)}{5} = \$2,000$$

With an **accelerated depreciation** method, more depreciation expense is recognized in the early years of an asset's life and less depreciation expense in the later years. Thus, accelerated depreciation results in lower net income in the early years of an asset's life and higher net income in the later years, compared to straight-line depreciation. One often-used accelerated depreciation method is the **double-declining balance** (DDB) method:

$$\text{DDB depreciation in year } x =$$

$$\frac{2}{\text{depreciable life in years}} \times \text{book value at beginning of year } x$$

Note that salvage value is not in the formula for double-declining balance depreciation. However, once the carrying (book) value of the asset reaches the salvage value, no additional depreciation expense is recognized.

EXAMPLE: Calculating double-declining balance depreciation expense

Littlefield Company recently purchased a machine at a cost of \$12,000. The machine is expected to have a residual value of \$2,000 at the end of its useful life in five years. Calculate depreciation expense for all five years using the double-declining balance method.

Answer:

The depreciation expense using the double declining balance method is:

- Year 1: $(2 / 5)(\$12,000) = \$4,800$
- Year 2: $(2 / 5)(\$12,000 - \$4,800) = \$2,880$
- Year 3: $(2 / 5)(\$12,000 - \$7,680) = \$1,728$

In years 1 through 3, the company has recognized cumulative depreciation expense of \$9,408. Since the total depreciation expense is limited to \$10,000 ($\$12,000 - \$2,000$ salvage value), the depreciation in year 4 is limited to \$592, rather than the $(2 / 5)(\$12,000 - \$9,408) = \$1,036.80$ using the DDB formula.

Year 5: Depreciation expense is \$0, since the asset is fully depreciated.

Note that the rate of depreciation is doubled ($2 / 5$) from straight-line, and the only thing that changes from year to year is the base amount (book value) used to calculate annual depreciation.



PROFESSOR'S NOTE

We've been discussing the "double" declining balance method, which uses a factor of two times the straight-line rate. Firms can compute declining balance depreciation based on any factor (e.g., 1.5, double, triple).

Depreciation under the **units-of-production method** is based on usage rather than time.

Depreciation expense is higher in periods of high usage.

units-of-production depreciation =

$$\frac{\text{original cost} - \text{salvage value}}{\text{life in output units}} \times \text{output units in the period}$$



PROFESSOR'S NOTE

The units-of-production method applied to natural resources is referred to as depletion.

Component Depreciation

IFRS requires firms to depreciate the components of an asset separately, thereby requiring useful life estimates for each component. For example, a building is made up of a roof, walls, flooring, electrical systems, plumbing, and many other components. Under **component depreciation**, the useful life of each component is estimated and depreciation expense is computed separately for each.

Component depreciation is allowed under U.S. GAAP but is seldom used.

EXAMPLE: Component depreciation

Global Airlines purchased a new airplane with an all-inclusive cost of \$50 million. The estimated life of the airplane is 30 years and the estimated salvage value is \$5 million. Global expects to replace the interior of the aircraft after 15 years. The component cost of the interior is estimated at \$3 million.

Calculate depreciation expense in year 1 using the straight-line method, both assuming the interior is a separate component and assuming the component method is not used.

Answer:

Straight-line depreciation using the component method:

Total aircraft cost	\$50,000,000
Interior cost	<u>(3,000,000)</u>
Aircraft component	\$47,000,000

Depreciation expense:

Aircraft component	\$1,400,000	$(\$47,000,000 - 5,000,000) / 30 \text{ years}$
Interior component	<u>200,000</u>	$(\$3,000,000 / 15 \text{ years})$
Year 1 expense	\$1,600,000	

Straight-line depreciation without the component method:

$$\text{Year 1 expense} \quad (\$50,000,000 - 5,000,000) / 30 \text{ years} = \$1,500,000$$

Depreciation expense is lower by \$100,000 each year (\$1,600,000 – \$1,500,000) for the first 15 years without the component method. However, at the end of year 15, Global will spend \$3,000,000 to replace the interior. Thus, additional depreciation expense of $\$3,000,000 / 15 \text{ years} = \$200,000$ each year is required for the last 15 years of the asset's life.

Under both scenarios, Global will have expended a total of \$53,000,000 and recognized \$48,000,000 of depreciation expense over the airplane's life:

$$\text{Component method} \quad \$1,600,000 \times 30 \text{ years} = \$48,000,000$$

$$\begin{aligned} \text{Non-component method} \quad & (\$1,500,000 \times 30 \text{ years}) + (\$200,000 \times 15 \text{ years}) \\ & = \$48,000,000 \end{aligned}$$

LOS 23.e: Describe how the choice of depreciation method and assumptions concerning useful life and residual value affect depreciation expense, financial statements, and ratios.

In the early years of an asset's life, compared to straight-line depreciation, using an accelerated depreciation method will result in higher depreciation expense and lower net income, total assets, and shareholders' equity. For a single long-lived asset, these effects reverse in the later years of its useful life. For a fast-growing firm, however, these effects will persist over time as long as the firm is acquiring more depreciable assets than it is derecognizing.

Return on assets and return on equity are higher with straight-line depreciation compared to accelerated methods. This is because the effect on the numerator (higher net income) is relatively larger than the effect on the denominators (higher assets and equity). Asset turnover ratios (revenue / average assets) are lower with straight-line depreciation. Assuming that the depreciation method for tax is unchanged, the choice of depreciation method for financial reporting will not affect cash flows.

Figure 23.2 summarizes the effects of depreciation methods on financial statements and ratios in the early years of an asset's useful life or for a fast-growing firm.

Figure 23.2: Effects of Depreciation Methods (Early Years or Fast-Growing Firm)

	Straight-Line	Accelerated
Depreciation expense	Lower	Higher
Net income	Higher	Lower
Total assets	Higher	Lower
Shareholders' equity	Higher	Lower
Return on assets	Higher	Lower
Return on equity	Higher	Lower
Asset turnover ratios	Lower	Higher
Cash flow	Same	Same

Useful Lives and Salvage Values

Calculating depreciation expense requires estimating an asset's useful life and its salvage (residual) value. Firms can manipulate depreciation expense, and therefore net income, by increasing or decreasing either of these estimates.

A longer estimated useful life decreases annual depreciation and increases reported net income, while a shorter estimated useful life will have the opposite effect. A higher estimate of the salvage value will also decrease depreciation and increase net income, while a lower estimate of the salvage value will increase depreciation and decrease net income.

A change in an accounting estimate, such as useful life or salvage value, is put into effect in the current period and prospectively. That is, the change in estimate is applied to the asset's carrying (book) value and depreciation is calculated going forward using the new estimate. The previous periods are not affected by the change.

EXAMPLE: Change in depreciation estimate

Alpine Company purchased machinery for \$20,000 with an estimated useful life of four years and a salvage value of \$4,000. Alpine uses the straight-line depreciation method. At the beginning of the third year, Alpine reduces its salvage value estimate to \$1,600. Determine the depreciation expense for each year.

Answer:

For the first two years, straight-line depreciation expense is $[(\$20,000 \text{ original cost} - \$4,000 \text{ salvage value}) / 4\text{-year life}] = \$4,000$ each year. At the beginning of the third year, the asset's carrying value on the balance sheet is $\$20,000 \text{ original cost} - \$8,000 \text{ accumulated depreciation} = \$12,000$.

To calculate straight-line depreciation expense for the remaining years, simply begin with the carrying value and depreciate over the remaining useful life using the new salvage value estimate. Depreciation expense for the last two years is $[(\$12,000 \text{ carrying value} - \$1,600 \text{ revised salvage value}) / 2 \text{ years remaining life}] = \$5,200$ each year.

Estimates are also involved when a manufacturing firm allocates depreciation expense between COGS and SG&A. While the allocation does not affect a firm's operating margin, it affects the firm's gross margin (which is computed before SG&A expense) and operating expenses.

LOS 23.g: Describe the different amortisation methods for intangible assets with finite lives and calculate amortisation expense.



PROFESSOR'S NOTE

We will address LOS 23.f later in this reading.

Intangible assets with finite lives are amortized over their useful lives. Amortization is identical to the depreciation of tangible assets. The same methods, straight-line, accelerated, and units-of-production, are permitted. The calculation of amortization expense also requires estimates of useful lives and salvage values. However, estimating useful lives is complicated by many legal, regulatory, contractual, competitive, and economic factors that may limit the use of the intangible assets.

As with depreciation, the total amount of amortization is the same under all of the methods. Timing of the amortization expense in the income statement is the only difference.

Another example of an intangible asset with an indefinite life is a trademark that may have a specific expiration date, but can be renewed at minimal cost. In this case, the trademark is considered to have an indefinite life and no amortization is required.

EXAMPLE: Calculating amortization expense

At the beginning of this year, Brandon Corporation entered into business acquisition. As a result of the acquisition, Brandon reported the following intangible assets:

Patent	\$480,000
Franchise agreement	\$350,000
Copyright	\$150,000
Goodwill	<u>\$550,000</u>
	\$1,530,000

The patent expires in 12 years. The franchise agreement expires in 7 years but can be renewed indefinitely at a minimal cost. The copyright is expected to be sold at the end of 20 years for \$30,000. Use the straight-line amortization method to calculate the total carrying value of Brandon's intangible assets at the end of the year.

Answer:

Goodwill is an indefinite-lived asset and is not amortized. Because the franchise agreement can be renewed indefinitely at minimal cost, it is also considered an indefinite-lived asset and is not amortized.

Using the straight-line method, amortization expense is \$46,000 as follows:

Patent	$\$40,000 = \$480,000 / 12 \text{ years}$
Copyright	$\underline{6,000} = (\$150,000 - 30,000) / 20 \text{ years}$
Amortization expense	\$46,000

Thus, the carrying value at the end of the first year is \$1,484,000 as follows:

Intangible assets, at cost	\$1,530,000
Accumulated amortization	<u>(46,000)</u>
Intangible assets, net	\$1,484,000

LOS 23.h: Describe how the choice of amortisation method and assumptions concerning useful life and residual value affect amortisation expense, financial statements, and ratios.

The choice of amortization method will affect expenses, assets, equity, and financial ratios in exactly the same way that the choice of depreciation method will. Just as with the depreciation of tangible assets, increasing either the estimate of an asset's useful life or the estimate of its

residual value will reduce annual amortization expense, which will increase net income, assets, ROE, and ROA for the typical firm.



MODULE QUIZ 23.2

1. East Company purchased a new truck at the beginning of this year for \$30,000. The truck has a useful life of eight years or 150,000 miles, and an estimated salvage value of \$3,000. If the truck is driven 16,500 miles this year, how much depreciation will East report under the double-declining balance (DDB) method and the units-of-production (UOP) method?

<u>DDB</u>	<u>UOP</u>
A. \$7,500	\$2,970
B. \$7,500	\$3,300
C. \$6,750	\$2,970
2. In the early years of an asset's life, a firm using the double-declining balance method, as compared to a firm using straight-line depreciation, will report lower:
 - A. depreciation expense.
 - B. operating cash flow.
 - C. retained earnings.
3. Which of the following statements about indefinite-lived intangible assets is *most accurate*?
 - A. They are amortized on a straight-line basis over a period not to exceed 40 years.
 - B. They are reported on the balance sheet indefinitely.
 - C. They never appear on the balance sheet unless they are internally developed.
4. At the beginning of this year, Fairweather Corp. incurred \$200,000 of research costs and \$100,000 of development costs to create a new patent. The patent is expected to have a useful life of 40 years with no salvage value. Calculate the carrying value of the patent at the end of this year, assuming Fairweather follows U.S. GAAP.
 - A. \$0.
 - B. \$97,500.
 - C. \$292,500.
5. Which of the following is *least likely* considered in determining the useful life an intangible asset?
 - A. Initial cost.
 - B. Legal, regulatory, or contractual provisions.
 - C. Provisions for renewal or extension.

MODULE 23.3: IMPAIRMENT AND REVALUATION



LOS 23.i: Describe the revaluation model.

Under U.S. GAAP, most long-lived assets are reported on the balance sheet at depreciated cost (original cost less accumulated depreciation and any impairment charges).

Video covering this content is available online.

There is no fair value alternative for asset reporting under U.S. GAAP. Under IFRS, most long-lived assets are also reported at depreciated cost (the *cost model*). IFRS provides an alternative, the **revaluation model**, that permits a long-lived asset to be reported at its fair value, as long as an active market exists for the asset so its fair value can be reliably (and somewhat objectively) estimated. Firms must choose the same treatment for similar assets (e.g., land and buildings) so they cannot revalue only specific assets that are more likely to increase than decrease in value. The revaluation model is rarely used in practice by IFRS reporting firms.

Under the revaluation model, an asset is carried at its depreciated cost, but at each revaluation date, the balance sheet value is adjusted to fair value. Between revaluation dates, depreciation is recorded for the asset. Revaluation to fair value must be done sufficiently often that the reported value is not significantly different from fair value. Whether a revaluation affects the income statement, rather than only affecting equity, depends on previous revaluations. Consider the following situations.

First Revaluation Date

When there have been no prior revaluations and fair value is less than the carrying value (cost minus accumulated depreciation as of the first revaluation date), a loss is recorded on the income statement, much like an impairment charge. If fair value at the first revaluation date is greater than the carrying value of the asset, the difference is recorded as **revaluation surplus**, a component of equity, so net income is not affected.

Subsequent Revaluation Dates

On a revaluation date after the first revaluation, if fair value is greater than the carrying value, a gain is first reported on the income statement to the extent it reverses any previously recorded loss from revaluation. If the revaluation gain is greater than prior losses reported in the income statement that have not been reversed, the excess is reported in the revaluation surplus account.

If fair value on a revaluation date after the first revaluation date is less than the carrying value, the difference first goes to reduce any existing balance in the revaluation surplus account. Any remaining difference in excess of the balance in the revaluation surplus account is reported on the income statement as a loss.

EXAMPLE: Revaluation model

On December 31, 20X1, Parsons PLC reports a carrying value for an asset class as £30 million and a fair value of £29 million and reports a revaluation loss on the income statement of £1 million.

On December 31, 20X2, Parsons determines that the fair market value of the asset class is £1.5 million above its carrying value. This upward revaluation will be reported as a £1 million gain on the income statement (to reverse the previously recorded loss on the income statement) and a £0.5 million increase in the revaluation surplus account that does not affect income, but increases equity directly.

On December 31, 20X3, Parsons determines that the fair market value of the asset class is £1 million below its carrying value. This loss first goes to reduce the existing revaluation surplus of £0.5 million, and the other £0.5 million of the downward revaluation is reported as a loss on the income statement.

LOS 23.j: Explain the impairment of property, plant, and equipment and intangible assets.

Both IFRS and U.S. GAAP require firms to write down impaired assets by recognizing a loss in the income statement. However, there are differences in applying the standards.



PROFESSOR'S NOTE

The following discussion applies to both tangible and intangible long-lived assets with finite lives that are held for use.

Impairments Under IFRS

Under IFRS, the firm must annually assess whether events or circumstances indicate an **impairment** of an asset's value has occurred. For example, there may have been a significant decline in the market value of the asset or a significant change in the asset's physical condition. If so, the asset's value must be tested for impairment.

An asset is impaired when its carrying value (original cost less accumulated depreciation) exceeds the **recoverable amount**. The recoverable amount is the greater of its fair value less any selling costs and its **value in use**. The value in use is the present value of its future cash flow stream from continued use.

If impaired, the asset's value must be written down on the balance sheet to the recoverable amount. An impairment loss, equal to the excess of carrying value over the recoverable amount, is recognized in the income statement.

Under IFRS, an impairment loss on an identifiable long-lived asset can be reversed if the asset's value recovers in the future. However, the loss reversal is limited to the original impairment loss.

Impairments Under U.S. GAAP

Under U.S. GAAP, an asset is tested for impairment only when events and circumstances indicate the firm may not be able to recover the carrying value through future use.

Determining an impairment and calculating the loss potentially involves two steps. In the first step, the asset is tested for impairment by applying a **recoverability test**. If the asset is impaired, the second step involves measuring the loss.

Recoverability. An asset is considered impaired if the carrying value (original cost less accumulated depreciation) is greater than the asset's future *undiscounted* cash flow stream. Because the recoverability test is based on estimates of future undiscounted cash flows, tests for impairment involve considerable management discretion.

Loss measurement. If impaired, the asset's value is written down to fair value on the balance sheet and a loss, equal to the excess of carrying value over the fair value of the asset (or the *discounted* value of its future cash flows if the fair value is not known), is recognized in the income statement.

Under U.S. GAAP, loss recoveries are typically not permitted.



PROFESSOR'S NOTE

The difference between testing for impairment and measuring the impairment loss can be confusing. In testing for impairment, undiscounted cash flows are used. Once impairment has been detected, the loss is based on fair value or the discounted expected future cash flows. Using undiscounted cash flows to test for impairment keeps PP&E assets from becoming "impaired" by increases in the discount rate when interest rates increase.

EXAMPLE: Asset impairment

Information related to equipment owned by Brownfield Company follows:

Original cost	\$900,000
Accumulated depreciation to date	\$100,000
Expected future cash flows	\$795,000
Fair value	\$790,000
Value in use	\$785,000
Selling costs	\$30,000

Assuming Brownfield will continue to use the equipment, test the asset for impairment under both IFRS and U.S. GAAP and discuss the results.

Answer:

The carrying value of the equipment is \$900,000 original cost – \$100,000 accumulated depreciation = \$800,000, and the recoverable amount under IFRS is \$785,000 (greater of \$785,000 value in use and \$760,000 fair value less selling costs). Under IFRS, the asset is written down on the balance sheet to the \$785,000 recoverable amount, and a \$15,000 loss (\$800,000 carrying value – \$785,000 recoverable amount) is recognized in the income statement.

Under U.S. GAAP, the asset is impaired because the \$795,000 expected future cash flows is less than the \$800,000 carrying value. The asset is written down on the balance sheet to its \$790,000 fair value, and a \$10,000 loss (\$800,000 carrying value – \$790,000 fair value) is recognized on the income statement.

Intangible Assets With Indefinite Lives

Intangible assets with indefinite lives are not amortized; rather, they are tested for impairment at least annually. An impairment loss is recognized when the carrying amount exceeds fair value.



PROFESSOR'S NOTE

The details of impairment for indefinite-lived intangibles, such as goodwill, are covered at Level II.

Long-Lived Assets Held for Sale

If a firm intends to sell an asset, it is probable that the asset will be sold, and the asset is available for immediate sale, then it must be reclassified from *held-for-use* to *held-for-sale*. When a firm reclassifies an asset as held-for-sale, the asset is tested for impairment. At this point, the asset is no longer depreciated or amortized. The held-for-sale asset is impaired if its carrying value exceeds its fair value less selling costs. If impaired, the asset is written down to net realizable value and the loss is recognized in the income statement.

For long-lived assets held for sale, the loss can be reversed under IFRS and U.S. GAAP if the value of the asset recovers in the future. However, the loss reversal is limited to the original impairment loss. Thus, the carrying value of the asset after reversal cannot exceed the carrying value before the impairment was recognized.

LOS 23.k: Explain the derecognition of property, plant, and equipment and intangible assets.

Eventually, long-lived assets are removed from the balance sheet. **Derecognition** occurs when assets are sold, exchanged, or abandoned.

When a long-lived asset is sold, the asset is removed from the balance sheet and the difference between the sale proceeds and the carrying value of the asset is reported as a gain or loss in the income statement. The carrying value is equal to original cost minus accumulated depreciation and any impairment charges.

The gain or loss is usually reported in the income statement as a part of other gains and losses, or reported separately if material. Also, if the firm presents its cash flow statement using the indirect method, the gain or loss is removed from net income to compute cash flow from operations because the proceeds from selling a long-lived asset are an investing cash inflow.

If a long-lived asset is abandoned, the treatment is similar to a sale, except there are no proceeds. In this case, the carrying value of the asset is removed from the balance sheet and a loss of that amount is recognized in the income statement.

If a long-lived asset is exchanged for another asset, a gain or loss is computed by comparing the carrying value of the old asset with fair value of the old asset (or the fair value of the new asset if that value is clearly more evident). The carrying value of the old asset is removed from the balance sheet and the new asset is recorded at its fair value.

LOS 23.f: Explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios.

Impairment reduces an asset's carrying value on the balance sheet. An impairment charge is recognized as a loss in the income statement, reducing assets and equity (retained earnings). In the year of impairment, ROA and ROE will decrease because the impairment charge reduces net income.

In subsequent periods, net income will be higher than it would have been without the impairment charge because depreciation will be lower (the asset has a lower depreciable value). Both ROA and ROE will increase in periods after the impairment charge because both equity and assets will fall as a result of the impairment charge. Asset turnover will increase in the period in which the impairment charge is taken, and in subsequent periods as well.

Asset impairment has no impact on cash flow because the impairment does not reduce taxable income; it is an unrealized loss until the asset is disposed of.

Analysis of Impairments

An impairment loss is an indication that the firm has not recognized sufficient depreciation or amortization expense, and has overstated earnings as a result.

The judgment required in determining asset impairments gives management considerable discretion about the timing and amounts of impairment charges. Consequently, impairment decisions present an opportunity for management to manipulate earnings. Waiting to recognize an impairment loss until a period of relatively high earnings would tend to smooth earnings.

Alternatively, existing managements may take more impairment charges in periods when earnings will be poor due to external (macroeconomic or industry) factors. New managements may also choose to take more or greater impairment charges when they take over. In either case, the resulting low earnings might not be perceived as the "fault" of management, and lower values for assets and equity give a boost to ROA and ROE going forward.

Revaluation

Under U.S. GAAP, most long-lived assets are reported on the balance sheet at depreciated cost using the *cost model* (original cost less accumulated depreciation and impairment charges). Revaluing long-lived assets upward is generally prohibited. One exception relates to long-lived assets held for sale, for which prior impairment losses can be reversed.

Under IFRS, firms can choose to use the **revaluation model** and report long-lived assets at their fair values. Firms can choose depreciated cost for some asset classes and fair value for others.

PROFESSOR'S NOTE

Do not confuse the revaluation model with fair-value reporting of trading securities or some types of inventory, under which all gains and losses flow through the income statement.

Revaluing an asset's value upward will result in:

- Higher total assets and higher shareholders' equity.
- Lower leverage ratios as measured by the debt ratio (total debt / total assets) and the debt-to-equity ratio (higher denominators).
- Higher depreciation expense and thus lower profitability in periods after revaluation.
- Lower ROA and ROE in periods after revaluation (lower numerators and higher denominators). However, if the increase in the asset value is the result of higher operating capacity, such higher capacity should result in higher revenues and thus higher earnings.

An analyst should check the source of the appraisal that supports the revaluation. Appraisals from independent sources are usually more reliable than appraisals by management.

Derecognition of Assets

Derecognition of an asset refers to its disposal by sale, exchange for another asset, or abandonment. Under the cost model, the carrying (book) value of a long-lived asset is its historical cost minus accumulated depreciation or amortization, adjusted for any impairment charges taken. Under the revaluation model, the carrying value of an asset is its value as of the last revaluation date, less any subsequent depreciation or amortization.

The difference between the sale price and the carrying value is reported as a gain or loss on the income statement. Such gains and losses may be reported in *other income or losses* or as a separate line item if the amount is material. When an asset is abandoned, the treatment is the same, but the sale price is zero and the loss is equal to the carrying value.

If an asset is exchanged for another asset, the sale price is taken to be the fair value of the asset exchanged, or the value of the asset acquired if that value is more readily available. There are no cash flows with an exchange of assets, but gains or losses, based on the difference between the carrying value and the "price" received for the exchanged asset, are reported on the income statement (as they are with a sale). If the fair values of the assets cannot be reliably estimated, the price of the acquired asset is taken to be the carrying value of the exchanged asset and no gains or losses are recorded.

MODULE QUIZ 23.3

1. Two years ago, Metcalf Corp. purchased machinery for \$800,000. At the end of last year, the machinery had a fair value of \$720,000. Assuming Metcalf uses the revaluation model, what amount, if any, is recognized in Metcalf's net income this year if the machinery's fair value is

- \$810,000?
- \$0.
 - \$80,000.
 - \$90,000.
- According to U.S. GAAP, an asset is impaired when:
 - the firm cannot fully recover the carrying amount of the asset through operations.
 - accumulated depreciation plus salvage value exceeds acquisition cost.
 - the present value of future cash flows from an asset exceeds its carrying value.
 - In the year after an impairment charge on a finite-lived identifiable intangible asset, compared to not taking the charge, net income is *most likely* to be:
 - lower.
 - higher.
 - unaffected.
 - A firm recently recognized a \$15,000 loss on the sale of machinery used in its manufacturing operation. The original cost of the machinery was \$100,000 and the accumulated depreciation at the date of sale was \$60,000. What amount did the firm receive from the sale?
 - \$25,000.
 - \$45,000.
 - \$85,000.
 - Other things equal, which of the following actions related to property, plant, and equipment will *most likely* decrease a firm's return on assets in future periods?
 - Impairment.
 - Derecognition.
 - Upward revaluation.

MODULE 23.4: FIXED ASSET DISCLOSURES



LOS 23.l: Describe the financial statement presentation of and disclosures relating to property, plant, and equipment and intangible assets.

Video covering this content is available online.

IFRS Disclosures

Under IFRS, the firm must disclose the following for each class of property, plant, and equipment (PP&E):

- Basis for measurement (usually historical cost).
- Useful lives or depreciation rate.
- Gross carrying value and accumulated depreciation.
- Reconciliation of carrying amounts from the beginning of the period to the end of the period.

The firm must also disclose:

- Title restrictions and assets pledged as collateral.
- Agreements to acquire PP&E in the future.

If the revaluation (fair value) model is used, the firm must disclose:

- The revaluation date.
- How fair value was determined.
- Carrying value using the historical cost model.

Under IFRS, the disclosure requirements for intangible assets are similar to those for PP&E, except that the firm must disclose whether the useful lives are finite or indefinite.

For impaired assets, the firm must disclose:

- Amounts of impairment losses and reversals by asset class.
- Where the losses and loss reversals are recognized in the income statement.
- Circumstances that caused the impairment loss or reversal.

U.S. GAAP Disclosures

Under U.S. GAAP, the PP&E disclosures include:

- Depreciation expense by period.
- Balances of major classes of assets by nature and function, such as land, improvements, buildings, machinery, and furniture.
- Accumulated depreciation by major classes or in total.
- General description of depreciation methods used.

Under U.S. GAAP, the disclosure requirements for intangible assets are similar to those for PP&E. In addition, the firm must provide an estimate of amortization expense for the next five years.

For impaired assets, the firm must disclose:

- A description of the impaired asset.
- Circumstances that caused the impairment.
- How fair value was determined.
- The amount of loss.
- Where the loss is recognized in the income statement.

LOS 23.m: Analyze and interpret financial statement disclosures regarding property, plant, and equipment and intangible assets.

Financial statement disclosures provide an analyst considerable information about a company's fixed assets and depreciation (amortization) methods. An analyst can use these data to estimate the average age of the firm's assets. The average age is useful for two reasons:

1. Older, less-efficient assets may make a firm less competitive.
2. The average age of assets helps an analyst to estimate the timing of major capital expenditures and a firm's future financing requirements.

The level of detail provided in footnote disclosures regarding fixed assets and depreciation varies across firms. Because assets are often grouped by their useful lives, the following methods of estimating the average age, economic life, and remaining useful life of a firm's assets do not produce precise values, but they can highlight issues for further investigation.

Three useful calculations (in years) for an analyst are:

Average Age

$$\text{average age} = \frac{\text{accumulated depreciation}}{\text{annual depreciation expense}}$$

This calculation is more accurate for a firm that uses straight-line depreciation. The calculation can be significantly affected by the mix of assets.

Total Useful Life

$$\text{total useful life} = \frac{\text{historical cost}}{\text{annual depreciation expense}}$$

Historical cost is gross PP&E before deducting accumulated depreciation.

Remaining Useful Life

$$\text{remaining useful life} = \frac{\text{ending net PP&E}}{\text{annual depreciation expense}}$$

Net PP&E is equal to original cost (gross PP&E) minus accumulated depreciation.



PROFESSOR'S NOTE

The remaining useful life can also be approximated by subtracting the average age from the average depreciable life.

EXAMPLE: Calculating average age and total useful life

At the end of 20X8, a company has gross PP&E of \$3 million and accumulated depreciation of \$1 million. During the year, depreciation expense was \$500,000.

What is the average age, total useful life, and remaining useful life of the company's PP&E?

Answer:

$$\text{average age} = \frac{\text{accumulated depreciation}}{\text{depreciation expense}} = \frac{\$1,000,000}{\$500,000} = 2 \text{ years}$$

$$\text{total useful life} = \frac{\text{historical cost}}{\text{depreciation expense}} = \frac{\$3,000,000}{\$500,000} = 6 \text{ years}$$

$$\text{remaining useful life} = \frac{\text{ending net PP&E}}{\text{depreciation expense}} = \frac{\$2,000,000}{\$500,000} = 4 \text{ years}$$

Another popular metric is the ratio of annual capital expenditures to depreciation expense. This ratio provides information about whether the firm is replacing its PP&E at the same rate as its assets are being depreciated.

LOS 23.n: Compare the financial reporting of investment property with that of property, plant, and equipment.

Under IFRS, property that a firm owns for the purpose of collecting rental income, earning capital appreciation, or both, is classified as **investment property**. U.S. GAAP does not distinguish investment property from other kinds of long-lived assets.

IFRS gives firms the choice of using a cost model or a fair value model when valuing investment property, if a fair value for the property can be established reliably. A firm generally must use

the same valuation model (cost or fair value) for all of its investment properties.

The cost model for investment property is the same as the cost model for valuing property, plant, and equipment, but the fair value model is different from the revaluation model we described earlier. Recall that under the revaluation model, an increase in an asset's carrying value is recognized as revaluation surplus in owners' equity (unless it reverses a previously recognized loss). For investment property, however, an upward revaluation is recognized as a gain on the income statement.

Firms are required to disclose which valuation model they use for investment property. Firms that use the fair value model must state how they determine the fair value of investment property and reconcile its beginning and ending values. Firms that use the cost model must disclose the fair value of their investment property, along with the disclosures that are required for other types of long-lived assets (e.g., useful lives, depreciation methods used).

In some cases, a firm may change its use of a property such that it becomes investment property or is no longer classified as investment property. For example, a firm may move its offices out of a building it owns and begin renting the space to others. If the firm uses the fair value model, the financial statement treatment of the asset's value depends on the nature of the change, as summarized in Figure 23.3. If the firm uses the cost model, the property's carrying amount does not change when it is transferred into or out of investment property.

Figure 23.3: Transfers To or From Investment Property (Fair Value Model)

Transfer From	Transfer To	Financial Statement Treatment
Owner-occupied	Investment property	Treat as revaluation: recognize gain only if it reverses previously recognized loss
Inventory	Investment property	Recognize gain or loss if fair value is different from carrying amount
Investment property	Owner-occupied or inventory	Fair value of asset at date of transfer will be its cost under new classification

MODULE QUIZ 23.4



1. Which of the following disclosures would *least likely* be found in the financial statement footnotes of a firm?
 - A. Accumulated depreciation.
 - B. Carrying values by asset class.
 - C. Average age of assets.
2. Metallurgy, Inc., reported depreciation expense of \$15 million for the most recent year. Beginning-of-year gross PP&E and accumulated depreciation were \$287 million and \$77 million, respectively. If end-of year gross PP&E and accumulated depreciation were \$300 million and \$80 million, the estimated remaining useful life of PP&E is *closest* to:
 - A. 10 years.
 - B. 15 years.
 - C. 20 years.
3. A firm owns a warehouse that it rents out. Under IFRS, the firm may report the value of this asset on its balance sheet using:
 - A. the cost model or the fair value model.
 - B. the cost model or the revaluation model.
 - C. the revaluation model or the fair value model.

KEY CONCEPTS

LOS 23.a

When an asset is expected to provide benefits for only the current period, its cost is expensed on the income statement for the period. If an asset is expected to provide benefits over multiple periods, it is capitalized rather than expensed.

LOS 23.b

The cost of a purchased finite-lived intangible asset is amortized over its useful life. Indefinite-lived intangible assets are not amortized, but are tested for impairment at least annually. The cost of internally developed intangible assets is expensed.

Under IFRS, research costs are expensed but development costs may be capitalized. Under U.S. GAAP, both research and development costs are expensed as incurred, except in the case of software created for sale to others.

The acquisition method is used to account for assets acquired in a business combination. The purchase price is allocated to the fair value of identifiable assets of the acquired firm less its liabilities. Any excess of the purchase price above the fair value of the acquired firm's net assets is recorded as goodwill, an unidentifiable intangible asset that cannot be separated from the business itself.

LOS 23.c

With capitalization, the asset value is put on the balance sheet and the cost is expensed through the income statement over the asset's useful life through either depreciation or amortization. Compared to expensing the asset cost, capitalization results in:

- Lower expense and higher net income in period of acquisition, higher expense (depreciation or amortization) and lower net income in each of the remaining years of the asset's life.
- Higher assets and equity.
- Lower CFI and higher CFO because the cost of a capitalized asset is classified as an investing cash outflow.
- Higher ROE and ROA in the initial period, and lower ROE and ROA in subsequent periods because net income is lower and both assets and equity are higher.
- Lower debt-to-assets and debt-to-equity ratios because assets and equity are higher.

LOS 23.d

Depreciation methods:

- Straight-line: Equal amount of expense each period.
- Accelerated (declining balance): Higher depreciation expense in the early years and lower depreciation expense in the later years of an asset's life.
- Units-of-production: Expense based on percentage usage rather than time.

Straight-line method:

$$\text{depreciation expense} = \frac{\text{original cost} - \text{salvage value}}{\text{depreciable life}}$$

Double-declining balance (DDB), an accelerated depreciation method:

DDB depreciation in year x =

$$\frac{2}{\text{depreciable life in years}} \times \text{book value at beginning of year } x$$

Units of production method:

$$\frac{\text{original cost} - \text{salvage value}}{\text{life in output units}} \times \text{output units used in the period}$$

IFRS requires component depreciation, in which significant parts of an asset are identified and depreciated separately.

LOS 23.e

In the early years of an asset's life, accelerated depreciation results in higher depreciation expense, lower net income, and lower ROA and ROE compared to straight-line depreciation. Cash flow is the same assuming tax depreciation is unaffected by the choice of method for financial reporting.

Firms can reduce depreciation expense and increase net income by using longer useful lives and higher salvage values.

LOS 23.f

Impairment charges decrease net income, assets, and equity, which results in lower ROA and ROE and higher debt-to-equity and debt-to-assets ratios for a typical firm.

Upward revaluation increases assets and equity, and thereby decreases debt-to-assets and debt-to-equity ratios. A downward revaluation has opposite effects. The effect on net income and related ratios depends on whether the revaluation is to a value above or below cost.

Derecognition of assets can result in either a gain or loss on the income statement. A loss will reduce net income and assets, while a gain will increase net income and assets.

LOS 23.g

Amortization methods for intangible assets with finite lives are the same as those for depreciation: straight line, accelerated, or units of production. Calculation of amortization expense for such assets is the same as with depreciation expense.

LOS 23.h

The choice of amortization method will affect expenses, assets, equity, and financial ratios in exactly the same way that the choice of depreciation method will. Just as with the depreciation of tangible assets, increasing either the estimate of an asset's useful life or the estimate of its residual value will reduce annual amortization expense, which will increase net income, assets, ROE, and ROA for a typical firm.

LOS 23.i

Under IFRS, firms have the option to revalue assets based on fair value under the revaluation model. U.S. GAAP does not permit revaluation.

The impact of revaluation on the income statement depends on whether the initial revaluation resulted in a gain or loss. If the initial revaluation resulted in a loss (decrease in carrying value), the initial loss would be recognized in the income statement and any subsequent gain would be recognized in the income statement only to the extent of the previously reported loss. Revaluation gains beyond the initial loss bypass the income statement and are recognized in shareholders' equity as a revaluation surplus.

If the initial revaluation resulted in a gain (increase in carrying value), the initial gain would bypass the income statement and be reported as a revaluation surplus. Later revaluation losses would first reduce the revaluation surplus.

LOS 23.j

Under IFRS, an asset is impaired when its carrying value exceeds the recoverable amount. The recoverable amount is the greater of fair value less selling costs and the value in use (present value of expected cash flows). If impaired, the asset is written down to the recoverable amount. Loss recoveries are permitted, but not above historical cost.

Under U.S. GAAP, an asset is impaired if its carrying value is greater than the asset's undiscounted future cash flows. If impaired, the asset is written down to fair value. Subsequent recoveries are not allowed for assets held for use.

Asset impairments result in losses in the income statement. Impairments have no impact on cash flow as they have no tax or other cash flow effects until disposal of the asset.

LOS 23.k

When a long-lived asset is *sold*, the difference between the sale proceeds and the carrying (book) value of the asset is reported as a gain or loss in the income statement.

When a long-lived asset is *abandoned*, the carrying value is removed from the balance sheet and a loss is recognized in that amount.

If a long-lived asset is *exchanged* for another asset, a gain or loss is computed by comparing the carrying value of the old asset with fair value of the old asset (or fair value of the new asset if more clearly evident).

LOS 23.1

There are many differences in the disclosure requirements for tangible and intangible assets under IFRS and U.S. GAAP. However, firms are generally required to disclose:

- Carrying values for each class of asset.
- Accumulated depreciation and amortization.
- Title restrictions and assets pledged as collateral.
- For impaired assets, the loss amount and the circumstances that caused the loss.
- For revalued assets (IFRS only), the revaluation date, how fair value was determined, and the carrying value using the historical cost model.

LOS 23.m

Analysts can use disclosures of the historical cost, accumulated depreciation (amortization), and annual depreciation (amortization) expense to estimate average age of assets, total useful life of assets, and remaining useful life of assets. These estimates are more accurate for firms that use straight-line depreciation.

$$\text{Average age} = \frac{\text{accumulated depreciation}}{\text{annual depreciation expense}}$$

$$\text{Total useful life} = \frac{\text{historical cost}}{\text{annual depreciation expense}}$$

$$\text{Remaining useful life} = \frac{\text{ending net PP&E}}{\text{annual depreciation expense}}$$

LOS 23.n

Under IFRS (but not U.S. GAAP), investment property is defined as property owned for the purpose of earning rent, capital appreciation, or both. Firms can account for investment property using the cost model or the fair value model. Unlike the revaluation model for property, plant, and equipment, increases in the fair value of investment property above its historical cost are recognized as gains on the income statement if the firm uses the fair value model.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 23.1

1. A If the future economic benefits of a purchase are highly uncertain, a company should expense the purchase in the period it is incurred. (LOS 23.a)
2. A The cost of an intangible asset is amortized if the asset has a finite life and was purchased or acquired in a business combination. Development costs for internally generated intangible assets may be capitalized under IFRS, but research costs are expensed as incurred. (LOS 23.b)
3. C As compared to a firm that capitalizes its expenditures, a firm that immediately expenses expenditures will report lower assets. Thus, asset turnover (revenue / average assets) will be higher for the expensing firm (lower denominator). (LOS 23.c)

Module Quiz 23.2

1. A Double-declining balance = \$30,000 book value \times (2/8) = \$7,500.
Units-of-production = $(\$30,000 \text{ cost} - \$3,000 \text{ salvage value}) / 150,000 \text{ miles} = \0.18 per mile .
16,500 miles driven \times \$0.18 per mile = \$2,970.
(LOS 23.d)
2. C In the early years, accelerated depreciation will result in higher depreciation expense; thus, lower net income. Lower net income will result in lower retained earnings. (LOS 23.e)
3. B Indefinite-lived intangible assets are not amortized; rather, they are reported on the balance sheet indefinitely unless they are impaired. (LOS 23.g)

4. **A** Under U.S. GAAP, research and development costs are expensed as incurred. Thus, the entire \$300,000 of R&D is expensed this year. The result is a zero carrying value. (LOS 23.g)
5. **A** Initial cost has nothing to do with the useful life of an intangible asset. (LOS 23.h)

Module Quiz 23.3

1. **B** Under the revaluation method, Metcalf reports the equipment on the balance sheet at fair value. At the end of last year, an \$80,000 loss was recognized (from \$800,000 to \$720,000) in the income statement. Any recovery is recognized in the income statement to the extent of the loss. Any remainder is recognized in shareholders' equity as revaluation surplus. Thus, at the end of this year, an \$80,000 gain is recognized in the income statement, and a \$10,000 revaluation surplus is recognized in shareholders' equity. (LOS 23.i)
2. **A** An asset is impaired when the firm cannot recover the carrying value. Under U.S. GAAP, recoverability is tested based on undiscounted future cash flows. (LOS 23.j)
3. **B** Because a finite-lived identifiable intangible asset would be amortized, amortization expense in the year after the reduction from the impairment charge would be lower (the carrying value of the asset would most likely be lower), increasing net income. (LOS 23.j)
4. **A** Gain or loss is equal to the sale proceeds minus the carrying value (cost minus accumulated depreciation) at the time of sale. Given the loss of \$15,000 and carrying value of \$40,000 (\$100,000 – \$60,000), we can solve for the proceeds of \$25,000 ($-15,000 + 40,000$). (LOS 23.k)
5. **C** An upward revaluation will increase the book value of assets and increase depreciation expense in future periods (decreasing net income), both of which reduce ROA. Impairment would have the opposite effects, decreasing future depreciation and book values. Derecognizing an asset may increase, decrease, or not affect ROA in future periods. (LOS 23.f)

Module Quiz 23.4

1. **C** The average age is not a required disclosure. However, it can be calculated given other disclosures. (LOS 23.l)
2. **B** The remaining useful life can be estimated as ending net PP&E value divided by annual depreciation.

$$(300 - 80)/15 = 14.66 \text{ years}$$

(LOS 23.m)
3. **A** Under IFRS, the warehouse is classified as investment property because it is owned primarily for rental income. Investment property may be reported using either the cost model or the fair value model. (LOS 23.n)

Reading 24

INCOME TAXES

EXAM FOCUS

In many countries, financial reporting standards and tax reporting standards differ. Candidates should be aware of the terminology that relates to each set of standards, notably taxes payable, which are the taxes actually due to the government, and income tax expense, which is reported on the income statement and reflects taxes payable plus any deferred income tax expense. The timing of revenue and expense recognition in the income statement and the tax return may lead to the creation of deferred tax liabilities, which the company may have to pay in the future, or deferred tax assets, which may provide benefits in the future. For the exam, you should know that some differences between taxable and pretax income are temporary, while some are permanent and will never reverse. Be prepared to calculate taxes payable, tax expense, deferred tax liabilities and assets, and be able to make the necessary adjustments for analytical purposes.

MODULE 24.1: TAX TERMS



LOS 24.a: Describe the differences between accounting profit and taxable income and define key terms, including deferred tax assets, deferred tax liabilities, valuation allowance, taxes payable, and income tax expense.

Video covering this content is available online.

Financial accounting standards (IFRS and U.S. GAAP) are often different than income tax laws and regulations. As a result, the amount of income tax expense recognized in the income statement may differ from the actual taxes owed to the taxing authorities.

Tax Return Terminology

- **Taxable income.** Income subject to tax based on the tax return.
- **Taxes payable.** The tax liability caused by *taxable income*. This is also known as current tax expense, but do not confuse this with *income tax expense* (see below).
- **Income tax paid.** The actual cash flow for income taxes including payments or refunds from other years.
- **Tax loss carryforward.** A current or past loss that can be used to reduce taxable income (thus, taxes payable) in the future. Can result in a deferred tax asset.
- **Tax base.** Net amount of an asset or liability used for tax reporting purposes.

Financial Reporting Terminology

- **Accounting profit.** Pretax financial income based on financial accounting standards. Also known as *income before tax* and *earnings before tax*.

- **Income tax expense.** Expense recognized in the income statement that includes taxes payable and *changes* in deferred tax assets and liabilities (DTA and DTL). The income tax expense equation is:

$$\text{income tax expense} = \text{taxes payable} + \Delta\text{DTL} - \Delta\text{DTA}$$

- **Deferred tax liabilities.** Balance sheet amounts that result from an excess of income tax expense over taxes payable that are expected to result in future cash outflows.
- **Deferred tax assets.** Balance sheet amounts that result from an excess of taxes payable over income tax expense that are expected to be recovered from future operations. Can also result from tax loss carryforwards.
- **Valuation allowance.** Reduction of deferred tax assets based on the likelihood the assets will not be realized.
- **Carrying value.** Net balance sheet value of an asset or liability.
- **Permanent difference.** A difference between taxable income (tax return) and pretax income (income statement) that will not reverse in the future.
- **Temporary difference.** A difference between the tax base and the carrying value of an asset or liability that will result in either taxable amounts or deductible amounts in the future. Several examples of how temporary differences arise are presented later in this review.

MODULE 24.2: DEFERRED TAX LIABILITIES AND ASSETS



Video covering this content is available online.

LOS 24.b: Explain how deferred tax liabilities and assets are created and the factors that determine how a company's deferred tax liabilities and assets should be treated for the purposes of financial analysis.

Differences between the treatment of an accounting item for tax reporting and for financial reporting can occur when:

- The timing of revenue and expense recognition in the income statement and the tax return differ.
- Certain revenues and expenses are recognized in the income statement but never on the tax return or vice-versa.
- Assets and/or liabilities have different carrying amounts and tax bases.
- Gain or loss recognition in the income statement differs from the tax return.
- Tax losses from prior periods may offset future taxable income.
- Financial statement adjustments may not affect the tax return or may be recognized in different periods.

Deferred Tax Liabilities

A **deferred tax liability** is created when income tax expense (income statement) is greater than taxes payable (tax return) due to temporary differences. Deferred tax liabilities occur when:

- Revenues (or gains) are recognized in the income statement before they are included on the tax return due to temporary differences.
- Expenses (or losses) are tax deductible before they are recognized in the income statement.

Deferred tax liabilities are expected to reverse (i.e., they are caused by temporary differences) and result in future cash outflows when the taxes are paid.

A deferred tax liability is most often created when an accelerated depreciation method is used on the tax return and straight-line depreciation is used on the income statement.

Deferred Tax Assets

A **deferred tax asset** is created when taxes payable (tax return) are greater than income tax expense (income statement) due to temporary differences. Deferred tax assets occur when:

- Revenues (or gains) are taxable before they are recognized in the income statement.
- Expenses (or losses) are recognized in the income statement before they are tax deductible.
- Tax loss carryforwards are available to reduce future taxable income.

Similar to deferred tax liabilities, deferred tax assets are expected to reverse through future operations. However, deferred tax assets are expected to provide future tax savings, while deferred tax liabilities are expected to result in future cash outflows. A firm that has taxable losses in excess of its taxable income can carry those excess losses forward and use them to reduce taxable income (and taxes) in future periods.

Post-employment benefits, warranty expenses, and tax loss carryforwards are typical causes of deferred tax assets.

Treatment for Analytical Purposes

If deferred tax liabilities are expected to reverse in the future, they are best classified by an analyst as liabilities. If, however, they are not expected to reverse in the future, they are best classified as equity (DTL decreased and equity increased by the same amount). The key question is, "When or will the total deferred tax liability be reversed in the future?" In practice, the treatment of deferred taxes for analytical purposes varies. An analyst must decide on the appropriate treatment on a case-by-case basis.

LOS 24.c: Calculate income tax expense, income taxes payable, deferred tax assets, and deferred tax liabilities, and calculate and interpret the adjustment to the financial statements related to a change in the income tax rate.



Video covering this content is available online.



PROFESSOR'S NOTE

The effects of changing tax rates on deferred tax assets and liabilities are explained later in this reading.

EXAMPLE: Deferred tax liabilities

Assume the original cost of an asset is \$600,000. The asset has a 3-year life and no salvage value is expected. For tax purposes, the asset is depreciated using an accelerated depreciation method with tax return depreciation of \$300,000 in year 1, \$200,000 in year 2, and \$100,000 in year 3. The firm recognizes straight-line (SL) depreciation expense of \$200,000 each year in its income statements. Earnings before interest, taxes, depreciation, and amortization (EBITDA) is \$500,000 each year. The firm's tax rate is 40%. Calculate the firm's income tax expense, taxes payable, and deferred tax liability for each year of the asset's life.

Answer:

The following tables illustrate the calculation of taxes payable reported on the tax return and income tax expense reported in the income statement.

Tax Return (40% Tax Rate, Accelerated Depreciation)

	Year 1	Year 2	Year 3	Total 1–3
EBITDA	\$500,000	\$500,000	\$500,000	\$1,500,000
Depreciation	\$300,000	\$200,000	\$100,000	\$600,000
Taxable income	\$200,000	\$300,000	\$400,000	\$900,000
Tax rate	× 0.40	× 0.40	× 0.40	× 0.40
Tax payable	\$80,000	\$120,000	\$160,000	\$360,000

Income Statement (40% Tax Rate, SL Depreciation)

	Year 1	Year 2	Year 3	Total 1–3
EBITDA	\$500,000	\$500,000	\$500,000	\$1,500,000
Depreciation	\$200,000	\$200,000	\$200,000	\$600,000
Pre-tax income	\$300,000	\$300,000	\$300,000	\$900,000
Tax rate	× 0.40	× 0.40	× 0.40	× 0.40
Income tax expense	\$120,000	\$120,000	\$120,000	\$360,000

In year 1, the firm recognizes \$120,000 of income tax expense on the income statement but taxes payable (tax return) are only \$80,000. So, income tax expense is *initially* higher than taxes payable. The \$40,000 difference is deferred to a future period by using an accelerated depreciation method for tax purposes. The \$40,000 is reported on the balance sheet by creating a DTL.

The tax base and the carrying value of the asset are used to calculate the *balance* of the DTL. At the end of year 1, the carrying value of the asset is \$400,000 and the tax base of the asset is \$300,000. By multiplying the \$100,000 difference by the 40% tax rate, we get the *balance* of the DTL of \$40,000.

We can reconcile income tax expense and taxes payable with the *change* in the DTL. In this example, the DTL increased \$40,000 (from zero to \$40,000) during year 1. Thus, income tax expense in year 1 is \$120,000 (\$80,000 taxes payable + \$40,000 *change* in the DTL).

In year 2, depreciation expense is the same on the tax return and the income statement. Thus, taxable income is equal to pretax income and there is no change in the DTL. Income tax expense in year 2 is \$120,000 (\$120,000 taxes payable + zero *change* in the DTL).

In year 3, the firm recognizes income tax expense of \$120,000 on the income statement but \$160,000 in taxes payable (tax return). The \$40,000 deferred tax liability recognized at the end of year 1 has reversed as a result of lower depreciation expense using the accelerated method on the tax return. In year 3, income tax expense is \$120,000 [\$160,000 taxes payable + (-\$40,000 *change* in DTL)].

Note that over the useful life of the asset, total depreciation, total taxable (and pretax) income, and total taxes payable (and income tax expense) are the same on the financial statements and the tax return. Also, at the end of year 3, both the tax base and the carrying value of the asset are equal to zero. By using accelerated depreciation for tax purposes, the firm *deferred* \$40,000 of taxes from year 1 to year 3.

EXAMPLE: Deferred tax assets

Consider warranty guarantees and associated expenses. Pretax income (financial reporting) includes an accrual for warranty expense, but warranty cost is not deductible for taxable income until the firm has made actual expenditures to meet warranty claims. Suppose:

- A firm has sales of \$5,000 for each of two years.
- The firm estimates that warranty expense will be 2% of annual sales (\$100).
- The actual expenditure of \$200 to meet all warranty claims was not made until the second year.
- Assume a tax rate of 40%.

Calculate the firm's income tax expense, taxes payable, and deferred tax assets for year 1 and year 2.

Answer:

For tax reporting, taxable income and taxes payable for two years are:

Tax Reporting—Warranty Expense

	Year 1	Year 2
Revenue	\$5,000	\$5,000
Warranty expense	0	200
<i>Taxable income</i>	\$5,000	\$4,800
<i>Taxes payable</i>	2,000	1,920
Net income	\$3,000	\$2,880

For financial reporting, pretax income and tax expense are:

Financial Reporting—Warranty Expense

	Year 1	Year 2
Revenue	\$5,000	\$5,000
Warranty expense	100	100
<i>Pretax income</i>	\$4,900	\$4,900
<i>Tax expense</i>	1,960	1,960
Net income	\$2,940	\$2,940

In year 1, the firm reports \$1,960 of tax expense in the income statement, but \$2,000 of taxes payable are reported on the tax return. In this example, taxes payable are *initially* higher than tax expense and the \$40 difference is reported on the balance sheet by creating a DTA.

The tax base and the carrying value of the warranty liability are used to calculate the *balance* of the DTA. At the end of year 1, the carrying value of the warranty liability is \$100 (the warranty expense has been recognized in the income statement but it has not been paid), and the tax base of the liability is zero (the warranty expense has not been recognized on the tax return). By multiplying the \$100 difference by the 40% tax rate, we get the *balance* of the DTA of \$40 [(\$100 carrying value – zero tax base) × 40%].

We can reconcile income tax expense and taxes payable with the *change* in the DTA. In this example, the DTA increased \$40 (from zero to \$40) during year 1. Thus, income tax expense in year 1 is \$1,960 (\$2,000 taxes payable – \$40 increase in the DTA).

In year 2, the firm recognizes \$1,960 of tax expense in the income statement but only \$1,920 is reported on the tax return (taxes payable). The \$40 deferred tax asset recognized at the end of year 1 has reversed as a result of the warranty expense recognition on the tax return. So, in year 2, income tax expense is \$1,960 (\$1,920 taxes payable + \$40 decrease in DTA).



PROFESSOR'S NOTE

To summarize, if taxable income (on the tax return) is less than pretax income (on the income statement) and the difference is expected to reverse in future years, a deferred tax liability is created. If taxable income is greater than pretax income and the difference is expected to reverse in future years, a deferred tax asset is created.

LOS 24.d: Calculate the tax base of a company's assets and liabilities.

Tax Base of Assets

An asset's **tax base** is the amount that will be deducted (expensed) on the tax return in the future as the economic benefits of the asset are realized. The **carrying value** is the value of the asset reported on the financial statements, net of depreciation and amortization.

Following are a few examples of calculating the tax bases of various assets.

Depreciable equipment. The cost of equipment is \$100,000. In the income statement, depreciation expense of \$10,000 is recognized each year for ten years. On the tax return, the asset is depreciated at \$20,000 per year for five years.

At the end of the first year, the tax base is \$80,000 (\$100,000 cost – \$20,000 accumulated tax depreciation) and the carrying value is \$90,000 (\$100,000 cost – \$10,000 accumulated financial depreciation). A deferred tax liability ($\$10,000 \times \text{tax rate}$) is created to account for the timing difference from different depreciation for tax and for financial reporting.

Sale of the machine for \$100,000, for example, would result in a gain of \$10,000 on the income statement and a gain of \$20,000 on the tax return. This would reverse the deferred tax liability.

Research and development. At the beginning of this year, \$75,000 of R&D was expensed in the income statement. On the tax return, the R&D was capitalized and is amortized on a straight-line basis over three years.

At the end of the first year, the tax base is \$50,000 (\$75,000 cost – \$25,000 accumulated tax amortization) and the asset has no carrying value (does not appear on the balance sheet) because the entire cost was expensed. Note that amortization for tax here leads to a deferred tax asset, since earnings before tax are less than taxable income.

Accounts receivable. Gross receivables totaling \$20,000 are outstanding at year-end. Because collection is uncertain, the firm recognizes bad debt expense of \$1,500 in the income statement. For tax purposes, bad debt expense cannot be deducted until the receivables are deemed worthless.

At the end of the year, the tax base of the receivables is \$20,000 since no bad debt expense has been deducted on the tax return. The carrying value is \$18,500 (\$20,000 – \$1,500 bad debt expense). Again, a deferred tax asset is the result.

Tax Base of Liabilities

A liability's tax base is the carrying value of the liability minus any amounts that will be deductible on the tax return in the future. The tax base of revenue received in advance is the carrying value minus the amount of revenue that will *not* be taxed in the future.

Following are a few examples of calculating the tax bases of various liabilities.

Customer advance. At year-end, \$10,000 was received from a customer for goods that will be shipped next year. On the tax return, revenue received in advance is taxable when collected.

The carrying value of the liability is \$10,000. The carrying value will be reduced when the goods are shipped next year. For revenue received in advance, the tax base is equal to the carrying

value minus any amounts that will *not* be taxed in the future. Since the customer advance has already been taxed, \$10,000 will not be taxed in the future. Thus, the customer advance liability has a tax base of zero (\$10,000 carrying value – \$10,000 revenue not taxed in the future). Since the \$10,000 has been taxed but not yet reported as revenue on the income statement, a deferred tax asset is created.

Warranty liability. At year-end, a firm estimates that \$5,000 of warranty expense will be required on goods already sold. On the tax return, warranty expense is not deductible until the warranty work is actually performed. The warranty work will be performed next year.

The carrying value of the warranty liability is \$5,000. The tax base is equal to the carrying value minus the amount deductible in the future. Thus, the warranty liability has a tax base of zero (\$5,000 carrying value – \$5,000 warranty expense deductible in the future). Delayed recognition of this expense for tax results in a deferred tax asset.

Note payable. The firm has an outstanding promissory note with a principal balance of \$30,000. Interest accrues at 10% and is paid at the end of each quarter.

The promissory note is treated the same way on the tax return and in the financial statements. Thus, the carrying value and the tax base are both \$30,000. Interest paid is included in both pre-tax income on the income statement and in taxable income on the tax return. With no timing difference, no deferred tax items are created.



MODULE QUIZ 24.1, 24.2

1. Which of the following tax definitions is *least accurate*?
 - A. Taxable income is income based on the rules of the tax authorities.
 - B. Taxes payable are the amount due to the government.
 - C. Pretax income is income tax expense divided by one minus the statutory tax rate.
2. Which of the following statements is *most accurate*? The difference between taxes payable for the period and the tax expense recognized on the financial statements results from differences:
 - A. in management control.
 - B. between basic and diluted earnings.
 - C. between financial and tax accounting.
3. An analyst is comparing a firm to its competitors. The firm has a deferred tax liability that results from accelerated depreciation for tax purposes. The firm is expected to continue to grow in the foreseeable future. How should the liability be treated for analysis purposes?
 - A. It should be treated as equity at its full value.
 - B. It should be treated as a liability at its full value.
 - C. The present value should be treated as a liability with the remainder being treated as equity.

Use the following data to answer Questions 4 through 9.

- A firm acquires an asset for \$120,000 with a 4-year useful life and no salvage value.
 - The asset will generate \$50,000 of cash flow for all four years.
 - The tax rate is 40% each year.
 - The firm will depreciate the asset over three years on a straight-line (SL) basis for tax purposes and over four years on a SL basis for financial reporting purposes.
4. Taxable income in year 1 is:
 - A. \$6,000.
 - B. \$10,000.

- C. \$20,000.
5. Taxes payable in year 1 are:
- A. \$4,000.
 - B. \$6,000.
 - C. \$8,000.
6. Pretax income in year 4 is:
- A. \$6,000.
 - B. \$10,000.
 - C. \$20,000.
7. Income tax expense in year 4 is:
- A. \$4,000.
 - B. \$6,000.
 - C. \$8,000.
8. Taxes payable in year 4 are:
- A. \$4,000.
 - B. \$6,000.
 - C. \$20,000.
9. At the end of year 2, the firm's balance sheet will report a deferred tax:
- A. asset of \$4,000.
 - B. asset of \$8,000.
 - C. liability of \$8,000.
10. If the tax base of an asset exceeds the asset's carrying value and a reversal is expected in the future:
- A. a deferred tax asset is created.
 - B. a deferred tax liability is created.
 - C. neither a deferred tax asset nor a deferred tax liability is created.
11. The author of a new textbook received a \$100,000 advance from the publisher this year. \$40,000 of income taxes were paid on the advance when received. The textbook will not be finished until next year. Determine the tax base of the advance at the end of this year.
- A. \$0.
 - B. \$40,000.
 - C. \$100,000.

MODULE 24.3: CHANGE IN TAX RATES



LOS 24.e: Evaluate the effect of tax rate changes on a company's financial statements and ratios.

Video covering this content is available online.

When the income tax rate changes, deferred tax assets and liabilities are adjusted to reflect the new rate. The adjustment can also affect income tax expense.

An increase in the tax rate will increase both deferred tax liabilities and deferred tax assets. A decrease in the tax rate will decrease both deferred tax liabilities and deferred tax assets.

DTL and DTA values on the balance sheet must be changed because the new tax rate is the rate expected to be in force when the associated reversals occur. If there is an increase (decrease) in the tax rate, when previously deferred income is recognized for tax, the tax due will be higher (lower), and when expense items previously reported in the financial statements are recognized for tax, the benefit will be greater (less).

Changes in the balance sheet values of DTLs and DTAs to account for a change in the tax rate will affect income tax expense in the current period.

$$\text{income tax expense} = \text{taxes payable} + \Delta\text{DTL} - \Delta\text{DTA}$$

If tax rates increase, the increase in the DTL is added to taxes payable and the increase in the DTA is subtracted from taxes payable to arrive at income tax expense.

If tax rates decrease, the decrease in the DTL would result in lower income tax expense and the decrease in the DTA would result in higher income tax expense. In the case of the DTL we are adding a negative change, and in the case of the DTA we are subtracting a negative change.

The following example illustrates the effects of a change in the tax rate.

EXAMPLE: Accounting effects of a change in a firm's tax rate

A firm owns equipment with a carrying value of \$200,000 and a tax base of \$160,000 at year-end. The tax rate is 40%. In this case, the firm will report a DTL of \$16,000 $[(\$200,000 \text{ carrying value} - \$160,000 \text{ tax base}) \times 40\%]$. The firm has recognized a bad debt expense of \$10,000 in its financial statements which is not yet deductible for tax purposes. The bad debt expense created a DTA of \$4,000 $[(\$10,000 \text{ tax base} - \text{zero carrying value}) \times 40\%]$. Calculate the effect on the firm's income tax expense if the tax rate decreases to 30%.

Answer:

As a result of the decrease in tax rate, the balance of the DTL is reduced to \$12,000 $[(\$200,000 \text{ carrying value} - \$160,000 \text{ tax base}) \times 30\%]$. Thus, due to the lower tax rate, the change in the DTL is -\$4,000 (\$16,000 reported DTL - \$12,000 adjusted DTL).

The balance of the DTA is reduced to \$3,000 $[(\$10,000 \text{ tax base} - \text{zero carrying value}) \times 30\%]$. Thus, due to the lower tax rate, the DTA decreases by \$1,000 (\$4,000 reported DTA - \$3,000 adjusted DTA).

Using the income tax equation, we can see that income tax expense decreases by \$3,000 (income tax expense = taxes payable + $\Delta\text{DTL} - \Delta\text{DTA}$).

The effects of a change in the tax rate can also be calculated based on the difference between the tax base and carrying value for an asset or a liability, as in the following example.

EXAMPLE: Deferred tax liability with a change in the tax rate

A firm purchases equipment for \$24,000 that has a three-year useful life. The firm depreciates the equipment using the straight-line method for financial reporting. For tax reporting, the firm uses double-declining balance in the first two years and switches to straight-line in Year 3. The tax rate is 40% in the first year, but in the second year, the expected tax rate decreases to 35%. Calculate the deferred tax liability in each year of the asset's life.

Answer:

The carrying value and tax base of the asset are shown in the following table:

	Year 1	Year 2	Year 3
Straight-line depreciation	\$8,000	\$8,000	\$8,000
Carrying value	\$16,000	\$8,000	0
DDB depreciation	\$16,000	\$5,333	\$2,667
Tax base	\$8,000	\$2,667	0
Carrying value minus tax base	\$8,000	\$5,333	0

The deferred tax liability is equal to the expected tax rate times the difference between the carrying value and the tax base. In Year 1, the expected tax rate was 40% and the DTL was $40\% \times \$8,000 = \$3,200$. In Year 2, the expected tax rate changed to 35% and the DTL was $35\% \times \$5,333 = \$1,867$. The DTL returns zero at the end of Year 3.



MODULE QUIZ 24.3

1. An increase in the tax rate causes the balance sheet value of a deferred tax asset to:
 - A. decrease.
 - B. increase.
 - C. remain unchanged.
2. Which one of the following statements is *most accurate*? Under the liability method of accounting for deferred taxes, a decrease in the tax rate at the beginning of the accounting period will:
 - A. increase taxable income in the current period.
 - B. increase a deferred tax asset.
 - C. reduce a deferred tax liability.

MODULE 24.4: PERMANENT DIFFERENCES



LOS 24.f: Identify and contrast temporary versus permanent differences in pre-tax accounting income and taxable income.

Video covering this content is available online.

A **permanent difference** is a difference between taxable income and pretax income that will not reverse in the future. Permanent differences do not create deferred tax assets or deferred tax liabilities. Permanent differences can be caused by revenue that is not taxable, expenses that are not deductible, or tax credits that result in a direct reduction of taxes. For example, in the United States, interest received on municipal bonds is typically not taxable (but appears on the financial statements as pretax income), and the cost of life insurance on key company officers is typically not tax-deductible (but appears on the financial statements as a pretax expense).

Permanent differences will cause the firm's **effective tax rate** to differ from the **statutory tax rate**. The statutory rate is the tax rate of the jurisdiction where the firm operates. The effective tax rate is derived from the income statement.

$$\text{effective tax rate} = \frac{\text{income tax expense}}{\text{pretax income}}$$

The statutory rate and effective rate may also differ if the firm is operating in more than one tax jurisdiction.

A **temporary difference** refers to a difference between the tax base and the carrying value of an asset or liability that will result in taxable amounts or deductible amounts in the future. If the temporary difference is expected to reverse in the future and the balance sheet item is expected to provide future economic benefits, a DTA or DTL is created.

Temporary differences can be **taxable temporary differences** that result in expected future taxable income or **deductible temporary differences** that result in expected future tax deductions.

Temporary differences leading to DTLs can arise from an investment in another firm (e.g., subsidiaries, affiliates, branches, and joint ventures) when the parent company recognizes earnings from the investment before dividends are received. However, if the parent company can control the timing of the future reversal and it is probable the temporary difference will not reverse, no DTL is reported.

A temporary difference from an investment will result in a DTA only if the temporary difference is expected to reverse in the future, and sufficient taxable profits are expected to exist when the reversal occurs.

LOS 24.g: Explain recognition and measurement of current and deferred tax items.

The measurement of deferred tax items depends on the tax rate expected to be in force when the underlying temporary difference reverses. We previously noted the effects of a change in the income tax rate on deferred tax assets and liabilities. In some circumstances, the applicable tax rate will depend on how the temporary difference will be settled. As an example, consider a tax jurisdiction that has a capital gains tax rate that is lower than the marginal tax rate. If, given its tax base, the currently unrealized gains on an asset will be taxed at the capital gains rate when the asset is disposed of, that rate should be used to calculate the deferred tax liability.

Another issue with the measurement of deferred tax items is whether a change in asset value is recorded on the income statement or taken directly to equity. In a case where the change that leads to a deferred tax item is taken directly to equity, the deferred tax item should also be taken directly to equity.

Consider a company reporting under IFRS that revalues PP&E upward. The revaluation gain is taken directly to equity without affecting either pretax income (on the income statement) or taxes payable (the gain is unrealized) so balance sheet deferred tax liabilities are not affected. Because the revaluation gain is taken directly to equity, the related future tax liability should be taken to equity as well. The adjustment is to reduce the amount of the gain added to revaluation surplus by the amount of the future tax liability on the revaluation gain.

The upward revaluation of the asset on the balance sheet will increase depreciation in subsequent periods, but will not affect the deferred tax liability. The tax liability on the increase in book value is incorporated into the recognition of the increase in revaluation surplus (rather than increasing the reported DTL). In each subsequent period, an amount equal to the additional depreciation from the upward revaluation of the asset, less the tax liability on that portion of the revaluation, is transferred from revaluation surplus to retained earnings. The previously unrealized gain in the asset's value is *realized* over time through use of the asset. The addition to retained earnings just offsets the after-tax decrease in net income (and retained earnings) from the additional depreciation resulting from the upward revaluation of the asset's carrying value.

LOS 24.h: Describe the valuation allowance for deferred tax assets—when it is required and what effect it has on financial statements.

Although deferred taxes are created from temporary differences that are expected to reverse in the future, neither deferred tax assets nor deferred tax liabilities are carried on the balance sheet at their discounted present value. However, deferred tax assets are assessed at each balance sheet date to determine the likelihood of sufficient future taxable income to recover the tax assets. Without future taxable income, a DTA is worthless.

According to U.S. GAAP, if it is more likely than not (greater than a 50% probability) that some or all of a DTA will not be realized (insufficient future taxable income to recover the tax asset), then the DTA must be reduced by a **valuation allowance**. The valuation allowance is a contra account that reduces the net balance sheet value of the DTA. Increasing the valuation allowance will decrease the net balance sheet DTA, increasing income tax expense and decreasing net income. Under IFRS, a similar calculation is made but only the net amount of the DTA is presented on the balance sheet. The amount of the valuation allowance is not separately disclosed.

If circumstances change, the net DTA can be increased by decreasing the valuation allowance. This would result in higher earnings.

It is up to management to defend the recognition of all deferred tax assets. If a company has order backlogs or existing contracts which are expected to generate future taxable income, a valuation allowance might not be necessary. However, if a company has cumulative losses over the past few years or a history of inability to use tax loss carryforwards, then the company would need to use a valuation allowance to reflect the likelihood that a deferred tax asset will never be realized.

Because an increase (decrease) in the valuation allowance will decrease (increase) earnings, management can manipulate earnings by changing the valuation allowance.

Whenever a company reports substantial deferred tax assets, an analyst should review the company's financial performance to determine the likelihood that those assets will be realized. Analysts should also scrutinize changes in the valuation allowance to determine whether those changes are economically justified.



PROFESSOR'S NOTE

A valuation allowance account is only used for deferred tax assets. Under U.S. GAAP, deferred tax assets and deferred tax liabilities appear separately on the balance sheet, and they are not typically netted.

LOS 24.i: Analyze disclosures relating to deferred tax items and the effective tax rate reconciliation and explain how information included in these disclosures affects a company's financial statements and financial ratios.

Companies are required to disclose details on the source of the temporary differences that cause the deferred tax assets and liabilities reported on the balance sheet. Changes in those balance sheet accounts are reflected in income tax expense on the income statement. Here are some common examples of temporary differences you may encounter:

- A deferred tax liability results from using accelerated *depreciation* for tax purposes and straight-line depreciation for the financial statements. The analyst should consider the firm's growth rate and capital spending levels when determining whether the difference will actually reverse.
- *Impairments* generally result in a deferred tax asset since the write-down is recognized immediately in the income statement, but the deduction on the tax return is generally not allowed until the asset is sold or disposed of.

- *Restructuring* generates a deferred tax asset because the costs are recognized for financial reporting purposes when the restructuring is announced, but not deducted for tax purposes until actually paid. Note that restructuring usually results in significant cash outflows (net of the tax savings) in the years after the restructuring costs are reported.
- In the United States, firms that use LIFO for their financial statements are required to use LIFO for tax purposes, so no temporary differences result. However, in countries where this is not a requirement, temporary differences can result from the *choice of inventory cost-flow method*.
- *Post-employment benefits* and *deferred compensation* are both recognized for financial reporting when earned by the employee but not deducted for tax purposes until actually paid. These can result in a deferred tax asset that will be reversed when the benefits or compensation are paid.
- A deferred tax adjustment is made to stockholders' equity to reflect the future tax impact of unrealized gains or losses on *available-for-sale marketable securities* that are taken directly to equity. No DTL is added to the balance sheet for the future tax liability when gains/losses are realized.

Typically, the following deferred tax information is disclosed:

- Deferred tax liabilities, deferred tax assets, any valuation allowance, and the net change in the valuation allowance over the period.
- Any unrecognized deferred tax liability for undistributed earnings of subsidiaries and joint ventures.
- Current-year tax effect of each type of temporary difference.
- Components of income tax expense.
- Reconciliation of reported income tax expense and the tax expense based on the statutory rate.
- Tax loss carryforwards and credits.

EXAMPLE: Analyzing deferred tax item disclosures

WCCO, Inc.'s income tax expense has consistently been larger than taxes payable over the last three years. WCCO disclosed in the footnotes to its 20X5 financial statements the major items recorded as deferred tax assets and liabilities (in millions of dollars), as shown in the following table.

Deferred Tax Disclosures in Footnotes to WCCO, Inc., Financial Statements

	20X5	20X4	20X3
Employee benefits	\$278	\$310	\$290
International tax loss carryforwards	101	93	115
Subtotal	379	403	405
Valuation allowance	(24)	(57)	(64)
Deferred tax asset	355	346	341
Property, plant, and equipment	452	361	320
Unrealized gains on available-for-sale securities	67	44	24
Deferred tax liability	519	405	343
Deferred income taxes	\$164	\$59	\$2

Use the table above to explain why income tax expense has exceeded taxes payable over the last three years. Also explain the effect of the change in the valuation allowance on WCCO's earnings for 20X5.

Answer:

The company's deferred tax asset balance results from international tax loss carryforwards and employee benefits (most likely pension and other post-retirement benefits), offset by a valuation allowance. The company's deferred tax liability balance results from property, plant, and equipment (most likely from using accelerated depreciation methods for tax purposes and straight-line on the financial statements) and unrealized gains on securities classified as available-for-sale (because the unrealized gain is not taxable until realized).

Income tax expense is equal to taxes payable plus deferred income tax expense. Because deferred tax liabilities have been growing faster than deferred tax assets, deferred income tax expense has been positive, resulting in income tax expense being higher than taxes payable.

Management decreased the valuation allowance by \$33 million in 20X5. This resulted in a reduction in deferred income tax expense and an increase in reported earnings for 20X5.

Analyzing the Effective Tax Rate Reconciliation

Some firms' reported income tax expense differs from the amount based on the statutory income tax rate. Recall that the statutory rate is the tax rate of the jurisdiction where the firm operates. The differences are generally the result of:

- Different tax rates in different tax jurisdictions (countries).
- Permanent tax differences: tax credits, tax-exempt income, nondeductible expenses, and tax differences between capital gains and operating income.
- Changes in tax rates and legislation.
- Deferred taxes provided on the reinvested earnings of foreign and unconsolidated domestic affiliates.
- Tax holidays in some countries (watch for special conditions such as termination dates for the holiday or a requirement to pay the accumulated taxes at some point in the future).

Understanding the differences between reported income tax expense and the amount based on the statutory income tax rate will enable the analyst to better estimate future earnings and cash flow.

When estimating future earnings and cash flows, the analyst should understand each element of the reconciliation, including its relative impact, how it has changed with time, and how it is

likely to change in the future.

In analyzing trends in tax rates, it is important to only include reconciliation items that are continuous in nature rather than those that are sporadic. Items including different rates in different countries, tax-exempt income, and non-deductible expenses tend to be continuous. Other items are almost always sporadic, such as the occurrence of large asset sales and tax holiday savings. The disclosures of each financial statement should be reviewed based on the footnotes and management discussion and analysis.

EXAMPLE: Analyzing the tax rate reconciliation

Novelty Distribution Company (NDC) does business in the United States and abroad. The company's reconciliation between effective and statutory tax rates for three years is provided in the following figure. Analyze the trend in effective tax rates over the three years shown.

Statutory U.S. Federal Income Tax Rate Reconciliation

	20X3	20X4	20X5
Statutory U.S. federal income tax rate	35.0%	35.0%	35.0%
State income taxes, net of related federal income tax benefit	2.1%	2.2%	2.3%
Benefits and taxes related to foreign operations	(6.5%)	(6.3%)	(2.7%)
Tax rate changes	0.0%	0.0%	(2.0%)
Capital gains on sale of assets	0.0%	(3.0%)	0.0%
Special items	(1.6%)	8.7%	2.5%
Other, net	0.8%	0.7%	(1.4%)
Effective income tax rates	29.8%	37.3%	33.7%

Answer:

For some trend analysis, the analyst may want to convert the reconciliation from percentages to absolute numbers. However, for this example, the trends can be analyzed simply by using the percentages.

The effective tax rate is upward trending over the 3-year period. Contributing to the upward trend is an increase in the state income tax rate and the loss of benefits related to taxes on foreign income. In 20X4, a loss related to the sale of assets partially offset an increase in taxes created by special items. In 20X3 and 20X5, the special items and the other items also offset each other. The fact that the special items and other items are so volatile over the 3-year period suggests that it will be difficult for an analyst to forecast the effective tax rate for NDC for the foreseeable future without additional information. This volatility also reduces comparability with other firms.

LOS 24.j: Identify the key provisions of and differences between income tax accounting under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (GAAP).

Accounting for income taxes under U.S. GAAP and IFRS is similar in most respects. However, there are some differences. Many differences relate to the different tax laws and regulations of the different countries. Figure 24.1 is a summary of a few of the more important differences.

Figure 24.1: Tax Accounting Differences, IFRS vs. U.S. GAAP

	U.S. GAAP	IFRS
Revaluation of fixed assets and intangible assets	Not applicable, no revaluation allowed.	Deferred taxes are recognized in equity.
Undistributed profit from an investment in a subsidiary	No deferred taxes for foreign subsidiaries that meet the indefinite reversal criterion. No deferred taxes for domestic subsidiaries if the amounts are tax free.	Deferred taxes are recognized unless the parent is able to control the distribution of profit and it is probable the temporary difference will not reverse in the future.
Undistributed profit from an investment in a joint venture (JV)	No deferred taxes for foreign corporate JVs that meet the indefinite reversal criterion.	Deferred taxes are recognized unless the venturer is able to control the sharing of profit and it is probable the temporary difference will not reverse in the future.
Undistributed profit from an investment in an associate firm	Deferred taxes are recognized from temporary differences.	Deferred taxes are recognized unless the investor is able to control the sharing of profit and it is probable the temporary difference will not reverse in the future.
Deferred tax asset recognition	Recognized in full and then reduced if "more likely than not" that some or all of the tax asset will not be realized.	Recognized if "probable" that sufficient taxable profit will be available to recover the tax asset.
Tax rate used to measure deferred taxes	Enacted tax rate only.	Enacted or substantively enacted tax rate.

MODULE QUIZ 24.4



- While reviewing a company, an analyst identifies a permanent difference between taxable income and pretax income. Which of the following statements *most accurately* identifies the appropriate financial statement adjustment?
 - The amount of the tax implications of the difference should be added to the deferred tax liabilities.
 - The present value of the amount of the tax implications of the difference should be added to the deferred tax liabilities.
 - No financial statement adjustment is necessary.
- A U.S. GAAP reporting firm reports an increased valuation allowance at the end of the current period. What effect will this have on the firm's income tax expense in the current period?
 - Increase.
 - Decrease.
 - No effect.
- KLH Company reported the following:
 - Gross DTA at the beginning of the year \$10,500
 - Gross DTA at the end of the year \$11,250
 - Valuation allowance at the beginning of the year \$2,700
 - Valuation allowance at the end of the year \$3,900

Which of the following statements *best* describes the expected earnings of the firm? Earnings are expected to:

- increase.
- decrease.
- remain relatively stable.

4. According to IFRS, the deferred tax consequences of revaluing held-for-use equipment upward is reported on the balance sheet:
- as an asset.
 - as a liability.
 - in stockholders' equity.

KEY CONCEPTS

LOS 24.a

Deferred tax terminology:

- **Taxable income.** Income subject to tax based on the tax return.
- **Accounting profit.** Pretax income from the income statement based on financial accounting standards.
- **Deferred tax assets.** Balance sheet asset value that results when taxes payable (tax return) are greater than income tax expense (income statement) and the difference is expected to reverse in future periods.
- **Deferred tax liabilities.** Balance sheet liability value that results when income tax expense (income statement) is greater than taxes payable (tax return) and the difference is expected to reverse in future periods.
- **Valuation allowance.** Reduction of deferred tax assets (contra account) based on the likelihood that the future tax benefit will not be realized.
- **Taxes payable.** The tax liability from the tax return. Note that this term also refers to a liability that appears on the balance sheet for taxes due but not yet paid.
- **Income tax expense.** Expense recognized in the income statement that includes taxes payable and changes in deferred tax assets and liabilities.

LOS 24.b

A *deferred tax liability* is created when income tax expense (income statement) is higher than taxes payable (tax return). Deferred tax liabilities occur when revenues (or gains) are recognized in the income statement before they are taxable on the tax return, or expenses (or losses) are tax deductible before they are recognized in the income statement.

A *deferred tax asset* is created when taxes payable (tax return) are higher than income tax expense (income statement). Deferred tax assets are recorded when revenues (or gains) are taxable before they are recognized in the income statement, when expenses (or losses) are recognized in the income statement before they are tax deductible, or when tax loss carryforwards are available to reduce future taxable income.

Deferred tax liabilities that are not expected to reverse, typically because of expected continued growth in capital expenditures, should be treated for analytical purposes as equity. If deferred tax liabilities are expected to reverse, they should be treated for analytical purposes as liabilities.

LOS 24.c

If taxable income is less than pretax income and the cause of the difference is expected to reverse in future years, a DTL is created. If taxable income is greater than pretax income and the

difference is expected to reverse in future years, a DTA is created.

The balance of the DTA or DTL is equal to the difference between the tax base and the carrying value of the asset or liability, multiplied by the tax rate.

Income tax expense and taxes payable are related through the change in the DTA and the change in the DTL:

$$\text{income tax expense} = \text{taxes payable} + \Delta\text{DTL} - \Delta\text{DTA}.$$

LOS 24.d

An asset's tax base is its value for tax purposes. The tax base for a depreciable fixed asset is its cost minus any depreciation or amortization previously taken on the tax return. When an asset is sold, the taxable gain or loss on the sale is equal to the sale price minus the asset's tax base.

A liability's tax base is its value for tax purposes. When there is a difference between the book value of a liability on a firm's financial statements and its tax base that will result in future taxable gains or losses when the liability is settled, the firm will recognize a deferred tax asset or liability to reflect this future tax or tax benefit.

LOS 24.e

When a firm's income tax rate increases (decreases), deferred tax assets and deferred tax liabilities are both increased (decreased) to reflect the new rate. Changes in these values will also affect income tax expense.

An increase in the tax rate will increase both a firm's DTL and its income tax expense. A decrease in the tax rate will decrease both a firm's DTL and its income tax expense.

An increase in the tax rate will increase a firm's DTA and decrease its income tax expense. A decrease in the tax rate will decrease a firm's DTA and increase its income tax expense.

LOS 24.f

A temporary difference is a difference between the tax base and the carrying value of an asset or liability that will result in taxable amounts or deductible amounts in the future.

A permanent difference is a difference between taxable income and pretax income that will not reverse in the future. Permanent differences do not create DTAs or DTLs.

LOS 24.g

Measurement of deferred tax items depends on the tax rate expected to be in force when the underlying temporary difference reverses. The applicable tax may depend on how the temporary difference will be settled (e.g., if a capital gains tax rate will apply). If a change that leads to a deferred tax item is taken directly to equity, such as an upward revaluation, the deferred tax item should also be taken directly to equity.

LOS 24.h

If it is more likely than not that some or all of a DTA will not be realized (because of insufficient future taxable income to recover the tax asset), then the DTA must be reduced by a valuation allowance. The valuation allowance is a contra account that reduces the DTA value on the balance sheet. Increasing the valuation allowance will increase income tax expense and reduce

earnings. If circumstances change, the DTA can be revalued upward by decreasing the valuation allowance, which would increase earnings.

LOS 24.i

Firms are required to reconcile their effective income tax rate with the applicable statutory rate in the country where the business is domiciled. Analyzing trends in individual reconciliation items can aid in understanding past earnings trends and in predicting future effective tax rates. Where adequate data is provided, they can also be helpful in predicting future earnings and cash flows or for adjusting financial ratios.

LOS 24.j

The accounting treatment of income taxes under U.S. GAAP and their treatment under IFRS are similar in most respects. One major difference relates to the revaluation of fixed assets and intangible assets. U.S. GAAP prohibits upward revaluations, but they are permitted under IFRS and any resulting effects on deferred tax are recognized in equity.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 24.1, 24.2

1. **C** Pretax income and income tax expense are not always linked because of temporary and permanent differences. (LOS 24.a)
2. **C** The difference between taxes payable for the period and the tax expense recognized on the financial statements results from differences between financial and tax accounting. (LOS 24.b)
3. **A** The DTL is not expected to reverse in the foreseeable future because a growing firm is expected to continue to increase its investment in depreciable assets, and accelerated depreciation for tax on the newly acquired assets delays the reversal of the DTL. The liability should be treated as equity at its full value. (LOS 24.b)
4. **B** Annual depreciation expense for tax purposes is $(\$120,000 \text{ cost} - \$0 \text{ salvage value}) / 3 \text{ years} = \$40,000$. Taxable income is $\$50,000 - \$40,000 = \$10,000$. (LOS 24.c)
5. **A** Taxes payable is taxable income \times tax rate = $\$10,000 \times 40\% = \$4,000$. (The \$10,000 was calculated in the previous question.) (LOS 24.c)
6. **C** Annual depreciation expense for financial purposes is $(\$120,000 \text{ cost} - \$0 \text{ salvage value}) / 4 \text{ years} = \$30,000$. Pretax income is $\$50,000 - \$30,000 = \$20,000$. (LOS 24.c)
7. **C** Because there has been no change in the tax rate, income tax expense is pretax income \times tax rate = $\$20,000 \times 40\% = \$8,000$. (The \$20,000 was calculated in the previous question.) (LOS 24.c)
8. **C** Note that the asset was fully depreciated for tax purposes after year 3, so taxable income is \$50,000. Taxes payable for year 4 = taxable income \times tax rate = $\$50,000 \times 40\% = \$20,000$. (LOS 24.c)

9. **C** At the end of year 2, the tax base is \$40,000 (\$120,000 cost – \$80,000 accumulated tax depreciation) and the carrying value is \$60,000 (\$120,000 cost – \$60,000 accumulated financial depreciation). Since the carrying value exceeds the tax base, a DTL of \$8,000 $[(\$60,000 \text{ carrying value} - \$40,000 \text{ tax base}) \times 40\%]$ is reported. (LOS 24.c)
10. **A** If the tax base of an asset exceeds the carrying value, a deferred tax asset is created. Taxable income will be lower in the future when the reversal occurs. (LOS 24.d)
11. **A** For revenue received in advance, the tax base is equal to the carrying value minus any amounts that will *not* be taxed in the future. Since the advance has already been taxed, \$100,000 will *not* be taxed in the future. Thus, the textbook advance liability has a tax base of \$0 (\$100,000 carrying value – \$100,000 revenue not taxed in the future). (LOS 24.d)

Module Quiz 24.3

1. **B** If tax rates increase, the balance sheet value of a deferred tax asset will also increase. (LOS 24.e)
2. **C** If the tax rate decreases, balance sheet DTL and DTA are both reduced. Taxable income is unaffected. (LOS 24.e)

Module Quiz 24.4

1. **C** No analyst adjustment is needed. If a permanent difference between taxable income and pretax income is identifiable, the difference will be reflected in the firm's effective tax rate. (LOS 24.f)
2. **A** Recognizing a greater valuation allowance reduces the net value of a deferred tax asset, which increases income tax expense in the current period. (LOS 24.h)
3. **B** The valuation allowance account increased from \$2,700 to \$3,900. The most likely explanation is the future earnings are expected to decrease, thereby reducing the value of the DTA. (LOS 24.i)
4. **C** The deferred tax consequences of revaluing an asset upward under IFRS are reported in stockholders' equity. (LOS 24.j)

Reading 25

NON-CURRENT (LONG-TERM) LIABILITIES

EXAM FOCUS

Candidates must understand the financial statement effects of issuing a bond at par, at a discount, or at a premium. You must be able to calculate the book value of the bond and interest expense at any point in time using the effective interest rate method. Also, be able to calculate the gain or loss from retiring a bond before its maturity date. You must thoroughly understand how the classification of a lease as either an operating or finance lease affects the balance sheet, income statement, and cash flow statement from both the lessee and lessor perspectives. Be able to distinguish between the two types of pension plans and identify the financial statement reporting of a defined benefit plan. Finally, be able to evaluate a firm's solvency using the various leverage and coverage ratios.

MODULE 25.1: BOND ISSUANCE



A **bond** is a contractual promise between a borrower (the bond issuer) and a lender (the bondholder) that obligates the bond issuer to make payments to the bondholder over the term of the bond. Typically, two types of payments are involved: (1) periodic interest payments, and (2) repayment of principal at maturity.

Video covering
this content is
available online.

Bond Terminology

- The **face value**, also known as the **maturity value** or **par value**, is the amount of principal that will be paid to the bondholder at maturity. The face value is used to calculate the coupon payments.
- The **coupon rate** is the interest rate stated in the bond that is used to calculate the coupon payments.
- The **coupon payments** are the periodic interest payments to the bondholders and are calculated by multiplying the face value by the coupon rate.
- The **effective rate of interest** is the interest rate that equates the present value of the future cash flows of the bond and the issue price. The effective rate is the market rate of interest required by bondholders and depends on the bond's risks (e.g., default risk, liquidity risk), as well as the overall structure of interest rates and the timing of the bond's cash flows. *Do not confuse the market rate of interest with the coupon rate.* The coupon rate is typically fixed for the term of the bond. The market rate of interest on a firm's bonds, however, will likely change over the bond's life, which changes the bond's market value as well.

- The **balance sheet liability** of a bond is equal to the present value of its remaining cash flows (coupon payments and face value), discounted at the market rate of interest *at issuance*. At maturity, the liability will equal the face value of the bond. The balance sheet liability is also known as the book value or carrying value of the bond.
- The interest expense reported in the income statement is calculated by multiplying the book value of the bond liability at the beginning of the period by the market rate of interest of the bond when it was issued.

At the date of issuance, the market rate of interest may be equal to, less than, or greater than the coupon rate.

- When the market rate is equal to the coupon rate, the bond is a par bond (priced at face value).
- When the market rate is greater than the coupon rate, the bond is a discount bond (priced below par).
- When the market rate is less than the coupon rate, the bond is a premium bond (priced above par).

LOS 25.a: Determine the initial recognition, initial measurement and subsequent measurement of bonds.

Bonds Issued at Par

When a bond is issued at par, the bond's yield at issuance is equal to the coupon rate. In this case, the present value of the coupon payments plus the present value of the face amount is equal to the par value. The effects on the financial statements are straightforward:

- On the balance sheet, assets and liabilities increase by the bond proceeds (face value). The book value of the bond liability will not change over the term of the bond.
- On the income statement, interest expense for the period is equal to the coupon payment because the yield at issuance and the coupon rate are the same.
- On the cash flow statement, the issue proceeds are reported as a cash inflow from financing activities and the coupon payments are reported as cash outflows from operating activities (under U.S. GAAP; they may be reported as CFO or CFF outflows under IFRS). At maturity, repayment of the face value is reported as a cash outflow from financing activities.

Bonds Issued at a Discount or Premium

When the bond's yield at issuance is not equal to the coupon rate, the proceeds received (the present value of the coupon payments plus the present value of the face value) are not equal to par value. In this case, the bond is issued at a *premium* or a *discount*. The premium or discount at the issue date is usually relatively small for coupon bonds.

If the coupon rate is less than the bond's yield, the proceeds received will be less than the face value. The difference is known as a *discount*. The coupon rate is lower than the coupon rate that would make the market price of the bond equal to its par value. Investors will pay less than face value because of the lower coupon rate. Such bonds are known as *discount bonds*.

If the coupon rate is greater than the bond's yield, the bond price and the proceeds received will be greater than face value. We refer to such bonds as **premium bonds**. In this case, investors will pay more for the above-market coupon payments.

Balance Sheet Measurement

When a company issues a bond, assets and liabilities both initially increase by the bond proceeds. At any point in time, the book value of the bond liability will equal the present value of the remaining future cash flows (coupon payments and face value) discounted at the bond's yield at issuance.



PROFESSOR'S NOTE

Interest expense and the book value of a bond liability are calculated using the bond's yield at the time it was issued, not its yield today. This is a critical point.

A premium bond is reported on the balance sheet at more than its face value. As the premium is amortized (reduced), the book value of the bond liability will decrease until it reaches the face value of the bond at maturity.

A discount bond is reported on the balance sheet at less than its face value. As the discount is amortized, the book value of the bond liability will increase until it reaches face value at maturity.

MODULE 25.2: DISCOUNT AND PREMIUM BONDS



LOS 25.b: Describe the effective interest method and calculate interest expense, amortisation of bond discounts/premiums, and interest payments.

Video covering this content is available online.

For a bond issued at a premium or discount, interest expense and coupon interest payments are not equal. Interest expense includes amortization of any discount or premium. Using the **effective interest rate method**, interest expense is equal to the book value of the bond liability at the beginning of the period, multiplied by the bond's yield at issuance.

- For a premium bond, interest expense is less than the coupon payment ($\text{yield} < \text{coupon rate}$). The difference between interest expense and the coupon payment is the amortization of the premium. The premium amortization is subtracted each period from the bond liability on the balance sheet. Thus, interest expense will decrease over time as the bond liability decreases.
- For a discount bond, interest expense is greater than the coupon payment ($\text{yield} > \text{coupon rate}$). The difference between interest expense and the coupon payment is the amortization of the discount. The amortization of the discount each period is added to the bond liability on the balance sheet. Therefore, interest expense will increase over time as the bond liability increases.



PROFESSOR'S NOTE

In the case of a discount bond, the coupon is too low relative to the required rate of return of the market. The purposes of amortizing the discount are to (1) increase the book value of the bond

liability over time, and (2) increase interest expense so that the coupon payment plus discount amortization is approximately equal to the interest expense that would have prevailed had the bond been issued at par. Conversely, amortizing a premium decreases the book value of the bond liability over time and decreases interest expense.

The effective interest rate method of amortizing a discount or premium is required under IFRS. Under U.S. GAAP, the effective interest rate method is preferred, but the straight-line method is allowed if the results are not materially different. The straight-line method is similar to straight-line depreciation in that the total discount or premium at issuance is amortized by equal amounts each period over the life of the bond.

While coupon interest is paid in cash, amortization is a noncash item. When presenting the cash flow statement using the indirect method, net income must be adjusted to remove the effects of any amortization of a discount or premium in order to calculate cash flow from operations.

Firms that follow U.S. GAAP must report cash interest paid in the cash flow statement as an operating cash flow. Firms that follow IFRS can report cash interest paid as either an operating or financing cash flow. Therefore, it may be necessary to reclassify interest paid when comparing firms that follow different standards.



PROFESSOR'S NOTE

Some analysts believe classifying interest expense as an operating activity is inconsistent with treating the bond proceeds as a financing activity. In addition, treating interest expense as an operating activity incorrectly describes the economics of a bond issued at a premium or discount. For bonds issued at a discount, cash flow from operations is overstated. This is because the coupon payment is reported as an operating cash flow, while the discount, when paid (as part of a bond's maturity payment), is reported as a financing cash flow. Stated differently, had the firm issued the bond at par, the coupon payment would have been higher to match the market rate of interest. Reclassifying interest as a financing activity in the cash flow statement corrects this inconsistent treatment.

EXAMPLE: Book values and cash flows

On December 31, 20X2, a company issued a 3-year, 10% annual coupon bond with a face value of \$100,000. Calculate the book value of the bond at year-end 20X2, 20X3, and 20X4, and the interest expense for 20X3, 20X4, and 20X5, assuming the bond was issued at a market rate of interest of (a) 10%, (b) 9%, and (c) 11%.

Answer:

(a) *Bond issued at par.* If the market rate of interest at issuance is 10%, the book value of the bonds will always be \$100,000, and the interest expense will always be \$10,000, which is equal to the coupon payment of $0.10 \times \$100,000$. There is no discount or premium to amortize.

(b) *Premium bond.* If the market rate of interest is 9%, the present value of the cash payments (a 3-year annuity of \$10,000 and a payment in three years of \$100,000) is \$102,531:

$$N = 3; PMT = 10,000; FV = 100,000; I/Y = 9; CPT \rightarrow PV = -\$102,531$$



PROFESSOR'S NOTE

The present value computed in this manner will have a minus sign.

The following table shows the interest expense and book value at the end of each year.

Interest Expense and Book Value for a Premium Bond

Year	(1) Beginning Book Value	(2) Interest Expense (1) \times 9%	(3) Coupon	(4) Ending Book Value (1) + (2) - (3)
20X3	\$102,531	\$9,228	\$10,000	\$101,759
20X4	101,759	9,158	10,000	100,917
20X5	100,917	9,083	10,000	100,000

The premium amortization for 20X3 is $10,000 - 9,228 = \$772$. For 20X4, the amortization is $10,000 - 9,158 = \$842$. Finally, for 20X5, premium amortization is \$917. Note that the premium has been fully amortized upon maturity so that the book value of the bond equals par value.

(c) *Discount bond*. If the market rate of interest is 11%, the present value of the cash payments (a 3-year annuity of \$10,000 and a payment in three years of \$100,000) is \$97,556.

$$N = 3; PMT = 10,000; FV = 100,000; I/Y = 11; CPT \rightarrow PV = \$97,556$$

The following table shows the interest expense and book value at the end of each year.

Interest Expense and Book Value for a Discount Bond

Year	(1) Beginning Book Value	(2) Interest Expense (1) \times 11%	(3) Coupon	(4) Ending Book Value (1) + (2) - (3)
20X3	\$97,556	\$10,731	\$10,000	\$98,287
20X4	98,287	10,812	10,000	99,099
20X5	99,099	10,901	10,000	100,000

Again, the pattern of discount amortization is such that the discount is fully amortized upon maturity, when the book value of the bond equals par value.

Zero-coupon bonds make no periodic interest payments. A zero-coupon bond, also known as a *pure-discount bond*, is issued at a discount from its par value and its annual interest expense is implied, but not explicitly paid. The actual interest payment is included in the face value that is paid at maturity. The effects of zero-coupon bonds on the financial statements are qualitatively the same as any discount bond, but the impact is larger because the discount is larger.

EXAMPLE: Accounting for a zero-coupon bond

On December 31, 20X0, Vine Corp. issues a three-year, zero-coupon bond with a par value of \$1,000 when the market interest rate is 12%. Using the effective interest method and an annual periodicity, calculate interest expense and the book value of the bond liability at the end of 20X1, 20X2, and 20X3.

Answer:

Cash received at issuance—and the initial book value of the bond liability—is \$711.78: $N = 3; I/Y = 12; PMT = 0; FV = 1,000; CPT \rightarrow PV = -711.78$

Interest expense each year is equal to market interest rate at issuance times the book value of the liability at the beginning of the year. Because the bond does not pay cash coupon interest, the entire amount of interest expense amortizes the discount.

	(1) Beginning Book Value	(2) Interest Expense (1) \times 12%	(3) Ending Book Value (1) + (2)
20X1	\$711.78	\$85.41	\$797.19
20X2	797.19	95.66	892.85
20X3	892.85	107.15	1,000.00



MODULE QUIZ 25.1, 25.2

1. According to U.S. GAAP, the coupon payment on a bond is reported as:
 - A. an operating cash outflow.
 - B. a financing cash outflow.
 - C. part operating cash outflow and part financing cash outflow.
 2. Using the effective interest rate method, the reported interest expense of a bond issued at a premium will:
 - A. decrease over the term of the bond.
 - B. increase over the term of the bond.
 - C. remain unchanged over the term of the bond.
- Use the following data to answer Questions 3 through 9.
- A firm issues a \$10 million bond with a 6% coupon rate, 4-year maturity, and annual interest payments when market interest rates are 7%.
3. The bond can be classified as a:
 - A. discount bond.
 - B. par bond.
 - C. premium bond.
 4. The annual coupon payments will each be:
 - A. \$600,000.
 - B. \$676,290.
 - C. \$700,000.
 5. Total of all cash payments to the bondholders is:
 - A. \$12,400,000.
 - B. \$12,738,721.
 - C. \$12,800,000.
 6. The initial book value of the bonds is:
 - A. \$9,400,000.
 - B. \$9,661,279.
 - C. \$10,000,000.
 7. For the first period the interest expense is:
 - A. \$600,000.
 - B. \$676,290.
 - C. \$700,000.
 8. If the market rate changes to 8% and the bonds are carried at amortized cost, the book value of the bonds at the end of the first year will be:
 - A. \$9,484,581.
 - B. \$9,661,279.
 - C. \$9,737,568.
 9. The total interest expense reported by the issuer over the life of the bond will be:
 - A. \$2,400,000.
 - B. \$2,738,721.
 - C. \$2,800,000.

MODULE 25.3: ISSUANCE COST, DERECOGNITION, AND DISCLOSURES

Issuance Costs



Video covering
this content is
available online.

Issuing a bond involves legal and accounting fees, printing costs, sales commissions, and other fees. Under IFRS and U.S. GAAP, the initial bond liability on the balance sheet (the proceeds from issuing the bond) is reduced by the amount of issuance costs, increasing the bond's effective interest rate. In effect, issuance costs are treated as unamortized discount.

Before 2016, under U.S. GAAP, issuance costs were capitalized as an asset and allocated to the income statement over the life of the bond. Although the treatment of issuance costs has now converged, U.S. GAAP still permits the earlier treatment.

Under both U.S. GAAP and IFRS, bond issuance costs (an outflow) are usually netted against the bond proceeds (an inflow) and reported on the cash flow statement as a financing cash flow.

Fair Value Reporting Option

Recall that the book value of a bond liability is based on its market yield *at issuance*. So, as long as the bond's yield does not change, the bond liability represents fair (market) value. However, if the yield changes, the balance sheet liability is no longer equal to fair value.

An *increase* in the bond's yield will result in a *decrease* in the fair value of the bond liability. Conversely, a *decrease* in the bond's yield *increases* its fair value. Changes in yield result in a divergence between the book value of the bond liability and the fair value of the bond. The fair value of the bond is the economic liability at a point in time.

IFRS and U.S. GAAP give firms the irrevocable option to report debt at fair value. Under this option, gains (decreases in bond liability) and losses (increases in bond liability) that result from changes in bonds' market yields are reported in the income statement.

For analysis, the market value of a firm's debt may be more appropriate than its book value. For example, a firm that issued a bond when interest rates were low is relatively better off when interest rates increase. This is because the firm could repurchase the bond at its now-lower market value. Decreasing the bond liability on the balance sheet to market value increases equity and decreases the debt-to-assets and debt-to-equity ratios. If interest rates have decreased since issuance, adjusting debt to its market value will have the opposite effects.

Summary of Financial Statement Effects of Issuing a Bond

Figure 25.1: Cash Flow Impact of Issuing a Bond

	Cash Flow From Financing	Cash Flow From Operations
Issuance of debt	Increased by cash received (Present value of the bond at the market interest rate)	No effect
Periodic interest payments	No effect	Decreased by interest paid [(coupon rate) × (face or par value)]
Payment at maturity	Decreased by face (par) value	No effect

Figure 25.2: Income Statement Impact of Issuing a Bond

$$\text{interest expense} = \left(\frac{\text{market rate}}{\text{at issue}} \right) \times \left(\frac{\text{balance sheet value of}}{\text{liability at beginning of period}} \right)$$

Issued at Par	Issued at a Premium	Issued at a Discount
Market rate = coupon rate	Market rate < coupon rate	Market rate > coupon rate
Interest expense = coupon rate \times face value = cash paid	Interest expense = cash paid – amortization of premium	Interest expense = cash paid + amortization of discount
Interest expense is constant	Interest expense decreases over time	Interest expense increases over time

Figure 25.3: Balance Sheet Impact of Issuing a Bond

Issued at Par	Issued at a Premium	Issued at a Discount
Carried at face value	Carried at face value plus premium The liability decreases as the premium is amortized to interest expense	Carried at face value less discount The liability increases as the discount is amortized to interest expense

LOS 25.c: Explain the derecognition of debt.

When bonds mature, no gain or loss is recognized by the issuer. At maturity, any original discount or premium has been fully amortized; thus, the book value of a bond liability and its face value are the same. The cash outflow to repay a bond is reported in the cash flow statement as a financing cash flow.

A firm may choose to **redeem** bonds before maturity because interest rates have fallen, because the firm has generated surplus cash through operations, or because funds from the issuance of equity make it possible (and desirable).

When bonds are redeemed before maturity, a gain or loss is recognized by subtracting the redemption price from the book value of the bond liability at the reacquisition date. For example, consider a firm that reacquires \$1 million face amount of bonds at 102% of par when the carrying value of the bond liability is \$995,000. The firm will recognize a loss of \$25,000 (\$995,000 carrying value – \$1,020,000 redemption price). Had the carrying value been greater than the redemption price, the firm would have recognized a gain.

If the redeemed bonds' issuance costs were capitalized, any remaining unamortized costs must be written off and included in the gain or loss calculation. No separate entry is necessary if the issuance costs were accounted for in the initial bond liability, because in that case no separate asset representing unamortized issuance costs would have been created.

Any gain or loss from redeeming debt is reported in the income statement, usually as a part of continuing operations, and additional information is disclosed separately. Redeeming debt is usually not a part of the firm's day-to-day operations; thus, analysts often eliminate the gain or loss from the income statement for analysis and forecasting.

When presenting the cash flow statement using the indirect method, any gain (loss) is subtracted from (added to) net income in calculating cash flow from operations. The redemption price is reported as an outflow from financing activities.

LOS 25.d: Describe the role of debt covenants in protecting creditors.

Debt covenants are restrictions imposed by the lender on the borrower to protect the lender's position. Debt covenants can reduce default risk and thus reduce borrowing costs. The restrictions can be in the form of affirmative covenants or negative covenants.

With **affirmative covenants**, the borrower promises to do certain things, such as:

- Make timely payments of principal and interest.
- Maintain certain ratios (such as the current, debt-to-equity, and interest coverage ratios) in accordance with specified levels.
- Maintain collateral, if any, in working order.

With **negative covenants**, the borrower promises to refrain from certain activities that might adversely affect its ability to repay the outstanding debt, such as:

- Increasing dividends or repurchasing shares.
- Issuing more debt.
- Engaging in mergers and acquisitions.

The bondholders can demand immediate repayment of principal if the firm violates a covenant (referred to as **technical default**). Analyzing the covenants is a necessary component of the credit analysis of a bond. Bond covenants are typically discussed in the financial statement footnotes.

Covenants protect bondholders from actions the firm may take that would harm the value of the bondholders' claims to the firm's assets and earnings (i.e., decrease credit quality). To the extent that covenants restrict, for example, the firm's ability to invest, take on additional debt, or pay dividends, analysis of covenants can be important when valuing the firm's equity (especially involving its growth prospects) as well as when analyzing and valuing its debt securities.



PROFESSOR'S NOTE

Debt covenants are described further in Fixed Income.

LOS 25.e: Describe the financial statement presentation of and disclosures relating to debt.

Firms will often report all of their outstanding long-term debt on a single line on the balance sheet. The portion that is due within the next year is reported as a current liability. The firm separately discloses more detail about its long-term debt in the footnotes. These disclosures are useful in determining the timing and amounts of future cash outflows. The footnote disclosure usually includes a discussion of:

- The nature of the liabilities.
- Maturity dates.
- Stated and effective interest rates.
- Call provisions and conversion privileges.
- Restrictions imposed by creditors.
- Assets pledged as security.

- The amount of debt maturing in each of the next five years.

A discussion of the firm's long-term debt is also found in the Management Discussion and Analysis section. This discussion is both quantitative, such as identifying obligations and commitments that are due in the future, and qualitative, such as discussing capital resource trends and material changes in the mix and cost of debt.

MODULE QUIZ 25.3

1. A firm has bonds with a \$10.0 million face value outstanding. The book value of the bond liability is \$10.2 million when the firm redeems the bonds for face value. Redeeming the bonds will result in:
 - a loss on the income statement and a financing cash outflow.
 - a gain on the income statement and a financing cash outflow.
 - a loss on the income statement and an operating cash outflow.
2. The purpose of debt covenants is *best* described as:
 - limiting issuance costs.
 - preventing technical default.
 - protecting the interests of creditors.
3. Which of the following is *least likely* to be disclosed in the financial statements of a bond issuer?
 - The amount of debt that matures in each of the next five years.
 - Collateral pledged as security in the event of default.
 - The market rate of interest on the balance sheet date.

MODULE 25.4: LEASE AND PENSION ACCOUNTING



Video covering this content is available online.

LOS 25.f: Explain motivations for leasing assets instead of purchasing them.

Instead of purchasing an asset, a firm may choose to lease the asset. With a lease, a firm (the **lessee**) essentially purchases the right to use an asset from another firm (the **lessor**) for a specified period, which can range from a month to many years. The lessee makes periodic payments to the lessor for the use of the asset. Thus, a lease can be considered an alternative to financing the purchase of an asset.

To be a lease, a contract must meet the following three requirements:

1. It must refer to a specific asset.
2. It must give the lessee effectively all the asset's economic benefits during the term of the lease.
3. It must give the lessee the right to determine how to use the asset during the term of the lease.

The advantages of leasing rather than purchasing an asset may include the following:

- *Less initial cash outflow.* Typically, a lease requires only a small down payment, if any.
- *Less costly financing.* Because a lease is effectively secured by the leased asset if the lessee defaults, the interest rate implicit in a lease contract may be less than the interest rate would be on a loan to purchase the asset.

- *Less risk of obsolescence.* At the end of a lease, the lessee often returns the leased asset to the lessor and therefore does not bear the risk of an unexpected decline in the asset's end-of-lease value.

LOS 25.g: Explain the financial reporting of leases from a lessee's perspective.

Under IFRS and U.S. GAAP, any lease in which both the benefits of ownership and the risks of ownership are substantially transferred to the lessee is classified as a **finance lease**. If either the benefits or the risks of ownership are not substantially transferred to the lessee, a lease is classified as an **operating lease**. Any given lease will be classified the same way by the lessee and the lessor.

Financial reporting standards require a lease to be classified as a finance lease if it meets any of the following five conditions:

1. Ownership of the leased asset transfers to the lessee.
2. The lessee has an option to buy the asset and is expected to exercise it.
3. The lease is for most of the asset's useful life.
4. The present value of the lease payments is greater than or equal to the asset's fair value.
5. The lessor has no other use for the asset.

Leases that are not classified as finance leases are classified as operating leases.

Under IFRS, all leases, except those that are short term (up to 12 months) or are of low value (up to \$5,000), require the lessee to record a **right-of-use asset** and a **lease liability** (both equal to the present value of the lease payments) on the balance sheet. The right-of-use asset is amortized over the term of the lease, with the amortization amount each period recorded on the income statement. The lease liability is reduced each period by the decrease in the principal portion outstanding that results from each lease payment. So while the lease asset and the liability both begin with the same value and reach zero at the end of the lease, they can have different values during the life of the lease, as the following example illustrates.

EXAMPLE: Accounting for a finance lease

Affordable Company leases a machine for its own use for four years with annual payments of \$10,000. At the end of the lease, which is also the end of the machine's useful life, Affordable will return the machine to the lessor. The interest rate implicit in the lease is 5%. Assuming that the right-of-use asset is amortized on a straight-line basis over the term of the lease, calculate the impact of the lease on Affordable's financial statements for each of the four years.

Answer:

The lease is classified as a finance lease because the lease is in effect for the asset's useful life.

The present value of the lease payments is:

$$N = 4; I/Y = 5; PMT = -10,000; FV = 0; CPT PV = 35,460$$

This amount will be recognized on the balance sheet as a right-of-use asset and as a lease liability.

The right-of-use asset will be amortized straight-line over the four years, decreasing each year by $\$35,460 / 4 = \$8,865$. This amount will be recognized each year on the income statement as amortization expense.

The lease liability will be treated as if it were an amortizing loan.

Year	(1) Beginning Liability	(2) Interest Expense $= (1) \times 5\%$	(3) Lease Payment	(4) Principal Repayment $= (3) - (2)$	(4) Ending Lease Liability $= (1) - (4)$	Book Value of Right-of- Use Asset
0					35,460	35,460
1	35,460	1,773	10,000	8,227	27,233	26,595
2	27,233	1,362	10,000	8,638	18,595	17,730
3	18,595	930	10,000	9,070	9,525	8,865
4	9,525	475	10,000	9,525	0	0

Interest expense will be recognized each year on the income statement, separately from the amortization expense for the right-of-use asset. On the balance sheet, the right-of-use asset value decreases by \$8,865 each year, and the lease liability is reduced by the principal repayment from column 4. Note that the book value of the right-of-use asset is less than the book value of the lease liability during the life of the lease. This is because the principal repayment in the early years of the lease is less than the straight-line amortization of the right-of-use asset. In the later years, the principal repayment is greater than the straight-line amortization, so that at the end of the lease, both the asset and the liability reach zero.

On the cash flow statement, repayment of principal will be classified as a cash outflow from financing. Under IFRS, the interest portion of each payment may be classified as either an operating or a financing cash outflow. Under U.S. GAAP, the interest portion is classified as an operating cash outflow.

Under U.S. GAAP, other than these differences in cash flow classification, a finance lease (that is not short term) is reported just as we have described for leases under IFRS.

For an *operating lease* (that is not short term) under U.S. GAAP, a lease liability is also recorded and amortized as under IFRS. However, the right-of-use asset is not amortized straight-line. Instead, it is amortized by the same amount each period as the decrease in the lease liability, so that the asset and the liability are equal in each period of the lease. On the income statement, interest expense and amortization of the right-of-use asset are not reported separately as they are for a finance lease; they are combined and reported as lease expense. On the cash flow statement, the full lease payment is classified as an operating cash outflow.

EXAMPLE: Accounting for an operating lease under U.S. GAAP

Using the same data from the previous example, calculate the impact of the lease on Affordable's financial statements if Affordable reports under U.S. GAAP and the lease is classified as an operating lease.

Answer:

The lease liability will be treated the same way as it is for a finance lease, but the book value of the right-of-use asset will remain equal to the book value of the lease liability.

Year	(1) Beginning Liability	(2) Interest Expense $= (1) \times 5\%$	(3) Lease Payment	(4) Principal Repayment $= (3) - (2)$	(4) Ending Lease Liability $= (1) - (4)$	Book Value of Right-of- Use Asset
0					35,460	35,460
1	35,460	1,773	10,000	8,227	27,233	27,233
2	27,233	1,362	10,000	8,638	18,595	18,595
3	18,595	930	10,000	9,070	9,525	9,525
4	9,525	475	10,000	9,525	0	0

On the balance sheet, the book value of the right-of-use asset is amortized by the same amount each period as the lease liability.

On the income statement, lease expense will equal interest plus amortization of the right-of-use asset. Because amortization is equal to the principal repayment each period, and the principal repayment equals the lease payment minus interest, lease expense each period is equal to the lease payment of \$10,000.

On the cash flow statement, the entire \$10,000 cash outflow is classified as cash from operations.

For short-term or low-value leases under IFRS, and for short-term leases under U.S. GAAP, no lease asset or liability is reported on the balance sheet. Each period, the lease payment is simply reported as rental expense on the income statement.

LOS 25.h: Explain the financial reporting of leases from a lessor's perspective.

Under both IFRS and U.S. GAAP, there are two lease classifications for lessors, finance leases and operating leases, just as for lessees. At the initiation of a finance lease, the lessor removes the leased asset from its balance sheet and adds a **lease receivable** asset, equal to the value of the expected lease payments. If this value is different from the asset's book value, the lessor will recognize a gain or a loss. Over the term of the lease, the lessor will use the effective interest method (the same method we have just seen for lessees) to amortize the lease receivable and will report the interest portion of the lease payments as income. This interest income is included in the lessor's revenues for the period if leasing is one of its primary business activities. On the cash flow statement, the entire cash inflow is classified as cash from operations.

For an operating lease, the lessor does not remove the leased asset from its balance sheet. The lessor will continue to record the depreciation expense over the life of the asset. On the income statement, the lessor reports the lease payments as income, while depreciation and other costs associated with leasing the asset are reported as expenses. As with a finance lease, the entire cash inflow is classified as cash from operations.

LOS 25.i: Compare the presentation and disclosure of defined contribution and defined benefit pension plans.



A **pension** is a form of deferred compensation earned over time through employee service. The most common pension arrangements are defined contribution plans and defined benefit plans.

Video covering this content is available online.

A **defined contribution plan** is a retirement plan in which the firm contributes a sum each period to the employee's retirement account. The firm's contribution can be based on any number of factors, including years of service, the employee's age, compensation, profitability, or even a percentage of the employee's contribution. In any event, the firm makes no promise to the employee regarding the future value of the plan assets. The investment decisions are left to the employee, who assumes all of the investment risk.

The financial reporting requirements for defined contribution plans are straightforward. Pension expense is simply equal to the employer's contribution. There is no future obligation to report on the balance sheet as a liability.

In a **defined benefit plan**, the firm promises to make periodic payments to employees after retirement. The benefit is usually based on the employee's years of service and the employee's compensation at, or near, retirement. For example, an employee might earn a retirement benefit of 2% of her final salary for each year of service. Consequently, an employee with 20 years of

service and a final salary of \$100,000, would receive \$40,000 ($\$100,000 \text{ final salary} \times 2\% \times 20 \text{ years of service}$) each year upon retirement until death. Because the employee's future benefit is defined, the employer assumes the investment risk.

A company that offers defined pension benefits typically funds the plan by contributing assets to a separate legal entity, usually a trust. The plan assets are managed to generate the income and principal growth necessary to pay the pension benefits as they come due.

Financial reporting for a defined benefit plan is much more complicated than for a defined contribution plan because the employer must estimate the value of the future obligation to its employees. The obligation involves forecasting a number of variables, such as future compensation levels, employee turnover, average retirement age, mortality rates, and an appropriate discount rate.

For a defined benefit plan, the **net pension asset** or **net pension liability** is a key element for analysis. If the fair value of the plan's assets is greater than the estimated pension obligation, the plan is said to be *overfunded* and the sponsoring firm records a net pension asset on its balance sheet. If the fair value of the plan's assets is less than the estimated pension obligation, the plan is *underfunded* and the firm records a net pension liability.

The change in the net pension asset or liability is recognized on the firm's financial statements each year. Some components are included in net income while others are recorded as other comprehensive income. The treatments of these cost components are similar under IFRS and U.S. GAAP. Component costs that go directly to equity as other comprehensive income are amortized to the income statement under U.S. GAAP, but under IFRS they are not amortized.



PROFESSOR'S NOTE

Accounting for defined benefit pension plans is addressed in more detail at Level II.

For manufacturing companies, under either IFRS or U.S. GAAP, pension expense is allocated to inventory and cost of goods sold for employees who provide direct labor to production and to salary or administrative expense for other employees. As a result, pension expense does not appear separately on the income statement for manufacturing companies. An analyst must examine the financial statement notes to find the details of these companies' pension expense.

LOS 25.j: Calculate and interpret leverage and coverage ratios.

Analysts use solvency ratios to measure a firm's ability to satisfy its long-term obligations. In evaluating solvency, analysts look at leverage ratios and coverage ratios.

Leverage Ratios

Leverage ratios focus on the balance sheet by measuring the amount of debt in a firm's capital structure. For calculating these ratios, "debt" refers to interest-bearing obligations. Non-interest-bearing liabilities, such as accounts payable, accrued liabilities, and deferred taxes, are not considered debt.

- *Debt-to-assets ratio* = total debt / total assets.
 - Measures the percentage of total assets financed with debt.
- *Debt-to-capital ratio* = total debt / (total debt + total equity).

Measures the percentage of total capital financed with debt. Debt-to-capital is similar to the debt-to-assets ratio, except that total capital excludes non-interest-bearing liabilities. Recall the balance sheet equation $A = L + E$. Thus, total assets and total capital differ by the firm's non-interest-bearing liabilities.

- *Debt-to-equity ratio* = total debt / total equity.

Measures the amount of debt financing relative to the firm's equity base. A firm whose debt-to-equity ratio is 1.0 has equal amounts of debt and equity. Stated differently, its debt-to-capital ratio is 50%.

- *Financial leverage ratio* = average total assets / average total equity.

Measure of leverage used in the DuPont formula.

All of these leverage ratios are interpreted similarly; that is, the higher the ratio, the higher the leverage. When comparing firms, analysts must remember that in some countries, debt financing is more popular than equity financing. Firms in these countries will have higher leverage.

Coverage Ratios

Coverage ratios focus on the income statement by measuring the sufficiency of earnings to repay interest and other fixed charges when due. Two popular coverage ratios are the interest coverage ratio and the fixed charge coverage ratio.

- *Interest coverage* = EBIT / interest payments.

A firm with lower interest coverage will have more difficulty meeting its interest payments.

- *Fixed charge coverage* = (EBIT + lease payments) / (interest payments + lease payments).

Similar to interest coverage ratio but more inclusive because operating lease payments are added to the numerator and denominator. Significant operating lease payments will reduce this ratio as compared to interest coverage. Fixed charge coverage is more meaningful for firms that engage in significant operating leases.

EXAMPLE: Leverage and coverage ratios

Westcliff Corporation is a hardware wholesaler. The following table shows selected information from Westcliff's most recent financial statements.

Liabilities and Equity	20X9	20X8
Accounts payable	\$360,000	\$310,000
Notes payable	385,200	321,100
Current maturities of long-term debt	60,000	60,000
Accrued liabilities	<u>90,800</u>	<u>117,600</u>
Total current liabilities	\$896,000	\$808,700
Long-term debt	740,000	800,000
Shareholders' equity	<u>727,600</u>	<u>588,700</u>
Total liabilities and equity	\$2,363,600	\$2,197,400

Partial Income Statement	20X9	20X8
Gross profit	\$610,500	\$580,800
Administrative expense	187,000	177,200
Lease expense	<u>24,000</u>	<u>22,800</u>
Earnings before interest and taxes	\$399,500	\$380,800
Interest expense	\$168,000	\$116,100

Discuss Westcliff's solvency using the appropriate leverage and coverage ratios.

Answer:

When evaluating solvency, accounts payable and accrued liabilities are not considered debt. Debt only includes interest-bearing obligations:

	20X9	20X8
Notes payable	\$385,200	\$321,100
Current maturities of long-term debt	60,000	60,000
Long-term debt	<u>740,000</u>	<u>800,000</u>
Total debt	\$1,185,200	\$1,181,100

Westcliff's leverage and coverage ratios are calculated as follows:

Debt-to-assets 2009: 1,185,200 debt / 2,363,600 assets = 50.1%

Debt-to-assets 2008: 1,181,100 debt / 2,197,400 assets = 53.7%

Debt-to-equity 2009: 1,185,200 debt / 727,600 equity = 1.6

Debt-to-equity 2008: 1,181,100 debt / 588,700 equity = 2.0

Debt-to-total capital 2009: 1,185,200 debt / 1,912,800 total capital = 62.0%

Debt-to-total capital 2008: 1,181,100 debt / 1,769,800 total capital = 66.7%

(Note: Total capital = total debt + shareholders' equity.)

Interest coverage 2009: 399,500 EBIT / 168,000 interest expense = 2.4

Interest coverage 2008: 380,800 EBIT / 116,100 interest expense = 3.3

Fixed charge coverage 2009:

(399,500 EBIT + 24,000 lease payments) / (168,000 interest expense + 24,000 lease payments) = 2.2

Fixed charge coverage 2008:

(380,800 EBIT + 22,800 lease payments) / (116,100 interest expense + 22,800 lease payments) = 2.9

Leverage declined in 20X9 using all three measures, mainly as a result of an increase in shareholders' equity. On the other hand, both coverage ratios declined in 20X9 as a result of higher interest expense. One possible explanation for the increase in interest expense, given lower leverage, is that interest rates are increasing.

MODULE QUIZ 25.4

- 
- Compared to purchasing a long-lived asset using debt financing, leasing the asset *most likely*:
 - is more costly to the lessee.
 - requires a greater initial cash outflow from the lessee.
 - allows the lessee to avoid the risk of obsolescence.
 - During the life of a long-term lease under IFRS, the lessee recognizes:
 - interest expense only.
 - amortization expense and interest expense.
 - neither amortization expense nor interest expense.
 - For the lessor, cash flows from a lease are classified as:

- A. operating.
 - B. investing.
 - C. financing.
4. A net pension asset or liability can be associated with:
- A. defined benefit pension plans only.
 - B. defined contribution pension plans only.
 - C. either defined benefit or defined contribution pension plans.
5. At the end of last year, Maui Corporation's assets and liabilities were as follows:

Total assets	\$98,500
Accrued liabilities	\$5,000
Short-term debt	\$12,000
Bonds payable	\$39,000

Maui's debt-to-equity ratio is *closest* to:

- A. 1.2.
- B. 1.3.
- C. 1.4.

KEY CONCEPTS

LOS 25.a

When a bond is issued, assets and liabilities both initially increase by the bond proceeds. At any point in time, the book value of the bond liability is equal to the present value of the remaining future cash flows (coupon payments and maturity value) discounted at the market rate of interest at issuance. The proceeds are reported in the cash flow statement as an inflow from financing activities.

A premium bond (coupon rate > market yield at issuance) is reported on the balance sheet at a value greater than its face value. As the premium is amortized (reduced), the book value of the bond liability will decrease until it reaches its face value at maturity.

A discount bond (market yield at issuance > coupon rate) is reported on the balance sheet at less than its face value. As the discount is amortized, the book value of the bond liability will increase until it reaches its face value at maturity.

LOS 25.b

Interest expense includes amortization of any discount or premium at issuance. Using the effective interest rate method, interest expense is equal to the book value of the bond liability at the beginning of the period multiplied by the bond's yield at issuance.

For a premium bond, interest expense is less than the coupon payment (yield < coupon rate). The difference between interest expense and the coupon payment is subtracted from the bond liability on the balance sheet.

For a discount bond, interest expense is greater than the coupon payment (yield > coupon rate). The difference between interest expense and the coupon payment is added to the bond liability on the balance sheet.

LOS 25.c

When bonds are redeemed before maturity, a gain or loss is recognized equal to the difference between the redemption price and the carrying (book) value of the bond liability at the reacquisition date.

LOS 25.d

Debt covenants are restrictions on the borrower that protect the bondholders' interests, thereby reducing both default risk and borrowing costs. Covenants can include restrictions on dividend payments and share repurchases; mergers and acquisitions; sale, leaseback, and disposal of certain assets; and issuance of new debt in the future. Other covenants require the firm to maintain ratios or financial statement items at specific levels.

LOS 25.e

The firm separately discloses details about its long-term debt in the footnotes. These disclosures are useful for determining the timing and amount of future cash outflows. The disclosures usually include a discussion of the nature of the liabilities, maturity dates, stated and effective interest rates, call provisions and conversion privileges, restrictions imposed by creditors, assets pledged as security, and the amount of debt maturing in each of the next five years.

LOS 25.f

Advantages of leasing rather than purchasing an asset may include a smaller initial cash outflow, lower-cost financing, and less risk of obsolescence.

LOS 25.g

A finance lease is a lease that transfers the benefits and risks of ownership to the lessee. A lease that does not transfer these benefits and risks is an operating lease.

Under IFRS, for both finance and operating leases, except for short-term leases, a lessee reports a right-of-use asset and a lease liability on its balance sheet, both equal to the present value of the promised lease payments. The interest portion of each lease payment is reported as interest expense, while the principal-repayment portion of each payment reduces the lease liability. For short-term or low-value leases, rent expense is reported on the income statement and no balance sheet entries are required.

Under U.S. GAAP, reporting for a finance lease is the same as under IFRS. For an operating lease, the reporting is the same as under IFRS except that the entire lease payment is recorded as a lease expense. For short-term leases, rent expense is reported on the income statement and no balance sheet entries are required.

LOS 25.h

A lease that is classified as finance or operating by the lessee is classified the same way by the lessor.

Under both IFRS and U.S. GAAP, with a finance lease, the lessor removes the leased asset from its balance sheet and adds a lease receivable asset. The lessor reports the interest portion of the lease payments as income. For an operating lease, the lessor keeps the leased asset on its balance sheet, reports lease payments as income, and reports depreciation and other costs as expenses.

LOS 25.i

A firm reports a net pension liability on its balance sheet if the fair value of a defined benefit plan's assets is less than the estimated pension obligation, or a net pension asset if the fair value of the plan's assets is greater than the estimated pension obligation. The change in the net pension asset or liability is reflected in a firm's comprehensive income each year.

Pension expense for a defined contribution pension plan is equal to the employer's contributions.

LOS 25.j

Analysts use solvency ratios to measure a firm's ability to satisfy its long-term obligations. In evaluating solvency, analysts look at leverage ratios and coverage ratios.

Leverage ratios, such as debt-to-assets, debt-to-capital, debt-to-equity, and the financial leverage ratio, focus on the balance sheet.

Debt-to-assets ratio = total debt / total assets

Debt-to-capital ratio = total debt / (total debt + total equity)

Debt-to-equity ratio = total debt / total equity

Financial leverage ratio = average total assets / average total equity

Coverage ratios, such as interest coverage and fixed charge coverage, focus on the income statement.

Interest coverage = EBIT / interest payments

Fixed charge coverage = (EBIT + lease payments) / (interest payments + lease payments)

Page

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 25.1, 25.2

1. A The actual coupon payment on a bond is reported as operating cash outflow under U.S. GAAP. (Module 25.1, LOS 25.a)
2. A Interest expense is based on the book value of the bond. As the premium is amortized, the book value of the bond decreases until it reaches face value. (Module 25.2, LOS 25.b)
3. A This bond is issued at a discount since the coupon rate < market rate. (Module 25.1, LOS 25.a)
4. A Coupon payment = (coupon rate × face value of bond) = 6% × \$10,000,000 = \$600,000. (Module 25.1, LOS 25.b)
5. A Four coupon payments and the face value = (\$600,000 × 4) + \$10,000,000 = \$12,400,000. (Module 25.1, LOS 25.b)
6. B The present value of a 4-year annuity of \$600,000 plus a 4-year lump sum of \$10 million, all valued at a discount rate of 7%, equals \$9,661,279. Choice C can be eliminated because

the bond was issued at a discount. (Module 25.2, LOS 25.b)

7. **B** Market interest rate × book value = $7\% \times \$9,661,279 = \$676,290$. (Module 25.2, LOS 25.b)

8. **C** The new book value = beginning book value + interest expense – coupon payment = $\$9,661,279 + \$676,290 - \$600,000 = \$9,737,569$. The interest expense was calculated in question 7. Alternatively, changing N from 4 to 3 and calculating the PV will yield the same result. The change in market rates will not affect amortized costs. (Module 25.2, LOS 25.b)

9. **B** Coupon payments + discount interest = coupon payments + (face value – issue value) = $\$2,400,000 + (\$10,000,000 - \$9,661,279) = \$2,738,721$. (Module 25.2, LOS 25.b)

Module Quiz 25.3

1. **B** The cash paid to redeem the bonds is a CFF outflow. Because the redemption price is less than the book value of the liability, the firm will recognize a gain. (LOS 25.c)
2. **C** Debt covenants exist to protect the interests of creditors. A bond is considered to be in “technical” default if the borrower violates the bond’s covenants. (LOS 25.d)
3. **C** The market rate on the balance sheet date is not typically disclosed. The amount of debt principal scheduled to be repaid over the next five years and collateral pledged (if any) are generally included in the footnotes to the financial statements. (LOS 25.e)

Module Quiz 25.4

1. **C** Avoiding the risk of obsolescence is one of the advantages of leasing assets instead of purchasing them. At the end of a lease, the lessee often returns the leased asset to the lessor, and therefore does not bear the risk of an unexpected decline in the asset’s end-of-lease value. The interest rate implicit in a lease contract may be less than the interest rate on a loan to purchase the asset. The terms of a lease may not require all the covenants typically included in loan agreements or bond indentures. (LOS 25.f)
2. **B** At lease inception, the lessee records a right-of-use asset and a lease liability, both equal to the present value of the lease payments. In each period over the life of the lease, the lessee recognizes interest expense for the interest portion of the lease payments and amortization expense on the right-of-use asset. (LOS 25.g)
3. **A** Cash flows from a lease are operating cash inflows for the lessor. (LOS 25.h)
4. **A** Defined benefit pension plans can be overfunded and result in a net pension asset, or they can be underfunded and result in a net pension liability. Defined contribution plans do not result in balance sheet assets or liabilities because they are neither owned by the sponsoring firm or obligations of the sponsoring firm. (LOS 25.i)
5. **A** Because $A - L = E$, shareholders’ equity is $98,500 \text{ assets} - 5,000 \text{ accrued liabilities} - 12,000 \text{ short-term debt} - 39,000 \text{ bonds payable} = \$42,500$. Thus, debt-to-equity is $(12,000 \text{ short-term debt} + 39,000 \text{ bonds payable}) / 42,500 \text{ equity} = 1.2$. Only interest-bearing liabilities are considered debt. Accrued liabilities are not interest bearing. (LOS 25.j)

Reading 26

FINANCIAL REPORTING QUALITY

EXAM FOCUS

Here we cover the quality of a firm's financial statements, which, together with the quality of reported earnings, determines what is defined as the overall quality of the firm's financial reports. Candidates must understand how choices about several accounting methods and estimates can affect reported earnings, financial position, and the classification of cash flows. You should learn the additional disclosures required when non-GAAP measures are reported and the warning signs that management may be manipulating financial reports through biased reporting choices and estimates.

MODULE 26.1: REPORTING QUALITY



LOS 26.a: Compare and contrast financial reporting quality with the quality of reported results (including quality of earnings, cash flow, and balance sheet items).

Video covering this content is available online.

Financial reporting quality refers to the characteristics of a firm's financial statements. The primary criterion for judging financial reporting quality is adherence to generally accepted accounting principles (GAAP) in the jurisdiction in which the firm operates. However, given that GAAP provide choices of methods, estimates, and specific treatment of many items, compliance with GAAP by itself does not necessarily result in financial reporting of the highest quality.

High quality financial reporting must be *decision useful*. Two characteristics of decision-useful financial reporting are *relevance* and *faithful representation*. Relevance refers to the fact that information presented in the financial statements is useful to users of financial statements in making decisions. Relevant information must also be *material* in that knowledge of it would likely affect the decisions of users of financial statements. Faithful representation encompasses the qualities of *completeness*, *neutrality*, and the absence of errors. We develop the concept of neutrality of financial reports later in this reading.

The **quality of earnings** is, in many respects, a separate issue. The quality of reported earnings (not the quality of earnings reports) can be judged based on the sustainability of the earnings as well as on their level. Sustainability can be evaluated by determining the proportion of reported earnings that can be expected to continue in the future. Increases in reported earnings resulting from changes in exchange rates or by sales of assets that have appreciated over many periods are not typically sustainable, whereas higher profits from increased efficiency or increasing market share would generally be considered sustainable.

One dollar of high-quality earnings is expected to add more value to a company than one dollar of low-quality earnings, based on the criterion of sustainability. The higher probability that high-quality earnings will continue in future periods increases their impact on the value of the

firm, calculated as the present value of expected future earnings. At the other end of the sustainability spectrum, a one-time gain of a dollar from favorable currency exchange rate movements is not likely to be repeated and, therefore, has a smaller impact on estimates of a company's value.

The importance of the level of earnings is that reported earnings must be high enough to sustain the company's operations and existence over time, as well as high enough to provide an adequate return to the company's investors. Both of these concerns are important in determining the quality of a company's reported earnings. Sustainability of reported cash flows is also a consideration in determining the quality of reported earnings, as are the value of items reported on the balance sheet. Inadequate accruals for probable liabilities and overstatement of asset values can both decrease the quality of reported earnings and bring sustainability into question.

From our discussion here, we can see that it is quite possible that a firm has high financial reporting quality but a low quality of reported earnings. Reported earnings may be GAAP-compliant and relevant, represent the company's economic activities faithfully, and be decision useful as a result, but have low sustainability or be low enough in amount that the provision of adequate investor returns or the sustainability of the company itself are called into question.

LOS 26.b: Describe a spectrum for assessing financial reporting quality.

Combining both financial reporting quality and the quality of reported earnings, we can categorize the quality of financial reports along a spectrum from best to worst. At the high-quality end of the spectrum, we have financial reports that are compliant with GAAP, decision useful, and report earnings that are sustainable and represent an adequate return on invested capital. At the opposite end of the spectrum are financial reports that are essentially fictitious (fraudulent). When reporting quality is that low, the quality of the reported earnings themselves is impossible to assess. We can identify several levels of quality between these two extremes.

Here is one possible categorization of the quality levels of financial reports, from best to worst:

1. Reporting is compliant with GAAP and decision useful; earnings are sustainable and adequate.
2. Reporting is compliant with GAAP and decision useful, but earnings quality is low (earnings are not sustainable or not adequate).
3. Reporting is compliant with GAAP, but earnings quality is low and reporting choices and estimates are biased.
4. Reporting is compliant with GAAP, but the amount of earnings is actively managed to increase, decrease, or smooth reported earnings.
5. Reporting is not compliant with GAAP, although the numbers presented are based on the company's actual economic activities.
6. Reporting is not compliant and includes numbers that are essentially fictitious or fraudulent.

LOS 26.c: Explain the difference between conservative and aggressive accounting.

Ideally, financial statements should be neutral or unbiased in order to offer the most value to analysts. In general, we describe the choices made within GAAP with respect to reported

earnings as **conservative accounting** if they tend to decrease the company's reported earnings and financial position (on the balance sheet) for the current period. We describe choices that increase reported earnings or improve the financial position for the current period as **aggressive accounting**.

Aggressive accounting often results in decreased earnings in future periods, while conservative accounting will tend to increase future period earnings. Both these types of bias are sometimes used by management, for different periods, in an attempt to smooth earnings over time because greater earnings volatility tends to reduce the value of a company's shares. Often **earnings smoothing** is accomplished through adjustment of accrued liabilities that are based on management estimates. During periods of higher-than-expected earnings, management may employ a conservative bias by adjusting an accrued liability upward to reduce reported earnings for that period. This effectively allows deferral of the recognition of these earnings to a future period for which earnings are less than expected. In such a future period, the accrued liability is adjusted downward to increase reported earnings in that period, perhaps to meet market expectations. Deferral of reported earnings through conservative bias in financial reporting so they can be used opportunistically in a future period is sometimes referred to as putting earnings in the "cookie jar" (presumably to be "enjoyed" later).

Some examples of conservative versus aggressive financial reporting based on management choices and estimates are shown in Figure 26.1.

Figure 26.1: Aggressive and Conservative Accounting

Aggressive	Conservative
Capitalizing current period costs	Expensing current period costs
Longer estimates of the lives of depreciable assets	Shorter estimates of the lives of depreciable assets
Higher estimates of salvage values	Lower estimates of salvage values
Straight-line depreciation	Accelerated depreciation
Delayed recognition of impairments	Early recognition of impairments
Less accrual of reserves for bad debt	More accrual of reserves for bad debt
Smaller valuation allowances on deferred tax assets	Larger valuation allowances on deferred tax assets

Bias can also be present in the way that financial results are presented. A company may present transparent financial statements that help analysts and investors to understand the results and the activities that led to them. Alternatively, a company may provide minimal disclosure in an attempt to emphasize positive developments and obscure information about negative developments.

We should avoid thinking about conservatism in financial reporting as "good" and aggressive reporting as "bad." Conservative bias can also be considered as a deviation from neutral reporting or faithful representation that reduces the usefulness of financial statements to analysts and investors.

Sometimes GAAP themselves can introduce conservatism by imposing a higher standard of verification for revenue and profit than for expenses and accrual of liabilities. For example:

- Research costs are typically expensed in the period incurred because of the uncertainty about the future benefits to be provided from research activities, while the associated revenue is not recognized until some future period.
- Accruals for legal liabilities are recorded when a future payment becomes “probable,” while the standard for recognizing increasing accrued asset value is stricter.
- Under U.S. GAAP, write-downs of inventory values are required when their future value is likely impaired, but increases in inventory value may not be recorded until the inventory is actually sold.

While conservative bias in financial reporting is not ideal for users of financial statements, it may be beneficial in reducing the probability of future litigation from users claiming they were misled, in reducing current period tax liability (when deductions for tax must also be deducted in the financial statements), and in protecting the interests of those who have less complete information than company management, such as buyers of the company’s debt.

LOS 26.d: Describe motivations that might cause management to issue financial reports that are not high quality.

One important motivation for aggressive accounting choices is to meet or exceed a benchmark number for earnings per share. Specifically, managers may be motivated to report earnings that are greater than:

- Earnings guidance offered earlier by management.
- Consensus analyst expectations.
- Those of the same period in the prior year.

The manager’s motivation here may be career oriented, seeking to enhance her reputation and improve future career opportunities. Because beating certain benchmarks is very important to subsequent stock price movements, managers may be motivated by incentive compensation (bonuses) that depends on stock returns. Other possible motivations are to gain credibility with equity market investors or improve the way the company is viewed by its customers and suppliers.

For companies that are highly leveraged and unprofitable, aggressive accounting may be motivated by a desire to avoid violating debt covenants.

When earnings exceed benchmark levels, managers may make conservative accounting choices in ways that allow these earnings to be shown in future periods, increasing the probability that future period earnings will meet or exceed the relevant benchmark amount.

LOS 26.e: Describe conditions that are conducive to issuing low-quality, or even fraudulent, financial reports.

Three factors that typically exist in cases where management provides low-quality financial reporting are motivation, opportunity, and a rationalization of the behavior. So to the sources of *motivation* previously listed, we can add conditions that increase the *opportunity* to present low-quality financial reports. Circumstances in which low-quality, or even fraudulent, financial reporting is more probable are:

- The company has weak internal controls.

- The board of directors provides inadequate oversight.
- Applicable accounting standards provide a large range of acceptable accounting treatments, provide for inconsequential penalties in the case of accounting fraud, or both.

The third likely element of low-quality financial reporting is *rationalization* by management for less-than-ethical actions. Most people who do something they know is wrong tell themselves a story that seems (at least to them) to justify breaking the rules. Whether the story is “I’ll fix it next period” or “I have to do it to get my bonus and pay for my parents’ care,” the resulting behavior is the same.

LOS 26.f: Describe mechanisms that discipline financial reporting quality and the potential limitations of those mechanisms.

Each country has its own regulatory body responsible for publicly traded securities and the markets in which they trade. For example, in the United States, the regulatory body is the Securities and Exchange Commission (SEC). In the U.K., it is the Financial Conduct Authority (FCA). The International Organization of Securities Commissions (IOSCO) coordinates securities regulation on an international basis with over 200 members, such as national securities regulators, stock exchanges, and regional authorities. One such regional authority, the European Securities and Markets Authority (ESMA), coordinates policy among the securities regulators of countries in the European Union.

Securities regulations typically require:

- A registration process for the issuance of new publicly traded securities.
- Specific disclosure and reporting requirements, including periodic financial statements and accompanying notes.
- An independent audit of financial reports.
- A statement of financial condition (or management commentary) made by management.
- A signed statement by the person responsible for the preparation of the financial reports.
- A review process for newly registered securities and periodic reviews after registration.

Enforcement actions by securities regulators may include fines, suspension of participation in issuance and trading of securities, and public disclosure of the results of disciplinary proceedings. Regulators may also pursue criminal prosecution of fraudulent or otherwise illegal activities.

In addition to the audit opinion, a requirement for securities that trade in the United States is that management must include an assessment of the effectiveness of the firm’s internal controls.

Note that an unqualified or “clean” audit opinion is not a guarantee that no fraud has occurred but only offers reasonable assurance that the financial reports (prepared under the direction of management) have been “fairly reported” with respect to the applicable GAAP. The auditor is selected and paid by the firm being audited.

Another source of discipline on financial reporting quality is private contracts, such as those with lenders. Such contracts will often specify how financial measures referenced in the loan covenants will be calculated. The counterparties to private contracts with the firm have an incentive to see that the firm produces high-quality financial reports.

LOS 26.g: Describe presentation choices, including non-GAAP measures, that could be used to influence an analyst's opinion.

Firms will sometimes report accounting measures that are not defined or required under GAAP. Such **non-GAAP measures** typically exclude some items in order to make the firm's performance look better than it would using measures defined and required by GAAP. The claim is often made that certain items are excluded because they are one-time or nonoperating costs that will not affect operating earnings going forward, because the items are non-cash charges, or to "improve comparability with companies that use different accounting methods" for depreciation or restructuring charges.

In the United States, companies that report non-GAAP measures in their financial statements are required to:

- Display the most comparable GAAP measure with equal prominence.
- Provide an explanation by management as to why the non-GAAP measure is thought to be useful.
- Reconcile the differences between the non-GAAP measure and the most comparable GAAP measure.
- Disclose other purposes for which the firm uses the non-GAAP measure.
- Include, in any non-GAAP measure, any items that are likely to recur in the future, even those treated as nonrecurring, unusual, or infrequent in the financial statements.

IFRS require that firms using non-IFRS measures in financial reports must:

- Define and explain the relevance of such non-IFRS measures.
- Reconcile the differences between the non-IFRS measure and the most comparable IFRS measure.

Overall, the supposition is that firms use non-GAAP measures in an attempt to control the metrics on which they are evaluated and to reduce the focus of analysts and investors on GAAP measures.



MODULE QUIZ 26.1

1. A firm reports net income of \$40 million. The firm's financial statements disclose in Management's Discussion and Analysis that \$30 million of net income is attributable to a gain on the sale of assets. Based only on this information, for this period, the firm is *best* described as having high quality of:
 - A. financial reporting only.
 - B. both earnings and financial reporting.
 - C. neither earnings nor financial reporting.
2. Which of the following financial reports are considered to be of the lowest quality? Financial reports that reflect:
 - A. unsustainable earnings.
 - B. biased accounting choices.
 - C. departures from accounting principles.
3. Financial reporting is *most likely* to be decision useful when management's accounting choices are:
 - A. neutral.
 - B. aggressive.
 - C. conservative.

4. Which of the following is *least likely* to be a motivation to overreport earnings?
 - A. Reduce tax obligations.
 - B. Meet analyst expectations.
 - C. Remain in compliance with bond covenants.
5. With respect to conditions that may lead to low-quality financial reporting, ineffective internal controls are *best* described as a(n):
 - A. motivation.
 - B. opportunity.
 - C. rationalization.
6. A limitation on the effectiveness of auditing in ensuring financial reporting quality is that:
 - A. detecting fraud is not the objective of audits.
 - B. public firms are not required to obtain audit opinions.
 - C. auditors may only issue a qualified or unqualified opinion but do not explain why.
7. Under IFRS, a firm that presents a nonstandard financial measure is *least likely* required to:
 - A. provide the same measure for at least two prior periods.
 - B. explain the reasons for presenting the nonstandard measure.
 - C. reconcile the nonstandard measure to a comparable standard measure.



MODULE 26.2: ACCOUNTING CHOICES AND ESTIMATES

LOS 26.h: Describe accounting methods (choices and estimates) that could be used to manage earnings, cash flow, and balance sheet items.

Revenue Recognition

One example of how a firm's choices affect the timing of revenue recognition is the choice of where in the shipping process the customer actually takes title to the goods. A firm may choose terms with their customer of **free-on-board** (FOB) at the shipping point (the firm's loading dock) or FOB at the destination (the customer's location). Choosing terms of FOB at the shipping point will mean that revenue is recognized earlier compared to FOB at the destination.

Firms can also manage the timing of revenue recognition by accelerating or delaying the shipment of goods. If additional revenue is required to meet targets, firms can offer discounts or special financing terms to increase orders in the current period, or ship goods to distributors without receiving an order. Overloading a distribution channel with more goods than would normally be sold during a period is referred to as **channel stuffing**. In periods where high earnings are expected, management may wish to delay recognition of revenue to the next period and hold or delay customer shipments to achieve this.

In a **bill-and-hold transaction**, the customer buys the goods and receives an invoice but requests that the firm keep the goods at their location for a period of time. The use of fictitious bill-and-hold transactions can increase earnings in the current period by recognizing revenue for goods that are actually still in inventory. Revenue for future periods will be decreased as real customer orders for these bill-and-hold items are filled but not recognized in revenue, offsetting the previous overstatement of revenue.

Estimates of Credit Losses

One example of accounting choices that affect financial reports is the estimation of losses from uncollectable customer credit accounts. On the balance sheet, the reserve for uncollectible debt is an offset to accounts receivable. If management determines the probability that accounts receivable will be uncollectible is lower than their current estimate, a decrease in the reserve for uncollectible accounts will increase net receivables reported on the balance sheet, reduce expenses on the income statement, and increase net income. An increase in the allowance for bad debt would have the opposite effect, decreasing net receivables on the balance sheet, increasing expenses, and decreasing net income.

A firm that simply underestimates the percentage of receivables that will be uncollectible will report higher receivables and higher net income as a result. At some point, when actual uncollectible accounts exceed the low estimate, the firm will report an additional expense that will reduce net income and net receivables.

Management can adjust the bad-debt reserve in order to smooth earnings. In periods of high earnings, the allowance for bad debt is increased to reduce reported earnings, in effect storing these earnings for later use. In subsequent periods, if earnings are below benchmark values, the bad-debt reserve can be reduced to meet earnings targets.

Other reserves recorded by a company, such as a reserve for warranty expense, can also be changed to manage reported earnings. A decrease in the estimated warranty expense as a percentage of sales will increase earnings, while an increase in the reserve for warranty expense will decrease earnings for the period.

Valuation Allowance

Another example of a contra account that can be used to manage earnings is a valuation allowance. Recall that a valuation allowance reduces the carrying value of a deferred tax asset based on the probability it will not be realized. Similar to the effects of an allowance for bad debt, increasing a valuation allowance will decrease the net deferred tax asset on the balance sheet and reduce net income for the period, while a decrease in the valuation allowance will increase the net deferred tax asset and increase net income for the period.

As with the contra account for bad debt, the valuation allowance can be understated to show higher asset values and it can also be adjusted over time to smooth earnings.

Depreciation Methods and Estimates

Compared to straight-line depreciation, using an accelerated method of depreciation increases expenses, and decreases net income, in the early years of an asset's life. In the later years of an asset's life, expenses are lower and net income higher when an accelerated depreciation method is used. The carrying value of a depreciable asset on the balance sheet will decrease more rapidly with accelerated depreciation than with straight-line depreciation.

Estimates of the useful life of a depreciable asset and its salvage value upon disposal can also affect net income and the carrying value of the asset. A greater salvage value will slow depreciation so the carrying value of the asset is greater, depreciation expense is less, and net income is higher. A smaller salvage value will have the opposite effects. If the salvage value of an asset is set higher than the actual sale price at the end of the asset's life, a loss on the sale of the asset will decrease net income in the period in which the asset is sold.

Using a longer estimated useful life of a depreciable asset decreases the periodic depreciation expense and increases net income in the early years of the asset's life compared to using a shorter estimated useful life.

Amortization and Impairment

Management choices and estimates regarding amortization of purchased intangible assets are similar to those for depreciation of tangible assets. The intangible asset *goodwill* is not amortized but is subject to a test for impairment. By ignoring or delaying recognition of an impairment charge for goodwill, management can increase earnings in the current period.

Inventory Method

The choice of inventory cost flow methods can have significant effects on both reported earnings and the balance sheet value of inventory. Consider the choice between FIFO and weighted-average inventory costing methods. During periods of rising prices, COGS under the FIFO method will be less than COGS under the weighted-average costing method. Gross profit, gross margin, and earnings will all be greater under the FIFO method than under the weighted-average method as a result. Balance sheet inventory value will be greater under FIFO than under the weighted-average method.

During periods of decreasing prices, the opposite is true; FIFO COGS are greater than weighted-average COGS and FIFO gross profits, gross margin, and earnings less than under the weighted-average method. With decreasing prices, balance sheet inventory will be less under FIFO than under the weighted-average cost method.

In terms of relevance, in an environment of either increasing or decreasing prices, FIFO results in more accurate balance sheet inventory values because inventory value is closer to current replacement cost than under the weighted average cost method. Conversely, COGS are closer to current (replacement) cost under the weighted-average cost method so that gross profit and margin better reflect economic reality. Gross profit under FIFO is distorted in that it includes gains from rising prices (or losses from decreasing prices), so the weighted-average cost method produces "better" information on the income statement. Financial reports that are transparent and provide users with the information needed to understand how the choice of inventory costing method affects income statement and balance sheet values are considered to be higher quality.

Related-Party Transactions

If a public firm does business with a supplier that is private and controlled by management, adjusting the price of goods supplied can shift profits either to or from the private company to manage the earnings reported by the public company.

Capitalization

Any expense that can be capitalized creates an asset on the balance sheet, and the impact of the expense on net income can be spread over many years. Consider a firm that has a marketing expense of \$1.5 million and chooses to capitalize this expense and amortize it over three years. In the period in which the expense is incurred, capitalization will reduce the expense on the income statement from \$1.5 million to \$0.5 million, increasing pretax income by \$1 million. At

the end of the year, the related balance sheet asset is \$1 million, and an amortization expense of \$0.5 million will be taken (and reduce net income) in each of the following two years. Greater capitalization of research and development costs will shift net income into the current period in the same way.

Capitalization also affects cash flow classifications. If an expense is capitalized, the entire amount is classified as an investing cash outflow so that operating cash flow is increased by that amount.

Other Cash Flow Effects

Management can affect the classification of cash flows through other methods, primarily with the goal of increasing reported cash flow from operations. Taking longer to pay suppliers increases operating cash flows and is referred to as **stretching payables**. Delaying payments that would normally be made near the end of a reporting period until the beginning of the next accounting period will increase operating cash flow in the current period and reduce it in the subsequent period. There is no effect on reported earnings in the current period from stretching payables.

Capitalizing interest expense will decrease cash flow from investing and increase cash flow from operations, along with its effects on the pattern of earnings from depreciating the interest expense over time rather than expensing it all in the current period. More generally, the ability under IFRS to classify interest and dividends paid as either CFO or CFF, and interest and dividends received as either CFO or CFI, gives management an additional way to manage reported operating cash flow.



MODULE 26.3: WARNING SIGNS

LOS 26.i: Describe accounting warning signs and methods for detecting manipulation of information in financial reports.

Below is a list of several warning signs that analysts should look for. The presence of these issues does not indicate fraud or even earnings manipulation, but in each case, the presence of one or more warning signs requires more analysis in order to determine whether there is a real business reason for the item or if earnings manipulation or fraud is driving the decisions and results. Avoiding investment in the company is one alternative when analysts and investors cannot obtain satisfactory answers to the questions raised when multiple warning signs are present.

Revenue Recognition

- Changes in revenue recognition methods.
- Use of bill-and-hold transactions.
- Use of barter transactions.
- Use of rebate programs that require estimation of the impact of rebates on net revenue.
- Lack of transparency with regard to how the various components of a customer order are recorded as revenue.

- Revenue growth out of line with peer companies.
- Receivables turnover is decreasing over multiple periods.
- Decreases in total asset turnover, especially when a company is growing through acquisition of other companies.
- Inclusion of nonoperating items or significant one-time sales in revenue.

Inventories

- Declining inventory turnover ratio.
- LIFO liquidations—drawing down inventory levels when LIFO (U.S. GAAP only) inventory costing is used so that COGS reflects the lower costs of items acquired in past periods, which increases current period earnings.

Capitalization Policies

- Firm capitalizes costs that are not typically capitalized by firms in their industry.

Relationship of Revenue and Cash Flow

- The ratio of operating cash flow to net income is persistently less than one or declining over time.

Other Warning Signs

- Depreciation methods, estimated asset lives, or estimates of salvage values are out of line with those of peer companies in the industry.
- Fourth-quarter earnings show a pattern (either high or low) compared to the seasonality of earnings in the industry or seasonality of revenue for the firm.
- The firm has significant transactions with related parties (entities controlled by management).
- Certain expenses are classified as nonrecurring but appear regularly in financial reports.
- Gross or operating profit margins are noticeably higher than are typical for the industry and peer companies.
- Management typically provides only minimal financial reporting information and disclosure.
- Management typically emphasizes non-GAAP earnings measures and uses special or nonrecurring designations aggressively for charges.
- Growth by purchasing a large number of businesses can provide many opportunities to manipulate asset values and future depreciation and amortization and make comparisons to prior period earnings problematic.

Analysts should consider adjusting prior-period earnings when large restructuring or impairment charges are recognized. Analysts sometimes take such events to be good news because they anticipate better firm performance going forward when poorly performing assets are disposed of. Because the charges represent, to some extent, “corrections” of previously understated expenses and overstated asset values, analysts should consider spreading these

costs across prior periods and restating prior earnings to give a more realistic picture of true earnings trends.

MODULE QUIZ 26.2, 26.3

1. For the current period, inappropriate capitalization is *most likely* to:
 - A. overstate revenues.
 - B. understate liabilities.
 - C. understate expenses.
2. A potential warning sign that a firm is engaging in channel stuffing is an unusual increase in the firm's:
 - A. receivables turnover.
 - B. days of sales outstanding.
 - C. number of days of payables.

KEY CONCEPTS

LOS 26.a

Financial reporting quality refers to the characteristics of a firm's financial statements. High-quality financial reporting adheres to generally accepted accounting principles (GAAP) and is decision useful in terms of relevance and faithful representation.

Quality of reported results refers to the level and sustainability of a firm's earnings, cash flows, and balance sheet items. High-quality earnings are high enough to provide the firm's investors with an adequate return and are sustainable in future periods.

LOS 26.b

A spectrum for assessing financial reporting quality considers both the quality of a firm's financial statements and the quality of its earnings. One such spectrum, from highest quality to lowest, is the following:

- Reporting is compliant with GAAP and decision useful; earnings are sustainable and adequate.
- Reporting is compliant and decision useful, but earnings quality is low.
- Reporting is compliant, but earnings quality is low and reporting choices and estimates are biased.
- Reporting is compliant, but earnings are actively managed.
- Reporting is not compliant, but the numbers presented are based on the company's actual economic activities.
- Reporting is not compliant and includes numbers that are fictitious or fraudulent.

LOS 26.c

Biased accounting choices that can be made within GAAP include conservative and aggressive accounting. Conservative accounting choices tend to decrease the company's reported earnings and financial position for the current period. Aggressive accounting choices tend to increase reported earnings or improve the financial position for the current period.

Some managers employ conservative bias during periods when earnings are above target and aggressive bias during poor periods of below-target earnings to artificially smooth earnings.

LOS 26.d

Motivations for firm managers to issue low-quality financial reports may include pressure to meet or exceed earnings targets, career considerations, increasing their compensation, improving perceptions of the firm among customers and suppliers, or meeting the terms of debt covenants.

LOS 26.e

Conditions that are often present when managers issue low-quality financial reports include motivations, opportunities, and rationalizations. Weak internal controls, inadequate oversight by the board of directors, and wide ranges of acceptable accounting treatments are among the factors that may provide opportunities for low-quality reporting.

LOS 26.f

Mechanisms that help to discipline financial reporting quality include regulation, auditing, and private contracts. Regulators typically require public companies to provide periodic financial statements and notes, including management commentary, and obtain independent audits.

A clean audit opinion offers reasonable assurance that financial statements are free from material errors but does not guarantee the absence of error or fraud. The fact that firms select and pay their auditors may limit the effectiveness of auditing to discipline financial reporting quality.

LOS 26.g

Firms may attempt to influence analysts' valuations by presenting non-GAAP measures, such as earnings that exclude certain nonrecurring items. IFRS requires firms to define and explain the relevance of any non-GAAP measures and reconcile them to the most comparable IFRS measure. Similar requirements apply to U.S. public firms.

LOS 26.h

Accounting choices and estimates that can be used to manage earnings include:

- Revenue recognition choices such as shipping terms (FOB shipping point versus FOB destination), accelerating shipments (channel stuffing), and bill-and-hold transactions.
- Estimates of reserves for uncollectible accounts or warranty expenses.
- Valuation allowances on deferred tax assets.
- Depreciation methods, estimates of useful lives and salvage values, and recognition of impairments.
- Inventory cost flow methods.
- Capitalization of expenses.
- Related-party transactions.

LOS 26.i

Accounting warning signs that indicate a need for closer analysis may include:

- Revenue growth out of line with comparable firms, changes in revenue recognition methods, or lack of transparency about revenue recognition.
- Decreases over time in turnover ratios (receivables, inventory, total asset).
- Bill-and-hold, barter, or related-party transactions.
- Net income not supported by operating cash flows.
- Capitalization decisions, depreciation methods, useful lives, salvage values out of line with comparable firms.
- Fourth-quarter earnings patterns not caused by seasonality.
- Frequent appearance of nonrecurring items.
- Emphasis on non-GAAP measures, minimal information and disclosure in financial reports.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 26.1

1. **A** Because a large proportion of net income is due to a one-time gain, this period's earnings are likely not sustainable and the firm may be said to have low quality of earnings for the period. Clear disclosure of this fact in the financial statements suggests high quality of financial reporting. (LOS 26.a)
2. **C** In the spectrum of financial reporting quality, financial reports that depart from generally accepted accounting principles are considered to be of lower quality than those that reflect biased accounting choices. Financial reports that reflect unsustainable earnings, such as one-time gains, can still be of high quality if they state the situation clearly. (LOS 26.b)
3. **A** Financial reporting is most likely to be decision useful when accounting choices are neutral. Either aggressive or conservative accounting choices by management may be viewed as biases. (LOS 26.c)
4. **A** Reducing tax obligations would be a reason to *underreport* earnings. The other choices are motivations to *overreport* earnings. (LOS 26.d)
5. **B** Ineffective internal controls are a condition that provides an opportunity for low-quality financial reporting. (LOS 26.e)
6. **A** The objective of audits is to provide reasonable assurance that financial statements are presented fairly. A firm that is engaging in accounting fraud may deceive its auditor. Regulators in most countries require publicly traded firms to obtain independent audits of their financial statements. Auditors may issue a qualified opinion noting certain aspects of financial statements that are inconsistent with accounting principles or an adverse opinion if they find that financial statements are materially misstated and do not conform with GAAP. (LOS 26.f)
7. **A** IFRS require a firm that presents a nonstandard financial measure to reconcile that measure to an IFRS measure and explain why the firm believes the nonstandard measure is

relevant to users of the financial statements. Presenting the nonstandard measure for prior periods is not a requirement. (LOS 26.g)

Module Quiz 26.2, 26.3

1. **C** Management may make inappropriate capitalization decisions to understate expenses by creating balance sheet assets for items that should instead be recognized as expenses in the current period, increasing net income in the current period. Revenues and liabilities are unlikely to be affected by capitalization decisions. (Module 26.2, LOS 26.h)
2. **B** Channel stuffing, which includes activities such as accelerating deliveries to distributors or sending customers unordered merchandise, would likely increase accounts receivable as a percentage of revenues. This would decrease the receivables turnover ratio and increase days of sales outstanding. Payables would not be affected. (Module 26.3, LOS 26.i)

Reading 27

APPLICATIONS OF FINANCIAL STATEMENT ANALYSIS

EXAM FOCUS

In this reading, we will apply the analytic methods detailed in the reading on Financial Analysis Techniques. Pay special attention to the method outlined for forecasting cash flows. Memorize the four types of items important in the determination of credit quality. Lastly, analyst adjustments to financial statements are covered one more time. Understand the reasons for all the adjustments covered and how the adjustments will affect financial ratios used for valuation and credit analysis.

MODULE 27.1: FORECASTING



LOS 27.a: Evaluate a company's past financial performance and explain how a company's strategy is reflected in past financial performance.

Video covering this content is available online.

In the review of Financial Analysis Techniques, we introduced a number of financial ratios that can be used to assess a company's profitability, leverage, solvency, and operational efficiency. The analyst can evaluate trends in these ratios, as well as their levels, to evaluate how the company has performed in these areas.

Trends in financial ratios and differences between a firm's financial ratios and those of its competitors or industry averages can indicate important aspects of a firm's business strategy. Consider two firms in the personal computer business. One builds relatively high-end computers with cutting-edge features, and one competes primarily on price and produces computers with various configurations using readily available technology. What differences in their financial statements would we expect to find?

Premium products are usually sold at higher gross margins than less differentiated commodity-like products, so we should expect cost of goods sold to be a higher proportion of sales for the latter. We might also expect the company with cutting-edge features and high quality to spend a higher proportion of sales on research and development, which may be quite minimal for a firm purchasing improved components from suppliers rather than developing new features and capabilities in-house. The ratio of gross profits to operating profits will be larger for a firm that spends highly on research and development or on advertising.

In general, it is important for an analyst to understand a subject firm's business strategy. If the firm claims it is going to improve earnings per share by cutting costs, examination of operating ratios and gross margins over time will reveal whether the firm has actually been able to implement such a strategy and whether sales have suffered as a result.

LOS 27.b: Demonstrate how to forecast a company's future net income and cash flow.

A forecast of future net income and cash flow often begins with a forecast of future sales. Over shorter horizons, the “top down” approach to forecasting sales is used. The analyst begins with a forecast of GDP growth, often supplied by outside research or an in-house economics group. Historical relationships can be used to estimate the relationship between GDP growth and the growth of industry sales. If the subject firm’s market share is expected to remain the same, the growth of firm sales will be the same as the growth in industry sales. If the analyst has reason to believe the firm’s market share will increase or decrease next period, the market share can be adjusted for this change and then multiplied by estimated industry sales for the next period to get the forecast of firm sales for the period.

In a simple forecasting model, some historical average or trend-adjusted measure of profitability (operating margin, EBT margin, or net margin) can be used to forecast earnings. In complex forecasting models, each item on an income statement and balance sheet can be estimated based on separate assumptions about its growth in relation to revenue growth. For multi-period forecasts, the analyst typically employs a single estimate of sales growth at some point that is expected to continue indefinitely.

To estimate cash flows, the analyst must make assumptions about future sources and uses of cash. The most important of these will be increases in working capital, capital expenditures on new fixed assets, issuance or repayments of debt, and issuance or repurchase of stock. A typical assumption is that noncash working capital as a percentage of sales remains constant. A first-pass model might indicate a need for cash in future periods, and these cash requirements can then be met by projecting necessary borrowing in future periods. For consistency, interest expense in future periods must also be adjusted for any increase in debt.

Figure 27.1 illustrates this method. This projection assumes the company’s sales increase 5% per year, its cost of goods sold is 35% of sales, and operating expenses are 55% of sales. It also assumes noncash working capital stays constant at 85% of sales, and fixed capital requirements will be 5% of sales in each year. Net income is projected to increase over the forecast period, but the analysis reveals that cash is expected to decrease, suggesting a need for financing.

Figure 27.1: Income and Cash Flow Projection

	20X0	20X1	20X2
Sales @ +5% per year	86,145	90,452	94,975
Cost of goods sold @ 35% of sales	30,151	31,658	33,241
Operating expenses @ 55% of sales	47,380	49,749	52,236
Pretax income	8,614	9,045	9,497
Taxes @ 35%	3,015	3,166	3,324
Net income	5,599	5,879	6,173
Cash	8,615	6,311	3,891
Noncash working capital @ 85% of sales	73,223	76,884	80,729
Current assets	81,838	83,195	84,620
Net income	5,599	5,879	6,173
- Investment in working capital	3,478	3,661	3,844
- Investment in fixed capital @ 5% of sales	4,307	4,523	4,749
Change in cash	(2,186)	(2,304)	(2,420)
Beginning cash	10,801	8,615	6,311
Ending cash	8,615	6,311	3,891

MODULE 27.2: CREDIT AND EQUITY ANALYSIS



LOS 27.c: Describe the role of financial statement analysis in assessing the credit quality of a potential debt investment.

Video covering this content is available online.

Traditionally, credit analysts have spoken of the “three Cs,” “four Cs,” or even the “five Cs” of credit analysis. One version of the three Cs includes: Character, Collateral, and Capacity to repay. Character refers to the firm management’s professional reputation and the firm’s history of debt repayment. The ability to pledge specific collateral reduces lender risk. It is the third C, the capacity to repay, that requires close examination of a firm’s financial statements and ratios. Since some debt is for periods of 30 years or longer, the credit analyst must take a very long-term view of the firm’s prospects.

Credit rating agencies such as Moody’s and Standard and Poor’s employ formulas that are essentially weighted averages of several specific accounting ratios and business characteristics. The specific items used in the formula and their weights vary from industry to industry, but the types of items considered can be separated into four general categories:

1. *Scale and diversification.* Larger companies and those with a wider variety of product lines and greater geographic diversification are better credit risks.
2. *Operational efficiency.* Such items as operating ROA, operating margins, and EBITDA margins fall into this category. Along with greater vertical diversification, high operating efficiency is associated with better debt ratings.
3. *Margin stability.* Stability of the relevant profitability margins indicates a higher probability of repayment (leads to a better debt rating and a lower interest rate). Highly variable operating results make lenders nervous.

4. *Leverage.* Ratios of operating earnings, EBITDA, or some measure of free cash flow to interest expense or total debt make up the most important part of the credit rating formula. Firms with greater earnings in relation to their debt and in relation to their interest expense are better credit risks.



PROFESSOR'S NOTE

We discuss credit quality in more detail in our reading on Fundamentals of Credit Analysis in Fixed Income.

LOS 27.d: Describe the use of financial statement analysis in screening for potential equity investments.

In many cases, an analyst must select portfolio stocks from the large universe of potential equity investments. Whether the object is to select growth stocks, income stocks, or value stocks, accounting items and ratios can be used to identify a manageable subset of available stocks for further analysis.

Some investment strategies even have financial ratios in their names, such as low price/earnings and low price/sales investing. Multiple criteria are used because a screen based on a single factor can include firms with other undesirable characteristics. For example, a company with a low price/earnings ratio may also have operating losses, declining sales prospects, or very high leverage.

Analysts should be aware that their equity screens will likely include and exclude many or all of the firms in particular industries. A screen to identify firms with low P/E ratios will likely exclude growth companies from the sample. A low price-to-book or high dividend screen will likely include an inordinate proportion of financial services companies.

Backtesting refers to using a specific set of criteria to screen historical data to determine how portfolios based on those criteria would have performed. There is, of course, no guarantee that screening criteria that have identified stocks that outperformed in the past will continue to do so. Analysts must also pay special attention to the potential effects of survivorship bias, data-mining bias, and look-ahead bias (see the reading on Sampling and Estimation in Quantitative Methods) when evaluating the results of backtesting.

LOS 27.e: Explain appropriate analyst adjustments to a company's financial statements to facilitate comparison with another company.

Because different companies choose different accounting methods, an analyst must be prepared to adjust the financial statements of one company to make them comparable to those of another company or group of companies. Differences in accounting methods chosen by firms subject to the same standards, as well as differences in accounting methods due to differences in local accounting standards, can make comparisons between companies problematic.

Consider two companies in the same industry that have different depreciation schedules. One company has selected straight-line depreciation even though physical assets in its industry tend to lose most of their productive value early in their economic lives. The analyst would need to adjust the depreciation of that firm so that the net income figures for the firms are comparable. A change in a firm's financial statement depreciation would lead to changes in gross profit, operating profit, and so on, down to net profit and earnings per share.

Differences between U.S. GAAP and IFRS require an analyst to adjust the financial statements of firms from different countries before comparing their financial results. Important differences between the two include their treatments of the effect of exchange rate changes, certain securities held by the firm, and inventory cost flows.

Several adjustments to improve the comparability of firms' financial statements and ratios are as follows.

Investments in Securities

Because the classification of a firm's investment securities affects how changes in their values are recorded, it can significantly affect reported earnings and assets. Recall that unrealized gains and losses on held-for-trading securities are recorded in income, while those on available-for-sale or held-to-maturity securities are not. Additionally, while unrealized gains and losses on held-for-trading and available-for-sale securities are reflected in balance sheet asset values, for held-to-maturity securities they are not.

When these differences in classifications lead to significant differences in reported net income or balance sheet asset values for otherwise comparable companies, an analyst can use disclosures to adjust net income and assets of one firm to what they would have been had their classifications been the same.

Inventory Accounting Differences

As we covered in the reading on Inventories, a firm using LIFO (permitted only under U.S. GAAP) will report higher cost of goods sold, lower income, and lower inventory compared to FIFO inventory accounting when costs are rising. The **LIFO reserve**, which all LIFO firms must report, can be used to adjust LIFO cost of goods and inventory to their FIFO-equivalent values.

Differences in Depreciation Methods and Estimates

Disclosures related to depreciation are not specific enough to permit adjustments to ensure comparability. However, some qualitative information for comparing companies' methods can be obtained.

Over an asset's life, differences between depreciation methods, estimates of useful lives, and estimates of salvage values used by otherwise comparable firms can lead to significant differences in reported income and balance sheet asset values. A firm that is aggressive in using higher estimates of useful asset lives or asset salvage values will report lower annual depreciation expense and higher net income, compared to a more conservative firm that uses lower estimates of useful lives or salvage values. If the analyst concludes that a firm's aggressive assumptions regarding asset lives, for example, are increasing balance sheet net asset values and reported net income, an adjustment to net income and fixed asset carrying values may be appropriate.

Note as well that upward revaluation of fixed asset values is permitted under IFRS but not under U.S. GAAP. Such a revaluation will increase assets and equity, and in a case where the upward revaluation reverses a previous downward revaluation, the increase in value is also reported on the income statement.

An analyst can estimate the number of years' worth of depreciation a firm has recognized by dividing accumulated depreciation from the balance sheet by depreciation expense from the income statement. The result can be interpreted as the average age of the firm's assets. Similarly, an analyst can estimate the **average useful life** of a firm's assets (gross property, plant, and equipment divided by depreciation expense) and their **average remaining useful life** (net property, plant, and equipment divided by depreciation expense). Comparing average ages and useful lives of assets within an industry may reveal differences in firms' future capital spending needs.

Goodwill

Two companies with identical assets, but where one has grown through acquisition of some business units while the other has grown internally by creating such business units, will show different balance sheet values for the same assets. For the company that has grown through acquisition:

- Tangible assets of the acquired units will be recorded at fair value as of the acquisition date, rather than at historical cost net of accumulated depreciation.
- Identifiable intangible assets of the acquired units will be valued at their acquisition cost, rather than not being included in balance sheet assets.
- Goodwill, the excess of acquisition price over the fair value of acquired net assets, will be shown on the balance sheet.

Two adjustments are typically made to goodwill to improve comparability in such a case. First, goodwill should be subtracted from assets when calculating financial ratios. Second, any income statement expense from impairment of goodwill in the current period should be reversed, increasing reported net income.

In calculating price to book value of equity per share, an analyst can remove goodwill from assets and recalculate a lower adjusted book value, resulting in a price to adjusted book value ratio that is greater.



MODULE QUIZ 27.1, 27.2

1. The table below shows selected data from a company's financial statements.

	20X6	20X7	20X8	20X9
Sales	8,614	9,217	9,862	10,553
COGS	5,304	5,622	6,072	6,679
Purchases	5,257	5,572	6,018	6,620
Inventory	2,525	2,475	2,421	2,362
Accounts receivable	3,491	3,728	3,928	4,352
Accounts payable	1,913	2,102	2,311	2,539

Based on these results, what was this company's *most likely* strategy for improving its operating activity during this period?

- A. Improve its inventory management.
- B. Change its credit and collections policies with its customers.
- C. Change the degree to which it uses trade credit from suppliers.

2. An analyst who is projecting a company's net income and cash flows is *least likely* to assume a constant relationship between the company's sales and its:
 - A. interest expenses.
 - B. cost of goods sold.
 - C. noncash working capital.
3. Credit analysts are likely to consider a company's credit quality to be improving if the company reduces its:
 - A. scale and diversification.
 - B. margin stability.
 - C. leverage.
4. Which of the following stock screens is *most likely* to identify stocks with high earnings growth rates?
 - A. Dividend payout ratio greater than 30%.
 - B. Price to cash flow per share ratio less than 12.
 - C. Book value to market value ratio less than 25%.
5. An analyst needs to compare the financial statements of Firm X and Firm Y. Which of the following differences in the two firms' financial reporting is *least likely* to require the analyst to make an adjustment?

<u>Firm X</u>	<u>Firm Y</u>
A. Straight-line depreciation	Accelerated depreciation
B. Direct method cash flows	Indirect method cash flows
C. IFRS financial reporting	U.S. GAAP financial reporting
6. When comparing a firm that uses LIFO inventory accounting to firms that use FIFO, an analyst should:
 - A. subtract the LIFO reserve from cost of sales.
 - B. add the change in the LIFO reserve to inventories.
 - C. subtract the change in the LIFO reserve from cost of sales.
7. The ratio of a firm's property, plant, and equipment, net of accumulated depreciation, to its annual depreciation expense is *best* interpreted as an estimate of the:
 - A. average age of the firm's assets.
 - B. average useful life of the firm's assets.
 - C. remaining useful life of the firm's assets.

KEY CONCEPTS

LOS 27.a

Trends in a company's financial ratios and differences between its financial ratios and those of its competitors or industry average ratios can reveal important aspects of its business strategy.

LOS 27.b

A company's future income and cash flows can be projected by forecasting sales growth and using estimates of profit margins and the increases in working capital and fixed assets necessary to support the forecast sales growth.

LOS 27.c

Credit analysis uses a firm's financial statements to assess its credit quality. Indicators of a firm's creditworthiness include its scale and diversification, operational efficiency, margin stability, and use of financial leverage.

LOS 27.d

Potentially attractive equity investments can be identified by screening a universe of stocks, using minimum or maximum values of one or more ratios. Which (and how many) ratios to use, what minimum or maximum values to use, and how much importance to give each ratio all present challenges to the analyst.

LOS 27.e

When companies use different accounting methods or estimates relating to areas such as inventory accounting, depreciation, capitalization, and off-balance-sheet financing, analysts must adjust the financial statements for comparability.

LIFO ending inventory can be adjusted to a FIFO basis by adding the LIFO reserve. LIFO cost of goods sold can be adjusted to a FIFO basis by subtracting the change in the LIFO reserve.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 27.1, 27.2

1. A To analyze this company's operating strategy, calculate its activity ratios:

	20X7	20X8	20X9
Inventory turnover	2.25	2.48	2.79
Receivables turnover	2.55	2.58	2.55
Payables turnover	2.78	2.73	2.73
Days of inventory on hand	162	147	131
Days of sales outstanding	143	142	143
Number of days of payables	132	134	134

The ratios that have changed most significantly are the ones related to inventory. Receivables and payables performance has remained steady, suggesting no change in the company's use of supplier credit or extension of customer credit. (Module 27.1, LOS 27.a)

2. A Projections of net income and cash flows are typically based on assumptions that cost of goods sold, operating expenses, and noncash working capital remain a constant percentage of sales. The projections then show whether additional borrowing is needed during the forecast period. If so, the analyst will adjust the interest expense to reflect the additional debt. (Module 27.1, LOS 27.b)
3. C Lower leverage improves a company's creditworthiness. Larger scale, more diversification, higher operating efficiency, and more stable margins also tend to indicate better credit quality. (Module 27.2, LOS 27.c)
4. C Firms with high growth rates will tend to have high market values relative to the book value of their equity. Low price to cash flow ratios would tend to identify value stocks rather than growth stocks. Screening for high dividend payout ratios would tend to identify mature firms with relatively few growth opportunities. (Module 27.2, LOS 27.d)

5. B Cash flows are the same under either method. Differences in depreciation methods and IFRS versus U.S. GAAP reporting can require an analyst to adjust financial statements to make them comparable. (Module 27.2, LOS 27.e)

6. C To adjust LIFO financial statement data to a FIFO basis, add the LIFO reserve to inventories on the balance sheet and subtract the change in the LIFO reserve from cost of sales on the income statement. Remember that the balance sheet is cumulative (use the full LIFO reserve) while the income statement refers to the most recent period (use the change for the period in the LIFO reserve). (Module 27.2, LOS 27.e)

7. C Remaining useful life = net PP&E / depreciation expense.

Average age of assets = accumulated depreciation / depreciation expense.

Average useful life = gross PP&E / depreciation expense.

(Module 27.2, LOS 27.e)

TOPIC QUIZ: FINANCIAL STATEMENT ANALYSIS

You have now finished the Financial Statement Analysis topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Financial Statement Analysis. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

FORMULAS

Activity Ratios:

$$\text{receivables turnover} = \frac{\text{annual sales}}{\text{average receivables}}$$

$$\text{days of sales outstanding} = \frac{365}{\text{receivables turnover}}$$

$$\text{inventory turnover} = \frac{\text{cost of goods sold}}{\text{average inventory}}$$

$$\text{days of inventory on hand} = \frac{365}{\text{inventory turnover}}$$

$$\text{payables turnover} = \frac{\text{purchases}}{\text{average trade payables}}$$

$$\text{number of days of payables} = \frac{365}{\text{payables turnover ratio}}$$

$$\text{total asset turnover} = \frac{\text{revenue}}{\text{average total assets}}$$

$$\text{fixed asset turnover} = \frac{\text{revenue}}{\text{average net fixed assets}}$$

$$\text{working capital turnover} = \frac{\text{revenue}}{\text{average working capital}}$$

Liquidity Ratios:

$$\text{current ratio} = \frac{\text{current assets}}{\text{current liabilities}}$$

$$\text{quick ratio} = \frac{\text{cash} + \text{marketable securities} + \text{receivables}}{\text{current liabilities}}$$

$$\text{cash ratio} = \frac{\text{cash} + \text{marketable securities}}{\text{current liabilities}}$$

$$\text{defensive interval} = \frac{\text{cash} + \text{marketable securities} + \text{receivables}}{\text{average daily expenditures}}$$

$$\text{cash conversion cycle} = \left(\frac{\text{days sales}}{\text{outstanding}} \right) + \left(\frac{\text{days of inventory}}{\text{on hand}} \right) - \left(\frac{\text{number of days}}{\text{of payables}} \right)$$

Solvency Ratios:

$$\text{debt-to-equity} = \frac{\text{total debt}}{\text{total shareholders' equity}}$$

$$\text{debt-to-capital} = \frac{\text{total debt}}{\text{total debt} + \text{total shareholders' equity}}$$

$$\text{debt-to-assets} = \frac{\text{total debt}}{\text{total assets}}$$

$$\text{financial leverage} = \frac{\text{average total assets}}{\text{average total equity}}$$

$$\text{interest coverage} = \frac{\text{earnings before interest and taxes}}{\text{interest payments}}$$

$$\text{fixed charge coverage} = \frac{\text{earnings before interest and taxes} + \text{lease payments}}{\text{interest payments} + \text{lease payments}}$$

Profitability Ratios:

$$\text{net profit margin} = \frac{\text{net income}}{\text{revenue}}$$

$$\text{gross profit margin} = \frac{\text{gross profit}}{\text{revenue}}$$

$$\text{operating profit margin} = \frac{\text{operating income}}{\text{revenue}} \text{ or } \frac{\text{EBIT}}{\text{revenue}}$$

$$\text{pretax margin} = \frac{\text{EBT}}{\text{revenue}}$$

$$\text{return on assets (ROA)} = \frac{\text{net income}}{\text{average total assets}}$$

$$\text{return on assets (ROA)} = \frac{\text{net income} + \text{interest expense} (1 - \text{tax rate})}{\text{average total assets}}$$

$$\text{return on assets (ROA)} = \frac{\text{operating income}}{\text{average total assets}} \text{ or } \frac{\text{EBIT}}{\text{average total assets}}$$

$$\text{return on total capital} = \frac{\text{EBIT}}{\text{average total capital}}$$

$$\text{return on equity} = \frac{\text{net income}}{\text{average total equity}}$$

$$\text{return on common equity} = \frac{\text{net income} - \text{preferred dividends}}{\text{average common equity}}$$

$$= \frac{\text{net income available to common}}{\text{average common equity}}$$

Free Cash Flow to the Firm:

$$\text{FCFF} = \text{net income} + \text{noncash charges} + [\text{cash interest paid} \times (1 - \text{tax rate})] - \text{fixed capital investment} - \text{working capital investment}$$

$$\text{FCFF} = \text{cash flow from operations} + [\text{cash interest paid} \times (1 - \text{tax rate})] - \text{fixed capital investment}$$

Free Cash Flow to Equity:

FCFE = cash flow from operations – fixed capital investment + net borrowing

common-size income statement ratios = $\frac{\text{income statement account}}{\text{sales}}$

common-size balance sheet ratios = $\frac{\text{balance sheet account}}{\text{total assets}}$

common-size cash flow ratios = $\frac{\text{cash flow statement account}}{\text{revenues}}$

original DuPont equation: ROE = $\left(\frac{\text{net profit}}{\text{margin}} \right) \left(\frac{\text{asset turnover}}{\text{turnover}} \right) \left(\frac{\text{leverage ratio}}{\text{ratio}} \right)$

extended DuPont equation:

ROE = $\left(\frac{\text{net income}}{\text{EBT}} \right) \left(\frac{\text{EBT}}{\text{EBIT}} \right) \left(\frac{\text{EBIT}}{\text{revenue}} \right) \left(\frac{\text{revenue}}{\text{total assets}} \right) \left(\frac{\text{total assets}}{\text{total equity}} \right)$

basic EPS = $\frac{\text{net income} - \text{preferred dividends}}{\text{weighted average number of common shares outstanding}}$

diluted EPS =

$$\frac{\left[\text{net income} - \frac{\text{preferred dividends}}{\text{shares}} \right] + \left[\frac{\text{convertible preferred dividends}}{\text{shares}} \right] + \left(\frac{\text{convertible debt interest}}{\text{shares}} \right) (1 - t)}{\left(\frac{\text{weighted average shares}}{\text{shares}} \right) + \left(\frac{\text{shares from conversion of conv. pfd. shares}}{\text{conv. pfd. shares}} \right) + \left(\frac{\text{shares from conversion of conv. debt}}{\text{conv. debt}} \right) + \left(\frac{\text{shares issuable from stock options}}{\text{stock options}} \right)}$$

Coefficients of Variation:

CV sales = $\frac{\text{standard deviation of sales}}{\text{mean sales}}$

CV operating income = $\frac{\text{standard deviation of operating income}}{\text{mean operating income}}$

CV net income = $\frac{\text{standard deviation of net income}}{\text{mean net income}}$

Inventories:

ending inventory = beginning inventory + purchases – COGS

FIFO COGS = LIFO COGS – (ending LIFO reserve – beginning LIFO reserve)

Long-Lived Assets:

$$\text{straight-line depreciation} = \frac{\text{cost} - \text{salvage value}}{\text{useful life}}$$

$$\text{DDB depreciation} = \left(\frac{2}{\text{useful life}} \right) (\text{cost} - \text{accumulated depreciation})$$

units-of-production depreciation =

$$\frac{\text{original cost} - \text{salvage value}}{\text{life in output units}} \times \text{output units in the period}$$

$$\text{average age} = \frac{\text{accumulated depreciation}}{\text{annual depreciation expense}}$$

$$\text{total useful life} = \frac{\text{historical cost}}{\text{annual depreciation expense}}$$

$$\text{remaining useful life} = \frac{\text{ending net PP&E}}{\text{annual depreciation expense}}$$

Deferred Taxes:

$$\text{income tax expense} = \text{taxes payable} + \Delta \text{DTL} - \Delta \text{DTA}$$

Debt Liabilities:

$$\text{interest expense} = \left(\frac{\text{market rate}}{\text{at issue}} \right) \times \left(\frac{\text{balance sheet value of the liability}}{\text{at the beginning of the period}} \right)$$

Performance Ratios:

$$\text{cash flow-to-revenue} = \frac{\text{CFO}}{\text{net revenue}}$$

$$\text{cash return-on-assets} = \frac{\text{CFO}}{\text{average total assets}}$$

$$\text{cash return-on-equity} = \frac{\text{CFO}}{\text{average total equity}}$$

$$\text{cash-to-income} = \frac{\text{CFO}}{\text{operating income}}$$

$$\text{cash flow per share} = \frac{\text{CFO} - \text{preferred dividends}}{\text{weighted average number of common shares}}$$

Coverage Ratios:

$$\text{debt coverage} = \frac{\text{CFO}}{\text{total debt}}$$

$$\text{interest coverage} = \frac{\text{CFO} + \text{interest paid} + \text{taxes paid}}{\text{interest paid}}$$

$$\text{reinvestment} = \frac{\text{CFO}}{\text{cash paid for long-term assets}}$$

$$\text{debt payment} = \frac{\text{CFO}}{\text{cash long-term debt repayment}}$$

$$\text{dividend payment} = \frac{\text{CFO}}{\text{dividends paid}}$$

$$\text{investing and financing} = \frac{\text{CFO}}{\text{cash outflows from investing and financing activities}}$$

INDEX

A

accelerated depreciation, 31, 187
accounting equation, 2
accounts payable, 60
accounts receivable, 26, 58
accrual accounting, 16
accrued liabilities, 60
accumulated other comprehensive income, 68
activity ratios, 117
adverse opinion, 4
aggressive accounting, 263
allowance for doubtful accounts, 58
amortization, 32, 180
amortized cost, 15
antidilutive securities, 39
assets, 2
audit, 4
auditor's opinion, 4
authorized shares, 68
available-for-sale securities, 46, 64
average age, 201, 286
average depreciable life, 201
average remaining useful life, 286
average useful life, 286

B

bad debt expense, 58
balance sheet, 2, 55, 56
basic EPS, 37
bill-and-hold transaction, 269
bond, balance sheet liability, 236
book value, 186, 236

business risk, 136
business segment, 137

C

capital adequacy, 136
capitalization, 179, 271
capital structure, 2
carrying value, 186, 212, 236
cash and cash equivalents, 58
cash conversion cycle, 121
cash flow from financing activities (CFF), 3, 80, 88, 95, 98
cash flow from investing activities (CFI), 3, 80, 88, 89, 98
cash flow from operating activities (CFO), 3, 80, 87, 98
cash flow per share, 102
cash flow statement, 79
cash flow-to-revenue ratio, 102
cash ratio, 120
cash return-on-assets ratio, 102
cash return-on-equity ratio, 102
change in accounting estimate, 34
change in accounting policy, 34
channel stuffing, 269
classified balance sheet, 56
coefficient of variation, 136
common-size balance sheet, 69, 112
common-size cash flow statement, 99
common-size income statement, 44, 112
complex capital structure, 37
component depreciation, 188
comprehensive income, 46
conservative accounting, 263
contra accounts, 58
contract, 26
contributed capital, 67
cost model, 60, 193
cost of goods sold (COGS), 30, 145
coupon payments, 235
coupon rate, 235
coverage ratios, 121

credit analysis, 137
credit quality, 283
current assets, 57
current cost, 15
current liabilities, 57
current portion of long-term debt, 60
current ratio, 120

D

days of inventory on hand, 118
days of sales outstanding, 118
debt covenants, 244
debt payment ratio, 103
debt ratios, 121
debt-to-assets ratio, 122
debt-to-capital ratio, 121
debt-to-equity ratio, 121
declining balance method, 31
deductible temporary differences, 223
defensive interval, 121
deferred tax assets, 61, 212, 213
deferred tax liabilities, 66, 212, 213
defined benefit plan, 250
defined contribution plan, 250
depreciation, 31, 123, 180, 186
derecognition, 196
derivative instruments, 64
development costs, 182
diluted EPS, 39, 40
dilutive securities, 39
direct method, 82, 87
disclaimer of opinion, 4
discontinued operation, 33
discount bonds, 237
dividend payment ratio, 103
dividends, cash flow classification, 81, 82
double-declining balance method, 32, 187
DuPont system, 129

E

earnings before interest, taxes, depreciation, and amortization (EBITDA) , 123
earnings guidance, 5
earnings per share (EPS), 37, 134
earnings smoothing, 263
economic depreciation, 186
effective interest rate method, 237
effective tax rate, 45, 222
effective tax rate reconciliation, 226
expenses, 2, 24
expensing, 179
extended DuPont equation, 131

F

face value, 235
fair value, 15
fair value model, 61
features for preparing financial statements, 16
finance lease, 246
Financial Accounting Standards Board, 12
financial assets, 63
Financial Conduct Authority, 12
financial leverage ratio, 122
financial reporting, 1
financial reporting quality, 261
financial statement analysis, 1
financial statement analysis framework, 6
financial statement notes, 3
financing cash flows, 3
first in, first out (FIFO), 30, 147
fixed asset turnover, 119
fixed charge coverage ratio, 122
footnotes, 3
free cash flow, 100
free cash flow to equity (FCFE), 101
free cash flow to the firm (FCFF), 100
free-on-board, 268

G

gains, 24
geographic segment, 137
going concern assumption, 4, 16
goodwill, 63, 181, 182
gross investment, 201
gross profit, 25
gross profit margin, 45, 123
growth in same-store sales, 135

H

held to maturity, 46
held-to-maturity securities, 64
historical cost, 15, 61, 186
horizontal common-size balance sheet or income statement, 114

I

identifiable intangible assets, 62, 181
impairment, 61, 194
income statement, 2, 23
income tax expense, 212
indirect method, 83
intangible assets, 32, 62, 181
interest, cash flow classification, 81, 82
interest coverage ratio, 122
interest expense, 236
internal controls, 4
International Accounting Standards Board, 12
International Organization of Securities Commissions, 12, 266
inventories, 30, 59
inventory disclosures, 164
inventory turnover, 118
investing and financing ratio, 103
investing cash flows, 3
investment property, 61, 202
issued shares, 68

L

last in, first out (LIFO), 30, 148
lease liability, 247
lease receivable, 249
lessee, 246
lessor, 246
leverage ratio, 122
liabilities, 2
LIFO liquidation, 159
LIFO reserve, 157, 285
line graph, 115
liquid asset requirement, 136
liquidity, 56
liquidity-based format, 56
liquidity ratios, 71, 117, 120
long-term financial liabilities, 66
losses, 24
lower of cost or market, 161

M

management's commentary, 3
Management's Discussion and Analysis (MD&A), 3
marketable securities, 58
market rate of interest, 236
mark-to-market, 64
matching principle, 29
maturity value, 235
measurement base, 15
measurement date, 33
minority interest, 24, 68
multi-step income statement, 25

N

net income per employee, 135
net interest margin, 136
net profit margin, 45, 123

net realizable value, 15, 58, 161
net revenue, 24
noncash investing and financing activities, 81
noncontrolling interest, 24, 68
noncurrent assets, 58
noncurrent liabilities, 58
non-GAAP measures, 267
notes payable, 60
number of days in inventory, 118
number of days of payables, 119

O

operating cash flows, 3
operating cycle, 57
operating lease, 246
operating profit, 25
operating profitability ratios, 123
operating profit margin, 46, 123
operating return on assets, 124
other comprehensive income (OCI), 46
other current assets, 59
outstanding shares, 68
owners' equity, 2, 67

P

par bond, 236
par value, 67, 235
payables turnover, 118
performance obligation, 26
performance ratios, 102
period costs, 30, 146
periodic inventory system, 151
permanent difference, 222, 223
perpetual inventory system, 151
phaseout period, 33
potentially dilutive securities, 37
preferred stock, 68

premium bonds, 237
prepaid expenses, 59
present value, 15
pretax margin, 46, 123
price-to-earnings (P/E) ratio, 134
prior-period adjustment, 34
product costs, 146
profitability ratios, 117, 122
property, plant, and equipment, 60
prospective application, 34
proxy statements, 5
purchase method, 182
pure-discount bonds, 240

Q

qualified opinion, 4
quality of earnings, 262
quick ratio, 120

R

receivables turnover, 117
recoverability test, 194
recoverable amount, 61, 194
redemption, 244
regulatory authorities, 12
reinvestment ratio, 103
related-party transactions, 271
remaining useful life, 202
required financial statements, 16
required reporting elements, 15
research and development costs, 182
research costs, 182
reserve requirements, 136
retail method, 59
retained earnings, 46, 68
retrospective application, 34
return on assets (ROA), 123

return on common equity, 124
return on equity (ROE), 124
return on total capital (ROTC), 124
revaluation model, 60, 197,
198
revaluation surplus, 193
revaluation to fair value, 197
revenue, 2, 24
right-of-use asset, 247

S

sales per employee, 135
sales per square foot, 136
scenario analysis, 138
screening for potential equity investments, 284
Securities and Exchange Commission, 12, 266
sensitivity analysis, 138
simple capital structure, 37
simulation, 138
software development costs, 182
solvency, 56
solvency ratios, 71, 117, 121
specific identification method, 30, 147
stacked column graph, 115
standard costing, 59
standard-setting bodies, 12
statement of cash flows, 3
statement of changes in equity, 2, 69
statement of comprehensive income, 2
statutory tax rate, 222
stock dividend, 37
stock split, 38
straight-line depreciation, 31, 186
stretching payables, 271
structure and content of financial statements, 17
sustainable growth rate, 135

T

taxable temporary differences, 223
tax base, 214, 217
taxes payable, 60, 212
tax loss carryforward, 212
tax rate changes, 220
tax return terminology, 212
technical default, 244
temporary difference, 222, 223
total asset turnover, 119
trading securities, 46, 64
transaction price, 27
treasury stock, 68
treasury stock method, 42

U

unearned revenue, 26, 60
unidentifiable intangible assets, 62, 181
units-of-production method, 187
unqualified opinion, 4
unusual or infrequent items, 33

V

valuation allowance, 212, 224, 270
valuation ratios, 117
value-at-risk, 136
value in use, 61, 194
vertical common-size balance sheet, 112
vertical common-size income statement, 112

W

weighted average cost method, 30, 148
weighted average number of common shares, 37
working capital, 58
working capital turnover, 120

Z

zero-coupon bonds, 240



2023 CFA® Exam Prep

SchweserNotes™

Corporate Issuers and
Equity Investments



LEVEL I BOOK 3

KAPLAN SCHWEISER

Book 3: Corporate Issuers and Equity Investments

SchweserNotes™ 2023

Level I CFA®



SCHWESERNOTES™ 2023 LEVEL I CFA® BOOK 3: CORPORATE ISSUERS AND EQUITY INVESTMENTS

©2022 Kaplan, Inc. All rights reserved.

Published in 2022 by Kaplan, Inc.

Printed in the United States of America.

ISBN: 978-1-0788-2588-7

These materials may not be copied without written permission from the author. The unauthorized duplication of these notes is a violation of global copyright laws and the CFA Institute Code of Ethics. Your assistance in pursuing potential violators of this law is greatly appreciated.

Required CFA Institute disclaimer: Kaplan Schweser is a CFA Institute Prep Provider. Only CFA Institute Prep Providers are permitted to make use of CFA Institute copyrighted materials which are the building blocks of the exam. We are also required to update our materials every year and this is validated by CFA Institute.

CFA Institute does not endorse, promote, review or warrant the accuracy or quality of the product and services offered by Kaplan Schweser. CFA Institute®, CFA® and “Chartered Financial Analyst®” are trademarks owned by CFA Institute.

Certain materials contained within this text are the copyrighted property of CFA Institute. The following is the copyright disclosure for these materials: “Copyright, 2022, CFA Institute. Reproduced and republished from 2023 Learning Outcome Statements, Level I, II, and III questions from CFA® Program Materials, CFA Institute Standards of Professional Conduct, and CFA Institute’s Global Investment Performance Standards with permission from CFA Institute. All Rights Reserved.”

Disclaimer: The SchweserNotes should be used in conjunction with the original readings as set forth by CFA Institute in their 2023 Level I CFA Study Guide. The information contained in these Notes covers topics contained in the readings referenced by CFA Institute and is believed to be accurate. However, their accuracy cannot be guaranteed nor is any warranty conveyed as to your ultimate exam success. The authors of the referenced readings have not endorsed or sponsored these Notes.

Contents

Learning Outcome Statements (LOS)

CORPORATE ISSUERS

READING 28

Corporate Structures and Ownership

Exam Focus

Module 28.1: Business Structures

Key Concepts

Answer Key for Module Quiz

READING 29

Introduction to Corporate Governance and Other ESG Considerations

Exam Focus

Module 29.1: Stakeholder Management

Module 29.2: Factors Affecting Corporate Governance

Key Concepts

Answer Key for Module Quizzes

READING 30

Business Models & Risks

Exam Focus

Module 30.1: Business Models

Key Concepts

Answer Key for Module Quiz

READING 31

Capital Investments

Exam Focus

Module 31.1: Capital Allocation Principles

Module 31.2: Net Present Value and Internal Rate of Return

Module 31.3: Capital Allocation Pitfalls and Real Options

Key Concepts

Answer Key for Module Quizzes

READING 32

Working Capital & Liquidity

Exam Focus

Module 32.1: Sources of Capital

Key Concepts

Answers to Module Quiz

READING 33

Cost of Capital—Foundational Topics

Exam Focus

Module 33.1: Weighted Average Cost of Capital

Module 33.2: Beta Estimation and Flotation Costs

Key Concepts

Answer Key for Module Quizzes

READING 34

Capital Structure

Exam Focus

Module 34.1: Capital Structure Theories

Module 34.2: Capital Structure Decisions

Key Concepts

Answer Key for Module Quizzes

READING 35

Measures of Leverage

Exam Focus

Module 35.1: Measures of Leverage

Key Concepts

Answer Key for Module Quiz

Topic Quiz: Corporate Issuers

EQUITY INVESTMENTS

READING 36

Market Organization and Structure

Exam Focus

Module 36.1: Markets, Assets, and Intermediaries

Module 36.2: Positions and Leverage

Module 36.3: Order Execution and Validity

Key Concepts

Answer Key for Module Quizzes

READING 37

Security Market Indexes

Exam Focus

Module 37.1: Index Weighting Methods

Module 37.2: Uses and Types of Indexes

Key Concepts

Answer Key for Module Quizzes

READING 38

Market Efficiency

Exam Focus

Module 38.1: Market Efficiency

Key Concepts

Answer Key for Module Quiz

READING 39

Overview of Equity Securities

Exam Focus

Module 39.1: Types of Equity Investments

Module 39.2: Foreign Equities and Equity Risk

Key Concepts

Answer Key for Module Quizzes

READING 40

Introduction to Industry and Company Analysis

Exam Focus

Module 40.1: Industry Analysis

Module 40.2: Pricing Power and Company Analysis

Key Concepts

Answer Key for Module Quizzes

READING 41

Equity Valuation: Concepts and Basic Tools

Exam Focus

Module 41.1: Dividends, Splits, and Repurchases

Module 41.2: Dividend Discount Models

Module 41.3: Relative Valuation Measures

Key Concepts

Answer Key for Module Quizzes

Topic Quiz: Equity Investments

Formulas

Index

Learning Outcome Statements (LOS)

28. Corporate Structures and Ownership

The candidate should be able to:

- a. compare business structures and describe key features of corporate issuers.
- b. compare public and private companies.
- c. compare the financial claims and motivations of lenders and owners.

29. Introduction to Corporate Governance and Other ESG Considerations

The candidate should be able to:

- a. describe a company's stakeholder groups and compare their interests.
- b. describe the principal-agent relationship and conflicts that may arise between stakeholder groups.
- c. describe corporate governance and mechanisms to manage stakeholder relationships and mitigate associated risks.
- d. describe both the potential risks of poor corporate governance and stakeholder management and the benefits from effective corporate governance and stakeholder management.
- e. describe environmental, social, and governance considerations in investment analysis.
- f. describe environmental, social, and governance investment approaches.

30. Business Models & Risks

The candidate should be able to:

- a. describe key features and types of business models.
- b. describe expected relations between a company's external environment, business model, and financing needs.
- c. explain and classify types of business and financial risks for a company.

31. Capital Investments

The candidate should be able to:

- a. describe types of capital investments made by companies.
- b. describe the capital allocation process and basic principles of capital allocation.
- c. demonstrate the use of net present value (NPV) and internal rate of return (IRR) in allocating capital and describe the advantages and disadvantages of each method.
- d. describe common capital allocation pitfalls.
- e. describe expected relations among a company's investments, company value, and share price.
- f. describe types of real options relevant to capital investment.

32. Working Capital & Liquidity

The candidate should be able to:

- a. compare methods to finance working capital.
- b. explain expected relations between working capital, liquidity, and short-term funding needs.
- c. describe sources of primary and secondary liquidity and factors affecting a company's liquidity position.
- d. compare a company's liquidity position with that of peers.
- e. evaluate short-term funding choices available to a company.

33. Cost of Capital—Foundational Topics

The candidate should be able to:

- a. calculate and interpret the weighted average cost of capital (WACC) of a company.
- b. describe how taxes affect the cost of capital from different capital sources.
- c. calculate and interpret the cost of debt capital using the yield-to-maturity approach and the debt-rating approach.
- d. calculate and interpret the cost of noncallable, nonconvertible preferred stock.
- e. calculate and interpret the cost of equity capital using the capital asset pricing model approach and the bond yield plus risk premium approach.
- f. explain and demonstrate beta estimation for public companies, thinly traded public companies, and nonpublic companies.
- g. explain and demonstrate the correct treatment of flotation costs.

34. Capital Structure

The candidate should be able to:

- a. explain factors affecting capital structure.
- b. describe how a company's capital structure may change over its life cycle.
- c. explain the Modigliani-Miller propositions regarding capital structure.
- d. describe the use of target capital structure in estimating WACC, and calculate and interpret target capital structure weights.
- e. describe competing stakeholder interests in capital structure decisions.

35. Measures of Leverage

The candidate should be able to:

- a. define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk.
- b. calculate and interpret the degree of operating leverage, the degree of financial leverage, and the degree of total leverage.
- c. analyze the effect of financial leverage on a company's net income and return on equity.
- d. calculate the breakeven quantity of sales and determine the company's net income at various sales levels.
- e. calculate and interpret the operating breakeven quantity of sales.

36. Market Organization and Structure

The candidate should be able to:

- a. explain the main functions of the financial system.
- b. describe classifications of assets and markets.
- c. describe the major types of securities, currencies, contracts, commodities, and real assets that trade in organized markets, including their distinguishing characteristics and major subtypes.
- d. describe types of financial intermediaries and services that they provide.
- e. compare positions an investor can take in an asset.
- f. calculate and interpret the leverage ratio, the rate of return on a margin transaction, and the security price at which the investor would receive a margin call.
- g. compare execution, validity, and clearing instructions.
- h. compare market orders with limit orders.
- i. define primary and secondary markets and explain how secondary markets support primary markets.
- j. describe how securities, contracts, and currencies are traded in quote-driven, order-driven, and brokered markets.
- k. describe characteristics of a well-functioning financial system.
- l. describe objectives of market regulation.

37. Security Market Indexes

The candidate should be able to:

- a. describe a security market index.
- b. calculate and interpret the value, price return, and total return of an index.
- c. describe the choices and issues in index construction and management.
- d. compare the different weighting methods used in index construction.
- e. calculate and analyze the value and return of an index given its weighting method.
- f. describe rebalancing and reconstitution of an index.
- g. describe uses of security market indexes.
- h. describe types of equity indexes.
- i. compare types of security market indexes.
- j. describe types of fixed-income indexes.
- k. describe indexes representing alternative investments.

38. Market Efficiency

The candidate should be able to:

- a. describe market efficiency and related concepts, including their importance to investment practitioners.
- b. contrast market value and intrinsic value.
- c. explain factors that affect a market's efficiency.
- d. contrast weak-form, semi-strong-form, and strong-form market efficiency.
- e. explain the implications of each form of market efficiency for fundamental analysis, technical analysis, and the choice between active and passive portfolio management.
- f. describe market anomalies.
- g. describe behavioral finance and its potential relevance to understanding market anomalies.

39. Overview of Equity Securities

The candidate should be able to:

- a. describe characteristics of types of equity securities.
- b. describe differences in voting rights and other ownership characteristics among different equity classes.
- c. compare and contrast public and private equity securities.
- d. describe methods for investing in non-domestic equity securities.
- e. compare the risk and return characteristics of different types of equity securities.
- f. explain the role of equity securities in the financing of a company's assets.
- g. contrast the market value and book value of equity securities.
- h. compare a company's cost of equity, its (accounting) return on equity, and investors' required rates of return.

40. Introduction to Industry and Company Analysis

The candidate should be able to:

- a. explain uses of industry analysis and the relation of industry analysis to company analysis.
- b. compare methods by which companies can be grouped.
- c. explain the factors that affect the sensitivity of a company to the business cycle and the uses and limitations of industry and company descriptors such as "growth," "defensive," and "cyclical".
- d. describe current industry classification systems, and identify how a company should be classified, given a description of its activities and the classification system.
- e. explain how a company's industry classification can be used to identify a potential "peer group" for equity valuation.
- f. describe the elements that need to be covered in a thorough industry analysis.
- g. describe the principles of strategic analysis of an industry.
- h. explain the effects of barriers to entry, industry concentration, industry capacity, and market share stability on pricing power and price competition.
- i. describe industry life-cycle models, classify an industry as to life-cycle stage, and describe limitations of the life-cycle concept in forecasting industry performance.
- j. describe macroeconomic, technological, demographic, governmental, social, and environmental influences on industry growth, profitability, and risk.
- k. compare characteristics of representative industries from the various economic sectors.
- l. describe the elements that should be covered in a thorough company analysis.

41. Equity Valuation: Concepts and Basic Tools

The candidate should be able to:

- a. evaluate whether a security, given its current market price and a value estimate, is overvalued, fairly valued, or undervalued by the market.
- b. describe major categories of equity valuation models.
- c. describe regular cash dividends, extra dividends, stock dividends, stock splits, reverse stock splits, and share repurchases.
- d. describe dividend payment chronology.
- e. explain the rationale for using present value models to value equity and describe the dividend discount and free-cash-flow-to-equity models.
- f. explain advantages and disadvantages of each category of valuation model.
- g. calculate the intrinsic value of a non-callable, non-convertible preferred stock.
- h. calculate and interpret the intrinsic value of an equity security based on the Gordon (constant) growth dividend discount model or a two-stage dividend discount model, as appropriate.
- i. identify characteristics of companies for which the constant growth or a multistage dividend discount model is appropriate.
- j. explain the rationale for using price multiples to value equity, how the price to earnings multiple relates to fundamentals, and the use of multiples based on comparables.
- k. calculate and interpret the following multiples: price to earnings, price to an estimate of operating cash flow, price to sales, and price to book value.
- l. describe enterprise value multiples and their use in estimating equity value.
- m. describe asset-based valuation models and their use in estimating equity value.

READING 28

CORPORATE STRUCTURES AND OWNERSHIP

EXAM FOCUS

Nothing difficult here, just learn the types of business organizations, the difference between public and private companies, and why the difference between the claims of equity holders and debt holders leads to a conflict of interests on some corporate issues.

MODULE 28.1: BUSINESS STRUCTURES



LOS 28.a: Compare business structures and describe key features of corporate issuers.

Video covering this content is available online.

Business structures refer to how businesses are set up from a legal and organizational point of view. Key features of business structures include:

- The legal relationship between the business and its owners.
- Whether the owners of the business also operate the business, and if not, the nature of the relationship between its owners and operators.
- Whether the owners' liability for the actions and debts of the business is limited or unlimited.
- The tax treatment of profits or losses from the business.

To understand these features, we can compare them among four commonly used types of business structures: sole proprietorships, general partnerships, limited partnerships, and corporations.

A **sole proprietorship** is a business owned and operated by an individual. Legally the business is an extension of the owner, who is personally responsible for claims against the business (i.e., unlimited liability), including taxes, and has the only claim on the net profits from the business. Sole proprietorships tend to be small in scale because they can only expand within the limits of the individual owner's ability to secure financing.

To do business on a scale that exceeds that of a sole proprietorship, two or more individuals can form a **general partnership**. In this structure, the **partnership agreement** specifies each partner's responsibilities for business operations and their shares of the partnership profits or losses. As with a sole proprietorship, the partners have unlimited liability for claims against the business, and profits from the business allocated to each partner are taxed as personal income.

A **limited partnership** involves two levels of partners. One or more general partners operate the business and have unlimited liability, as in a general partnership, but this structure also has **limited partners** who are liable only for the amount they invest in the partnership (i.e., **limited liability**) and have claims to its profits that are proportionate to their investments. Limited partners typically are not involved in appointing or removing general partners. How the profits are divided among the general and limited partners is specified in the partnership agreement. Because they are responsible for managing the business, the general partners typically receive a larger portion of profits than the limited partners. Profits allocated to the partners are taxed as personal income to each partner. As we explain in the Alternative Investments topic area, most private capital firms and hedge funds are structured as limited partnerships. Many large providers of professional services, such as legal and accounting firms, also use limited partnership structures.

The feature that distinguishes a **corporation**, or **limited company**, from the other business structures is that a corporation is a legal entity separate from its owners and managers. In this case, *all* a corporation's shareholders have limited liability. An owner can lose his entire investment if the company goes bankrupt and the value of his shares goes to zero, but has no legal liability for any further claims against the corporation. A corporation may, but is not required to, distribute its profits to its owners. Most large firms are corporations because that structure gives them the greatest access to capital, both debt (**borrowed capital**) and equity (**ownership capital**).

Another distinguishing feature of corporations is the separation of its owners and managers. An investor who buys shares of a corporation does not directly influence the company's day-to-day operations. Instead, the owners appoint a **board of directors** that is responsible for hiring the senior managers to operate the company. The board and the managers it hires are responsible for acting in the interests of the shareholders.



PROFESSOR'S NOTE

We will examine the voting rights of shareholders in the Equity Investments topic area.

Corporations can be for-profit or not-for-profit. The purpose of a **not-for-profit corporation** is to produce a particular social benefit or pursue a charitable goal on an ongoing basis. A nonprofit corporation may generate profits, but must reinvest any profits toward its mission rather than distributing them to owners. Nonprofit corporations are usually not taxed.

For-profit corporations may be public or private. In many countries, including the United States, a **public corporation** is one that has shares that are sold to the public and trade in an organized market. Other countries may consider a corporation public if it has at least a designated number of owners, even if its shares are not traded on an exchange or in a dealer market. A for-profit corporation that does not meet these definitions is a **private corporation**.

Depending on the country, a corporation's profits may be subject to **double taxation** if the government taxes companies on their earnings and also taxes dividends (which are distributions of earnings to owners) as personal income. For example, if corporations pay 30% tax on gross profits and individuals pay 20% tax on dividends received, the effective tax rate on profits distributed as dividends is 30% plus 20% of the remaining (1 - 30%), which equals 44%.



PROFESSOR'S NOTE

The Level II curriculum discusses how double taxation affects companies' dividend policies and some of the ways countries address this issue.

LOS 28.b: Compare public and private companies.

Key differences between public and private companies include how they issue shares to owners, how owners can transfer their shares, and the disclosure requirements to which the companies are subject.

A company can become public by issuing shares in an **initial public offering (IPO)**, after which its shares typically trade on an exchange. Once the shares are listed on an exchange, owners can sell shares, and new owners can buy shares, without dealing directly with the company.



PROFESSOR'S NOTE

We describe the mechanics of IPOs in the Equity Investments topic area.

Regulators require public companies to periodically, typically quarterly or semiannually, report their financial results in compliance with accepted accounting principles, and disclose other relevant information such as share purchases and sales by company executives.

Private companies can raise equity capital through a **private placement** of shares to **accredited investors**, typically institutions or high net worth individuals. A **private placement memorandum (PPR)** includes information about the company and the risks of investing in it. Disclosure requirements are less strict than those that apply to public companies. For example, private companies are not required to report financial results publicly or to a regulatory authority.

Investors in private companies tend to have long time horizons. They cannot sell their shares readily or without the company's approval, as shareholders of public companies can. Rather, they typically hold these investments until the company goes public or is acquired by another company. Returns on these investments may be greater on average than investments in public companies, especially for owners that invest early in a company's life.

Besides an IPO or being acquired by a public company, two other ways a private company can become public are through a direct listing or a special purpose acquisition company. In a **direct listing**, a stock exchange agrees to list a private company's existing shares. This differs from an IPO in that a direct listing does not raise any new capital for the company, but has advantages in that it can be done more quickly than an IPO and without involving an underwriter.

A **special purpose acquisition company (SPAC)** is a corporate structure set up to acquire a private company in the future. The SPAC raises capital through an IPO and puts the funds into a trust that it must use to make an acquisition within a specified period of time. The acquired company does not have to be identified at the time of the IPO. For this reason, SPACs are also known as **blank check companies**.

Some transactions result in a public company becoming private. These include a **leveraged buyout (LBO)**, in which outside investors buy all of the company's outstanding shares and

remove its stock exchange listing, and a **management buyout (MBO)** in which the company's managers do so.

LOS 28.c: Compare the financial claims and motivations of lenders and owners.

A company's lenders (debt holders) have a legal, contractual claim to the interest and principal payments the company has promised to make. Owners (equity holders) have a **residual claim** to the company's net assets (i.e., what remains after all other claims have been paid). That is, lenders have a higher priority of claims than equity owners.

Both debt holders and equity holders can potentially lose their entire investment if a company fails, but their losses cannot exceed the amounts they have invested. A key difference between debt and equity investments is their upside potential. Regardless of a company's success, the best result debt holders can achieve is to receive the interest and principal payments promised by the company. Equity, on the other hand, has a theoretically unlimited upside if a company succeeds and grows over time.

Because of this difference in their risk profiles, the interests of debt holders may conflict with the interests of equity holders. Debt holders are primarily concerned with a company's ability to repay its obligations, and less concerned with its growth prospects. Equity holders may favor actions that increase a company's potential growth, but also increase its risk level, such as adding financial leverage by issuing new debt. A company's existing debt holders may oppose such actions because increasing the company's risk (and the probability of defaulting on its debts) does not increase their expected return.



MODULE QUIZ 28.1

1. Which business structure has the largest degree of separation between the owners and operators of a business?
 - A. Corporation.
 - B. Limited partnership.
 - C. General partnership.
2. A public company can become a private company through a:
 - A. direct listing.
 - B. leveraged buyout.
 - C. special purpose acquisition company.
3. For a company that is financially sound, increasing the company's rate of growth is *most likely* to benefit:
 - A. equity holders, but not debt holders.
 - B. both debt holders and equity holders.
 - C. neither debt holders nor equity holders.

KEY CONCEPTS

LOS 28.a

A sole proprietorship is owned and operated by an individual. The owner has unlimited liability, profits are taxed as personal income, and the owner has the only residual claim on the net assets of the business.

In a general partnership, a partnership agreement states each partner's operating responsibilities and claims on partnership net assets. Partners have unlimited liability and profits are taxed as personal income.

A limited partnership has one or more general partners who manage the business and have unlimited liability, and multiple limited partners who each have limited liability. A partnership agreement specifies how profits are to be divided among the general and limited partners. Profits are taxed as personal income to all partners.

A corporation is a legal entity separate from its owners and managers. All owners have limited liability. Owners appoint a board of directors that hires managers to operate the company. A corporation may, but is not required to, distribute its profits to its owners as dividends. Dividends are subject to double taxation (corporate tax and personal income tax) in some countries.

LOS 28.b

A public company is one that (depending on the country) has shares listed on a stock exchange or has at least a minimum number of owners. Public companies are subject to greater disclosure requirements than private companies. Private companies can raise capital through private placements of securities, but only to accredited investors (not to the public).

A private company can become public by issuing shares in an initial public offering, carrying out a direct listing on a stock exchange, or being acquired by a public company, which may be a special purpose acquisition company. In a leveraged buyout or a management buyout, a public company is changed to a private company.

LOS 28.c

Debt holders have a legal claim to the interest and principal the company promises. Equity holders have a residual claim to the company's net assets after debt repayment; that is, debt holders have a higher priority of claims than equity holders.

Debt has limited upside potential; the best result for debt holders is to receive the promised principal and interest payments. Equity has (theoretically) unlimited upside potential. This difference may increase conflicts of interest between debt holders and equity holders.

ANSWER KEY FOR MODULE QUIZ

Module Quiz 28.1

- 1. A** In a corporation, owners are most often not directly involved in operating the business. Both general partnerships and limited partnerships have general partners who operate the business. (LOS 28.a)
- 2. B** Leveraged buyouts can result in a public company going private. Direct listings and special purpose acquisition companies are methods for a private company to go public. (LOS 28.b)
- 3. A** Assuming a company is repaying interest and principal in full and on time, debt holders have no further claims. Equity holders benefit from company growth. (LOS 28.c)

READING 29

INTRODUCTION TO CORPORATE GOVERNANCE AND OTHER ESG CONSIDERATIONS

EXAM FOCUS

Candidates should understand the idea of a firm's stakeholders, how conflicts can arise between stakeholders, how effective corporate governance can mitigate problems arising from these conflicts, and the rationale for incorporating environmental, social, and governance factors into the portfolio selection process. As of 2021, \$46 trillion of managed assets, 40% of the total, globally have a mandate to consider ESG characteristics and this will increase over time.¹ This seems reason enough to take a good look at what this actually means to various investors.

MODULE 29.1: STAKEHOLDER MANAGEMENT



LOS 29.a: Describe a company's stakeholder groups and compare their interests.

Video covering this content is available online.

Under **shareholder theory**, the primary focus of a system of corporate governance is the interests of the firm's shareholders, which are taken to be the maximization of the market value of the firm's common equity. Under this theory, corporate governance is primarily concerned with the conflict of interest between the firm's managers and its owners (shareholders).

The focus of corporate governance under **stakeholder theory** is broader; it considers conflicts among the several groups that have an interest in the activities and performance of the firm. These groups include shareholders, employees, suppliers, and customers, among others.

The following have been identified as the primary stakeholders of a corporation:

Shareholders have a residual interest in the corporation in that they have claim to the net assets of the corporation after all liabilities have been settled. Shareholders have voting rights for the election of the board of directors and for other important corporate matters, which gives them effective control of the firm and its management. They have an interest in the ongoing profitability and growth of the firm, both of which can increase the value of their ownership shares.

The **board of directors** has a responsibility to protect the interests of shareholders; to hire, fire, and set the compensation of the firm's senior managers; to set the strategic direction of the

firm; and to monitor financial performance and other aspects of the firm's ongoing activities.

Typically, the firm's executives (most-senior managers) serve on the board of directors, along with directors who are not otherwise employed by the firm. In a one-tier board structure, both company executives and non-executive board members serve on a single board of directors. In some countries, boards have a two-tier structure in which the non-executive board members serve on a supervisory board that oversees a management board, made up of company executives.

Senior managers typically receive compensation (remuneration) that is made up of a salary, a bonus based on some measure of company performance, and perquisites (e.g., expense accounts, use of company planes, special retirement benefits, vacation time off). Their interests can be expected to include continued employment and maximizing the total value of their compensation. Executive bonuses are typically tied to some measure of firm performance, giving senior managers a strong interest in the financial success of the firm.

Employees also have an interest in the sustainability and success of the firm. They have an interest in their rate of pay, opportunities for career advancement, training, and working conditions.

Creditors supply debt capital to the firm and are primarily owners of the firm's outstanding bonds and banks that have made loans to the firm. Providers of debt capital to the firm do not typically have a vote in firm management and do not participate in firm growth beyond receiving their promised interest and principal payments. The interests of creditors are protected to varying degrees by covenants in their debt agreements with the firm.

Suppliers of resources to the firm have an interest preserving an ongoing relationship with the firm, in the profitability of their trade with the firm, and in the growth and ongoing stability of the firm. As they are typically short-term creditors of the firm, they also have an interest in the firm's solvency and ongoing financial strength.

LOS 29.b: Describe the principal-agent relationship and conflicts that may arise between stakeholder groups.

The **principal-agent conflict** arises because an agent is hired to act in the interests of the principal, but an agent's interests may not coincide exactly with those of the principal. Consider an insurance agent who is paid a commission on policies written. It would be in the agent's interest to write insurance policies on people or property that are not good risks, in order to maximize commission income. The principal (the owner of the insurance company) does not want to issue policies that are bad risks as that is a money-losing proposition. Insurance companies mitigate this conflict by imposing underwriting standards for the policies they will issue and by continuing to work only with agents who consistently act in the company's best interest.

Conflicts of Interest Between Shareholders and Managers or Directors

In the context of a corporation, shareholders are the principals (owners), and firm management and board members (directors) are their agents. Managers and directors may choose a lower level of business risk than shareholders would. This conflict can arise because the risk of company managers and directors is more dependent of firm performance compared to the risk of shareholders, who hold diversified portfolios of stocks and are not dependent on the firm for employment.

Conflicts may also arise when directors who are also managers favor management interests at the expense of shareholders or when directors favor one group of shareholders at the expense of another.

There is also an **information asymmetry** between shareholders and managers because managers have more and better information about the functioning of the firm and its strategic direction than shareholders do. This decreases the ability of shareholders or non-executive directors to monitor and evaluate whether managers are acting in the best interests of shareholders.

Conflicts Between Groups of Shareholders

A single shareholder or group of shareholders may hold a majority of the votes and act against the interests of the minority shareholders. Some firms have different classes of common stock outstanding, some with more voting power than others. A group of shareholders may have effective control of the company although they have a claim to less than 50% of the earnings and assets of the company.

In the event of an acquisition of the company, controlling shareholders may be in a position to get better terms for themselves relative to the terms forced on minority shareholders. Majority shareholders may cause the company to enter into **related party transactions**, agreements or specific transactions that benefit entities in which they have a financial interest, to the detriment of minority shareholders.

Conflicts of Interest Between Creditors and Shareholders

Shareholders may prefer more business risk than creditors do because creditors have a limited upside from good results compared to shareholders. Equity owners could also act against the interests of creditors by issuing new debt that increases the default risk faced by existing debt holders, or by the company paying greater dividends to equity holders, thereby increasing creditors' risk of default.

Conflicts of Interest Between Shareholders and Other Stakeholders

The company may decide to raise prices or reduce product quality in order to increase profits to the detriment of customers. The company may employ strategies that significantly reduce the taxes they pay to the government.

LOS 29.c: Describe corporate governance and mechanisms to manage stakeholder relationships and mitigate associated risks.

In the CFA Institute publication, *The Corporate Governance of Listed Companies: A Manual for Investors*,² **corporate governance** is described as “the system of internal controls and procedures by which individual companies are managed. It provides a framework that defines the rights, roles, and responsibilities of various groups . . . within an organization. At its core, corporate governance is the arrangement of checks, balances, and incentives a company needs in order to minimize and manage the conflicting interests between insiders and external shareowners.”

Stakeholder management refers to the management of company relations with stakeholders and is based on having a good understanding of stakeholder interests and maintaining effective communication with stakeholders. With respect to the company’s relationship with shareholders, there are standard practices. These practices are required by corporate laws and similar in many jurisdictions, although there are some differences across countries.

We can classify mechanisms to manage stakeholder relationships as shareholder, creditor, board and management, employee, customer and supplier, and government mechanisms.

Shareholder Mechanisms

Corporations typically hold an **annual general meeting** after the end of the firm’s fiscal year. At the general meeting, company management provides shareholders with the audited financial statements for the year, addresses the company’s performance and significant actions over the period, and answers shareholder questions.

Corporate laws dictate when the annual general meeting may occur and how the meeting must be communicated to shareholders. Typically, anyone owning shares is permitted to attend the annual general meeting, to speak or ask questions, and to vote their shares. A shareholder who does not attend the annual general meeting can vote her shares by **proxy**, meaning she assigns her right to vote to another who will attend the meeting, often a director, member of management, or the shareholder’s investment advisor. A proxy may specify the shareholder’s vote on specific issues or leave the vote to the discretion of the person to whom the proxy is assigned.

Ordinary resolutions, such as approval of auditor and the election of directors, require a simple majority of the votes cast. Other resolutions, such as those regarding a merger or takeover, or that require amendment of corporate bylaws, are termed **special resolutions** and may require a supermajority vote for passage, typically two-thirds or three-fourths of the votes cast. Such special resolutions can also be addressed at **extraordinary general meetings**, which can be called anytime there is a resolution about a matter that requires a vote of the shareholders.

When there are multiple board member elections at one meeting, some companies use majority voting and some use cumulative voting. With **majority voting**, the candidate with the most votes for each single board position is elected. With **cumulative voting**, shareholders can cast all their votes (shares times number of board position elections) for a single board candidate or divide them among board candidates. Cumulative voting can result in greater minority shareholder representation on the board compared to majority voting. Minority shareholders may have special rights by law when the company is acquired by another company.

Activist shareholders pressure companies in which they hold a significant number of shares for changes, often changes they believe will increase shareholder value. They may bring

pressure for change by initiating shareholder lawsuits or by seeking representation on the board of directors. Other activist tactics include proposing shareholder resolutions for a vote and raising their issues to all shareholders or the public to gain wider support. Hedge funds have, more and more, engaged in shareholder activism to increase the market values of firms in which they hold significant stakes.

A group may initiate a **proxy fight**, in which they seek the proxies of shareholders to vote in favor of their alternative proposals. An activist group may make a **tender offer** for a specific number of shares of a company to gain enough votes to control the company.

Both senior managers and boards of directors can be replaced by shareholders when they believe company performance is poor and would be improved by change. The threat of a **hostile takeover**, one not supported by the company's management, can act as an incentive for company managements and boards to pursue policies better aligned with the interests of shareholders.

Creditor Mechanisms

When a company issues a bond, it specifies the rights of bondholders and the company's obligations in a legal document called a **bond indenture**. An indenture typically includes **covenants** that may require the company to take certain actions or restrict it from taking certain actions. A bond can be backed by **collateral**, which is a specific asset against which the bondholders will have a claim if the company defaults on the bond. A financial institution may act as a trustee to monitor the company's compliance with its bond covenants.



PROFESSOR'S NOTE

We explain covenants in more detail in the Fixed Income topic area.

Creditor committees may form among bondholders to protect their interests when an issuer experiences financial distress. Some countries require such committees when a company files for bankruptcy.

Board of Directors and Management Mechanisms

The **board of directors** is elected by shareholders to act in their interests. A board of directors typically has committees made up of board members with particular expertise. These committees report to the board, which retains the overall responsibility for the various board functions. The following are examples of typical board committees.

An **audit committee** is responsible for:

- Oversight of the financial reporting function and implementation of accounting policies.
- Effectiveness of the company's internal controls and the internal audit function.
- Recommending an external auditor and its compensation.
- Proposing remedies based on their review of internal and external audits.

A **governance committee** is responsible for:

- Oversight of the company's corporate governance code.
- Implementing the company's code of ethics and policies regarding conflicts of interest.

- Monitoring changes in relevant laws and regulations.
- Ensuring that the company is in compliance with all applicable laws and regulations, as well as with the company's governance policies.

A **nominations committee** proposes qualified candidates for election to the board, manages the search process, and attempts to align the board's composition with the company's corporate governance policies.

A **compensation committee** or **remuneration committee** recommends to the board the amounts and types of compensation to be paid to directors and senior managers. This committee may also be responsible for oversight of employee benefit plans and evaluation of senior managers.

A **risk committee** informs the board about appropriate risk policy and risk tolerance of the organization, and oversees the enterprise-wide risk management processes of the organization.

An **investment committee** reviews and reports to the board on management proposals for large acquisitions or projects, sale or other disposal of company assets or segments, and the performance of acquired assets and other large capital expenditures.

The number and size of board committees will depend on the size, complexity, and nature of the business. Regulations often require that firms have audit committees. Financial services firms are often required to have a risk committee as well. Some companies combine two functions into one committee. The composition of a board committee is often based on its function, with audit committees, compensation committees, and governance committees often made up of only non-executive or independent directors.

Employee, Customer, and Supplier Mechanisms

Labor laws, employment contracts, and the right to form unions are the primary mechanisms for employees to manage relationships with employers. Some countries have laws requiring that large companies' boards of directors include employee representatives. Employee stock ownership plans may help align company and employee interests. For customers and suppliers, contracts tend to be the mechanism through which they manage their relationships with companies. In recent years, customers and other stakeholders have increasingly used social media as a mechanism to influence company behavior.

Government Mechanisms

Governments enact and enforce regulations that govern companies' actions. They often do so by establishing agencies to regulate industries or sectors such as financial markets, or by monitoring specific issues such as workplace safety and environmental protection. In some countries regulators adopt corporate governance codes that companies must either adopt or explain why they have not done so.

An important factor that can affect stakeholder relationships is the legal environment within which a company operates. Shareholders' and creditors' interests are considered to be better protected in countries with a **common-law system** under which judges' rulings become law in some instances. In a **civil law system**, judges are bound to rule based only on specifically

enacted laws. In general, the rights of creditors are more clearly defined than those of shareholders and, therefore, are not as difficult to enforce through the courts.



MODULE QUIZ 29.1

1. The theory that deals with conflicts of interest between a company's owners and its creditors is *most appropriately* called:
 - A. structure theory.
 - B. stakeholder theory.
 - C. shareholder theory.
2. For which two of a company's stakeholders does information asymmetry *most likely* make monitoring more difficult?
 - A. Suppliers and employees.
 - B. Employees and managers.
 - C. Managers and shareholders.
3. The *least likely* item to be a requirement for good stakeholder management is:
 - A. maintaining effective communication with other stakeholders.
 - B. an understanding of the interests of several stakeholder groups.
 - C. the ability to put aside the interests of one's stakeholder group.
4. The type of voting that is *most likely* to allow minority stockholders a greater representation on the board of directors is:
 - A. majority voting.
 - B. supermajority voting.
 - C. cumulative voting.
5. The type of resolution *most likely* to require a supermajority of shareholder votes for passage is a resolution to:
 - A. acquire a company.
 - B. choose a board member.
 - C. approve the choice of an auditor.
6. The board of directors committee *most likely* to be responsible for monitoring the performance of a project that requires a large capital expenditure is:
 - A. the risk committee.
 - B. the audit committee.
 - C. the investment committee.

MODULE 29.2: FACTORS AFFECTING CORPORATE GOVERNANCE



Video covering this content is available online.

LOS 29.d: Describe both the potential risks of poor corporate governance and stakeholder management and the benefits from effective corporate governance and stakeholder management.

Risks of Poor Governance and Stakeholder Management

When corporate governance is weak, the control functions of audits and board oversight may be weak as well. The risk is that some stakeholders can gain an advantage, to the disadvantage of other stakeholders. Accounting fraud, or simply poor recordkeeping, will have negative implications for company performance and value.

When governance is weak and managers are not monitored, they may serve their own interests by choosing less-than-optimal risk, reducing company value. Without proper monitoring and oversight, management may be given incentive compensation that allows them to pursue their own benefit rather than the company's interests. If they are allowed to engage in related-party transactions that benefit their friends or family, this will also decrease company value.

Poor compliance procedures with respect to regulation and reporting can easily lead to legal and reputational risks. Violating stakeholder rights can lead to stakeholder lawsuits. A company's reputation can be damaged by failure to comply with governmental regulations. Failure to manage creditors' rights well can lead to debt default and bankruptcy.

Benefits of Effective Governance and Stakeholder Management

Effective corporate governance can improve operational efficiency by ensuring that management and board member incentives align their interests well with those of shareholders. Effective governance implies effective control and monitoring. Just as weak control can lead to abuses, a strong system of controls and compliance with laws and regulations can avoid many legal and regulatory risks.

Formal policies regarding conflicts of interest and related-party transactions can also lead to better operating results. Proper governance with respect to the interests of creditors can reduce the risk of debt default or bankruptcy, thereby reducing the cost of debt financing. Alignment of management interests with those of shareholders leads to better financial performance and greater company value.

LOS 29.e: Describe environmental, social, and governance considerations in investment analysis.

While the quality of corporate governance has long been a consideration in investment analysis, the consideration of environmental and social factors is a more recent development. The use of environmental, social, and governance factors in making investment decisions is referred to as **ESG investing**. Many issues can be considered in this context, including harm or potential harm to the environment, risk of loss due to environmental accidents, the changing demographics of the workforce, and reputational risks from corrupt practices or human rights abuses.

Some of the terms related to ESG investing include **responsible investing** (a broad term for considering ESG factors in investment decisions), **sustainable investing** (investing in companies or industries based on the perceived sustainability of their output), and **socially responsible investing** (choosing investments based on the investor's moral or social values).

Conflict may occur when integrating ESG considerations into portfolio construction when the manager has a fiduciary responsibility to act in the best financial interests of the account owner or beneficiaries. Choosing to construct a portfolio based on an environmental, social, or governance concern at the expense of investor returns would violate the manager's fiduciary duty. On the other hand, failing to consider the risks that may arise from negative ESG factors may also violate fiduciary duty.

LOS 29.f: Describe environmental, social, and governance investment approaches.

There are several approaches to integrating ESG factors into the portfolio management process. The following are some important examples:

Negative screening refers to excluding specific companies or industries from consideration for the portfolio based on their practices regarding human rights, environmental concerns, or corruption. Examples of industries where ESG factors might lead to exclusion are mining, oil extraction and transport, and tobacco. Specific companies that might be excluded are those with poor records on corruption and human rights (labor) practices. Company scores based on a range of ESG concerns are often used in negative screening to identify companies that should be excluded.

Under the **positive screening** approach, investors attempt to identify companies that have positive ESG practices. For example, a portfolio manager may focus on environmental sustainability, employee rights and safety, and overall governance practices. Often a scoring system across a set of ESG factors is used to identify companies for inclusion in portfolios. A related approach, the **relative/best-in-class** approach, seeks to identify companies within each industry group with the best ESG practices. By constructing portfolios of these companies, a manager can preserve the index sector weightings in the portfolio while still taking advantage of opportunities to profit from (or simply to support) positive ESG practices.

Full integration refers to the inclusion of ESG factors or ESG scores in traditional fundamental analysis. A company's ESG practices are included in the process of estimating fundamental variables, such as a company's cost of capital or future cash flows. To the extent that ESG practices will affect such variables, integrating them into the analysis can help in determining which companies are currently overpriced or underpriced.

Thematic investing refers to investing in sectors or companies in an attempt to promote specific ESG-related goals, such as more sustainable practices in agriculture, greater use of cleaner energy sources, improved management of water resources, or the reduction of carbon emissions.

Engagement/active ownership investing refers to using ownership of company shares or other securities as a platform to promote improved ESG practices. Share ownership is used to initiate or support (through share voting) positive ESG changes. Contact with senior management or board members to promote such changes is also an active ownership strategy. Recently, this strategy has been used to promote reduction in a company's carbon footprint, increased wages, or other social and environmental goals, which may or may not be associated with improved financial results over time.

Another approach to ESG investing is **green finance**. Green finance refers to producing economic growth in a more sustainable way by reducing emissions and better managing natural resource use. An important part of green finance is the issuance of **green bonds**, bonds for which the funds raised are used for projects with a positive environmental impact. Issuance of green bonds has increased significantly in recent years, led by issuance in the United States and in China, which is prioritizing improvement in environmental conditions.

- 
1. Benefits of effective corporate governance and stakeholder management *most likely* include:
 - A. reduced risk of default.
 - B. more efficient related-party transactions.
 - C. greater control exercised by the most-interested stakeholders.
 2. The method of ESG integration that does not exclude any sectors but seeks to invest in the companies with the best practices regarding employee rights and environmental sustainability is:
 - A. thematic investing.
 - B. positive screening.
 - C. negative screening.

KEY CONCEPTS

LOS 29.a

The primary stakeholders of a corporation include shareholders, the board of directors, senior management, employees, creditors, and suppliers.

LOS 29.b

The principal-agent relationship refers to owners employing agents to act in their interests. Conflicts can arise because the agent's incentives may not align with those of the owner or, more generally, because the interests of one group within a corporation are not the same as those of other groups.

LOS 29.c

Corporate governance refers to the internal controls and procedures of a company that delineate the rights and responsibilities of various groups and how conflicts of interest among the various groups are to be resolved.

Shareholders, creditors, boards of directors, employees, customers, suppliers, and government have different mechanisms with which to manage their stakeholder relationships with companies.

Proxy voting is the primary shareholder mechanism. Shareholders can remove senior managers and boards of directors if they believe company performance would improve with a change. Activist shareholders may engage in proxy fights or hostile takeovers.

Creditor mechanisms include bond indentures and creditor committees. Employee mechanisms include labor laws and unions. Contracts are the primary mechanism for customers and suppliers. Governments may enact regulations or appoint regulatory agencies.

Duties of a board of directors include:

- Selecting senior management, setting their compensation, and evaluating their performance.
- Setting the strategic direction for the company.
- Approving capital structure changes, significant acquisitions, and large investment expenditures.
- Reviewing company performance and implementing any necessary corrective steps.
- Planning for continuity of management and the succession of the CEO.
- Establishing, monitoring, and overseeing the firm's internal controls and risk management.
- Ensuring the quality of the firm's financial reporting and internal audit.

LOS 29.d

The risks of poor governance include weak control systems, poor decision making, legal risk, reputational risk, and default risk. Good corporate governance can improve operational efficiency and performance, reduce default risk, reduce the cost of debt, improve financial performance, and increase firm value.

LOS 29.e

The use of environmental, social, and governance (ESG) factors in making investment decisions is referred to as ESG investing. Many issues can be considered in this context, including harm or potential harm to the environment, risk of loss due to environmental accidents, the changing demographics of the workforce, and reputational risks from corrupt practices or human rights abuses.

LOS 29.f

Methods of integrating ESG concerns or factors in portfolio construction include:

- Negative screening.
- Positive screening.
- Relative/best-in-class investing.
- Full integration.
- Thematic investing.
- Engagement/active ownership.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 29.1

1. **B** Stakeholder theory focuses on the conflicts of interest among owners and several groups that have an interest in a company's activities, including creditors. (LOS 29.a)
2. **C** Information asymmetry can exist between a company's shareholders and its managers because the company's managers may be much more knowledgeable about the company's functioning and strategic direction. This makes it more difficult for shareholders to monitor the firm's managers and determine whether they are acting in shareholders' interests. (LOS 29.a)
3. **C** The ability to manage the conflicting interests of company relations with stakeholders requires good communication with stakeholders and a good understanding of their various interests. (LOS 29.b)
4. **C** With cumulative voting, shareholders get a vote for each share they own times the number of director elections each year and can give all their votes to a single candidate for the board. This helps minority stockholders to get more proportional representation on the board of directors. (LOS 29.c)
5. **A** Ordinary resolutions, such as those to appoint an auditor or elect a board member, require a simple majority. Acquisitions, mergers, takeovers, and amendments to the company bylaws often require a supermajority of more than 50% for passage. (LOS 29.c)

6. **C** The investment committee reviews proposals for large acquisitions or projects and also monitors the performance of acquired assets and of projects requiring large capital expenditures. (LOS 29.c)

Module Quiz 29.2

1. **A** Reduced risk of default is among the benefits of effective corporate governance. Risks from poor corporate governance include related-party transactions by managers and opportunities for some stakeholder groups to gain advantage at the expense of others. (LOS 29.d)
2. **B** Positive screening does not exclude any sectors but seeks to invest in the companies with the best practices. Negative screening typically excludes some sectors. Thematic investing refers to making an investment in a company or project in order to advance specific social or environmental goals. (LOS 29.e, 29.f)

¹ Barrons.com. "ESG Failed Its Big Test. A Reckoning is Coming." April 15, 2022.

²

www.cfainstitute.org/learning/products/publications/readings/Pages/the_corporate_governance_of_listed_companies_a_manual_for_investors.aspx

READING 30

BUSINESS MODELS & RISKS

EXAM FOCUS

A company's business model tells us how and why it intends to make profits and create value for its owners. Here we look at some common types of company business models and how they affect companies' business and financial risk as well as their financing needs. There are a lot of terms for candidates to absorb.

MODULE 30.1: BUSINESS MODELS



LOS 30.a: Describe key features and types of business models.

Video covering
this content is
available online.

A **business model** offers some detail about how a company proposes to make money. A successful firm must provide a product or service, find customers, deliver the product or service, and make a profit. The business model explains how a firm either does or proposes to do this.

A business model is not the same as a **financial plan**, which has detailed projections for revenue and expenses, as well as plans for financing the business.

"How we will provide it, sell it, and make a profit" is clearly an oversimplification, but this is the essence of a business model. In practice, the answers to these questions have many facets. Here, we present a framework that incorporates some of the complexities involved.

A business model should:

- *Identify the firm's potential customers*, how they are acquired, the cost of customer acquisition, and how the company will monitor and maintain customer satisfaction. Potential customers can be defined in innumerable ways, ranging from every consumer within a geographic area, to dog owners, to only the company's home-country military in the case of a weapon.
- *Describe the firm's product or service*, how it meets a need for its potential customers, and what differentiates its products from those of competitors (e.g., low price, premium quality, innovative features).
- *Explain how the firm will sell its product or service* (e.g., online, physical location, direct mail, trade shows, sales representatives); whether they will sell direct to the buyers (**direct sales**) or use intermediaries such as wholesalers, retailers, agents, or franchisees; and how will they deliver their product or service. The answers to these questions comprise a firm's **channel**

strategy. A strategy that includes both digital and physical channels, such as internet sales with delivery at a physical location, is referred to as an **omnichannel strategy**.

Firms that sell to other businesses are said to be **B2B** (business to business) firms, while firms that sell to consumers are said to be **B2C** (business to consumer) firms.

- *Describe the key assets and suppliers of the firm.* Key assets may be, for example, a patent, software, or skilled employees. Key suppliers may be a battery manufacturer for an electric vehicle company, a lithium miner for a battery maker, or a supplier of large excavation machinery for a lithium miner.
- *Explain its pricing strategy* and why buyers will pay that price for their product, given the competitive landscape of the market.

Value-based pricing refers to setting prices based on the value received (or perceived) by the buyer. **Cost-based pricing** refers to setting prices based on the costs of producing the firm's good or service (plus a profit).

Price discrimination refers to setting different prices for different customers or identifiable groups of customers. Common examples are **tiered pricing** (based on volume of purchases), **dynamic pricing** (depending on the time of day or day of the week), such as peak and off-peak pricing and low-priced airline tickets for very early or very late flights, and **auction pricing** (e.g., eBay).

Pricing models for multiple products include:

- **Bundling**—Where multiple products are complementary (e.g., a furnished apartment), bundling the products may be a profitable strategy.
- **Razors-and-blades**—A company may find it profitable to sell a piece of equipment for a relatively low price (low margins) and make profits by selling a consumable used with the equipment. Printers and ink cartridges and an e-reader and e-books are common examples.
- **Optional products**—Options or add-ons priced with high margins are added to the product after the purchase decision has been made. An example is the many pricey options that may be offered after a customer has decided to purchase an automobile.

Other pricing models include:

- **Penetration pricing**—A company offers a product at low margins or even at a loss for a period of time to grow market share and achieve greater scale of operations. Netflix followed this strategy to grow its subscriber base rapidly.
- **Freemium pricing**—Offer a product with basic functionality at no cost, but sell/unlock other functionality for a fee. Video game makers have used this strategy to encourage wide usage and then profit on sales of greater functionality (e.g., weapons).
- **Hidden revenue**—Online content may be “free” but generate revenue through ads. For example, internet search is free to the user with revenue coming from the sale of user data.

Models that offer alternatives to outright purchase include:

- **Subscription model**—Microsoft’s model for software has changed from selling the software to a subscription (paying monthly for access) to their Office suite of software.
- **Fractional ownership**—Time share companies sell condominium ownership by the week; use of private jets is sold for specific amounts of time.

- **Licensing**—For a biotech company that has developed a new and effective drug, it may be most profitable to license the production of the drug to an established drug maker with a large sales force and established distribution channels, rather than developing those resources itself for the single drug.
- **Franchising**—Similar to licensing, but a franchisee typically is permitted to sell in a specific area and pays a percentage of sales to the franchisor, which provides some level of product and marketing support.

A firm's **value proposition** refers to how customers will value the characteristics of the product or service, given the competing products and their prices. How the firm executes its value proposition is referred to as its **value chain**. A firm's value chain comprises the assets of the firm and how the organization of the firm will add value and exploit the firm's competitive advantage. A value chain should not be confused with a firm's **supply chain**, which includes every step in producing and delivering its products, even those that other firms perform.

In his 1985 book *Competitive Advantage*, Michael Porter presents five activities in which firms should strive to execute well:

- Inbound logistics
- Operations
- Outbound logistics
- Marketing
- Sales and service

Other business models include:

- **Private label manufacturers**—Companies produce products for others to market under their own brand name, for example, Costco's Kirkland branded products.
- **Licensing agreements**—A company brand is used by another company on its products for a fee, such as a lunch box branded with a Marvel character.
- **Value-added resellers**—Offer such things as installation, service, support, or customization for complex equipment.

E-commerce models for direct sales include:

- **Affiliate marketing**—Another company is paid a commission for measurable marketing results such as page views, leads, or sales.
- **Marketplace businesses**—Provide a platform for buyers and sellers but do not own the goods being sold. Ebay is a prime example of a marketplace business.
- **Aggregators**—Provide a marketplace but sell products and services under its own brand name. Spotify is an example.

Network effects refer to the increase in the value of a network as its user base grows. There are many examples of this including WhatsApp, eBay, and Facebook. Network effects support an initial strategy of penetration pricing.

Crowdsourcing models benefit from user contributions: content in the case of Wikipedia, traffic conditions and events in the case of Waze, and product improvements or new applications in the case of open-source software.

Hybrid business models incorporate both platform and traditional sales models.

LOS 30.b: Describe expected relations between a company's external environment, business model, and financing needs.

Providers of both debt and equity capital are concerned with firm risk and firm growth. Lenders like to see less uncertainty about earnings, cash flow, operating margins, and the like. Equity holders like earnings growth over time, but are also concerned with earnings volatility.

A firm's overall risk will depend on its business model and other risk factors, both firm-specific and external to the firm. The firm's overall risk, in turn, affects the cost and availability of both debt and equity capital. Here we briefly describe some significant external factors that can affect business risk.

- *Changes in economic conditions* (e.g., economic growth, inflation and interest rates) typically affect all firms to some extent, increasing firm risk. Some industries and sectors have predictable demand for their products and are less affected by economic cycles. Firms with large investments in fixed assets and which face demand that is sensitive to economic cycles can have large cyclical changes in earnings and cash flow.
- *Changing demographics* can affect the demand for some sectors' and firms' products, either positively or negatively. Consumer tastes change over time as well.
- The winds of *political, legal, and regulatory change* affect us all; businesses are no exception.

One firm-specific factor is the stage of firm development. A start-up firm that requires large amounts of capital to grow has different financial needs than a stable, mature firm. Another firm-specific factor is a firm's vulnerability to competition; more vulnerable firms have more business risk.

A firm's business model can have significant effects on its financing needs. Some businesses follow an **asset-light model** renting or leasing major assets, or having them owned by franchisees, to reduce capital requirements. Large companies lease hospitals, hotels, distribution warehouses, and data centers to reduce their capital needs. In the case of **lean start-ups** a firm "rents" its employees, outsourcing as much work as possible to reduce fixed employment costs. Firms with **pay-in-advance models**, such as insurance companies and online retailers, reduce their working capital needs and may use the advance payments to reduce their capital needs even more.

LOS 30.c: Explain and classify types of business and financial risks for a company.

Macro risk refers to the risk (to operating profit) arising from economic, political, and legal risk factors, as well as other risks that affect all businesses within a country or region, such as demographic changes over time. The primary macro risk for many companies is the risk of an economic slowdown or recession. The level of economic activity or growth may affect some companies strongly and we refer to such companies or their industries as *cyclical*. Other companies, such as utilities and health care providers, are not affected strongly by economic

cycles and we refer to them as *non-cyclical* or *defensive*. Multinational companies may face other macro risks such as political instability, political conflict, and changes in exchange rates.

Business risk refers to the variability of operating income (EBIT) that arises from both firm-specific risk factors and industry risk factors. Care must be taken when defining a firm's industry. Narrower definitions are, in general, better for identifying factors that affect firm profitability, but the definitions must be wide enough that data on demand and competition are available and trends can be identified.

Industry risk factors include:

- Revenue and earnings cyclicalities.
- Industry structure: Low concentration (many smaller firms) is associated with high competitive intensity.
- Competitive intensity: Higher competitive intensity in the industry typically reduces profitability.
- Competitive dynamics within the value chain: Profits are affected by actions of buyers, suppliers, and actual and potential competitors.
- Long-term growth and demand expectations: An industry with increasing demand and high long-term growth prospects is more attractive to investors, but may also attract more competition.
- Other industry risks are regulatory risks and other relevant external risks.

Firm-specific risk factors include:

- **Competitive risks** such as the erosion of an existing competitive advantage over time or the introduction of innovative business models that disrupt the industry. Competitive advantage results from cost advantages (including scale of operations), product differentiation, and positive network effects from greater product usage. High costs incurred by a customer to change to a different supplier (**switching costs**) increase the competitive advantage of existing firms. Firms always have **execution risk**, as some managements can find a way to fail with even the best of business plans.
- **Product market risk:** For firms early in their life cycles, expectations of growth in demand may decrease over time, consumer preferences may change, products may become obsolescent, and patents may expire. Firms with many products typically face less product risk.
- **Capital investment risk** refers to investing firm assets in opportunities that do not produce returns above the firm's cost of capital. Many acquisitions (e.g., Time Warner) turn out to be quite ill-advised, while some (e.g., YouTube) turn out to be brilliant.
- **ESG risk** measures often focus on corporate governance risk, but the risk of running afoul of current expectations for environmentally and socially progressive company policies can damage a company's reputation and bottom line (or not, e.g., Volkswagen).
- Business risk is increased by higher **operating leverage** that results from higher percentages of fixed costs, relative to variable costs, in a firm's cost structure. The effect of sales variability on operating income is magnified by higher operating leverage.
- **Financial risk** refers to the increase in the variability of net income and cash flows that results from using debt in a firm's capital structure, which increases financial leverage.

Financial leverage magnifies the effects of business risk on profits. Fixed costs related to leases and underfunded pension obligations also increase financial risk. Higher levels of debt in a firm's capital structure increase the risk of financial distress, default, or even insolvency.



PROFESSOR'S NOTE

Our reading on Measures of Leverage examines the effects and calculation of operating leverage and financial leverage in some detail.



MODULE QUIZ 30.1

1. A business model is *least likely* to include details about a company's:
 - A. largest customers.
 - B. workforce characteristics.
 - C. revenue and expense estimates.
2. Which of the following is *most likely* referred to as a firm-specific risk? A firm's:
 - A. competitive position.
 - B. exchange rate uncertainty.
 - C. exposure to demographic trends.
3. For an online seller of bouquets of flowers with same-day delivery through association with local florists, a decline in sales due to an increasing age of the population is an example of:
 - A. macro risk.
 - B. industry risk.
 - C. firm-specific risk.
4. A local pizza chain, Gino's, has an exclusive relationship with a local cheesemaker for premium mozzarella and has contracted with a food delivery company for prompt delivery of its pies. Which of the following statements is *most* accurate?
 - A. The delivery service is part of Gino's value chain.
 - B. The cheese maker is part of Gino's value proposition.
 - C. The delivery service is not part of Gino's supply chain.

KEY CONCEPTS

LOS 30.a

A business model should identify a firm's potential customers, describe its products or services and explain how it will sell them, describe its key assets and suppliers, and explain its pricing strategy.

Value proposition refers to how a firm's customers will value the characteristics of the product or service. Value chain refers to how a firm executes its value proposition.

LOS 30.b

A firm's overall risk affects the cost and availability of capital and depends on its business model as well as firm-specific and external risk factors.

LOS 30.c

Macro risk arises from economic, political, legal, regulatory, and demographic changes that may affect all businesses within a country or region.

Firm-specific risk factors comprise competitive, execution, product market, capital investment, and ESG risks.

Industry risk factors include cyclicalities, industry structure, competitive intensity and dynamics, and long-term growth expectations.

Business risk refers to the variability of operating income that arises from macro, industry, and firm-specific factors, and is magnified by operating leverage.

ANSWER KEY FOR MODULE QUIZ

Module Quiz 30.1

1. **C** Detailed forecasts of revenue and expenses would be in a financial plan, but typically not in a business model. A firm's largest customers and information about its workforce and its value are likely elements of a business model. (LOS 30.a)
2. **A** Firms with a weak competitive position have more risk than firms with competitive advantages of large scale and brand name recognition. Uncertainty about macroeconomic variables, such as an exchange rate, and the effects of demographic trends are risks considered to be external to the firm. (LOS 30.b)
3. **A** The effects of demographic changes are considered a macro risk. Macro risks may affect different industries or sectors differently, but are not considered industry risk. (LOS 30.c)
4. **B** Prompt delivery and premium ingredients are part of Gino's value proposition. The delivery company is part of Gino's supply chain as it is part of the process of getting the pies to the customers. However, the delivery service is not part of Gino's value chain, which refers only to what Gino's does itself. (LOS 30.a)

READING 31

CAPITAL INVESTMENTS

EXAM FOCUS

How a firm's management allocates the firm's capital to various activities is a vital determinant of the firm's financial results over time. Candidates should understand the process of evaluating investment opportunities. Calculating a project's IRR and NPV and understanding the implications of each measure are key components of capital allocation decisions. Understanding the types of real options and mistakes that are commonly encountered in capital allocation decisions will lead to better decisions.

MODULE 31.1: CAPITAL ALLOCATION PRINCIPLES



Video covering
this content is
available online.

LOS 31.a: Describe types of capital investments made by companies.

Capital investments may be divided into the categories of business maintenance and business growth. **Business maintenance investments** include going concern projects and regulatory/compliance projects.

- **Going concern projects** may be needed to maintain the business or reduce costs. Projects that maintain the business are normally made without detailed analysis. The only issues are whether the existing operations should continue and, if so, whether existing procedures or processes should be maintained. Projects to improve efficiency may involve determining whether equipment that is obsolete, but still usable, should be replaced. A fairly detailed analysis is necessary in this case.
- **Regulatory/compliance projects** may be required by a governmental agency or insurance company and often involve safety-related or environmental concerns. These projects typically generate little to no revenue.

Business growth investments include expansion projects and other projects that increase the size and scope of a company.

- Companies take on **expansion projects** to grow the business. These projects involve a complex decision-making process because they require an explicit forecast of future demand. A very detailed analysis is required.
- Other projects, such as new investments outside a company's existing lines of business, also entail a complex decision-making process that requires a detailed analysis due to the large amount of uncertainty involved.

LOS 31.b: Describe the capital allocation process and basic principles of capital allocation.

The **capital allocation process** is identifying and evaluating capital projects, that is, projects where the cash flows to the firm will be received over a period longer than a year. Any corporate decisions with an impact on future earnings can be examined using this framework. Decisions about whether to buy a new machine, expand business into another geographic area, move the corporate headquarters to Cleveland, or replace a delivery truck, to name a few, can be examined using a capital allocation analysis.

For a number of good reasons, capital allocation may be the most important responsibility that a financial manager has. First, because a capital allocation decision often involves the purchase of costly long-term assets with lives of many years, the decisions made may determine the future success of the firm. Second, the principles underlying the capital allocation process also apply to other corporate decisions, such as working capital management and making strategic mergers and acquisitions. Finally, making good capital allocation decisions is consistent with management's primary goal of maximizing shareholder value.

The capital allocation process has four administrative steps:

Step 1: Idea generation. The most important step in the capital allocation process is generating good project ideas. Ideas can come from a number of sources, including senior management, functional divisions, employees, or sources outside the company.

Step 2: Analyzing project proposals. Because the decision to accept or reject a capital project is based on the project's expected future cash flows, a cash flow forecast must be made for each project to determine its expected profitability.

Step 3: Create the firm-wide capital budget. Firms must prioritize profitable projects according to the timing of the project's cash flows, available company resources, and the company's overall strategic plan. Many projects that are attractive individually may not make sense strategically.

Step 4: Monitoring decisions and conducting a post-audit. It is important to follow up on all capital allocation decisions. An analyst should compare the actual results to the projected results, and project managers should explain why projections did or did not match actual performance. Because the capital allocation process is only as good as the estimates of the inputs into the model used to forecast cash flows, a post-audit should be used to identify systematic errors in the forecasting process and improve company operations.

Principles of Capital Allocation

The capital allocation process involves the following key assumptions:

- *Decisions are based on cash flows, not accounting income.* The relevant cash flows to consider as part of the capital allocation process are **incremental cash flows**, the changes in cash flows that will occur if the project is undertaken.

- *Cash flows are based on opportunity costs.* **Opportunity costs** are cash flows that a firm will lose by undertaking the project under analysis. These are cash flows generated by an asset the firm already owns that would be forgone if the project under consideration is undertaken. Opportunity costs should be included in project costs. For example, when building a plant, even if the firm already owns the land, the cost of the land should be charged to the project because it could be sold if not used.
- *The timing of cash flows is important.* Capital allocation decisions account for the time value of money, which means that cash flows received earlier are worth more than cash flows to be received later.
- *Cash flows are analyzed on an after-tax basis.* The impact of taxes must be considered when analyzing all capital allocation projects. Firm value is based on cash flows they get to keep, not those they send to the government.
- *Financing costs are reflected in the project's required rate of return.* Do not consider financing costs specific to the project when estimating incremental cash flows. The discount rate used in the capital allocation analysis takes account of the firm's cost of capital. Only projects that are expected to return more than the cost of the capital needed to fund them will increase the value of the firm.

Sunk costs are costs that cannot be avoided, even if the project is not undertaken. Because these costs are not affected by the accept/reject decision, they should not be included in the analysis. An example of a sunk cost is a consulting fee paid to a marketing research firm to estimate demand for a new product before making a decision on the project.

Externalities are the effects the acceptance of a project may have on other firm cash flows. The primary one is a negative externality called **cannibalization**, which occurs when a new project takes sales from an existing product. When considering externalities, the full implication of the new project (loss in sales of existing products) should be taken into account. An example of cannibalization is when a soft drink company introduces a diet version of an existing beverage. An analyst should subtract the lost sales of the existing beverage from the expected sales of the new diet version when estimating incremental project cash flows. A positive externality exists when doing the project would have a positive effect on sales of a firm's other product lines.

A project has a **conventional cash flow pattern** if the sign on the cash flows changes only once, with one or more cash outflows followed by one or more cash inflows. An **unconventional cash flow pattern** has more than one sign change. For example, a project might have an initial investment outflow, a series of cash inflows, and a cash outflow for asset retirement costs at the end of the project's life.

Independent projects are projects that can be evaluated solely on their own profitability. For example, if projects A and B are independent, and both projects are profitable, then the firm could accept both projects. Multiple projects are **mutually exclusive** if only one of them can be accepted so that profitability must be evaluated among the projects. If Projects A and B are mutually exclusive, either Project A or Project B can be accepted, but not both. Making a capital allocation decision to select one of two different stamping machines, each with different costs and outputs, is an example of ranking two mutually exclusive projects.

Some projects must be undertaken in a certain order, or sequence, so that investing in a project today creates the opportunity to invest in other projects in the future. For example, if a project

undertaken today is profitable, that may create the opportunity to invest in a second project a year from now. However, if the project undertaken today turns out to be unprofitable, the firm will not invest in the second project. The opportunity to undertake the second project, depending on the outcome of the first, is referred to as a *real option*.

If a firm has unlimited access to investment capital, the firm can undertake all projects with expected returns that exceed the cost of capital. Many firms have constraints on the amount of capital they can raise and must use *capital rationing*. If a firm's profitable project opportunities exceed the amount of funds available, the firm must ration, or prioritize, its capital expenditures with the goal of achieving the maximum increase in value for shareholders, given its available capital.



MODULE QUIZ 31.1

1. Which of the following is *most likely* a going concern project?
 - A. Opening a retail outlet in a new region.
 - B. Acquiring and merging with a supplier to secure a source for a key component.
 - C. Purchasing a new model of a factory machine that will decrease unit production costs.
2. In the capital allocation process, a post-audit is used to:
 - A. improve cash flow forecasts and stimulate management to improve operations and bring results into line with forecasts.
 - B. improve cash flow forecasts and eliminate potentially profitable but risky projects.
 - C. stimulate management to improve operations, bring results into line with forecasts, and eliminate potentially profitable but risky projects.
3. Which of the following statements concerning the principles underlying the capital allocation process is *most accurate*?
 - A. Cash flows should be based on opportunity costs.
 - B. Financing costs should be reflected in a project's incremental cash flows.
 - C. The net income for a project is essential for making a correct capital allocation decision.
4. A manufacturer of clothes washing machines decides to add matching clothes dryers to its product line. In this case, it is *most likely* important in the project analysis to consider:
 - A. cannibalization.
 - B. positive externalities.
 - C. sunk costs.

MODULE 31.2: NET PRESENT VALUE AND INTERNAL RATE OF RETURN



Video covering
this content is
available online.

LOS 31.c: Demonstrate the use of net present value (NPV) and internal rate of return (IRR) in allocating capital and describe the advantages and disadvantages of each method.

Net Present Value (NPV)

Net present value (NPV) is the sum of the present values of all the expected incremental cash flows if a project is undertaken. The discount rate used is the firm's cost of capital, adjusted for the risk level of the project. For a normal project, with an initial cash outflow followed by a

series of expected after-tax cash inflows, the NPV is the present value of the expected inflows minus the initial cost of the project.

$$NPV = CF_0 + \frac{CF_1}{(1+k)^1} + \frac{CF_2}{(1+k)^2} + \dots + \frac{CF_n}{(1+k)^n} = \sum_{t=0}^n \frac{CF_t}{(1+k)^t}$$

where:

CF_0 = initial investment outlay (a negative cash flow)

CF_t = after-tax cash flow at time t (can be positive or negative)

k = required rate of return for project

A positive NPV project is expected to increase shareholder wealth, a negative NPV project is expected to decrease shareholder wealth, and a zero NPV project has no expected effect on shareholder wealth.

For *independent* projects, the *NPV decision rule* is simply to accept any project with a positive NPV and to reject any project with a negative NPV.

EXAMPLE: Net present value

Using the project cash flows presented in Table 1, compute the NPV of the project and determine whether it should be accepted or rejected. Assume that the cost of capital is 9%.

Table 1: Expected Net After-Tax Cash Flows

Year	Cash Flow
0	-\$100
1	25
2	50
3	75

Answer:

$$NPV = -100 + \frac{25}{1.09} + \frac{50}{(1.09)^2} + \frac{75}{(1.09)^3} = 22.93$$

The project has a positive NPV, so it should be accepted.

You may calculate NPV directly by using the cash flow (CF) keys on your calculator. The process is illustrated in Table 2.

Table 2: Calculating NPV With the TI Business Analyst II Plus

Keystrokes	Explanation	Display
[CF] [2nd] [CLR WORK]	Clear memory registers	CF0 = 0.0000
100 [+/-] [ENTER]	Initial cash outlay	CF0 = -100.0000
[↓] 25 [ENTER]	Period 1 cash flow	C01 = 25.0000
[↓]	Frequency of cash flow 1	F01 = 1.0000
[↓] 50 [ENTER]	Period 2 cash flow	C02 = 50.0000
[↓]	Frequency of cash flow 2	F02 = 1.0000
[↓] 75 [ENTER]	Period 3 cash flow	C03 = 75.0000
[↓]	Frequency of cash flow 3	F03 = 1.0000
[NPV] 9 [ENTER]	9% discount rate	I = 9.0000
[↓] [CPT]	Calculate NPV	NPV = 22.9335

Internal Rate of Return (IRR)

For a normal project, the **internal rate of return (IRR)** is the discount rate that makes the present value of the expected incremental after-tax cash inflows just equal to the initial cost of the project. More generally, the IRR is the discount rate that makes the present value of a project's estimated cash inflows equal to the present value of the project's estimated cash outflows. That is, IRR is the discount rate that makes the following relationship hold:

$$PV(\text{inflows}) = PV(\text{outflows})$$

The IRR is also the discount rate for which the NPV of a project is equal to zero:

$$NPV = 0 = CF_0 + \frac{CF_1}{(1 + IRR)^1} + \frac{CF_2}{(1 + IRR)^2} + \dots + \frac{CF_n}{(1 + IRR)^n} = \sum_{t=0}^n \frac{CF_t}{(1 + IRR)^t}$$

To calculate the IRR, you may use the trial-and-error method. That is, just keep guessing IRRs until you get the right one, or you may use a financial calculator.

IRR decision rule: First, determine the required rate of return for a given project. This is usually the firm's cost of capital. Note that the required rate of return may be higher or lower than the firm's cost of capital to adjust for differences between the project's risk and the average risk of all of the firm's projects (which is reflected in the firm's current cost of capital).

If $IRR >$ the required rate of return, accept the project.

If $IRR <$ the required rate of return, reject the project.

For this reason, the minimum IRR, above which a project will be accepted, is often referred to as the **hurdle rate**. Projects with IRRs above this rate will be accepted, while those with IRRs below this rate will not be accepted.

EXAMPLE: Internal rate of return

Continuing with the cash flows presented in Table 1 for the previous example, compute the IRR of the project and determine whether it should be accepted or rejected. Assume that the required rate of return is 9%.

Answer:

$$0 = -100 + \frac{25}{(1 + IRR)} + \frac{50}{(1 + IRR)^2} + \frac{75}{(1 + IRR)^3}$$

The cash flows should be entered as in Table 2 (if you haven't changed or cleared them, they are still there from the previous calculation of NPV).

With the TI calculator, the IRR can be calculated with:

[IRR] [CPT] to get 19.4377%.

The project should be accepted because its IRR is greater than the 9% required rate of return.

The Relative Advantages and Disadvantages of the NPV and IRR Methods

A **key advantage of NPV** is that it is a direct measure of the expected increase in the value of the firm. NPV is theoretically the best method. Its main weakness is that it does not include any consideration of the size of the project. For example, an NPV of \$100 is great for a project costing \$100 but not so great for a project costing \$1 million.

A **key advantage of IRR** is that it measures profitability as a percentage, showing the return on each dollar invested. The IRR provides information on the margin of safety that the NPV does not. From the IRR, we can tell how much below the IRR (estimated return) the actual project return could fall, in percentage terms, before the project becomes uneconomic (has a negative NPV).

The *disadvantages* of the IRR method are (1) the possibility of producing rankings of mutually exclusive projects different from those from NPV analysis and (2) the possibility that a project has multiple IRRs or no IRR.



MODULE QUIZ 31.2

1. A company is considering the purchase of a copier that costs \$5,000. Assume a required rate of return of 10% and the following cash flow schedule:

- Year 1: \$3,000.
- Year 2: \$2,000.
- Year 3: \$2,000.

The project's NPV is *closest* to:

- A. -\$309.
- B. +\$883.
- C. +\$1,523.

2. A company is considering moving its manufacturing facilities to either Texas or South Carolina to decrease taxes and labor costs. After estimating all the relevant incremental after-tax cash flows of each move, an analyst estimates the IRR of a move to Texas to be 13% and the IRR of a move to South Carolina to be 15%. If the appropriate discount rate to evaluate the moves is 14%, the analyst:

- A. can conclude that the move to South Carolina should be undertaken.
- B. cannot conclude that the move to South Carolina should be undertaken because the two moves are mutually exclusive.
- C. may find that the move to Texas is preferable when projects are ranked by their NPVs.



MODULE 31.3: CAPITAL ALLOCATION PITFALLS AND REAL OPTIONS

Video covering this content is available online.

LOS 31.d: Describe common capital allocation pitfalls.

Common mistakes managers make when evaluating capital projects include the following:

- *Failing to incorporate economic responses into the analysis.* For example, if a profitable project is in an industry with low barriers to entry, competitors will likely undertake similar projects, reducing future profitability.
- *Misusing standardized templates.* Since managers may evaluate hundreds of projects in a given year, they often create templates to streamline the analysis process. However, the template may not be an exact match for the project, resulting in estimation errors.
- *Pet projects of senior management.* Projects that have the personal backing of influential members of senior management may contain overly optimistic projections that make the project appear more profitable than it really is. In addition, the project may not be subjected to the same level of analysis as other projects.
- *Basing investment decisions on EPS or ROE.* Managers whose incentive compensation is tied to increasing EPS or ROE may avoid positive long-term NPV investments that are expected to reduce EPS or ROE in the short run.
- *Using the IRR criterion for project decisions.* When comparing two mutually exclusive projects, one project may have a higher IRR, but a lower NPV. The NPV criterion is theoretically sound, accurately reflecting the goal of maximizing shareholder wealth, and should be used to choose between two projects that are both acceptable.
- *Poor cash flow estimation.* For a complex project, it is easy to double count or fail to include certain cash flows in the analysis. For example, the effects of inflation must be properly accounted for.
- *Misestimating overhead costs.* The cost of a project should include only the incremental overhead costs related to management time and information technology support. These costs are often difficult to quantify, and over- or underestimation can lead to incorrect investment decisions.
- *Using the incorrect discount rate.* The required rate of return on the project should reflect the project's risk. Simply using the company's WACC as a discount rate without adjusting it for the risk of the project may lead to significant errors when estimating the NPV of a project.
- *Politics involved with spending the entire capital budget.* Many managers try to spend their entire capital budget each year and ask for an increase for the following year. In a company with a culture of maximizing shareholder value, managers will return excess funds whenever there is a lack of positive NPV projects and make a case for expanding the budget when there are multiple positive NPV opportunities.
- *Failure to generate alternative investment ideas.* Generating investment ideas is a crucial step in the capital allocation process. However, once a manager comes up with a "good" idea, they may go with it rather than searching for an idea that is "better."

- *Improper handling of sunk and opportunity costs.* Managers should not consider sunk costs in the evaluation of a project because they are not incremental cash flows (they are incurred whether the project is undertaken or not). Managers should always consider opportunity costs because they are incremental. However, in practice, many managers do this incorrectly.
-

LOS 31.e: Describe expected relations among a company's investments, company value, and share price.

One way to approach the question of whether a company is creating value for its shareholders is to compare the return on the company's investment in assets to its cost of capital. A company's **return on invested capital (ROIC)**, or simply **return on capital**, is defined as its **net operating profit after tax (NOPAT)**, or simply after-tax net profit, over a period, divided by the average book value of its total capital over the period.

$$\text{return on invested capital} = \frac{\text{after-tax net profit}}{\text{average book value of total capital}}$$

or:

$$\text{return on invested capital} = \frac{\text{net operating profit after tax}}{\text{average book value of total capital}}$$

Because we want to measure the return to all sources of capital (both debt and equity), after-tax net profit is net income plus after-tax interest expense. For net *operating* profit after tax, we would subtract after-tax nonoperating income as well. The denominator is the sum of the average book values of debt, common stock, and preferred stock.

Because return on invested capital is a measure of the after-tax return on the amounts invested in the company over time by both equity investors and debtholders, we can compare it to the company's *weighted average cost of capital* (WACC), a weighted average of the required after-tax rates of return on the company's various sources of capital. If a firm's ROIC is greater than its WACC, then the company's management is increasing the value of the firm (and shareholders' wealth). Of course, the opposite is true as well, an ROIC less than a firm's WACC indicates that value is being reduced.



PROFESSOR'S NOTE

We will discuss WACC in detail in our reading on Cost of Capital—Foundational Topics.

Because the NPV method is a direct measure of the expected change in firm value from undertaking a capital project, the NPV is related to share value. In theory, a positive NPV project should cause a proportionate increase in a company's stock price.

In reality, the impact of a project on the company's stock price is more complicated. A company's stock price is a function of the present value of its expected future earnings stream. As a result, changes in the stock price will result more from changes in *expectations* about the profitability of a firm's future investments.

LOS 31.f: Describe types of real options relevant to capital investment.

Real options are future actions that a firm can take, given that they invest in a project today. Real options are similar to financial options (put and call options) in that they give the option holder the right, but not the obligation, to take a future action. The value of real options should be included in the calculation of project's NPV. Options never have negative values because if, in the future, the specified action will have a negative value, the option holder will not take the action.

Types of real options include the following:

- **Timing options** allow a company to delay making an investment because they expect to have better information in the future.
- **Abandonment options** are similar to put options (the option to sell an asset at a given price in the future). They allow management to abandon a project if the present value of the incremental cash flows from exiting a project exceeds the present value of the incremental cash flows from continuing the project.
- **Expansion options** are similar to call options (the option to buy an asset at a given price in the future). Expansion options allow a company to make additional investments in future projects if the company decides they will create value.
- **Flexibility options** give managers choices regarding the operational aspects of a project. The two main forms are price-setting and production flexibility options.
 - *Price-setting* options allow the company to change the price of a product. For example, the company may raise prices if demand for a product is high, in order to benefit from that demand without increasing production.
 - *Production-flexibility* options may include paying workers overtime, using different materials as inputs, or producing a different variety of product.
- **Fundamental options** are projects that are options themselves because the payoffs depend on the price of an underlying asset. For example, the payoff for a copper mine is dependent on the market price for copper. If copper prices are low, it may not make sense to open a copper mine, but if copper prices are high, opening the copper mine could be very profitable. The operator has the option to close the mine when prices are low and open it when prices are high.



MODULE QUIZ 31.3

1. An analyst is estimating the NPV of a project to introduce a new spicier version of its well-known barbecue sauce into its product line. A cost that should *most likely* be excluded from his analysis is:
 - A. \$200,000 to develop a recipe for the new sauce.
 - B. a \$150,000 decrease in sales of its current sauce as some current customers switch to the spicier sauce.
 - C. \$100,000 for a marketing survey that was conducted to determine demand for a spicier sauce.
2. Fullen Machinery is investing \$400 million in new industrial equipment. The present value of the future after-tax cash flows resulting from the equipment is \$700 million. Fullen currently has 200 million shares of common stock outstanding, with a current market price of \$36 per share. Assuming that this project is new information and is independent of other expectations about the company, what is the theoretical effect of the new equipment on Fullen's stock price? The stock price will:

- A. decrease to \$33.50.
 - B. increase to \$37.50.
 - C. increase to \$39.50.
3. Albert Duffy, a project manager at Crane Plastics, is considering taking on a new capital project. When presenting the project, Duffy shows members of Crane's executive management team that because the company has the ability to have employees work overtime, the project makes sense. The project Duffy is taking on would be *best* described as having:
- A. a fundamental option.
 - B. an expansion option.
 - C. a flexibility option.
4. A company is considering building a distribution center that will allow it to expand sales into a new region comprising three provinces. John Parker, a firm analyst, has argued that the current analysis fails to incorporate the amount they could get from selling the distribution center at the end of year 2, rather than operating it to the end of the project's assumed economic life. Parker is suggesting that:
- A. the assumed investment horizon is too long.
 - B. the analysis should include the value of a put option.
 - C. the analysis should include the value of a call option.

KEY CONCEPTS

LOS 31.a

Business maintenance investments include going concern projects to maintain a business or to reduce costs, and required regulatory/compliance projects. Business growth investments include expansion projects and other projects that increase the size and scope of a company.

LOS 31.b

Capital allocation is the process of evaluating capital projects, projects with cash flows over a period longer than one year.

Steps of the capital allocation process are: (1) Generate investment ideas; (2) Analyze project ideas; (3) Create a firm-wide capital budget; and (4) Monitor decisions and conduct a post-audit.

Capital allocation decisions should be based on incremental after-tax cash flows, the expected differences in after-tax cash flows if a project is undertaken.

Acceptable independent projects can all be undertaken, while a firm must choose between or among mutually exclusive projects.

Project sequencing concerns the opportunities for future capital projects that may be created by undertaking a current project.

If a firm cannot undertake all profitable projects because of limited ability to raise capital, the firm should choose that group of fundable positive NPV projects with the highest total NPV.

LOS 31.c

NPV is the sum of the present values of a project's expected cash flows and represents the change in firm value from undertaking a project. Positive NPV projects should be undertaken, but negative NPV projects should not because they are expected to decrease the value of the firm.

An IRR is the discount rate at which the present values of a project's expected cash inflows and cash outflows are equal (i.e., the discount rate for which the NPV of a project is zero). A project for which the IRR is greater (less) than the appropriate discount rate for the project will have an NPV that is positive (negative) and should be accepted (not accepted).

LOS 31.d

Common mistakes in the capital allocation process include the following:

- Failing to incorporate economic responses into the analysis
- Misusing standardized project evaluation templates
- Having overly optimistic assumptions for pet projects of senior management
- Basing long-term investment decisions on short-term EPS or ROE considerations
- Using the IRR criterion for choosing between mutually exclusive projects
- Poor cash flow estimation
- Misestimating overhead costs
- Using a discount rate that does not accurately reflect the project's risk
- Politics involved with spending the entire capital budget
- Failure to generate alternative investment ideas
- Improper handling of sunk and opportunity costs

LOS 31.e

Return on invested capital can be compared to a company's weighted average cost of capital to indicate whether the company has increased or decreased firm value over time.

NPV is a measure of the expected change in company value from undertaking a project. A firm's stock price may be affected to the extent that engaging in a project with a positive NPV was previously unanticipated by investors (not already reflected in the stock price).

LOS 31.f

Real options allow managers to make future decisions that change the value of capital allocation decisions made today.

- Timing options allow a company to delay making an investment.
- Abandonment options allow management to abandon a project if the present value of the incremental cash flows from exiting a project exceeds the present value of the incremental cash flows from continuing a project.
- Expansion options allow a company to make additional investments in a project if doing so creates value.
- Flexibility options give managers choices regarding the operational aspects of a project. The two main forms are price-setting and production flexibility options.
- Fundamental options are projects that are options themselves because the payoffs depend on the price of an underlying asset.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 31.1

1. **C** Going concern projects are those to maintain the business or to increase the efficiency of existing operations. The other two projects are business growth investments that increase the size of the company. (LOS 31.a)
2. **A** A post-audit identifies what went right and what went wrong. It is used to improve forecasting and operations. (LOS 31.b)
3. **A** Cash flows are based on opportunity costs. Financing costs are recognized in the project's required rate of return. Accounting net income, which includes non-cash expenses, is irrelevant; incremental cash flows are essential for making correct capital allocation decisions. (LOS 31.b)
4. **B** It is quite possible that offering a matching dryer will increase sales of their washers, because some consumers will prefer a matching set. The increased sales of their washers is a positive externality, and those incremental sales should be considered in the analysis. Cannibalization would be a consideration if the introduction of dryers was expected to decrease washer sales. Sunk costs should not be considered in project analysis. (LOS 31.b)

Module Quiz 31.2

1. **B** $CF_0 = -5,000; CF_1 = 3,000; CF_2 = 2,000; CF_3 = 2,000; I/Y = 10; NPV = \883 . (LOS 31.c)
2. **A** Based on the IRRs, the move to South Carolina will have a positive NPV (the IRR is greater than the discount rate) and the move to Texas will have a negative NPV. In this case, we can rank the two projects based on their IRRs. If the appropriate discount rate was less than both IRRs, for example, 10%, the IRR rankings could not reliably be used to choose between the two proposed moves. (LOS 31.c)

Module Quiz 31.3

1. **C** The cost of the marketing survey should not be included because it is a sunk cost; it will be incurred whether they decide to do the project or not. The decrease in sales of their current sauce if the spicier version is introduced (cannibalization) should be considered in the analysis. The cost of recipe development should be included because it will only be incurred if they decide to go ahead with the introduction of the new spicier sauce. (LOS 31.d)
2. **B** The NPV of the new equipment is $\$700 \text{ million} - \$400 \text{ million} = \$300 \text{ million}$. The NPV of this project is added to Fullen's current market value. On a per share basis, the addition is worth $\$300 \text{ million} / 200 \text{ million shares} = \1.50 . $\$36.00 + \$1.50 = \$37.50$. (LOS 31.e)
3. **C** The project described has a production-flexibility regarding the level of production. Other flexibility options might be to produce a different product or to use different inputs at some future date. Including the value of real options can improve the NPV estimates for individual projects. (LOS 31.f)
4. **B** The option to abandon the project and receive the market value of the facility if actual cash flows are less than expected over the first two years can be viewed as a valuable put option

that should be included in the calculation of the project's NPV. (LOS 31.f)

READING 32

WORKING CAPITAL & LIQUIDITY

EXAM FOCUS

This reading broadly covers a firm's operating cash flows, working capital cash flows, and various sources of short-term and long-term financing. The reading is relatively short but has a lot of useful (and testable) information. The requirements for candidates are clear in the LOS for this one, so focus on those as you study the material.

MODULE 32.1: SOURCES OF CAPITAL



LOS 32.a: Compare methods to finance working capital.

Video covering
this content is
available online.

Internal Sources of Funds

A company's **operating cash flow** can be estimated as net income plus depreciation less dividend payments. It is a measure of how much cash the firm is generating from its business operations that is available to finance and invest in its business.

Accounts receivable represent the amounts owed to a company from customer sales it has made on credit. Collection of accounts receivable is typically an important source of cash over the short term. Collecting accounts receivable sooner increases a company's available short-term funds.

Accounts payable represent the amounts a company owes to its suppliers. Delaying payment of accounts payable increases the cash a company has available, but doing so may have explicit costs. Often, credit terms offer a discount for earlier payment. Terms of "2/15 net 45," for example, mean that if an invoice is paid within 15 days, the customer receives a 2% discount. The full amount of the invoice must be paid within 45 days if the customer does not take advantage of the discount for early payment. The cost of delaying payment from day 15 to day 45 (for 30 days) is 2%, or more than 24% annualized, which may be a fairly expensive source of short-term funding.

Selling **inventory** generates funds, and there is a cost to having funds invested in inventory. A company may attempt to shorten production times and the time between getting goods ready for sale and the actual sales. On the other hand, carrying too little inventory may result in lost sales when orders cannot be filled immediately.

Liquid debt or equity securities that a company intends to sell within a year are classified as **marketable securities**. These can represent a quick source of cash for a company.

Financial Intermediaries

Bank **lines of credit** are used primarily by large, financially sound companies. Lines of credit may be uncommitted, committed, or revolving.

- *Uncommitted line of credit.* A bank extends an offer of credit for a certain amount but may refuse to lend if circumstances change, making this a less reliable source of funds.
- *Committed (regular) line of credit.* A bank extends an offer of credit that it “commits to” for some period of time. The fact that the bank has committed to extend credit in amounts up to the credit line makes this a more reliable source of short-term funding than an uncommitted line of credit. Banks charge a fee for making such a commitment. Loans under the agreement are typically for periods of less than a year, and interest charges are stated in terms of a short-term reference rate, plus a margin to compensate for the credit risk of the loan. Outside the United States, similar arrangements are referred to as *overdraft lines of credit*.
- *Revolving line of credit.* An even more reliable source of short-term financing than a committed line of credit, a revolving line of credit is typically for a longer term, sometimes years. With a revolving line of credit, companies may borrow and repay funds as their needs change over time. Along with committed lines of credit, revolving credit lines can be verified and can be listed in the footnotes of a firm’s financial statements as a source of liquidity.

Companies with weaker credit may have to pledge assets as collateral for bank borrowings. Fixed assets, inventory, and accounts receivable may all serve as collateral for loans. Short-term financing is typically collateralized by receivables or inventory and longer-term loans are secured with a claim to fixed (longer-term) assets. The bank may also have a **blanket lien**, which gives it a claim to all current and future firm assets as collateral in case the primary collateral is insufficient and the borrowing firm defaults. When a firm assigns its receivables to the bank making a loan, the company still services the receivables and remains responsible for any receivables that are not paid.

Secured (asset-backed) loans are backed by collateral, for example, fixed assets, receivables, or inventory. **Factoring** refers to the actual sale of receivables at a discount from their face value. The size of the discount will depend on the amount of time until the receivables are due, the creditworthiness of the firm’s customers, and the firm’s collection history on its receivables. The “factor” (the buyer of the receivables) takes on the responsibility for collecting receivables and the credit risk of the receivables portfolio.

Web-based and **non-bank lenders** typically lend to medium-to-small-size firms and typically charge fees in addition to interest charges.

In managing its short-term financing, a firm should focus on the objectives of having sufficient sources of funding for current, as well as foreseeable future, cash needs and should seek the most cost-effective rates available given its needs, assets, and creditworthiness. The firm should have the ability to prepay short-term borrowings when cash flow permits and have the flexibility to structure its short-term financing so that the debt matures without peaks and can be matched to expected cash flows. For larger borrowers, it is important that the firm has alternative sources of short-term funding and even alternative lenders for a particular type of financing. It is often worth having slightly higher overall short-term funding costs in order to have flexibility and redundant sources of financing.

Capital Markets Sources of Funds

Companies can issue **public debt** (trades on public markets) or **private debt** (provided by private entities and not actively traded). Debt payments have priority over payments to equity holders and the interest paid on debt is typically tax deductible. In addition to current market conditions, interest rates on a company's debt depend on its creditworthiness and the collateral, if any, pledged for repayment of the debt.

Large, creditworthy companies often issue short-term debt securities called **commercial paper**, which is sometimes issued by companies with lower credit ratings as well. Whether the firm sells the paper directly to investors (direct placement) or sells it through dealers (dealer-placed paper), the interest costs are typically lower than the rate they would get from a bank (in cases where that is an alternative). Maturities on commercial paper range from a few days to one year (272 days in the U.S.). The debt is unsecured and typically supplemented with a **backup line of credit** that will provide funds if markets for commercial paper are disrupted so that a company cannot replace maturing commercial paper with new paper.

Long-term debt often carries a fixed interest rate through maturity, which may be decades after issuance.

Companies sell **common equity** (common stock) both in public markets and privately. Equity owners have a lower priority to a firm's cash flows than debtholders, making a company's equity riskier than its debt.

LOS 32.b: Explain expected relations between working capital, liquidity, and short-term funding needs.

Companies must devote sufficient assets to working capital to meet their operating needs. They must keep enough inventory on hand so that customers receive products without undue delays. Companies that sell goods and services on credit will have accounts receivable outstanding. Liquidity in the form of cash and equivalents is important for managing day-to-day fluctuations in cash outflows.

A company's working capital needs depend largely on the nature of its business. For example, a personal services provider or a software developer may have little need to carry physical inventory, while an automobile dealer or a high-end jewelry retailer may need to carry significant inventory. Its need to carry inventory, its liquidity requirements (cash and equivalents), its extension of credit to customers (accounts receivable), and the degree to which it uses credit from trade vendors (accounts payable) all factor into a company's needs for working capital.

Companies have some latitude in how they manage working capital. Strategies fall on a spectrum from conservative to aggressive. In general, more conservative working capital management means current assets are higher as a percentage of sales, and a more aggressive approach means current assets are lower as a percentage of sales. The basic trade-off is between greater financial flexibility (conservative) and greater return on assets (aggressive).

Approaches to financing working capital (i.e., short-term funding) also range from conservative to aggressive. Financing current assets with equity or long-term debt is more conservative

because the costs tend to be stable and predictable, although typically higher than the costs of short-term debt. Financing with short-term debt is more aggressive in that it brings a risk that renewed financing (i.e., rolling over short-term debt) may become difficult to obtain or unavailable. One moderate strategy for short-term financing uses equity or long-term debt to finance current assets the company views as permanent, while using short-term debt to finance current assets the company views as seasonal or variable.

LOS 32.c: Describe sources of primary and secondary liquidity and factors affecting a company's liquidity position.

A company's **primary sources of liquidity** are the sources of cash it uses in its normal day-to-day operations. The company's *cash balances* result from selling goods and services, collecting receivables, and generating cash from other sources such as short-term investments. Typical sources of *short-term funding* include trade credit from vendors and lines of credit from banks. Effective *cash flow management* of a firm's collections and payments can also be a source of liquidity for a company.

Secondary sources of liquidity include liquidating short-term or long-lived assets, renegotiating debt agreements, or filing for bankruptcy and reorganizing the company. While using its primary sources of liquidity is unlikely to change the company's normal operations, resorting to secondary sources of liquidity such as these can change the company's financial structure and operations significantly and may indicate that its financial position is deteriorating.

In general, a company's liquidity position improves if it can get cash to flow in more quickly and flow out more slowly.

Drags on liquidity delay or reduce cash inflows, or increase borrowing costs. Examples include uncollected receivables and bad debts, obsolete inventory (takes longer to sell and can require significant price discounts), and limited short-term credit availability due to economic conditions.

Pulls on liquidity accelerate cash outflows. Examples include paying vendors sooner and changes in credit terms that accelerate the required payment of outstanding balances.

LOS 32.d: Compare a company's liquidity position with that of peers.

Some companies tend to have chronically weak liquidity positions, often due to specific factors that affect the company or its industry. These companies typically need to borrow against their long-lived assets to acquire working capital.

Liquidity ratios are employed by analysts to determine the firm's ability to pay its short-term liabilities.

- The *current ratio* is the best-known measure of liquidity:

$$\text{current ratio} = \frac{\text{current assets}}{\text{current liabilities}}$$

The higher the current ratio, the more likely it is that the company will be able to pay its short-term bills. A current ratio of less than one means that the company has negative working capital and is probably facing a liquidity crisis. Working capital equals current assets minus current liabilities.

- The *quick ratio* or *acid-test ratio* is a more stringent measure of liquidity because it does not include inventories and other assets that might not be very liquid:

$$\text{quick ratio} = \frac{\text{cash} + \text{short-term marketable securities} + \text{receivables}}{\text{current liabilities}}$$

The higher the quick ratio, the more likely it is that the company will be able to pay its short-term bills.

The current and quick ratios differ only in the assumed liquidity of the current assets that the analyst projects will be used to pay off current liabilities.

- A measure of accounts receivable liquidity is the *receivables turnover*:

$$\text{receivables turnover} = \frac{\text{credit sales}}{\text{average receivables}}$$

It is considered desirable to have a receivables turnover figure close to the industry norm.



PROFESSOR'S NOTE

This formula for the receivables turnover ratio uses credit sales in the numerator, rather than total sales, as shown in the Financial Statement Analysis reading on ratio analysis. While an analyst within a company will know what proportion of sales are credit or cash sales, an external analyst will likely not have this information but may be able to estimate it based on standard industry practice.

In most cases when a ratio compares a balance sheet account (such as receivables) with an income or cash flow item (such as sales), the balance sheet item will be the average of the account instead of simply the end-of-year balance. Averages are calculated by adding the beginning-of-year account value and the end-of-year account value, then dividing the sum by two.

- The inverse of the receivables turnover multiplied by 365 is the *number of days of receivables* (also called *average days' sales outstanding*), which is the average number of days it takes for the company's customers to pay their bills:

$$\begin{aligned}\text{number of days of receivables} &= \frac{365}{\text{receivables turnover}} \\ &= \frac{\text{average receivables}}{\text{average day's credit sales}}\end{aligned}$$

It is considered desirable to have a collection period (and receivables turnover) close to the industry norm. The firm's credit terms are another important benchmark used to interpret this ratio. A collection period that is too high might mean that customers are too slow in paying their bills, which means too much capital is tied up in assets. A collection period that is too low might indicate that the firm's credit policy is too rigorous, which might be hampering sales.

- A measure of a firm's efficiency with respect to its processing and inventory management is the *inventory turnover*:

$$\text{inventory turnover} = \frac{\text{cost of goods sold}}{\text{average inventory}}$$



PROFESSOR'S NOTE

Pay careful attention to the numerator in the turnover ratios. For inventory turnover, be sure to use cost of goods sold, not sales.

- The inverse of the inventory turnover multiplied by 365 is the *average inventory processing period* or *number of days of inventory*:

$$\begin{aligned}\text{number of days of inventory} &= \frac{365}{\text{inventory turnover}} \\ &= \frac{\text{average inventory}}{\text{average day's COGS}}\end{aligned}$$

As is the case with accounts receivable, it is considered desirable to have an inventory processing period (and inventory turnover) close to the industry norm. A processing period that is too high might mean that too much capital is tied up in inventory and could mean that the inventory is obsolete. A processing period that is too low might indicate that the firm has inadequate stock on hand, which could hurt sales.

- A measure of the use of trade credit by the firm is the *payables turnover ratio*:

$$\text{payables turnover ratio} = \frac{\text{purchases}}{\text{average trade payables}}$$

The inverse of the payables turnover ratio multiplied by 365 is the *payables payment period* or *number of days of payables*, which is the average amount of time it takes the company to pay its bills:

$$\begin{aligned}\text{number of days of payables} &= \frac{365}{\text{payables turnover ratio}} \\ &= \frac{\text{average payables}}{\text{average day's purchases}}\end{aligned}$$

- The **operating cycle**, the average number of days that it takes to turn raw materials into cash proceeds from sales, is:

$$\text{operating cycle} = \text{days of inventory} + \text{days of receivables}$$

- The *cash conversion cycle* or *net operating cycle* is the length of time it takes to turn the firm's cash investment in inventory (including raw materials for a manufacturer) back into cash, in the form of collections from the inventory sales. The cash conversion cycle is computed from the average receivables collection period, average inventory processing period, and the payables payment period:

$$\text{cash conversion cycle} = \left(\frac{\text{average days}}{\text{of receivables}} \right) + \left(\frac{\text{average days}}{\text{of inventory}} \right) - \left(\frac{\text{average days}}{\text{of payables}} \right)$$

High cash conversion cycles are considered undesirable. A conversion cycle that is too high implies that the company has an excessive amount of investment in working capital.

LOS 32.e: Evaluate short-term funding choices available to a company.

A company's short-term funding plan should ensure that it maintains sufficient borrowing capacity to meet its ongoing needs, including seasonal or other times of peak requirements.

The primary consideration when choosing a strategy for short-term funding is cost. Companies should stay up-to-date on the costs of the various funding sources available to them.

Even companies that rely primarily or exclusively on one particular type of funding may find it advantageous to work with more than one lender.

Companies that rely on significant short-term financing should use more than one type of financing and multiple lenders for a given type of financing, because markets and lender circumstances can change over time, sometimes suddenly.

Maintaining excess funding for unforeseen events or to take advantage of business opportunities is also an important consideration. Ensuring access to funds may increase borrowing costs.



MODULE QUIZ 32.1

1. A company has an agreement with its lender to borrow funds as they need up to a specified maximum amount and to repay its borrowings as they have funds available. This arrangement is *most appropriately* called:
 - A. a committed line of credit.
 - B. a revolving line of credit.
 - C. a capped line of credit.
2. A company receives an invoice of \$150,000 for machine tools with terms of “1.5/15 net 40.” The cost to the company of delaying payment of this receivable is *most appropriately* described as \$2,250 for the use of:
 - A. \$150,000 for 40 days.
 - B. \$150,000 for 25 days.
 - C. \$147,750 for 25 days.
3. Which of the following *most likely* represents conservative working capital management?
 - A. Decreasing inventory on hand to reduce insurance costs.
 - B. Financing an increase in receivables by increasing long-term borrowing.
 - C. Selling marketable securities and using the proceeds to acquire real estate.
4. An example of a primary source of liquidity is:
 - A. liquidating assets.
 - B. negotiating debt contracts.
 - C. short-term investment portfolios.
5. Firm P and Firm Q have the same current assets and current liabilities, but Firm P has a lower quick ratio than Firm Q. Compared with Firm Q, it is *most likely* that Firm P has:
 - A. greater inventory.
 - B. greater payables.
 - C. a higher receivables turnover ratio.
6. A company would *best* improve its cash conversion cycle by decreasing its:
 - A. receivables turnover.
 - B. payables turnover.
 - C. inventory turnover.

KEY CONCEPTS

LOS 32.a

Companies choose among their financing alternatives based on cost, risk, ease of access, flexibility, and current availability. Types of company financing are as follows:

- Internal sources: operating cash flows, accounts payable, accounts receivable, inventory, and marketable securities.
- Financial intermediaries: bank borrowing, asset-backed loans, loans from non-bank finance companies (more expensive), asset leasing.
- Capital markets: long-term debt (highest priority), common equity (has a residual claim to firm cash flows and assets).

LOS 32.b

A conservative approach to working capital management involves high levels of current assets financed with long-term debt and equity. Compared to a more aggressive approach, a conservative approach provides more liquidity and involves less financial risk, but has higher financing costs and will reduce returns.

An aggressive approach to working capital management involves lower levels of current assets and financing working capital needs with short-term debt. Compared to a more conservative approach, an aggressive approach typically has lower financing costs and results in higher returns, but decreases liquidity and increases financial risk.

LOS 32.c

Primary sources of liquidity are the sources of cash a company uses in its normal operations. If its primary sources are inadequate, a company can use secondary sources of liquidity, such as asset sales, debt renegotiation, and bankruptcy reorganization.

A company's liquidity position depends on the effectiveness of its cash flow management and is influenced by drags on its cash inflows (e.g., uncollected receivables, obsolete inventory) and pulls on its cash outflows (e.g., early payments to vendors, reductions in credit limits).

LOS 32.d

Measures of a company's short-term liquidity include:

- Current ratio = current assets / current liabilities.
- Quick ratio = (cash + marketable securities + receivables) / current liabilities.

Measures of how well a company is managing its working capital include:

- Receivables turnover = credit sales / average receivables.
- Number of days of receivables = 365 / receivables turnover.
- Inventory turnover = cost of goods sold / average inventory.
- Number of days of inventory = 365 / inventory turnover.
- Payables turnover = purchases / average trade payables.
- Number of days of payables = 365 / payables turnover.

The operating cycle and the cash conversion cycle (net operating cycle) are summary measures of the effectiveness of a company's working capital management. Cycles that are high relative to a company's peers suggest the company has too much cash tied up in working capital.

- Operating cycle = days of inventory + days of receivables.
- Cash conversion cycle = days of inventory + days of receivables - days of payables.

LOS 32.e

Companies must maintain sufficient short-term borrowing capacity to meet anticipated peak cash needs. They should also maintain some additional borrowing capacity to meet unanticipated needs and provide the flexibility to take advantage of investment opportunities that may arise.

The primary consideration when choosing short-term funding sources is cost. However, companies that rely on significant short-term borrowing should, at a minimum, use more than one type of debt and maintain a relationship with more than one lender, as circumstances and market conditions can change. It is important for company managers to keep up-to-date on the available types of short-term financing and their costs.

ANSWERS TO MODULE QUIZ

Module Quiz 32.1

1. **B** A line of credit where the borrower can draw funds as they need them and repay them when they have the funds available to do so is called a revolving line of credit. (LOS 32.a)
2. **C** The terms indicate that the company can pay $\$150,000(1 - 0.015) = \$147,750$ on day 15 (after the invoice date) or pay \$150,000 on day 40—effectively gaining the use of \$147,750 for 25 days at a cost of \$2,250. (LOS 32.a)
3. **B** Financing an increase in a current asset with long-term borrowing is an example of conservative working capital management. The other choices describe decreases in current assets and therefore more likely represent aggressive working capital management. (LOS 32.b)
4. **C** Primary sources of liquidity include ready cash balances, short-term funds (e.g., short-term investment portfolios), and effective cash flow management. Secondary sources of liquidity include renegotiating debt contracts, liquidating assets, and filing for bankruptcy protection and reorganization. (LOS 32.c)
5. **A** Firms P and Q will have the same current ratios, CA/CL. The quick ratio numerator is CA – inventory, so for firm P to have a smaller quick ratio than firm Q, it must have greater inventory. (LOS 32.d)
6. **B** A decrease in the payables turnover would increase days payables, which would decrease (improve) the firm's cash conversion cycle. A decrease in a company's receivables turnover would increase days receivables, and a decrease in a company's inventory turnover would increase its days inventory on hand. Both would increase the cash conversion cycle. (LOS 32.d)

READING 33

COST OF CAPITAL—FOUNDA TIONAL TOPICS

EXAM FOCUS

The firm must decide how to raise the capital to fund its business or finance its growth, dividing it among common equity, debt, and preferred stock. The mix that produces the minimum overall cost of capital will maximize the value of the firm (share price). From this reading, you must understand weighted average cost of capital and its calculation and be ready to calculate the costs of common equity, preferred stock, and the after-tax cost of debt.

MODULE 33.1: WEIGHTED AVERAGE COST OF CAPITAL



Video covering this content is available online.

LOS 33.a: Calculate and interpret the weighted average cost of capital (WACC) of a company.

LOS 33.b: Describe how taxes affect the cost of capital from different capital sources.

The capital allocation process involves discounted cash flow analysis. To conduct such analysis, we must know the appropriate discount rate to use. This reading discusses how, as an analyst, you can determine the proper rate at which to discount the cash flows associated with a project. This discount rate is the firm's **weighted average cost of capital (WACC)** and is also referred to as the **marginal cost of capital (MCC)**.

Basic definitions. On the right (liabilities and equity) side of a firm's balance sheet, we have debt, preferred equity, and common equity. These are normally referred to as the *capital components* of the firm. Any increase in a firm's total assets will have to be financed through an increase in at least one of these capital accounts. The cost of each of these components is called the *component cost* of capital.

Throughout this review, we focus on the following capital components and their component costs:

k_d The rate at which the firm can issue new debt. This is the yield to maturity on existing debt. This is also called the before-tax component cost of debt.

$k_d(1 - t)$ The after-tax cost of debt. Here, t is the firm's marginal tax rate. The after-tax component cost of debt, $k_d(1 - t)$, is used to calculate the WACC.

- k_{ps} The cost of preferred stock.
- k_{ce} The cost of common equity. It is the required rate of return on common stock and is generally difficult to estimate.

In many countries, the interest paid on corporate debt is tax deductible. Because we are interested in the after-tax cost of capital, we adjust the cost of debt, k_d , for the firm's marginal tax rate, t . Because there is typically no tax deduction allowed for payments to common or preferred stockholders, there is no equivalent deduction to k_{ps} or k_{ce} .

How a company raises capital and how it budgets or invests it are considered independently. Most companies have separate departments for the two tasks. The financing department is responsible for keeping costs low and using a balance of funding sources: common equity, preferred stock, and debt. Generally, it is necessary to raise each type of capital in large sums. The large sums may temporarily overweight the most recently issued capital, but in the long run, the firm will adhere to target weights. Because of these and other financing considerations, each investment decision must be made assuming a WACC, which includes each of the different sources of capital and is based on the long-run target weights. A company creates value by producing a return on assets that is higher than the required rate of return on the capital needed to fund those assets.

The WACC, as we have described it, is the cost of financing firm assets. We can view this cost as an opportunity cost. Consider how a company could reduce its costs if it found a way to produce its output using fewer assets, like less working capital. If we need less working capital, we can use the funds freed up to buy back our debt and equity securities in a mix that just matches our target capital structure. Our after-tax savings would be the WACC based on our target capital structure multiplied by the total value of the securities that are no longer outstanding.

For these reasons, any time we are considering a project that requires expenditures, comparing the return on those expenditures to the WACC is the appropriate way to determine whether undertaking that project will increase the value of the firm. This is the essence of the capital allocation decision. Because a firm's WACC reflects the average risk of the projects that make up the firm, it is not appropriate for evaluating all new projects. It should be adjusted upward for projects with greater-than-average risk and downward for projects with less-than-average risk.

The weights in the calculation of a firm's WACC are the proportions of each source of capital in a firm's capital structure.

Calculating a Company's Weighted Average Cost of Capital

The WACC is given by:

$$\text{WACC} = (w_d)[k_d(1 - t)] + (w_{ps})(k_{ps}) + (w_{ce})(k_{ce})$$

where:

w_d = percentage of debt in the capital structure

w_{ps} = percentage of preferred stock in the capital structure

w_{ce} = percentage of common stock in the capital structure

EXAMPLE: Computing WACC

Suppose Dexter, Inc.'s target capital structure is as follows:

$$w_d = 0.45, w_{ps} = 0.10, \text{ and } w_{ce} = 0.45$$

Its before-tax cost of debt is 7.5%, its cost of equity is 11.5%, its cost of preferred stock is 9.0%, and its marginal tax rate is 40%. Calculate Dexter's WACC.

Answer:

Dexter's WACC will be:

$$\text{WACC} = (w_d)(k_d)(1 - t) + (w_{ps})(k_{ps}) + (w_{ce})(k_{ce})$$

$$\text{WACC} = (0.45)(0.075)(0.6) + (0.10)(0.09) + (0.45)(0.115) = 0.081 = 8.1\%$$

LOS 33.c: Calculate and interpret the cost of debt capital using the yield-to-maturity approach and the debt-rating approach.

The **after-tax cost of debt**, $k_d(1 - t)$, is used in computing the WACC. It is the interest rate at which firms can issue new debt (k_d) net of the tax savings from the tax deductibility of interest, $k_d(t)$:

$$\text{after-tax cost of debt} = \text{interest rate} - \text{tax savings} = k_d - k_d(t) = k_d(1 - t)$$

$$\text{after-tax cost of debt} = k_d(1 - t)$$

EXAMPLE: Cost of debt

Dexter, Inc., is planning to issue new debt at an interest rate of 7.5%. Dexter has a 40% marginal federal-plus-state tax rate. What is Dexter's cost of debt capital?

Answer:

$$k_d(1 - t) = 7.5\%(1 - 0.4) = 4.5\%$$



PROFESSOR'S NOTE

It is important that you realize that the cost of debt is the market interest rate (YTM) on new (marginal) debt, not the coupon rate on the firm's existing debt. CFA Institute may provide you with both rates, and you need to select the current market rate.

In tax jurisdictions where interest payments are not deductible for tax purposes, t is zero, and the pretax and after-tax cost of debt are the same.

It may be that interest is tax deductible but the amount of the deduction is capped at some maximum. In this case, if the firm's interest costs will be less than the maximum tax-deductible amount if a project is undertaken, the appropriate cost of debt is the after-tax cost based on the firm's tax rate. For a project that would require debt interest payments in excess of the

maximum tax-deductible amount, the appropriate cost of debt is the pretax cost because no additional tax deductions for interest on debt are permitted.

If a market YTM is not available because the firm's debt is not publicly traded, the analyst may use the rating and maturity of the firm's existing debt to estimate the before-tax cost of debt. If, for example, the firm's debt carries a single-A rating and has an average maturity of 15 years, the analyst can use the yield curve for single-A rated debt to determine the current market rate for debt with a 15-year maturity. This approach is an example of **matrix pricing** or valuing a bond based on the yields of comparable bonds.

If any characteristics of the firm's anticipated debt would affect the yield (e.g., covenants or seniority), the analyst should make the appropriate adjustment to his estimated before-tax cost of debt. For firms that primarily employ floating-rate debt, the analyst should estimate the longer-term cost of the firm's debt using the current yield curve (term structure) for debt of the appropriate rating category.

LOS 33.d: Calculate and interpret the cost of noncallable, nonconvertible preferred stock.

The **cost of preferred stock** (k_{ps}) is:

$$k_{ps} = D_{ps} / P$$

where:

D_{ps} = preferred dividends

P = market price of preferred

EXAMPLE: Cost of preferred stock

Suppose Dexter, Inc., has preferred stock that pays an \$8.50 dividend per share and sells for \$100 per share. What is Dexter's cost of preferred stock?

Answer:

$$k_{ps} = D_{ps} / P$$

$$k_{ps} = \$8.50 / \$100 = 0.085 = 8.5\%$$

Note that the equation $k_{ps} = D_{ps} / P$ is just a rearrangement of the preferred stock valuation model $P = D_{ps} / k_{ps}$, where P is the market price.

LOS 33.e: Calculate and interpret the cost of equity capital using the capital asset pricing model approach and the bond yield plus risk premium approach.

The opportunity **cost of equity capital** (k_{ce}) is the required rate of return on the firm's common stock. The rationale here is that the firm could avoid part of the cost of common stock outstanding by using retained earnings to buy back shares of its own stock. The cost of (i.e., the required return on) common equity can be estimated using one of the following approaches:

1. The capital asset pricing model approach.

Estimate the risk-free rate, R_f . Yields on default risk-free debt such as U.S.

Step 1: Treasury notes are usually used. The most appropriate maturity to choose is one that is close to the useful life of the project.

Step 2: Estimate the stock's beta, β . This is the stock's risk measure.

Step 3: Estimate the expected rate of return on the market, $E(R_{mkt})$.

Use the capital asset pricing model (CAPM) equation to estimate the required rate of return

$$k_{ce} = R_f + \beta[E(R_{mkt}) - R_f]$$

EXAMPLE: Using CAPM to estimate k_{ce}

Suppose $R_f = 5\%$, $R_{mkt} = 11\%$, and Dexter has a beta of 1.1. Estimate Dexter's cost of equity.

Answer:

The required rate of return for Dexter's stock is:

$$k_{ce} = 5\% + 1.1(11\% - 5\%) = 11.6\%$$



PROFESSOR'S NOTE

If you are unfamiliar with the capital asset pricing model, you can find more detail and the basic elements of its derivation in the Portfolio Management topic area.

2. Bond yield plus risk premium approach.

Analysts often use an ad hoc approach to estimate the required rate of return. They add a risk premium (three to five percentage points) to the market yield on the firm's long-term debt.

$$k_{ce} = \text{bond yield} + \text{risk premium}$$

EXAMPLE: Estimating k_{ce} with bond yields plus a risk premium

Dexter's interest rate on long-term debt is 7.5%. Suppose the risk premium is estimated to be 4%. Estimate Dexter's cost of equity.

Answer:

Dexter's estimated cost of equity is:

$$k_{ce} = 7.5\% + 4.0\% = 11.5\%$$

Note that the two models gave us different estimates of k_{ce} . The CAPM estimate was 11.6% and the bond yield plus risk premium estimate was 11.5%. Analysts must use their judgment to decide which is most appropriate.

-  1. An analyst gathered the following data about a company:

Capital Structure	Required Rate of Return
30% debt	10% for debt
20% preferred stock	11% for preferred stock
50% common stock	18% for common stock

Assuming a 40% tax rate, what after-tax rate of return must the company earn on its investments?

- A. 13.0%.
 - B. 14.2%.
 - C. 18.0%.
2. A company is planning a \$50 million expansion. The expansion is to be financed by selling \$20 million in new debt and \$30 million in new common stock. The before-tax required return on debt is 9% and 14% for equity. If the company is in the 40% tax bracket, the company's marginal cost of capital is *closest* to:
- A. 7.2%.
 - B. 10.6%.
 - C. 12.0%.
3. A company has \$5 million in debt outstanding with a coupon rate of 12%. Currently, the yield to maturity (YTM) on these bonds is 14%. If the firm's tax rate is 40%, what is the company's after-tax cost of debt?
- A. 5.6%.
 - B. 8.4%.
 - C. 14.0%.
4. A company's \$100, 8% preferred is currently selling for \$85. What is the company's cost of preferred equity?
- A. 8.0%.
 - B. 9.4%.
 - C. 10.8%.
5. If the risk-free rate is 2% and the market rate of return is 5%, the cost of equity for a company with a beta of 0.8 is closest to:
- A. 2.4%.
 - B. 4.0%.
 - C. 4.4%.

MODULE 33.2: BETA ESTIMATION AND FLOTATION COSTS



Video covering this content is available online.

LOS 33.f: Explain and demonstrate beta estimation for public companies, thinly traded public companies, and nonpublic companies.

For a publicly traded company with an active market, the beta of a stock is estimated from the linear relationship between the returns on the stock (dependent variable) and the returns on a proxy for the market portfolio (independent variable). In the U.S., the proxy for the market portfolio is typically the S&P 500 Index, but other broad-based indexes can be used. The slope of an ordinary least squares regression line of a stock's returns on the returns of the market is used as an estimate of beta, a stock's systematic risk.

The estimate of beta is affected by the choice of the proxy for the overall market portfolio, by the length of the total time period selected, and by the time period for the individual returns. In practice, 60 observations of monthly returns are most often used.

A stock's beta estimated in this manner is referred to as a "raw" or "unadjusted" beta. Studies have shown that stock betas move toward one over time. To adjust for this reversion toward one, the following adjustment is used:

$$\text{Adjusted beta} = \frac{2}{3} \times \text{unadjusted beta} + \frac{1}{3}$$



PROFESSOR'S NOTE

The theory of the capital asset pricing model (CAPM) and using beta as an estimate of an asset's systematic risk is more fully developed in the topic area of Portfolio Management.

For a (target) stock that is not publicly traded or that is simply very thinly (infrequently) traded, we can estimate its beta from the returns on a comparable stock or group of comparable stocks that are actively traded. We can select a stock or group of stocks from the same industrial classification as the target company, or simply identify a stock or stocks that have similar business risks to our target company (a peer company or peer group of companies).

For companies that are comparable in their business risks, their individual betas will vary due to differences in their use of financial leverage and the tax rates they are subject to.

To account for these differences, we begin with the beta of a company or group of companies that are *purely* engaged in a business similar to that of the target company. Thus, using the beta of a conglomerate that is engaged in the same business as the target company would be inappropriate because its beta depends on its many different lines of business.

The beta of a firm is a function not only of the business risks of its lines of business but also of its financial structure. For a given set of projects, the greater a firm's reliance on debt financing, the greater its equity beta. For this reason, we must adjust the beta from a comparable company (or group of companies) for its leverage and tax rate (unlever it) to get the comparable company's asset beta (unlevered beta). We can then adjust the comparable company's asset beta (re-lever it) based on the financial leverage and the marginal tax rate of the target company. We can then use this equity beta in the CAPM equation to estimate the cost of equity of the target company.

To get the *unlevered beta* or *asset beta* of a comparable company, we use the following formula:

$$\beta_{\text{ASSET}} = \beta_{\text{EQUITY}} \left\{ \frac{1}{1 + [(1 - t) \frac{D}{E}]} \right\}$$

where:

D/E = comparable company's debt-to-equity ratio and t is its marginal tax rate

To get the equity beta for the target company, we use the target company's tax rate and debt-to-equity ratio to re-lever the comparable company's asset beta:

$$\beta_{\text{TARGET}} = \beta_{\text{ASSET}} \left\{ 1 + \left[(1 - t) \frac{D}{E} \right] \right\}$$

The following example illustrates this technique.

EXAMPLE: Estimating the cost of capital for a thinly traded stock

Acme, Inc., is in the food distribution business and its stock is thinly traded. It has a D/E ratio of 2 and a marginal tax rate of 40%. Balfor, a publicly traded firm that operates only in the food distribution business, has a D/E ratio of 1.5, a marginal tax rate of 30%, and an equity beta of 0.9. The risk-free rate is 5%, and the expected return on the market portfolio is 12%. Calculate Balfor's asset beta and Acme's adjusted equity beta, and estimate the cost of equity capital for Acme.

Answer:

Balfor's asset beta:

$$\beta_{\text{ASSET}} = 0.9 \left\{ \frac{1}{1 + [(1 - 0.3)1.5]} \right\} = 0.439$$

Unadjusted equity beta for Acme:

$$\text{Acme Beta}_{\text{unadjusted}} = 0.439[1 + (1 - 0.4)(2)] = 0.966$$

$$\text{Acme Beta}_{\text{adjusted}} = (2/3) 0.966 + 1/3 = 0.977$$

$$\text{Acme's cost of equity} = 5\% + 0.977(12\% - 5\%) = 11.839\%$$

If we use a peer group of companies, the method is the same, but for our estimate of the asset (unlevered) beta of the peer group stocks, we simply take the average of the asset betas calculated for each of our peer group firms.

LOS 33.g: Explain and demonstrate the correct treatment of flotation costs.

Flotation costs are the fees charged by investment bankers when a company raises external equity capital (issues new shares). Flotation costs can be substantial and often amount to between 2% and 7% of the total amount of equity capital raised, depending on the type of offering.

Many financial textbooks incorporate flotation costs directly into the cost of capital by increasing the cost of external equity; funds from the issuance of new shares. For example, if a company has a dividend of \$1.50 per share, a current price of \$30 per share, and an expected growth rate of 6%, the cost of internal equity (without flotation costs) would be:

$$r_e = \left(\frac{\$1.50(1 + 0.06)}{\$30} \right) + 0.06 = 0.1130, \text{ or } 11.30\%$$



PROFESSOR'S NOTE

This form of the cost of equity presents it as the expected dividend yield on shares, plus the expected growth rate of the firm's dividend over the foreseeable future.

If we incorporate flotation costs of 4.5% directly into the cost of equity computation, the cost of equity increases:

$$r_e = \left[\frac{\$1.50(1 + 0.06)}{\$30(1 - 0.045)} \right] + 0.06 = 0.1155, \text{ or } 11.55\%$$

This method of adjustment increases the WACC used to evaluate potential projects, as we have replaced the current share price with the net proceeds of a newly issued shares. This form helps us understand the impact of raising equity funds in excess of those available internally (retained earnings).

An alternative method of adjusting for flotation costs (the “correct” treatment according to the CFA curriculum) is to treat them as a cash outflow at the inception of a project. If we can identify the equity flotation costs specific to a project, we can simply add them to the initial project cost when calculating its NPV. It is argued that this is the more appropriate method because the flotation costs are not ongoing. They are incurred at the time new shares are issued and, therefore, should be allocated to the project under consideration.

EXAMPLE: Accounting for flotation costs

Omni Corporation is considering a project that requires a \$400,000 cash outlay and is expected to produce cash flows of \$150,000 per year for the next four years. Omni’s tax rate is 35%, and the before-tax cost of debt is 6.5%. The current share price for Omni’s stock is \$36 per share, and the expected dividend next year is \$2 per share. Omni’s expected growth rate is 5%. Assume that Omni finances the project with 50% debt and 50% equity capital and that flotation costs for equity are 4.5%. The appropriate discount rate for the project is the WACC.

Calculate the NPV of the project using the two alternative treatments of flotation costs and comment on the results.

Answer:

$$\text{after-tax cost of debt} = 6.5\%(1 - 0.35) = 4.23\%$$

$$\text{cost of equity} = \left(\frac{\$2}{\$36} \right) + 0.05 = 0.1055, \text{ or } 10.55\%$$

$$\text{WACC} = 0.50(0.0423) + 0.50(0.1055) = 7.39\%$$

Because the project is financed with 50% equity, the amount of equity capital raised is $0.50 \times \$400,000 = \$200,000$.

Flotation costs are 4.5%, which equates to a dollar cost of $\$200,000 \times 0.045 = \$9,000$.

$$\begin{aligned} \text{NPV} &= -\$400,000 - \$9,000 + \frac{\$150,000}{1.0739} + \frac{\$150,000}{(1.0739)^2} \\ &\quad + \frac{\$150,000}{(1.0739)^3} + \frac{\$150,000}{(1.0739)^4} = \$94,640 \end{aligned}$$

If we had adjusted the cost of equity for flotation costs, the cost of equity would have increased to $10.82\% \left(= \frac{\$2.00}{\$36(1 - 0.045)} + 0.05 \right)$, increasing the WACC to 7.53%. Using this method, the NPV of the project would have been:

$$\begin{aligned} \text{NPV} &= -\$400,000 + \frac{\$150,000}{1.0753} + \frac{150,000}{(1.0753)^2} + \frac{\$150,000}{(1.0753)^3} + \frac{150,000}{(1.0753)^4} \\ &= \$102,061 \end{aligned}$$

The two methods result in significantly different estimates for the project NPV. With the preferred method, the flotation costs are recovered over the four-year life of the project and the NPV is lower. With the second method the flotation costs are essentially spread over an infinite horizon, so the NPV for the four-year project is significantly higher than when flotation costs are treated as a cash expense at $t = 0$.

Note that flotation costs may be tax-deductible for some firms. In that case, the initial cash flow of the project should be adjusted by the after-tax flotation cost. In this example, Omni would have an after-tax flotation cost of $\$9,000(1 - 0.35) = \$5,850$, and the project NPV would be \$97,790.

MODULE QUIZ 33.2



1. An analyst is estimating an equity beta for Jay Company, which has a thinly traded stock. Jay is in the same line of business as Cass Company, which trades more frequently and has a beta of 1.2. Jay's debt-to-equity ratio is 2.0 and Cass's debt-to-equity ratio is 1.6. Both companies have a tax rate of 30%. For Jay, the analyst should use an adjusted equity beta *closest to*:
 - A. 0.57.
 - B. 1.24.
 - C. 1.36.
2. When the equity portion of the financing for new capital projects will be raised by issuing new equity shares, equity flotation costs should be:
 - A. used to adjust the estimated cost of equity capital.
 - B. treated as part of each project's initial cash outflow.
 - C. disregarded because they are unlikely to be material.

KEY CONCEPTS

LOS 33.a

$$\text{WACC} = (w_d)(k_d)(1 - t) + (w_{ps})(k_{ps}) + (w_{ce})(k_{ce})$$

The weighted average cost of capital, or WACC, is calculated using weights based on the market values of each component of a firm's capital structure and is the correct discount rate to use to discount the cash flows of projects with risk equal to the average risk of a firm's projects.

LOS 33.b

Interest expense on a firm's debt is tax deductible, so the pre-tax cost of debt must be reduced by the firm's marginal tax rate to get an after-tax cost of debt capital:

$$\text{after-tax cost of debt} = k_d (1 - \text{firm's marginal tax rate})$$

The pretax and after-tax capital costs are equal for both preferred stock and common equity because dividends paid by the firm are not tax deductible.

LOS 33.c

The before-tax cost of fixed-rate debt capital, k_d , is the rate at which the firm can issue new debt.

- The yield-to-maturity approach assumes the before-tax cost of debt capital is the YTM on the firm's existing publicly traded debt.
- If a market YTM is not available, the analyst can use the debt rating approach, estimating the before-tax cost of debt capital based on market yields for debt with the same rating and average maturity as the firm's existing debt.

LOS 33.d

The cost (and yield) of noncallable, nonconvertible preferred stock is simply the annual dividend divided by the market price of preferred shares.

LOS 33.e

The cost of equity capital, k_{ce} , is the required rate of return on the firm's common stock.

Approaches to estimating k_{ce} :

- CAPM approach: $k_{ce} = R_f + \beta[E(R_{mkt}) - R_f]$.
- Bond yield plus risk premium approach: add a risk premium of 3% to 5% to the market yield on the firm's long-term debt.

LOS 33.f

We can use the beta of a comparable company or group of companies to estimate the beta of a target company. This method involves the following steps:

1. Estimate the beta for a comparable company or companies.
2. Unlever the beta to get the asset beta using the marginal tax rate and debt-to-equity ratio of the comparable company:

$$\beta_{\text{ASSET}} = \beta_{\text{EQUITY}} \left\{ \frac{1}{1 + [(1 - t) \frac{D}{E}]} \right\}$$

3. Re-lever the beta using the marginal tax rate and debt-to-equity ratio for the target company:

$$\beta_{\text{TARGET}} = \beta_{\text{ASSET}} \left\{ 1 + [(1 - t) \frac{D}{E}] \right\}$$

4. Use the CAPM to estimate the target company's required return on equity.

LOS 33.g

The correct method to account for flotation costs of raising new equity capital is to increase a project's initial cash outflow by the flotation cost attributable to the project when calculating the project's NPV.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 33.1

1. **A** WACC = $(w_d)(k_d)(1 - t) + (w_{ps})(k_{ps}) + (w_{ce})(k_{ce}) = (0.3)(0.1)(1 - 0.4) + (0.2)(0.11) + (0.5)(0.18) = 13\%$ (LOS 33.a)
2. **B** $w_d = 20 / (20 + 30) = 0.4$, $w_{ce} = 30 / (20 + 30) = 0.6$ WACC = $(w_d)(k_d)(1 - t) + (w_{ce})(k_{ce}) = (0.4)(9)(1 - 0.4) + (0.6)(14) = 10.56\%$ = MCC (LOS 33.a, 33.b)
3. **B** $k_d(1 - t) = (0.14)(1 - 0.4) = 8.4\%$ (LOS 33.c)
4. **B** $k_{ps} = D_{ps} / P_{ps}$, $D_{ps} = \$100 \times 8\% = \8 , $k_{ps} = 8 / 85 = 9.4\%$ (LOS 33.d)
5. **C** $k_e = 2\% + 0.8(5\% - 2\%) = 4.4\%$. (LOS 33.e)

Module Quiz 33.2

1. **B** Unlevered beta for Cass = $1.2 \left\{ \frac{1}{1 + [(1 - 0.3)1.6]} \right\} = 0.566$ Unadjusted equity beta for Jay
= $0.566 \{1 + [(1 - 0.3)2.0]\} = 1.358$ Adjusted equity beta for Jay = $(2/3 \times 1.358) + 1/3 = 1.239$ (LOS 33.f)
2. **B** The correct treatment of flotation costs is to treat them as part of each project's initial cash outflow. (LOS 33.g)

READING 34

CAPITAL STRUCTURE

EXAM FOCUS

This reading explores various theories related to how a firm chooses its proportions of debt and equity financing. Candidates should know the Miller and Modigliani propositions and understand how taxes drive these results. Understand well the concept of an optimal capital structure based on the trade-off between the tax savings and additional risk of increased financial leverage. How firms make financing decisions in practice by developing a target capital structure is important for analysts. Finally, understand how capital structure decisions affect various firm stakeholders and the nature of conflicts of interest among them with regard to debt financing.

MODULE 34.1: CAPITAL STRUCTURE THEORIES



LOS 34.a: Explain factors affecting capital structure.

Video covering
this content is
available online.

A company's **capital structure** refers to how it finances its assets and operations. In a simple case, a company's capital structure might be described as 60% equity (stock) and 40% debt (bonds). In general, a company will choose a capital structure that minimizes its weighted average cost of capital, that is, its cost of capital given its proportions of debt and equity financing. Note that when we are discussing a company's debt-to-equity ratios in the context of capital structure, we use market values for both debt and equity, not accounting values from the balance sheet.

Capital structures vary considerably among companies, but tend to be similar within industries. We may view factors that affect capital structures as those that determine a company's capacity to service debt. These factors may be internal or external to a firm. Internal factors include the characteristics of the business or industry, the company's existing debt level, the corporate tax rate, the company's stage in its life cycle, and management's preferences. External factors include market and business cycle conditions, regulation, and industry norms.

Industry characteristics. In general, the more stable, predictable, and recurring are a company's revenues and cash flows, the higher proportion of debt it can have in its capital structure. With this in mind we can say, other things equal:

- Companies in non-cyclical industries are better able to support high proportions of debt than companies in cyclical industries.
- Companies with low fixed operating costs as a proportion of total costs (i.e., low operating leverage) are better able to support high proportions of debt than companies with high fixed

costs.

- Companies with subscription-based revenue models are better able to support high proportions of debt than companies with pay-per-use revenue models.

The types of assets companies use to generate revenues also affect their ability to issue and service debt. For example, creditors tend to view tangible assets as better collateral than intangible assets, especially when those assets can be sold for cash if necessary without losing significant value (i.e., are more liquid) or are more readily substituted for similar assets (i.e., are more fungible). A company that owns its productive assets outright, as opposed to using assets owned by others (such as a franchise model), has more collateral. This improves its access to debt financing and reduces its borrowing costs.

Some companies and industries have characteristics that may conflict in terms of supporting debt in the capital structure. For example, a manufacturer and seller of construction equipment operates in a cyclical industry and may have high fixed costs, but is also likely to own significant tangible assets that it can pledge as loan collateral.

Existing debt. If the factors we have listed so far contribute to a company's **debt capacity** (the extent to which it can finance itself with debt without risking insolvency), its existing financial leverage shows how much of its debt capacity it is already using. Analysts use financial ratios to estimate a company's debt capacity and usage. A higher current ratio (current assets/current liabilities) indicates a greater ability to repay short-term debt. High profitability ratios (operating income, EBIT, or EBITDA as a proportion of revenues) are also more favorable for a company's ability to service its debt.

Leverage and coverage ratios are key for analyzing debt capacity. For example, many analysts consider it a warning sign if a company's debt-to-EBITDA ratio is greater than three or if its interest coverage ratio is less than two.

Corporate tax rate. In countries where interest on debt is tax-deductible, a company's tax rate is a factor in determining its optimal capital structure. Higher tax rates increase the value to a company of the tax shield from interest expense, and therefore encourage higher proportions of debt.

Management preferences. Companies may have their own policies about how much debt financing to use, based on the covenants on their existing debt or on debt rating agency thresholds, and often include a safety margin in determining their preferred amount of debt.

Debt ratings from agencies such as Standard & Poor's and Moody's reflect the probability of default on a company's debt. The agencies perform an extensive analysis of a company's ability to make interest and principal payments through maturity and assign ratings based on their expectations of a bond's default risk. Lower debt ratings (higher risk of default) indicate higher risk for owners of both equity and debt securities, who, in turn, require higher expected returns.

Because the cost of capital is affected by debt ratings, many managers attempt to maintain a minimum debt rating, which places a limit on their borrowing. If a bond rating drops from investment grade (BBB- or above) to speculative grade (BB+ or below), the cost of debt is likely to increase significantly. Another motive for maintaining a minimum rating is that it gives the company the flexibility to issue additional debt when it identifies attractive investment opportunities.

Companies may wish to have their debt or equity securities included in published indexes. Index providers typically have minimum requirements for inclusion, such as market capitalization for equity or issue size for debt. Meeting these requirements may influence a company's capital structure decisions.

Market and business cycle conditions. Market conditions can influence a firm's financing decisions. When market interest rates or a company's credit spread above a benchmark rate decrease, or are thought to be temporarily low, firms will likely show a preference for issuing debt. Even when rates and spreads are not temporarily low, but have just fallen to historically low levels that are expected to persist, firms often increase their target proportion of debt to reflect its lower cost. Because credit spreads tend to narrow during business cycle expansions and widen during contractions, companies may issue more debt during expansions.

Regulation. Some companies and industries are subject to regulation with regard to their capital structure. In most countries, regulators require financial institutions to meet specific standards for capital adequacy (i.e., a minimum proportion of equity). Many public utilities are subject to similar regulations.

Industry norms. Because many of the factors that affect capital structure tend to be similar for companies in the same industry, within many industries the capital structures of most companies fall within a range that analysts view as "normal."

LOS 34.b: Describe how a company's capital structure may change over its life cycle.

Company characteristics that influence the proportion of debt in a company's capital structure include the following:

- *Growth and stability of revenue.* High growth of revenue or stability of significant revenue suggest continuing ability to service debt.
- *Growth and predictability of cash flow.* Growing cash flow increases the ability to service debt. Significant and stable cash flows indicate continuing ability to service debt.
- *Amount of business risk.* More business risk (operational risk and demand risk) means greater variability of earnings and cash flows, which decreases the ability to service debt.
- *Amount and liquidity of company assets.* Assets can be pledged as collateral to make a company's debt more attractive. When assets are more liquid (easier to turn into cash, values more stable), they can be pledged more readily.
- *Cost and availability of debt financing.* Companies find debt relatively more attractive when the cost of debt is lower and investors are more willing to lend to the company. Both of these are greater when the above characteristics support the issuance of debt.

We can see the effects of these various factors on the debt-to-equity ratios of companies at different stages of their company's life cycle. Consider the following three stages:

Start-up stage: Sales are just beginning and operating earnings and cash flows tend to be low or negative. Business risk is relatively high. Company debt is quite risky and, if issued, would require high interest rates. Assets, both accounts receivable and fixed assets, typically are low

and therefore not available as collateral for debt. For these reasons, start-up companies are financed almost exclusively with equity.

Growth stage: Revenue and cash flow are rising and business risk is somewhat reduced. Debt financing cost is somewhat reduced and investors may be willing to lend to the company, often with the loans secured by fixed assets or accounts receivable. Depending on the company, debt issuance may be as much as 20% of the firm's capital structure.

Mature stage: In this stage, revenue growth is slowing and business risk is much lower. Cash flow is significant and relatively stable. Debt financing is widely available at relatively low cost. Firms issue debt, both secured and unsecured, in amounts in excess of 20% of a firm's capital structure and sometimes significantly more than that. Over time, as the equity value of a mature company grows, the debt proportion will fall. Some companies may repurchase their debt, reducing its proportion in the capital structure (deleveraging) in addition to paying significant cash dividends.



PROFESSOR'S NOTE

We will address company and industry life cycles in the Equity Investments topic area.

LOS 34.c: Explain the Modigliani–Miller propositions regarding capital structure.

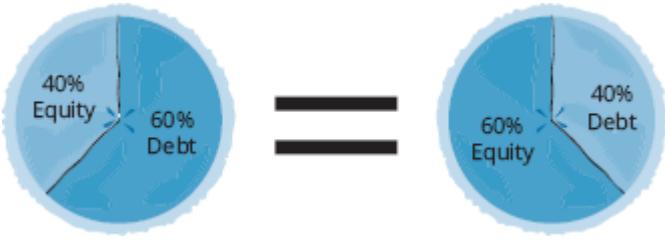
MM Proposition I (No Taxes): Capital Structure Irrelevance

In 1958, Nobel laureates Franco Modigliani and Merton Miller (we will refer to them as MM) published their seminal work on capital structure theory. In it, MM demonstrate that under certain assumptions, the value of a firm is unaffected by its capital structure. This result is referred to as MM I. The assumptions that lead to MM I are:

- *Capital markets are perfectly competitive:* There are no transactions costs, taxes, or bankruptcy costs.
- *Investors have homogeneous expectations:* They have the same expectations with respect to cash flows generated by the firm.
- *Riskless borrowing and lending:* Investors can borrow and lend at the risk-free rate.
- *No agency costs:* There are no conflicts of interest between managers and shareholders.
- *Investment decisions are unaffected by financing decisions:* Operating income is independent of how the firm is financed.

Intuitively, we can explain MM I in terms of a pie. MM I essentially concludes that the amount of pie available (value of the firm) does not depend on how it is sliced (the capital structure). The value of a firm does not change depending on how the claims to its earnings are divided. This idea is illustrated in Figure 34.1.

Figure 34.1: MM Capital Structure Irrelevance Proposition



Consider why the pie analogy holds. The operating earnings (EBIT) of a firm are available to providers of capital. In a company with no debt, all of the operating earnings are available to equity holders, and the value of the company is the discounted present value of these earnings, with a discount rate that depends on the risk of (uncertainty about) earnings. If a company is financed partly by debt and partly by equity, operating earnings are divided between debtholders and equity holders. If one entity purchased all the debt and all the equity of the firm, it would be entitled to all the operating earnings (EBIT) of the firm, so the total value of its holdings must equal the value of an all-equity firm, which also has a claim to all the operating earnings of the firm. Under the assumption that operating earnings are unaffected by financing decisions, the total value of debt and equity will be unaffected by the proportions of debt and equity in a firm's capital structure.

MM Proposition II: Cost of Equity and Leverage

MM's second proposition (MM II) is framed in terms of a firm's cost of capital, rather than firm value. Based on the same assumptions as MM I, MM II states that the cost of equity increases linearly as a company increases its proportion of debt financing. The cash flows promised to bondholders have priority, so that equity holders receive the remaining cash flows after the claims of debtholders have been met. Because the cash flows promised to debtholders are more certain (less risky) than the residual cash flows promised to equity holders, the cost of debt will be less than the cost of equity. The greater the amount of debt in a firm's capital structure, the more uncertain are the residual cash flows to equity holders. MM II tells us that as companies increase the proportion of debt financing, the risk of the cash flows to equity holders increases, which increases the cost of equity.

The conclusion of MM II is that the decrease in financing costs from using a larger proportion of (lower cost) debt is just offset by the increase in the cost of equity, resulting in no change in the firm's weighted average cost of capital (WACC).

Given MM II (that the value of the firm, and therefore its WACC, is unaffected by changes in the proportion of debt financing), we can state the relation that must hold between a company's debt-to-equity ratio and its cost of equity:

$$r_e = r_0 + \frac{D}{E}(r_0 - r_d)$$

where:

r_e = cost of equity

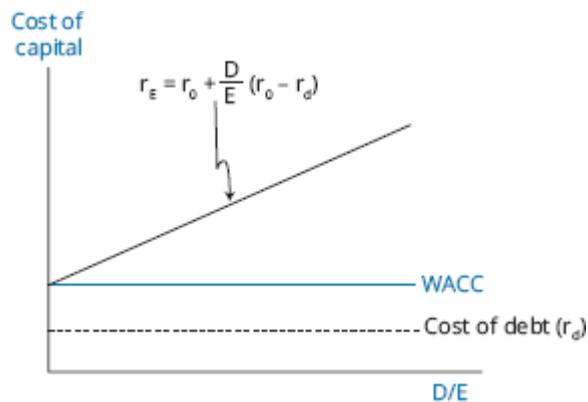
r_0 = cost of equity with no debt (all equity)

r_d = cost of debt

$\frac{D}{E}$ = debt-to-equity ratio

As leverage (the debt-to-equity ratio) increases, the cost of equity increases, but the cost of debt and WACC are unchanged. This relationship between the cost of equity financing and the debt-to-equity ratio is illustrated in Figure 34.2.

Figure 34.2: MM Proposition II (No Taxes)



MM II is consistent with MM I; if the benefits of greater use of lower-cost debt financing are just offset by the increased cost of equity, the proportions of debt versus equity in the firm's capital structure do not affect the firm's overall cost of capital or the value of the firm.

MM with Taxes: Value is Maximized at 100% Debt

As is often done in theoretical work, MM started with several simplifying assumptions. The next step is to examine the effects of relaxing some of those assumptions. Here we examine the MM propositions under the assumptions that earnings are taxed and that interest payments to debtholders are tax deductible. Under the tax code of most countries, interest payments are a pretax expense and are therefore tax deductible, while dividends paid to equity holders are not tax deductible.

This differential tax treatment encourages firms to use debt financing because debt provides a **tax shield** that adds value to the company. The tax shield is equal to the tax rate multiplied by the amount of debt in the capital structure. In other words, the value of a levered firm is equal to the value of an unlevered firm plus the value of the tax shield provided by debt financing.

To continue our analogy of a pie, with the introduction of taxes, the government gets a slice of the pie. When debt financing is used, the government's slice of the pie is smaller, so that the amount of pie available to debt and equity holders is greater. If a firm were 100% financed with debt, the taxes avoided would be at a maximum and the after-tax cash flows of the firm (size of the pie) would be maximized. The conclusion of *MM I with taxes* is that the value of the firm is maximized with 100% debt financing.

If we assume a positive tax rate, the formula to solve for return on equity can be used to illustrate *MM II with taxes*:

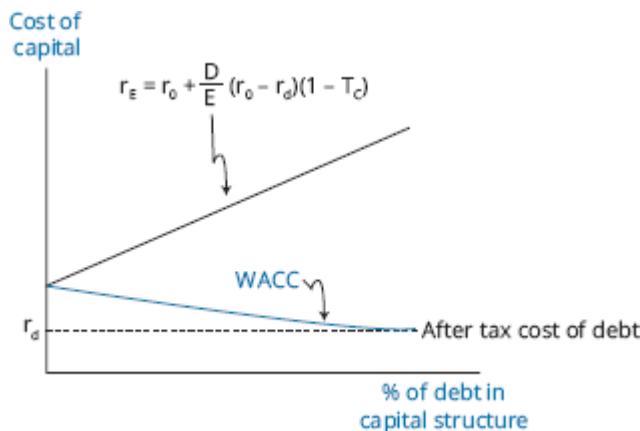
$$r_E = r_0 + \frac{D}{E}(r_0 - r_d)(1 - T_C)$$

where:

T_C = tax rate

Figure 34.3 illustrates that the tax shield provided by debt causes the WACC to decline as leverage increases. The value of the firm is maximized at the point where the WACC is minimized, which is 100% debt.

Figure 34.3: MM Proposition II (With Taxes)



In practice, we do not find that companies finance their assets with 100% debt, or close to it. MM suggested that differential investor tax rates on dividends and interest income could explain differences in capital structures.

Current theory suggests that differences in value-maximizing choices of how much financial leverage to use are the result of additional costs of using debt financing that we have not yet considered. One type of cost that can be expected to increase at higher levels of debt financing is **costs of financial distress**. Costs of financial distress are the increased costs a company faces when earnings decline to the point where the firm has trouble paying its fixed financing costs (interest on debt). The expected costs of financial distress for a firm have two components:

- *Costs of financial distress and bankruptcy* can be direct or indirect. Direct costs of financial distress include the cash expenses associated with the bankruptcy, such as legal fees and administrative fees. Indirect costs include foregone investment opportunities and the costs that result from losing the trust of customers, creditors, suppliers, and employees. Additionally, during periods of financial distress, conflicts of interest between managers (who represent equity owners) and debtholders impose additional costs referred to as the **agency costs of debt**.
- *Probability of financial distress* is related to the firm's use of operating and financial leverage. In general, higher amounts of financial leverage increase the probability of financial distress (higher probability that cash flows will fall to an amount that is insufficient to make their promised debt payments). Other factors to consider include the quality of a firm's management and the company's corporate governance structure. Lower-quality management and poor corporate governance lead to a higher probability of financial distress.

Higher expected costs of financial distress tend to discourage companies from using large proportions of debt in their capital structures, all else equal. The expected costs of financial distress shrink our pie as the proportion of debt financing is increased.

Static Trade-Off Theory

The **static trade-off theory** seeks to balance the costs of financial distress with the tax shield benefits from using debt. There is an amount of debt financing at which the increase in the value of the tax shield from additional borrowing is exceeded by the value reduction of higher expected costs of financial distress. This point represents the **optimal capital structure** for a firm, where the WACC is minimized and the value of the firm is maximized.

Accounting for the costs of financial distress, the expression for the value of a levered firm becomes:

$$V_L = V_U + (t \times \text{debt}) - PV(\text{costs of financial distress})$$

We illustrate this relationship in Figure 34.4.

Figure 34.4: Static Trade-Off Theory: Cost of Capital vs. Capital Structure

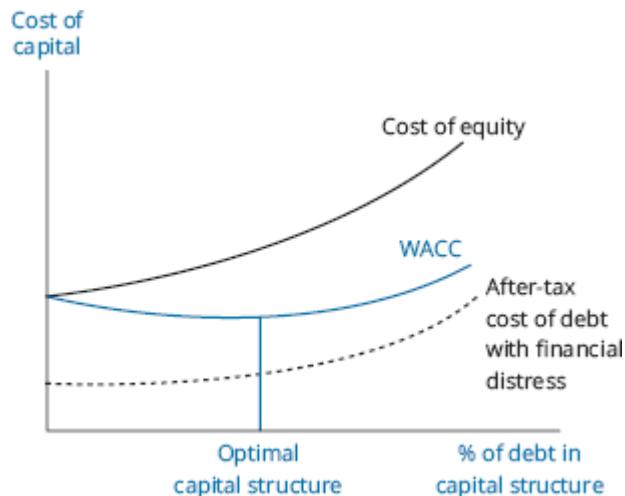
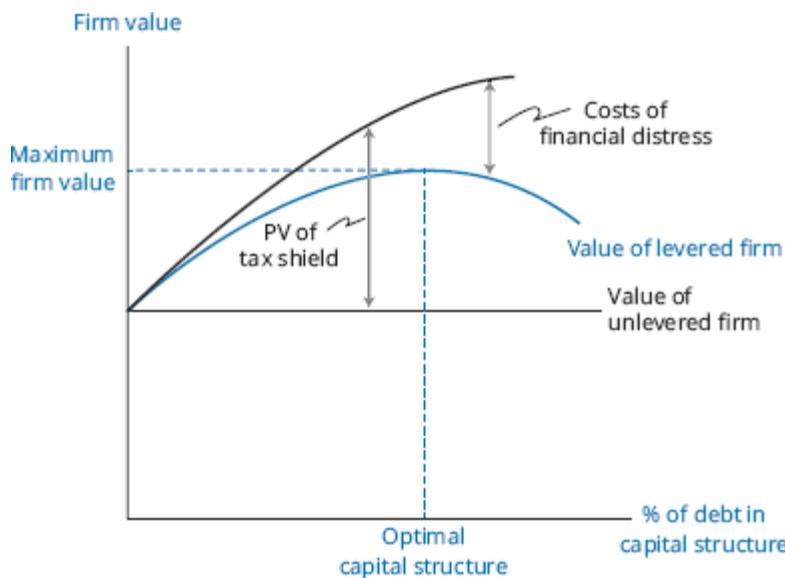


Figure 34.5 illustrates the trade-off between firm value and capital structure. Note that each firm's optimal capital structure depends on its business risk (operating risk and sales risk), tax rate, corporate governance, industry influences, and other factors.

Figure 34.5: Static Trade-Off Theory: Firm Value vs. Capital Structure



To summarize:

- *MM's propositions with no taxes or costs of financial distress* are that a company's capital structure is irrelevant because its WACC and its value (the discounted present value of its operating earnings) are unchanged by changes in capital structure.
- *MM's propositions with taxes but without costs of financial distress* are that a company's WAAC is minimized and its value is maximized with 100% debt.
- *Static trade-off theory* indicates that firm value initially increases (and WACC decreases) with additional debt financing, but company value decreases at some point when the increase in the expected value of financial distress outweighs the tax benefits of additional debt.



MODULE QUIZ 34.1

1. A company is *most likely* to be financed only by equity during its:
 - A. start-up stage.
 - B. growth stage.
 - C. mature stage.
2. A company's optimal capital structure:
 - A. maximizes firm value and minimizes the weighted average cost of capital.
 - B. minimizes the interest rate on debt and maximizes expected earnings per share.
 - C. maximizes expected earnings per share and maximizes the price per share of common stock.
3. Which of the following statements regarding Modigliani and Miller's Proposition I is *most accurate*?
 - A. A firm's cost of equity financing increases as the proportion equity in a firm's capital structure is increased.
 - B. A firm's cost of debt financing increases as firm's financial leverage increases.
 - C. A firm's weighted average cost of capital is not affected by its choice of capital structure.

MODULE 34.2: CAPITAL STRUCTURE DECISIONS



LOS 34.d: Describe the use of target capital structure in estimating WACC, and calculate and interpret target capital structure weights.

Video covering this content is available online.

Target capital structure is the capital structure that a firm seeks to achieve on average over time to maximize firm value. That is, it reflects management's beliefs about its optimal capital structure.

In practice, a firm's actual capital structure tends to fluctuate around the target capital structure for two reasons:

- *Management may choose to exploit opportunities in a specific financing source.* For example, a temporary rise in the firm's stock price may create a good opportunity to issue additional equity, which would result in a higher percentage of equity than the target.
- *Market values of a firm's equity and debt fluctuate.* Because capital structure weights are based on market values, market fluctuations (especially the market value of firm equity) may cause the firm's capital structure to vary from the target.

For analysis, the weights to use when estimating a firm's WACC should be based on its target capital structure, the proportions (based on market values) of debt, preferred stock, and equity that the firm expects to achieve over time. An analyst may use management's stated target proportions of equity and debt; however, most firms do not provide these.

In the absence of stated capital structure weights, an analyst must estimate a firm's target capital structure. Alternatives for estimating target capital structure include the following:

- An analyst may simply use the firm's current capital structure (based on market values) as the best indication of its target capital structure.
- If there has been a noticeable trend in the firm's capital structure, an analyst may incorporate this trend into her estimate of the firm's target.
- An analyst may use the average capital structure weights for a firm's industry.

Because firms may be opportunistic in making decisions about whether to issue debt or equity, investors may infer management's expectations for future cash flows, or whether management believes the firm's shares are overvalued or undervalued, based on their financing decisions.

Costs of asymmetric information arise from the fact that managers typically have more information about a company's prospects and future performance than owners or creditors. Firms with complex products or little transparency in financial statements tend to have higher costs of asymmetric information, which result in higher required returns on both debt and equity capital.

Because shareholders and creditors are aware that asymmetric information problems exist, these investors look for management behavior that signals what knowledge or opinions management may have about the firm's prospects. For example, taking on the commitment to make fixed interest payments through debt financing sends a signal that management is confident in the firm's ability to make these payments in the future. By contrast, issuing equity is typically viewed as a negative signal that managers believe a firm's stock is overvalued. The cost of asymmetric information increases with the proportion of equity in the capital structure.

Agency costs of equity are related to conflicts of interest between managers and owners. Managers who do not have a stake in the company do not bear the costs associated with excessive compensation or taking on too much (or too little) risk. Because shareholders are aware of this conflict, they take steps to reduce these costs. The result is called the **net agency cost of equity**. Net agency costs of equity have three components:

- *Monitoring costs* are associated with supervising management and include the expenses of reporting to shareholders and paying the board of directors. Strong corporate governance systems reduce monitoring costs.
- *Bonding costs* relate to assuring shareholders that the managers are working in the shareholders' best interest. Examples of bonding costs include premiums for insurance to guarantee performance and implicit costs associated with non-compete agreements.
- *Residual losses* may occur even with adequate monitoring and bonding provisions because such provisions do not provide a perfect guarantee.

According to agency theory, the use of debt forces managers to be disciplined with regard to how they spend cash because they have less free cash flow to use for their own benefit. It follows that greater amounts of financial leverage tend to reduce agency costs.

Pecking order theory, based on asymmetric information, is related to the signals management sends to investors through its financing choices. According to pecking order theory, managers prefer to make financing choices that are least likely to send negative signals to investors. Financing choices under pecking order theory follow a hierarchy based on visibility to investors. Internally generated capital is most preferred, debt is the next-best choice, and external equity is the least preferred financing option. Pecking order theory implies that the capital structure is a by-product of individual financing decisions.

LOS 34.e: Describe competing stakeholder interests in capital structure decisions.

A company's various stakeholders often have goals that differ from those of management and board members. They may have interests that conflict with those of management in terms of capital structure decisions, especially the issuance of debt. We highlight some of those possible conflicts here.

Public Debtholders

The best possible outcome for a company's debtholders is that their promised interest payments and principal repayment are made in full and as scheduled. That is, their upside is strictly limited with respect to company performance. Their maximum downside is a loss of 100% if the company is unable to repay any of the debt. We can contrast this with the upside and downside for common equity holders. While the downside for owners of common equity is also a loss of 100% of their investment, their upside could be some multiple of their investment if firm performance is exceptionally good.

This asymmetry of outcomes creates a situation where debtholders prefer that a company issues less debt, rather than more, to decrease the probability of financial distress that may lead to a default on the debt they own. Common stockholders, on the other hand, may prefer greater debt financing even though that increases company risk. In general, this conflict is greater for owners of long-term debt compared with owners of short-term debt. The conflict is less severe for owners of debt that has greater seniority (priority of payment) or is backed by valuable liquid assets.

Preferred Stockholders

The priority for payments of preferred stock dividends is after debtholders are paid but ahead of common stock dividends. Similar to debt, preferred stock has an upside that is limited to the timely payment of the promised dividends. However, if the firm cannot make the promised payments on its debt, preferred stock dividends will not be paid. Thus the conflict of interest between the owners of common shares and holders of preferred stock is similar to the conflict of interest between common stockholders and debtholders. Preferred stock has no maturity date and therefore can be considered very long term.

Private Equity or Controlling Shareholders

The interests of a shareholder who holds voting control may differ not only from the interests of debtholders but also from the interests of other common stock investors. The nature of the

conflict depends on the specific circumstances. A controlling shareholder may pursue personal interests that will not necessarily increase shareholder value, may have a short-term focus if they intend to sell their shares, and may oppose share issuance that would dilute their holdings or lead to a loss of control.

Banks and Private Lenders

Unlike holders of a firm's public debt (bonds), banks and private lenders often have access to nonpublic information about a firms' operations. These lenders can more easily restructure debt or adjust terms if problems arise, and they often work with management to ensure that their interests are considered in company decisions regarding investments and the issuance of additional debt.

Private lenders typically hold debt to maturity, while owners of public debt can usually sell that debt if they believe the company's financial health is deteriorating. Banks and private lenders to small or mid-size companies often work closely with them and have influence on company decisions, because the interests of both are somewhat aligned. That said, the risk tolerance of individual lenders can vary, and some provide more services and have more influence than others.

Customers and Suppliers

Customers of specialized products have an interest in the financial health and survival of firms that are their key suppliers, similar to the interests of debtholders.

Suppliers typically are short-term creditors of a firm and thus have an interest in the firm's continuing ability to meet its obligations. Some suppliers have invested time and capital in developing specialized products for a firm and will lose significant revenue if that firm fails. Suppliers in general, and suppliers of custom products for a specific company in particular, have significant interests in the customer firm's stability. Suppliers of more commodity-like products typically have many customers and less exposure to the financial problems of a single firm.

Employees

Employees sometimes own the company's common stock in retirement accounts or employee stock ownership plans. For most employees, the value of their employment with the company is much larger than their stock ownership. Employees who have specialized skills, such that they would face difficulties in finding alternative employment, have a stronger preference for less financial and operational risk, compared with employees who have skills that are more easily transferable.

Managers and Directors

Senior managers and board members may have significant ownership of a firm's common stock. Senior managers, in particular, may be compensated with stock options that affect their incentives. Owning stock options magnifies the benefits of increases in share values and can therefore increase management's appetite for risk. Because stock options increase in value by more than the underlying shares, managers with significant stock options may prefer a greater level of risk.

Managers may prefer a strategy of repurchasing stock to paying cash dividends. Stock repurchases decrease the number of shares outstanding and tend to increase share prices. Companies may issue debt to raise money for stock repurchases. This effectively increases the proportion of debt in the capital structure and the financial risk of the company.

Another consideration, however, is that senior managers and board members are often very highly paid for the work involved. Because of this, they may prefer that the company take on less risk as they focus primarily on the company's survival and their continued employment.

Regulators and Government

Some firms, especially financial institutions, are required to maintain minimum proportions of equity financing, effectively restricting their ability to issue debt or pay cash dividends.

Regulated utilities may be limited in the prices they can charge, with allowable prices set in reference to their cost of capital. There may be an advantage to utilities in using debt financing because the cost of debt capital is clear, while the cost of equity must be estimated.

We have seen in times of financial crises that governments may step in to rescue companies judged to be "too big" or "too essential" to fail. In these cases, governments often require the companies to raise more equity, forego dividend payments and share repurchases, and maintain minimum proportions of equity in their capital structures.



MODULE QUIZ 34.2

1. Which of the following is *least likely* an appropriate method for an analyst to estimate a firm's target capital structure?
 - A. Use the firm's current proportions of debt and equity based on market values, with an adjustment for recent trends in its capital structure.
 - B. Use average capital structure weights for the firm's industry, based on book values of debt and equity.
 - C. Use the firm's current capital structure, based on market values of debt and equity.
2. To determine their target capital structures in practice, it is *least likely* that firms will:
 - A. use the book value of their debt to make financing decisions.
 - B. match the maturities of their debt issues to specific firm investments.
 - C. determine an optimal capital structure based on the expected costs of financial distress.
3. The pecking order theory of financial structure decisions:
 - A. is based on information asymmetry.
 - B. suggests that debt is the first choice for financing an investment of significant size.
 - C. suggests that debt is the riskiest and least preferred source of financing.
4. Compared with managers who do not have significant compensation in the form of stock options, managers who have such compensation will be expected to favor:
 - A. less financial leverage.
 - B. greater firm risk.
 - C. issuance of common stock.

KEY CONCEPTS

LOS 34.a

Capital structures vary among companies but tend to be similar within industries. Internal factors that affect capital structures include the characteristics of the business, the company's

existing debt level, their corporate tax rate, the company's life cycle stage, and management's policies. External factors include market and business cycle conditions, regulation, and industry norms.

LOS 34.b

A company's ability to issue debt is greater with predictable cash flows sufficient to make required debt payments and with liquid tangible assets the company can pledge as collateral for debt.

New companies with few assets and negative or uncertain cash flows will use little to no debt. Growth companies with positive cash flows and decreasing business risk may use debt for up to 20% of their financing. Mature companies with predictable cash flows tend to use debt for more than 20% of their financing, sometimes significantly more.

LOS 34.c

MM's propositions with no taxes are that a company's capital structure is irrelevant, because its WACC and firm value (the discounted present value of its operating earnings) are unchanged by changes in capital structure.

MM's propositions with taxes are that a company's WAAC is minimized and its value maximized with 100% debt financing.

Static trade-off theory adds the expected costs of financial distress to the model. It indicates that firm value initially increases (and WACC decreases) with additional debt financing, but that company value decreases at some point with additional debt as the increase in the expected costs of financial distress outweigh the increase in tax benefits from additional debt.

LOS 34.d

For target capital structure weights (based on market values of debt and equity), analysts can use:

- a company's stated target weights.
- a company's current capital structure weights.
- a company's current capital structure weights adjusted for the observed trend in its capital structure weights.
- industry average capital structure weights.

In practice, a company's capital structure will fluctuate around the target due to management's exploitation of market opportunities and market value fluctuations (especially of equity) over time.

Pecking order theory is based on information asymmetry between firm management and investors and suggests that management's choice of financing method signals their beliefs about firm value. The theory concludes that retained earnings are the most preferred source of funds, followed by debt financing, and then issuing new equity.

Agency costs of equity, which arise because management and shareholders may have conflicting interests, are reduced by increased debt issuance.

LOS 34.e

A company's various stakeholders may have interests that conflict with those of management in terms of capital structure decisions, especially the issuance of debt.

While stakeholder issues can depend on specific circumstances, in general, compared with common shareholders:

- Debtholders, preferred stockholders, customers (especially those with specific requirements), suppliers (especially those of firm-specific items), employees (especially those with firm-specific skills), directors and managers without significant stock options, firm regulators who impose minimum equity requirements, and government entities may have a preference for less financial leverage (risk) and less operational risk.
- Managers and directors holding or receiving significant stock options, some private lenders, some private equity owners, and some shareholders with controlling interests may prefer more financial leverage.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 34.1

1. **B** For an analyst, target capital structure should always be based on market values of debt and equity. The other two choices are appropriate methods for estimating a firm's capital structure for analysis. (LOS 34.d)
2. **C** While it is a useful theoretical concept, in practice determining an optimal capital structure based on the cost savings of debt and the expected costs of financial distress is not feasible. Because debt rating companies often use book values of debt, firms use book values of debt when choosing financing sources. It is common for firms to match debt maturities to the economic lives of specific investments. (LOS 34.d)
3. **A** Pecking order theory is based on information asymmetry and the resulting signals that different financing choices send to investors. It suggests that retained earnings are the first choice for financing an investment and issuing new equity is the least preferred choice. (LOS 34.d)
4. **B** Given the asymmetric returns on stock options, we would expect managers with significant stock options in their compensation to favor greater financial leverage and issuance of debt to increase potential stock price gains. Issuing common stock could decrease the market price of shares, which would decrease the value of stock options. (LOS 34.e)

Module Quiz 34.2

1. **A** During the start-up stage a firm is unlikely to have positive earnings and cash flows or significant assets that can be pledged as debt collateral, so firms in this stage are typically financed by equity only. (LOS 34.b)
2. **A** The optimal capital structure minimizes the firm's WACC and maximizes the firm's value (stock price). (LOS 34.c)

3. C MM's Proposition II (with no taxes) states that capital structure is irrelevant because the decrease in a firm's WACC from additional debt financing is just offset by the increase in WACC from a decrease in equity financing. The cost of debt is held constant and the cost of equity financing increases as the proportion of *debt* in the capital structure is increased. (LOS 34.c)

READING 35

MEASURES OF LEVERAGE

EXAM FOCUS

Here we define and calculate various measures of leverage and the firm characteristics that affect the levels of operating and financial leverage. Operating leverage magnifies the effect of changes in sales on operating earnings. Financial leverage magnifies the effect of changes in operating earnings on net income (earnings per share). The breakeven quantity of sales is that quantity of sales for which total revenue just covers total costs. The operating breakeven quantity of sales is the quantity of sales for which total revenue just covers total operating costs. Be sure you understand how a firm's decisions regarding its operating structure and scale and its decisions regarding the use of debt and equity financing (its capital structure) affect its breakeven levels of sales and the uncertainty regarding its operating earnings and net income.

MODULE 35.1: MEASURES OF LEVERAGE



LOS 35.a: Define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk.

Video covering this content is available online.

Leverage, in the sense we use it here, refers to the amount of fixed costs a firm has. These fixed costs may be fixed operating expenses, such as building or equipment leases, or fixed financing costs, such as interest payments on debt. Greater leverage leads to greater variability of the firm's after-tax operating earnings and net income. A given change in sales will lead to a greater change in operating earnings when the firm employs operating leverage; a given change in operating earnings will lead to a greater change in net income when the firm employs financial leverage.



PROFESSOR'S NOTE

The British refer to leverage as "gearing."

Business risk refers to the risk associated with a firm's operating income and is the result of uncertainty about a firm's revenues and the expenditures necessary to produce those revenues. Business risk is the combination of sales risk and operating risk.

- **Sales risk** is the uncertainty about the firm's sales.
- **Operating risk** refers to the additional uncertainty about operating earnings caused by fixed operating costs. The greater the proportion of fixed costs to variable costs, the greater a firm's operating risk.

Financial risk refers to the additional risk that the firm's common stockholders must bear when a firm uses fixed cost (debt) financing. When a company finances its operations with debt, it takes on fixed expenses in the form of interest payments. The greater the proportion of debt in a firm's capital structure, the greater the firm's financial risk.

LOS 35.b: Calculate and interpret the degree of operating leverage, the degree of financial leverage, and the degree of total leverage.

The **degree of operating leverage (DOL)** is defined as the percentage change in operating income (EBIT) that results from a given percentage change in sales:

$$DOL = \frac{\text{percentage change in EBIT}}{\text{percentage change in sales}} = \frac{\frac{\Delta EBIT}{EBIT}}{\frac{\Delta Q}{Q}}$$

To calculate a firm's DOL for a particular level of unit sales, Q , DOL is:

$$DOL = \frac{Q(P - V)}{Q(P - V) - F}$$

where:

Q = quantity of units sold

P = price per unit

V = variable cost per unit

F = fixed costs

Multiplying, we have:

$$DOL = \frac{S - TVC}{S - TVC - F}$$

where:

S = sales

TVC = total variable costs

F = fixed costs

Note that in this form, the denominator is operating earnings (EBIT).

EXAMPLE: Degree of operating leverage

Atom Company produced 5,000 units last year that it sold for \$75 each. Atom's fixed costs were \$70,000, and its variable cost per unit was \$50. Calculate and interpret Atom's degree of operating leverage at this level of production.

Answer:

$$DOL = \frac{Q(P - V)}{Q(P - V) - F} = \frac{5,000(\$75 - \$50)}{5,000(\$75 - \$50) - \$70,000}$$

$$DOL = \frac{\$125,000}{\$55,000} = 2.2727$$

The result indicates that if Atom Company has a 3% increase in sales, its EBIT will increase by $2.2727 \times 3\% = 6.82\%$.

It is important to note that the degree of operating leverage for a company depends on the level of sales. For example, if Atom Company sells 10,000 units, the DOL is decreased:

$$\begin{aligned} DOL(\text{Atom}) &= \frac{Q(P - V)}{[Q(P - V) - F]} = \frac{10,000(\$75 - \$50)}{[10,000(\$75 - \$50) - \$70,000]} \\ &= \frac{250,000}{180,000} = 1.39 \end{aligned}$$

DOL is highest at low levels of sales and declines at higher levels of sales.

The **degree of financial leverage (DFL)** is interpreted as the ratio of the percentage change in net income (or EPS) to the percentage change in EBIT:

$$DFL = \frac{\text{percentage change in EPS}}{\text{percentage change in EBIT}}$$

For a particular level of operating earnings, DFL is calculated as:

$$DFL = \frac{EBIT}{EBIT - \text{interest}}$$



PROFESSOR'S NOTE

We use the terms “earnings per share” (EPS) and “net income” interchangeably in this reading.

EXAMPLE: Degree of financial leverage

From the previous example, Atom Company’s operating income from selling 5,000 units is \$55,000. Assume that Atom has an annual interest expense of \$20,000. Calculate and interpret Atom’s degree of financial leverage.

Answer:

$$DFL = \frac{EBIT}{EBIT - I} = \frac{\$55,000}{\$55,000 - \$20,000} = 1.5714$$

The result indicates that if Atom Company has a 3% increase in EBIT, earnings per share will increase by $1.5714 \times 3\% = 4.71\%$.



PROFESSOR'S NOTE

Look back at the formulas for DOL and DFL and convince yourself that if there are no fixed costs, DOL is equal to one, and that if there are no interest costs, DFL is equal to one. Values of one mean no leverage. No fixed costs, no operating leverage. No interest costs, no financial leverage. This should help tie these formulas to the concepts and help you know when you have the formulas right (or wrong). If you plug in zero for fixed costs, DOL should be one, and if you plug in zero for interest, DFL should be one.

The **degree of total leverage (DTL)** combines the degree of operating leverage and financial leverage. DTL measures the sensitivity of EPS to change in sales. DTL is computed as:

$$DTL = DOL \times DFL$$

$$DTL = \frac{\% \Delta EBIT}{\% \Delta \text{sales}} \times \frac{\% \Delta EPS}{\% \Delta EBIT} = \frac{\% \Delta EPS}{\% \Delta \text{sales}}$$

$$DTL = \frac{Q(P - V)}{Q(P - V) - F - I}$$

$$DTL = \frac{S - TVC}{S - TVC - F - I}$$

EXAMPLE: Degree of total leverage

Continuing with our previous example, calculate Atom Company's degree of total leverage and determine how much Atom's EPS will increase if its sales increase by 10%.

Answer:

$$DTL = DOL \times DFL = 2.2727 \times 1.5714 = 3.5713$$

Alternatively,

$$\begin{aligned} DTL &= \frac{S - TVC}{S - TVC - F - I} = \frac{\$375,000 - \$250,000}{\$375,000 - \$250,000 - \$70,000 - \$20,000} \\ &= 3.5714 \end{aligned}$$

The result indicates that if Atom Company has a 10% increase in sales, earnings per share will increase by $3.5714 \times 10\% = 35.714\%$.

LOS 35.c: Analyze the effect of financial leverage on a company's net income and return on equity.

The use of financial leverage significantly increases the risk and potential reward to common stockholders. The following examples involving Beta Company illustrate how financial leverage affects net income and shareholders' return on equity (ROE).

EXAMPLE 1: Beta Company financed with 100% equity

Assume that the Beta Company has \$500,000 in assets that are financed with 100% equity. Fixed costs are \$120,000. Beta is expected to sell 100,000 units, resulting in operating income of $[100,000 (\$4 - \$2)] - \$120,000 = \$80,000$. Beta's tax rate is 40%. Calculate Beta's net income and return on equity if its EBIT increases or decreases by 10%.

Answer:

Beta's Return on Equity With 100% Equity Financing

	EBIT Less 10%	Expected EBIT	EBIT Plus 10%
EBIT	\$72,000	\$80,000	\$88,000
Interest expense	0	0	0
Income before taxes	\$72,000	\$80,000	\$88,000
Taxes at 40%	28,800	32,000	35,200
Net income	\$43,200	\$48,000	\$52,800
Shareholders' equity	\$500,000	\$500,000	\$500,000
Return on equity (ROE)	8.64%	9.60%	10.56%

EXAMPLE 2: Beta Company financed with 50% equity and 50% debt

Continuing the previous example, assume that Beta Company is financed with 50% equity and 50% debt. The interest rate on the debt is 6%. Calculate Beta's net income and return on equity if its EBIT increases or decreases by 10%. Beta's tax rate is 40%.

Answer:

Beta's Return on Equity with 50% Equity Financing

	EBIT Less 10%	Expected EBIT	EBIT Plus 10%
EBIT	\$72,000	\$80,000	\$88,000
Interest expense at 6%	15,000	15,000	15,000
Income before taxes	\$57,000	\$65,000	\$73,000
Taxes at 40%	22,800	26,000	29,200
Net income	\$34,200	\$39,000	\$43,800
Shareholders' equity	\$250,000	\$250,000	\$250,000
Return on equity (ROE)	13.68%	15.60%	17.52%

The interest expense associated with using debt represents a fixed cost that reduces net income. However, the lower net income value is spread over a smaller base of shareholders' equity, serving to magnify the ROE. In all three of the scenarios shown in the two examples, ROE is higher using leverage than it is without leverage.

Further analyzing the differences between the examples, we can see that the use of financial leverage not only increases the *level* of ROE, it also increases the *rate of change* for ROE. In the unleveraged scenario, ROE varies directly with the change in EBIT. For an increase in EBIT of 10%, the ROE increases from 9.60% to 10.56%, for a rate of change of 10%. In the leveraged scenario, ROE is more volatile. For an increase in EBIT of 10%, the ROE increases from 15.60% to 17.52%, for a rate of change of 12.3%.

The use of financial leverage increases the risk of default but also increases the potential return for equity holders.



PROFESSOR'S NOTE

Recall how this relationship is reflected in the DuPont formula used to analyze ROE. One of the components of the DuPont formula is the equity multiplier (assets/equity), which captures the effect of financial leverage on ROE.

LOS 35.d: Calculate the breakeven quantity of sales and determine the company's net income at various sales levels.

LOS 35.e: Calculate and interpret the operating breakeven quantity of sales.

The level of sales that a firm must generate to cover all of its fixed and variable costs is called the **breakeven quantity**. The **breakeven quantity of sales** is the quantity of sales for which revenues equal total costs, so that net income is zero. We can calculate the breakeven quantity by simply determining how many units must be sold to just cover total fixed costs.

For each unit sold, the **contribution margin**, which is the difference between price and variable cost per unit, is available to help cover fixed costs. We can thus describe the breakeven quantity of sales, Q_{BE} , as:

$$Q_{BE} = \frac{\text{fixed operating costs} + \text{fixed financing costs}}{\text{price} - \text{variable cost per unit}}$$

EXAMPLE: Breakeven quantity of sales

Consider the prices and costs for Atom Company and Beta Company shown in the following table. Compute and illustrate the breakeven quantity of sales for each company.

Operating Costs for Atom Company and Beta Company

	Atom Company	Beta Company
Price	\$4.00	\$4.00
Variable costs	\$3.00	\$2.00
Fixed operating costs	\$10,000	\$80,000
Fixed financing costs	\$30,000	\$40,000

Answer:

For Atom Company, the breakeven quantity is:

$$Q_{BE}(\text{Atom}) = \frac{\$10,000 + \$30,000}{\$4.00 - \$3.00} = 40,000 \text{ units}$$

Similarly, for Beta Company, the breakeven quantity is:

$$Q_{BE}(\text{Beta}) = \frac{\$80,000 + \$40,000}{\$4.00 - \$2.00} = 60,000 \text{ units}$$

The breakeven quantity and the relationship between sales revenue, total costs, net income, and net loss are illustrated in Figure 35.1 and Figure 35.2.

Figure 35.1: Breakeven Analysis for Atom Company

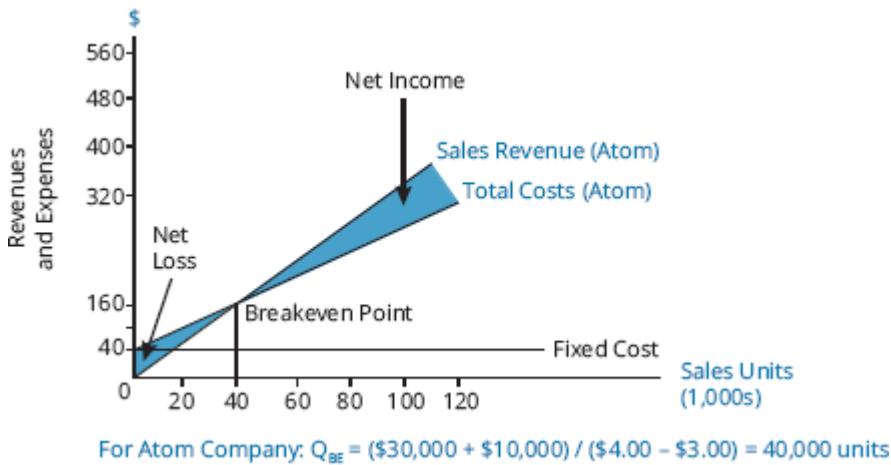
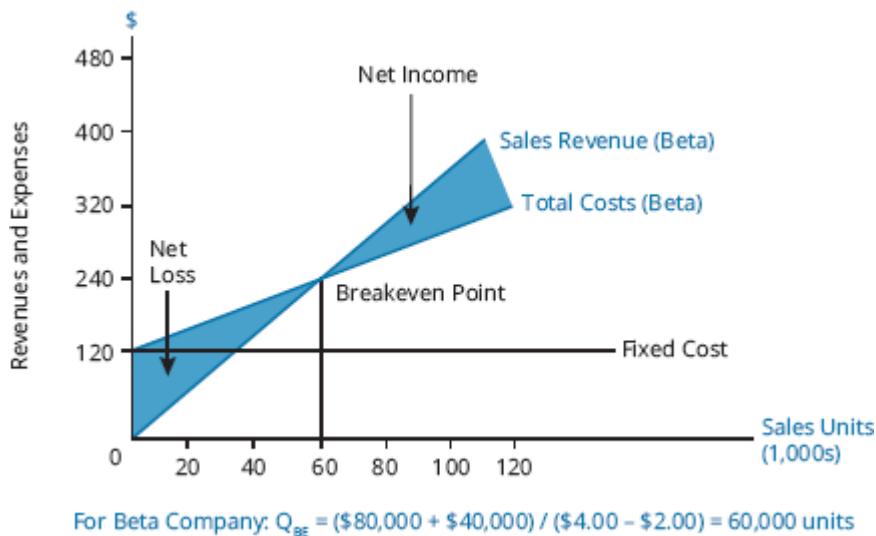


Figure 35.2: Breakeven Analysis for Beta Company



We can also calculate an **operating breakeven quantity of sales**. In this case, we consider only fixed operating costs and ignore fixed financing costs. The calculation is simply:

$$Q_{OBE} = \frac{\text{fixed operating costs}}{\text{price} - \text{variable cost per unit}}$$

EXAMPLE: Operating breakeven quantity of sales

Calculate the operating breakeven quantity of sales for Atom and Beta, using the same data from the previous example.

Answer:

For Atom, the operating breakeven quantity of sales is:

$$\$10,000 / (\$4.00 - \$3.00) = 10,000 \text{ units}$$

For Beta, the operating breakeven quantity of sales is:

$$\$80,000 / (\$4.00 - \$2.00) = 40,000 \text{ units}$$

We can summarize the effects of leverage on net income through an examination of Figure 35.1 and Figure 35.2. Other things equal, a firm that chooses operating and financial structures that result in greater total fixed costs will have a higher breakeven quantity of sales. Leverage of either type magnifies the effects of changes in sales on net income. The further a firm's sales are from its breakeven level of sales, the greater the magnifying effects of leverage on net income.

These same conclusions apply to operating leverage and the operating breakeven quantity of sales. One company may choose a larger scale of operations (larger factory), resulting in a greater operating breakeven quantity of sales and greater leverage, other things equal.

Note that the degree of total leverage is calculated for a particular level of sales. The slope of the net income line in Figure 35.1 and Figure 35.2 is related to total leverage but is not the same thing. The degree of total leverage is different for every level of sales.



MODULE QUIZ 35.1

1. Business risk is the combination of:
 - A. operating risk and financial risk.
 - B. sales risk and financial risk.
 - C. operating risk and sales risk.
2. Which of the following is a key determinant of operating leverage?
 - A. Level and cost of debt.
 - B. The competitive nature of the business.
 - C. The trade-off between fixed and variable costs.
3. Which of the following statements about capital structure and leverage is *most accurate*?
 - A. Financial leverage is directly related to operating leverage.
 - B. Increasing the corporate tax rate will not affect capital structure decisions.
 - C. A firm with low operating leverage has a small proportion of its total costs in fixed costs.
4. Jayco, Inc., sells blue ink for \$4 a bottle. The ink's variable cost per bottle is \$2. Ink has fixed operating costs of \$4,000 and fixed financing costs of \$6,000. What is Jayco's breakeven quantity of sales, in units?
 - A. 2,000.
 - B. 3,000.
 - C. 5,000.
5. Jayco, Inc., sells blue ink for \$4 a bottle. The ink's variable cost per bottle is \$2. Ink has fixed operating costs of \$4,000 and fixed financing costs of \$6,000. What is Jayco's operating breakeven quantity of sales, in units?
 - A. 2,000.
 - B. 3,000.
 - C. 5,000.
6. If Jayco's sales increase by 10%, Jayco's EBIT increases by 15%. If Jayco's EBIT increases by 10%, Jayco's EPS increases by 12%. Jayco's degree of operating leverage (DOL) and degree of total leverage (DTL) are *closest to*:
 - A. 1.2 DOL and 1.5 DTL.
 - B. 1.2 DOL and 2.7 DTL.
 - C. 1.5 DOL and 1.8 DTL.

Use the following data to answer Questions 7 and 8.

Jayco, Inc., sells 10,000 units at a price of \$5 per unit. Jayco's fixed costs are \$8,000, interest expense is \$2,000, variable costs are \$3 per unit, and EBIT is

\$12,000.

7. Jayco's degree of operating leverage (DOL) and degree of financial leverage (DFL) are *closest* to:
 - A. 2.50 DOL and 1.00 DFL.
 - B. 1.67 DOL and 2.00 DFL.
 - C. 1.67 DOL and 1.20 DFL.
8. Jayco's degree of total leverage (DTL) is *closest* to:
 - A. 2.00.
 - B. 1.75.
 - C. 1.50.
9. Vischer Concrete has \$1.2 million in assets that are currently financed with 100% equity. Vischer's EBIT is \$300,000, and its tax rate is 30%. If Vischer changes its capital structure (recapitalizes) to include 40% debt, what is Vischer's ROE before and after the change? Assume that the interest rate on debt is 5%.

<u>ROE at 100% equity</u>	<u>ROE at 60% equity</u>
A. 17.5%	26.8%
B. 25.0%	26.8%
C. 25.0%	37.5%

KEY CONCEPTS

LOS 35.a

Leverage increases the risk and potential return of a firm's earnings and cash flows.

Operating leverage increases with fixed operating costs.

Financial leverage increases with fixed financing costs.

Sales risk is uncertainty about the firm's sales.

Business risk refers to the uncertainty about operating earnings (EBIT) and results from variability in sales and expenses. Business risk is magnified by operating leverage.

Financial risk refers to the additional variability of EPS compared to EBIT. Financial risk increases with greater use of fixed cost financing (debt) in a company's capital structure.

LOS 35.b

The degree of operating leverage (DOL) is calculated as $\frac{Q(P - V)}{Q(P - V) - F}$ and is interpreted as $\frac{\% \Delta EBIT}{\% \Delta sales}$.

The degree of financial leverage (DFL) is calculated as $\frac{EBIT}{EBIT - I}$ and is interpreted as $\frac{\% \Delta EPS}{\% \Delta EBIT}$.

The degree of total leverage (DTL) is the combination of operating and financial leverage and is calculated as $DOL \times DFL$ and interpreted as $\frac{\% \Delta EPS}{\% \Delta sales}$.

LOS 35.c

Using more debt and less equity in a firm's capital structure reduces net income through added interest expense but also reduces net equity. The net effect can be to either increase or decrease ROE.

LOS 35.d

The breakeven quantity of sales is the amount of sales necessary to produce a net income of zero (total revenue just covers total costs) and can be calculated as:

$$\frac{\text{fixed operating costs} + \text{fixed financing costs}}{\text{price} - \text{variable cost per unit}}$$

Net income at various sales levels can be calculated as total revenue (i.e., price \times quantity sold) minus total costs (i.e., total fixed costs plus total variable costs).

LOS 35.e

The operating breakeven quantity of sales is the amount of sales necessary to produce an operating income of zero (total revenue just covers total operating costs) and can be calculated as:

$$\frac{\text{fixed operating costs}}{\text{price} - \text{variable cost per unit}}$$

ANSWER KEY FOR MODULE QUIZ

Module Quiz 35.1

1. **C** Business risk refers to the risk associated with a firm's operating income and is the result of uncertainty about a firm's revenues and the expenditures necessary to produce those revenues. Business risk is the combination of sales risk (the uncertainty associated with the price and quantity of goods and services sold) and operating risk (the leverage created by the use of fixed costs in the firm's operations). (LOS 35.a)
2. **C** The extent to which costs are fixed determines operating leverage. (LOS 35.b)
3. **C** If fixed costs are a small percentage of total costs, operating leverage is low. Operating leverage is separate from financial leverage, which depends on the amount of debt in the capital structure. Increasing the tax rate would make the after-tax cost of debt cheaper. (LOS 35.b)
4. **C** $Q_{BE} = \frac{\$4,000 + \$6,000}{\$4.00 - \$2.00} = 5,000 \text{ units}$ (LOS 35.d)
5. **A** $Q_{OBE} = \frac{\$4,000}{\$4.00 - \$2.00} = 2,000 \text{ units}$ (LOS 35.e)

6. C

$$DOL = \frac{15\%}{10\%} = 1.5$$

$$DFL = \frac{12\%}{10\%} = 1.2$$

$$DTL = DOL \times DFL = 1.5 \times 1.2 = 1.8$$

(LOS 35.b)

7. C

$$DOL = \frac{Q(P - V)}{[Q(P - V) - F]} = \frac{10,000(5 - 3)}{[10,000(5 - 3) - 8,000]} = 1.67$$

$$DFL = \frac{EBIT}{EBIT - I} = \frac{12,000}{12,000 - 2,000} = 1.2$$

(LOS 35.b)

8. A

$$DTL = \frac{Q(P - V)}{[Q(P - V) - F - I]} = \frac{10,000(5 - 3)}{[10,000(5 - 3) - 8,000 - 2,000]} = 2,$$

or because we calculated the components in Question 7, $DTL = DOL \times$

$$DFL = 1.67 \times 1.2 = 2.0 \text{ (LOS 35.b)}$$

9. A

With 100% equity:

EBIT	\$300,000
Interest expense	0
Income before taxes	\$300,000
Taxes at 30%	90,000
Net income	\$210,000
Shareholders' equity	\$1,200,000
ROE = NI/equity	17.5%

With 60% equity:

EBIT	\$300,000
Interest expense (\$480,000 at 5%)	24,000
Income before taxes	\$276,000
Taxes at 30%	82,800
Net income	\$193,200
Shareholders' equity	\$720,000
ROE = NI/equity	26.8%

(LOS 35.c)

TOPIC QUIZ: CORPORATE ISSUERS

You have now finished the Corporate Issuers topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Corporate Issuers. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

READING 36

MARKET ORGANIZATION AND STRUCTURE

EXAM FOCUS

There is a great deal of introductory material in this review. Almost all of the types of securities discussed are covered in detail elsewhere in the curriculum. We introduce the terminology you will need but leave many of the details to the readings specific to each security type. You should understand the concept of purchasing stock on margin and be able to calculate the return on an investment using margin. Be able to differentiate between market and limit orders as well as between quote-driven, order-driven, and brokered markets. Know that market regulation should increase informational, allocational, and operational market efficiency.

MODULE 36.1: MARKETS, ASSETS, AND INTERMEDIARIES



Video covering this content is available online.

LOS 36.a: Explain the main functions of the financial system.

The three main functions of the financial system are to:

1. Allow entities to save and borrow money, raise equity capital, manage risks, trade assets currently or in the future, and trade based on their estimates of asset values.
2. Determine the returns (i.e., interest rates) that equate the total supply of savings with the total demand for borrowing.
3. Allocate capital to its most efficient uses.

The financial system allows the transfer of assets and risks from one entity to another as well as across time. Entities who utilize the financial system include individuals, firms, governments, charities, and others.

Achievement of Purposes in the Financial System

The financial system allows entities to save, borrow, issue equity capital, manage risks, exchange assets, and to utilize information. The financial system is best at fulfilling these roles when the markets are liquid, transactions costs are low, information is readily available, and when regulation ensures the execution of contracts.

Savings. Individuals will save (e.g., for retirement) and expect a return that compensates them for risk and the use of their money. Firms save a portion of their sales to fund future expenditures. Vehicles used for saving include stocks, bonds, certificates of deposit, real assets, and other assets.

Borrowing. Individuals may borrow in order to buy a house, fund a college education, or for other purposes. A firm may borrow in order to finance capital expenditures and for other activities. Governments may issue debt to fund their expenditures. Lenders can require collateral to protect them in the event of borrower defaults, take an equity position, or investigate the credit risk of the borrower.

Issuing equity. Another method of raising capital is to issue equity, where the capital providers will share in any future profits. Investment banks help with issuance, analysts value the equity, and regulators and accountants encourage the dissemination of information.

Risk management. Entities face risks from changing interest rates, currency values, commodities values, and defaults on debt, among other things. For example, a firm that owes a foreign currency in 90 days can lock in the price of this foreign currency in domestic currency units by entering into a forward contract. Future delivery of the foreign currency is guaranteed at a domestic-currency price set at inception of the contract. In this transaction, the firm would be referred to as a *hedger*. This hedging allows the firm to enter a market that it would otherwise be reluctant to enter by reducing the risk of the transaction. Hedging instruments are available from exchanges, investment banks, insurance firms, and other institutions.

Exchanging assets. The financial system also allows entities to exchange assets. For example, Proctor and Gamble may sell soap in Europe but have costs denominated in U.S. dollars. Proctor and Gamble can exchange their euros from soap sales for dollars in the currency markets.

Utilizing information. Investors with information expect to earn a return on that information in addition to their usual return. Investors who can identify assets that are currently undervalued or overvalued in the market can earn extra returns from investing based on their information (when their analysis is correct).

Return Determination

The financial system also provides a mechanism to determine the rate of return that equates the amount of borrowing with the amount of lending (saving) in an economy. Low rates of return increase borrowing but reduce saving (increase current consumption). High rates of return increase saving but reduce borrowing. The **equilibrium interest rate** is the rate at which the amount individuals, businesses, and governments desire to borrow is equal to the amount that individuals, businesses, and governments desire to lend. Equilibrium rates for different types of borrowing and lending will differ due to differences in risk, liquidity, and maturity.

Allocation of Capital

With limited availability of capital, one of the most important functions of a financial system is to allocate capital to its most efficient uses. Investors weigh the expected risks and returns of different investments to determine their most preferred investments. As long as investors are well informed regarding risk and return and markets function well, this results in an allocation to capital to its most valuable uses.

LOS 36.b: Describe classifications of assets and markets.

Financial assets include securities (stocks and bonds), derivative contracts, and currencies. **Real assets** include real estate, equipment, commodities, and other physical assets.

Financial securities can be classified as debt or equity. **Debt securities** are promises to repay borrowed funds. **Equity securities** represent ownership positions.

Public (publicly traded) securities are traded on exchanges or through securities dealers and are subject to regulatory oversight. Securities that are not traded in public markets are referred to as **private securities**. Private securities are often illiquid and not subject to regulation.

Derivative contracts have values that depend on (are derived from) the values of other assets. **Financial derivative contracts** are based on equities, equity indexes, debt, debt indexes, or other financial contracts. **Physical derivative contracts** derive their values from the values of physical assets such as gold, oil, and wheat.

Markets for immediate delivery are referred to as **spot markets**. Contracts for the future delivery of physical and financial assets include forwards, futures, and options. Options provide the buyer the right, but not the obligation, to purchase (or sell) assets over some period or at some future date at predetermined prices.

The **primary market** is the market for newly issued securities. Subsequent sales of securities are said to occur in the **secondary market**.

Money markets refer to markets for debt securities with maturities of one year or less. **Capital markets** refer to markets for longer-term debt securities and equity securities that have no specific maturity date.

Traditional investment markets refer to those for debt and equity. **Alternative markets** refer to those for hedge funds, commodities, real estate, collectibles, gemstones, leases, and equipment. Alternative assets are often more difficult to value, illiquid, require investor due diligence, and therefore often sell at a discount.

LOS 36.c: Describe the major types of securities, currencies, contracts, commodities, and real assets that trade in organized markets, including their distinguishing characteristics and major subtypes.

Assets can be classified as securities, currencies, contracts, commodities, and real assets. Their characteristics and subtypes are as follows.

Securities

Securities can be classified as fixed-income or equity securities, and individual securities can be combined in pooled investment vehicles. Corporations and governments are the most common issuers of individual securities. The initial sale of a security is called an **issue** when the security is sold to the public.

Fixed-income securities typically refer to debt securities that are promises to repay borrowed money in the future. Short-term fixed-income securities generally have a maturity of less than one or two years; long-term term maturities are longer than five to ten years, and intermediate term maturities fall in the middle of the maturity range.

Although the terms are used loosely, *bonds* are generally long term, whereas *notes* are intermediate term. *Commercial paper* refers to short-term debt issued by firms. Governments issue *bills* and banks issue *certificates of deposit*. In *repurchase agreements*, the borrower sells a high-quality asset and has both the right and obligation to repurchase it (at a higher price) in the future. Repurchase agreements can be for terms as short as one day.

Convertible debt is debt that an investor can exchange for a specified number of equity shares of the issuing firm.

Equity securities represent ownership in a firm and include common stock, preferred stock, and warrants.

- **Common stock** is a residual claim on a firm's assets. Common stock dividends are paid only after interest is paid to debtholders and dividends are paid to preferred stockholders. Furthermore, in the event of firm liquidation, debtholders and preferred stockholders have priority over common stockholders and are usually paid in full before common stockholders receive any payment.
- **Preferred stock** is an equity security with scheduled dividends that typically do not change over the security's life and must be paid before any dividends on common stock may be paid.
- **Warrants** are similar to options in that they give the holder the right to buy a firm's equity shares (usually common stock) at a fixed exercise price prior to the warrant's expiration.

Pooled investment vehicles include mutual funds, depositories, and hedge funds. The term refers to structures that combine the funds of many investors in a portfolio of investments. The investor's ownership interests are referred to as *shares*, *units*, *depository receipts*, or *limited partnership interests*.

- **Mutual funds** are pooled investment vehicles in which investors can purchase shares, either from the fund itself (open-end funds) or in the secondary market (closed-end funds).
- **Exchange-traded funds (ETFs)** and **exchange-traded notes (ETNs)** trade like closed-end funds but have special provisions allowing conversion into individual portfolio securities, or exchange of portfolio shares for ETF shares, that keep their market prices close to the value of their proportional interest in the overall portfolio. These funds are sometimes referred to as *depositories*, with their shares referred to as *depository receipts*.
- **Asset-backed securities** represent a claim to a portion of a pool of financial assets such as mortgages, car loans, or credit card debt. The return from the assets is passed through to investors, with different classes of claims (referred to as *tranches*) having different levels of risk.
- **Hedge funds** are organized as limited partnerships, with the investors as the limited partners and the fund manager as the general partner. Hedge funds utilize various strategies and purchase is usually restricted to investors of substantial wealth and investment knowledge. Hedge funds often use leverage. Hedge fund managers are compensated based on the amount of assets under management as well as on their investment results.



PROFESSOR'S NOTE

Asset-backed securities are described in more detail in Fixed Income. Mutual funds and ETFs are discussed in Portfolio Management. Hedge funds are discussed in Alternative Investments.

Currencies

Currencies are issued by a government's central bank. Some are referred to as **reserve currencies**, which are those held by governments and central banks worldwide. These include the dollar and euro and, secondarily, the British pound, Japanese yen, and Swiss franc. In spot currency markets, currencies are traded for immediate delivery.

Contracts

Contracts are agreements between two parties that require some action in the future, such as exchanging an asset for cash. Financial contracts are often based on securities, currencies, commodities, or security indexes (portfolios). They include futures, forwards, options, swaps, and insurance contracts.

A **forward contract** is an agreement to buy or sell an asset in the future at a price specified in the contract at its inception. An agreement to purchase 100 ounces of gold 90 days from now for \$1,000 per ounce is a forward contract. Forward contracts are not traded on exchanges or in dealer markets.

Futures contracts are similar to forward contracts except that they are standardized as to amount, asset characteristics, and delivery time and are traded on an exchange (in a secondary market) so that they are liquid investments.

In a **swap contract**, two parties make payments that are equivalent to one asset being traded (swapped) for another. In a simple *interest rate swap*, floating rate interest payments are exchanged for fixed-rate payments over multiple settlement dates. A *currency swap* involves a loan in one currency for the loan of another currency for a period of time. An *equity swap* involves the exchange of the return on an equity index or portfolio for the interest payment on a debt instrument.

An **option contract** gives its owner the right to buy or sell an asset at a specific exercise price at some specified time in the future. A **call option** gives the option buyer the right (but not the obligation) to buy an asset. A **put option** gives the option buyer the right (but not the obligation) to sell an asset.

Sellers, or writers, of call (put) options receive a payment, referred to as the *option premium*, when they sell the options but incur the obligation to sell (buy) the asset at the specified price if the option owner chooses to exercise it.

Options on currencies, stocks, stock indexes, futures, swaps, and precious metals are traded on exchanges. Customized options contracts are also sold by dealers in the over-the-counter market.

An **insurance contract** pays a cash amount if a future event occurs. They are used to hedge against unfavorable, unexpected events. Examples include life, liability, and automobile

insurance contracts. Insurance contracts can sometimes be traded to other parties and often have tax-advantaged payouts.

Credit default swaps are a form of insurance that makes a payment if an issuer defaults on its bonds. They can be used by bond investors to hedge default risk. They can also be used by parties that will experience losses if an issuer experiences financial distress and by others who are speculating that the issuer will experience more or less financial trouble than is currently expected.

Commodities

Commodities trade in spot, forward, and futures markets. They include precious metals, industrial metals, agricultural products, energy products, and credits for carbon reduction.

Futures and forwards allow both hedgers and speculators to participate in commodity markets without having to deliver or store the physical commodities.

Real Assets

Examples of **real assets** are real estate, equipment, and machinery. Although they have been traditionally held by firms for their use in production, real assets are increasingly held by institutional investors both directly and indirectly.

Buying real assets directly often provides income, tax advantages, and diversification benefits. However, they often entail substantial management costs. Furthermore, because of their heterogeneity, they usually require the investor to do substantial due diligence before investing. They are illiquid because their specialization may result in a limited pool of investors for a particular real asset.

Rather than buying real assets directly, an investor may choose to buy them indirectly through an investment such as a *real estate investment trust (REIT)* or *master limited partnership (MLP)*. The investor owns an interest in these vehicles, which hold the assets directly. Indirect ownership interests are typically more liquid than ownership of the assets themselves. Another indirect ownership method is to buy the stock of firms that have large ownership of real assets.

LOS 36.d: Describe types of financial intermediaries and services that they provide.

Financial intermediaries stand between buyers and sellers, facilitating the exchange of assets, capital, and risk. Their services allow for greater efficiency and are vital to a well-functioning economy. Financial intermediaries include brokers and exchanges, dealers, securitizers, depository institutions, insurance companies, arbitrageurs, and clearinghouses.

Brokers, Dealers, and Exchanges

Brokers help their clients buy and sell securities by finding counterparties to trades in a cost efficient manner. They may work for large brokerage firms, for banks, or at exchanges.

Block brokers help with the placement of large trades. Typically, large trades are difficult to place without moving the market. For example, a large sell order might cause a security's price

to decrease before the order can be fully executed. Block brokers help conceal their clients' intentions so that the market does not move against them.

Investment banks help corporations sell common stock, preferred stock, and debt securities to investors. They also provide advice to firms, notably about mergers, acquisitions, and raising capital.

Exchanges provide a venue where traders can meet. Exchanges sometimes act as brokers by providing electronic order matching. Exchanges regulate their members and require firms that list on the exchange to provide timely financial disclosures and to promote shareholder democratization. Exchanges acquire their regulatory power through member agreement or from their governments.

Alternative trading systems (ATS), which serve the same trading function as exchanges but have no regulatory function, are also known as **electronic communication networks (ECNs)** or **multilateral trading facilities (MTFs)**. ATS that do not reveal current client orders are known as *dark pools*.

Dealers facilitate trading by buying for or selling from their own inventory. Dealers provide liquidity in the market and profit primarily from the spread (difference) between the price at which they will buy (bid price) and the price at which they will sell (ask price) the security or other asset.

Some dealers also act as brokers. **Broker-dealers** have an inherent conflict of interest. As brokers, they should seek the best prices for their clients, but as dealers, their goal is to profit through prices or spreads. As a result, traders typically place limits on how their orders are filled when they transact with broker-dealers.

Dealers that trade with central banks when the banks buy or sell government securities in order to affect the money supply are referred to as **primary dealers**.

Securitizers

Securitizers pool large amounts of securities or other assets and then sell interests in the pool to other investors. The returns from the pool, net of the securitizer's fees, are passed through to the investors. By securitizing the assets, the securitizer creates a diversified pool of assets with more predictable cash flows than the individual assets in the pool. This creates liquidity in the assets because the ownership interests are more easily valued and traded. There are also economies of scale in the management costs of large pools of assets and potential benefits from the manager's selection of assets.

Assets that are often securitized include mortgages, car loans, credit card receivables, bank loans, and equipment leases. The primary benefit of securitization is to decrease the funding costs for the assets in the pool. A firm may set up a *special purpose vehicle (SPV)* or *special purpose entity (SPE)* to buy firm assets, which removes them from the firm's balance sheet and may increase their value by removing the risk that financial trouble at the firm will give other investors a claim to the assets' cash flows.

The cash flows from securitized assets can be segregated by risk. The different risk categories are called *tranches*. The senior tranches provide the most certain cash flows, while the junior tranches have greater risk.

Depository Institutions

Examples of **depository institutions** include banks, credit unions, and savings and loans. They pay interest on customer deposits and provide transaction services such as checking accounts. These financial intermediaries then make loans with the funds, which offer diversification benefits. The intermediaries have expertise in evaluating credit quality and managing the risk of a portfolio of loans of various types.

Other intermediaries, such as payday lenders and factoring companies, lend money to firms and individuals on the basis of their wages, accounts receivable, and other future cash flows. These intermediaries often finance the loans by issuing commercial paper or other debt securities.

Securities brokers provide loans to investors who purchase securities on margin. When this margin lending is to hedge funds and other institutions, the brokers are referred to as *prime brokers*.

The equity owners (stockholders) of banks, brokers, and other intermediaries absorb any loan losses before depositors and other lenders. The more equity capital an intermediary has, the less risk for depositors. Poorly capitalized intermediaries (those with less equity) have less incentive to reduce the risk of their loan portfolios because they have less capital at risk.

Insurance Companies

Insurance companies are intermediaries, in that they collect insurance premiums in return for providing risk reduction to the insured. The insurance firm can do this efficiently because it provides protection to a diversified pool of policyholders, whose risks of loss are typically uncorrelated. This provides more predictable losses and cash flows compared to a single insurance contract, in the same way that a bank's diversified portfolio of loans diversifies the risk of loan defaults.

Insurance firms also provide a benefit to investors by managing the risks inherent in insurance: moral hazard, adverse selection, and fraud. **Moral hazard** occurs because the insured may take more risks once he is protected against losses. **Adverse selection** occurs when those most likely to experience losses are the predominant buyers of insurance. In **fraud**, the insured purposely causes damage or claims fictitious losses so he can collect on his insurance policy.

Arbitrageurs

In its pure (riskless) form, **arbitrage** refers to buying an asset in one market and reselling it in another at a higher price. By doing so, arbitrageurs act as intermediaries, providing liquidity to participants in the market where the asset is purchased and transferring the asset to the market where it is sold.

In markets with good information, pure arbitrage is rare because traders will favor the markets with the best prices. More commonly, arbitrageurs try to exploit pricing differences for similar instruments. For example, a dealer who sells a call option will often also buy the stock because the call and stock price are highly correlated. Likewise, arbitrageurs will attempt to exploit discrepancies in the pricing of the call and stock. Many (risk) arbitrageurs use complex models for valuation of related securities and for risk control. Creating similar positions using different

assets is referred to as *replication*. This is also a form of intermediation because similar risks are traded in different forms and in different markets.

Clearinghouses and Custodians

Clearinghouses act as intermediaries between buyers and sellers in financial markets and provide:

- Escrow services (transferring cash and assets to the respective parties).
- Guarantees of contract completion.
- Assurance that margin traders have adequate capital.
- Limits on the aggregate net order quantity (buy orders minus sell orders) of members.

Through these activities, clearinghouses limit **counterparty risk**, the risk that the other party to a transaction will not fulfill its obligation. In some markets, the clearinghouse ensures only the trades of its member brokers and dealers, who, in turn, ensure the trades of their retail customers.

Custodians also improve market integrity by holding client securities and preventing their loss due to fraud or other events that affect the broker or investment manager.

MODULE 36.2: POSITIONS AND LEVERAGE



LOS 36.e: Compare positions an investor can take in an asset.

Video covering
this content is
available online.

An investor who owns an asset, or has the right or obligation under a contract to purchase an asset, is said to have a **long position**. A **short position** can result from borrowing an asset and selling it, with the obligation to replace the asset in the future (a short sale). The party to a contract who must sell or deliver an asset in the future is also said to have a short position. In general, investors who are long benefit from an increase in the price of an asset and those who are short benefit when the asset price declines.

Hedgers use short positions in one asset to hedge an existing risk from a long position in another asset that has returns that are strongly correlated with the returns of the asset shorted. For example, wheat farmers may take a short position in (i.e., sell) wheat futures contracts. If wheat prices fall, the resulting increase in the value of the short futures position offsets, partially or fully, the loss in the value of the farmer's crop.



PROFESSOR'S NOTE

As a rule of thumb, hedgers must "do in the futures market what they must do in the future." Thus, the farmer who must sell wheat in the future can reduce the risk from wheat price fluctuations by selling wheat futures.

The buyer of an option contract is said to be long the option. The seller is short the option and is said to have written the option. Note that an investor who is long (buys) a call option on an asset profits when the value of the underlying asset increases in value, while the party short the option has losses. A long position in a put option on an asset has the right to sell the asset at a

specified price and profits when the price of the underlying asset falls, while the party short the option has losses.

In swaps, each party is long one asset and short the other, so the designation of the long and short side is often arbitrary. Usually, however, the side that benefits from an increase in the quoted price or rate is referred to as the long side.

In a currency contract, each party is long one currency and short the other. For example, the buyer of a euro futures contract priced in dollars is long the euro and short the dollar.

Short Sales and Positions

In a **short sale**, the short seller (1) simultaneously borrows and sells securities through a broker, (2) must return the securities at the request of the lender or when the short sale is closed out, and (3) must keep a portion of the proceeds of the short sale on deposit with the broker. Short sellers hope to profit from a fall in the price of the security or asset sold short, buying at a lower price in the future in order to repay the loan of the asset originally sold at a higher price. The repayment of the borrowed security or other asset is referred to as "covering the short position."

In a short sale, the short seller must pay all dividends or interest that the lender would have received from the security that has been loaned to the short seller. These payments are called **payments-in-lieu** of dividends or interest. The short seller must also deposit the proceeds of the short sale as collateral to guarantee the eventual repurchase of the security. The broker then earns interest on these funds and may return a portion of this interest to the short seller at a rate referred to as the **short rebate rate**. The short rebate rate is usually only provided to institutional investors and is typically 0.1% less than overnight interest rates. If the security is difficult to borrow, the short rebate rate may be lower or negative. The difference between the interest earned on the proceeds from the short sale and the short rebate paid is the return to the lender of the securities. A short sale may also require the short seller to deposit additional margin in the form of cash or short-term riskless securities.

Leveraged Positions

The use of borrowed funds to purchase an asset results in a **leveraged position** and the investor is said to be using leverage. Investors who use leverage to buy securities by borrowing from their brokers are said to buy on **margin** and the borrowed funds are referred to as a **margin loan**. The interest rate paid on the funds is the **call money rate**, which is generally higher than the government bill rate. The call money rate is lower for larger investors with better collateral.

At the time of a new margin purchase, investors are required to provide a minimum amount of equity, referred to as the **initial margin requirement**. This requirement may be set by the government, exchange, clearinghouse, or broker. Lower risk in an investor's portfolio will often result in the broker lending more funds.

The use of leverage magnifies both the gains and losses from changes in the value of the underlying asset. The additional risk from the use of borrowed funds is referred to as risk from **financial leverage**.

LOS 36.f: Calculate and interpret the leverage ratio, the rate of return on a margin transaction, and the security price at which the investor would receive a margin call.

The **leverage ratio** of a margin investment is the value of the asset divided by the value of the equity position. For example, an investor who satisfies an initial margin requirement of 50% equity has a 2-to-1 leverage ratio so that a 10% increase (decrease) in the price of the asset results in a 20% increase (decrease) in the investor's equity amount.

EXAMPLE: Margin transaction

Given the following information:

Shares purchased	1,000
Purchase price per share	\$100
Annual dividend per share	\$2.00
Initial margin requirement	40%
Call money rate	4%
Commission per share	\$0.05
Stock price after one year	\$110

Calculate (1) the leverage ratio and (2) the investor's return on the margin transaction (return on equity) if the stock is sold at the end of one year.

Answer:

1. The leverage ratio = $1 / 0.40 = 2.5$.
2. The total purchase price is $1,000 \times \$100 = \$100,000$. The investor must post initial margin of $40\% \times \$100,000 = \$40,000$. The remaining $\$60,000$ is borrowed. The commission on the purchase is $1,000 \times \$0.05 = \50 . Thus, the total initial equity investment is $\$40,050$.

At the end of one year, the stock value is $1,000 \times \$110 = \$110,000$, for a gain of $\$9,950$.

Dividends received are $1,000 \times \$2.00 = \$2,000$. Interest paid is $\$60,000 \times 4\% = \$2,400$. The commission on the sale is $1,000 \times \$0.05 = \50 .

The gain on the transaction in one year is $\$9,950 + \$2,000 - \$2,400 - \$50 = \$9,500$. The return on the equity investment is $\$9,500 / \$40,050 = 23.72\%$. The investor's net return is less than the asset total return (10% price appreciation + 2% dividend = 12%) multiplied by the leverage ratio ($12\% \times 2.5 = 30\%$) because of the loan interest and commissions.

We can also solve for the return on the margin transaction with the cash flow functions on a financial calculator. The initial cash outflow is the $\$40,000$ initial margin + $\$50$ purchase commission = $\$40,050$. The inflow after one year is the $\$110,000$ stock value + $\$2,000$ dividends - $\$60,000$ margin repayment - $\$2,400$ margin interest - $\$50$ sale commission = $\$49,550$. Using the cash flow functions: $CF_0 = -40,050$; $CF_1 = 49,550$; $CPT IRR = 23.72\%$.

To ensure that the loan is covered by the value of the asset, an investor must maintain a minimum equity percentage, called the **maintenance margin requirement**, in the account. This minimum is typically 25% of the current position value, but brokers may require a greater minimum equity percentage for volatile stocks.

If the percentage of equity in a margin account falls below the maintenance margin requirement, the investor will receive a **margin call**, a request to bring the equity percentage in the account back up to the maintenance margin percentage. An investor can satisfy this request by depositing additional funds or depositing other unmargined securities that will bring the equity percentage up to the minimum requirement. If the investor does not meet the margin call, the broker must sell the position.

The stock price which results in a margin call can be calculated by using the following formula:

$$\text{margin call price} = P_0 \left(\frac{1 - \text{initial margin}}{1 - \text{maintenance margin}} \right)$$

where:

P_0 = initial purchase price

EXAMPLE: Margin call price

If an investor purchases a stock for \$40 per share with an initial margin requirement of 50% and the maintenance margin requirement is 25%, at what price will the investor get a margin call?

Answer:

$$\frac{\$40(1 - 0.5)}{1 - 0.25} = \$26.67$$

A margin call is triggered at a price below \$26.67.

In a short sale, the investor must deposit initial margin equal to a percentage of the value of the shares sold short to protect the broker in case the share price increases. An increase in the share price can decrease the margin percentage below the maintenance margin percentage and generate a margin call.



MODULE QUIZ 36.1, 36.2

1. An investor who buys a government bond from a dealer's inventory is said to obtain:
 - A. a real asset in a primary market transaction.
 - B. a financial asset in a primary market transaction.
 - C. a financial asset in a secondary market transaction.
2. Daniel Ferramosco is concerned that a long-term bond he holds might default. He therefore buys a contract that will compensate him in the case of default. What type of contract does he hold?
 - A. Physical derivative contract.
 - B. Primary derivative contract.
 - C. Financial derivative contract.
3. A financial intermediary buys a stock and then resells it a few days later at a higher price. Which intermediary would this *most likely* describe?
 - A. Broker.
 - B. Dealer.
 - C. Arbitrageur.
4. Which of the following is *most* similar to a short position in the underlying asset?
 - A. Buying a put.
 - B. Writing a put.
 - C. Buying a call.

5. An investor buys 1,000 shares of a stock on margin at a price of \$50 per share. The initial margin requirement is 40% and the margin lending rate is 3%. The investor's broker charges a commission of \$0.01 per share on purchases and sales. The stock pays an annual dividend of \$0.30 per share. One year later, the investor sells the 1,000 shares at a price of \$56 per share. The investor's rate of return is *closest* to:

- A. 12%.
- B. 27%.
- C. 36%.

MODULE 36.3: ORDER EXECUTION AND VALIDITY

LOS 36.g: Compare execution, validity, and clearing instructions.



LOS 36.h: Compare market orders with limit orders.

Video covering
this content is
available online.

Securities dealers provide prices at which they will buy and sell shares. The **bid price** is the price at which a dealer will buy a security. The **ask or offer price** is the price at which a dealer will sell a security. The difference between the bid and ask prices is referred to as the **bid-ask spread** and is the source of a dealer's compensation. The bid and ask are quoted for specific trade sizes (**bid size** and **ask size**).



PROFESSOR'S NOTE

Calculations with bid and ask prices are unlikely to appear on the Level I exam but they do appear at Level II. If you need to work with bid and ask prices, just remember that the price you get will be the one that is *worse for you*.

- Securities: If you are buying, you must pay the higher price. If you are selling, you only receive the lower price.
- Currencies: The bid or ask price you get is the one that gives you less of the currency you are acquiring. This works regardless of which way the exchange rate is quoted.

The quotation in the market is the highest dealer bid and lowest dealer ask from among all dealers in a particular security. More liquid securities have market quotations with bid-ask spreads that are lower (as a percentage of share price) and therefore have lower transaction costs for investors. Traders who post bids and offers are said to *make a market*, while those who trade with them at posted prices are said to *take the market*.

When investors want to buy or sell, they must enter orders that specify the size of the trade and whether to buy or sell. The order can also include *execution instructions* that specify how to trade, *validity instructions* that specify when the order can be filled, and *clearing instructions* that specify how to settle the trade.

Execution Instructions

The most common orders, in terms of execution instructions, are market or limit orders. A **market order** instructs the broker to execute the trade immediately at the best possible price. A **limit order** places a minimum execution price on sell orders and a maximum execution price

on buy orders. For example, a buy order with a limit of \$6 will be executed immediately as long as the shares can be purchased for \$6 or less.

A market order is often appropriate when the trader wants to execute quickly, as when the trader has information she believes is not yet reflected in market prices. The disadvantage of market orders is that they may execute at unfavorable prices, especially if the security has low trading volume relative to the order size. A market buy order may execute at a high price or a market sell order may execute at a low price. Executing at an unfavorable price represents a concession by the trader for immediate liquidity. Unfortunately, these price concessions are unpredictable.

To avoid price execution uncertainty, a trader can place a limit order instead of the market order. The disadvantage of the limit order is that it might not be filled. For example, if a trader places a limit buy order of \$50 and no one is willing to sell at \$50, the order will not be filled. Furthermore, if the stock price rises over time, the trader misses out on the gains.

A limit buy order above the best ask or a limit sell order below the best bid are said to be *marketable* or *aggressively priced* because at least part of the order is likely to execute immediately. If the limit price is between the best bid and the best ask, a limit order is said to be *making a new market* or *inside the market*. Limit orders waiting to execute are called **standing limit orders**.

A limit buy order at the best bid or a limit sell order at the best ask are said to *make the market*. Again, the order might not be filled. A buy order with a limit price below the best bid, or a sell order with a limit price above the best ask, is said to be *behind the market*. It will likely not execute until security prices move toward the limit price. A limit buy order with a price considerably lower than the best bid, or a limit sell order with a price significantly higher than the best ask, is said to be *far from the market*.

Other execution instructions concern the volume of the trade. **All-or-nothing orders** execute only if the whole order can be filled. Orders can specify the minimum size of a trade, which is beneficial when trading costs depend on the number of executed trades rather than the size of the order.

Trade visibility can also be specified. **Hidden orders** are those for which only the broker or exchange knows the trade size. These are useful for investors that have a large amount to trade and do not want to reveal their intentions. Traders can also specify **display size**, where some of the trade is visible to the market, but the rest is not. These are also referred to as **iceberg orders** because part of most of the order is hidden from view. They allow the investor to advertise some of the trade, with the rest of the trade potentially executed once the visible part has executed. Sometimes entering trades for part of the position the trader wishes to establish is a way to estimate the liquidity of, or the buying interest in, the security in question.

Validity Instructions

Validity instructions specify when an order should be executed. Most orders are **day orders**, meaning they expire if unfilled by the end of the trading day. **Good til canceled orders** last until they are filled. **Immediate-or-cancel** orders are canceled unless they can be filled immediately. They are also known as **fill-or-kill** orders. **Good-on-close** orders are only filled at the end of the trading day. If they are market orders, they are referred to as **market-on-close**

orders. These are often used by mutual funds because their portfolios are valued using closing prices. There are also **good-on-open** orders.

Stop orders are those that are not executed unless the stop price has been met. They are often referred to as **stop loss orders** because they can be used to prevent losses or to protect profits. Suppose an investor purchases a stock for \$50. If the investor wants to sell out of the position if the price falls 10% to \$45, he can enter a **stop-sell order** at \$45. If the stock trades down to \$45 or lower, this *triggers* a market order to sell. There is no guarantee that the order will execute at \$45, and a rapidly falling stock could be sold at a price significantly lower than \$45.

A **stop-buy** is entered with at stop (trigger) above the current market price. There are two primary reasons a trader would enter a stop-buy order. (1) A trader with a short position could attempt to limit losses from an increasing stock price with a stop-buy order. (2) It is often said, "You don't get paid for being right until the market agrees with you." With this in mind, an investor who believes a stock is undervalued, but does not wish to own it until there are signs that market participants are being convinced of this undervaluation, may place a stop-buy order at a price some specific percentage above the current price.

Note that stop orders reinforce market momentum. Stop-sell orders execute when market prices are falling, and stop-buy orders execute when the market is rising. Execution prices for stop orders are therefore often unfavorable.

EXAMPLE: Using stop orders

Raymond Flowers believes that the shares of Acme Corp. that he owns are overvalued currently but knows that stocks often continue to increase above their intrinsic values for some time before correcting. What type of order should Flowers place if he wants to sell his shares when the price begins to fall a significant amount?

Answer:

Flowers should enter a good til canceled stop-sell order at a price some percentage below the current level. If, for example, the shares are trading at 40, he could enter a stop-sell order at 36, 10% below the current level. Investors sometimes move these stops up as a stock continues to increase in price. In response to a price increase to 42, Flowers might move his stop-sell order up to 37.80, 10% below the new price. Note that a limit order to sell with a limit price below the current market price would likely execute immediately.

Clearing Instructions

Clearing instructions tell the trader how to clear and settle a trade. They are usually standing instructions and not attached to an order. Retail trades are typically cleared and settled by the broker, whereas institutional trades may be settled by a custodian or another broker, which might be the trader's prime broker. Using two brokers allows the investor to keep one broker as her prime broker for margin and custodial services while using a variety of other brokers for specialized execution.

One important clearing instruction is whether a sell order is a short sale or long sale. In the former, the broker must confirm that the security can be borrowed and in the latter, that the security can be delivered.

LOS 36.i: Define primary and secondary markets and explain how secondary markets support primary markets.

Primary capital markets refer to the sale of newly issued securities. New equity issues involve either:

- New shares issued by firms whose shares are already trading in the marketplace. These issues are called **seasoned offerings** or **secondary issues**.
- First-time issues by firms whose shares are not currently publicly traded. These are called **initial public offerings (IPOs)**.

Secondary financial markets are where securities trade after their initial issuance. Placing a buy order on the London Stock Exchange is an order in the secondary market and will result in purchase of existing shares from their current owner.

Primary Market: Public Offerings

Corporate stock or bond issues are almost always sold with the assistance of an investment banking firm. The investment bank finds investors who agree to buy part of the issue. These are not actual orders but are referred to as **indications of interest**. When the number of shares covered by indications of interest are greater (less) than the number of shares to be offered, the offering price may be adjusted upward (downward). This process of gathering indications of interest is referred to as **book building**. In London, the book builder is referred to as the **book runner**. In Europe, an **accelerated book build** occurs when securities must be issued quickly. To build a book, the investment bank disseminates information about the firm's financials and prospects. The issuer must also make disclosures including how the funds will be used.

The most common way an investment bank assists with a security issuance is through an **underwritten offering**. Here, the investment bank agrees to purchase the entire issue at a price that is negotiated between the issuer and bank. If the issue is undersubscribed, the investment bank must buy the unsold portion. In the case of an IPO, the investment bank also agrees to make a market in the stock for a period after the issuance to provide price support for the issue.

An investment bank can also agree to distribute shares of an IPO on a **best efforts** basis, rather than agreeing to purchase the whole issue. If the issue is undersubscribed, the bank is not obligated to buy the unsold portion.

Note that investment banks have a conflict of interest in an underwritten offer. As the issuer's agents, they should set the price high to raise the most funds for the issuer. But, as underwriters, they would prefer that the price be set low enough that the whole issue sells. This also allows them to allocate portions of an undervalued IPO to their clients. This results in IPOs typically being underpriced. Issuers also could have an interest in underpricing the IPO because of the negative publicity when an undersubscribed IPO initially trades at a price below the IPO price investors pay. An IPO that is oversubscribed and has the expectation of trading significantly above its IPO price is referred to as a hot issue.

Primary Market: Private Placements and Other Transactions

In a **private placement**, securities are sold directly to qualified investors, typically with the assistance of an investment bank. Qualified investors are those with substantial wealth and investment knowledge. Private placements do not require the issuer to disclose as much information as they must when the securities are being sold to the public. The issuance costs are less with a private placement and the offer price is also lower because the securities cannot be resold in public markets, making them less valuable than shares registered for public trading.

In a **shelf registration**, a firm makes its public disclosures as in a regular offering but then issues the registered securities over time when it needs capital and when the markets are favorable.

A **dividend reinvestment plan (DRP or DRIP)** allows existing shareholders to use their dividends to buy new shares from the firm at a slight discount.

In a **rights offering**, existing shareholders are given the right to buy new shares at a discount to the current market price. Shareholders tend to dislike rights offerings because their ownership is diluted unless they exercise their rights and buy the additional shares. However, rights can be traded separately from the shares themselves in some circumstances.

In addition to firms issuing securities, governments issue short-term and long-term debt, either by auction or through investment banks.

Importance of the Secondary Market

Secondary markets are important because they provide liquidity and price/value information. Liquid markets are those in which a security can be sold quickly without incurring a discount from the current price. The better the secondary market, the easier it is for firms to raise external capital in the primary market, which results in a lower cost of capital for firms with shares that have adequate liquidity.

LOS 36.j: Describe how securities, contracts, and currencies are traded in quote-driven, order-driven, and brokered markets.

The trading of securities in the secondary market has encouraged the development of market structures to facilitate trading. Trading can be examined according to when securities are traded and how they are traded.

Securities markets may be structured as call markets or continuous markets. In **call markets**, the stock is only traded at specific times. Call markets are potentially very liquid when in session because all traders are present, but they are obviously illiquid between sessions. In a call market, all trades, bids, and asks are declared, and then one negotiated price is set that clears the market for the stock. This method is used in smaller markets but is also used to set opening prices and prices after trading halts on major exchanges.

In **continuous markets**, trades occur at any time the market is open. The price is set by either the auction process or by dealer bid-ask quotes.

Market Structures

There are three main categories of securities markets: *quote-driven markets* where investors trade with dealers, *order-driven markets* where rules are used to match buyers and sellers, and *brokered markets* where investors use brokers to locate a counterparty to a trade.

Quote-Driven Markets

In **quote-driven markets**, traders transact with dealers (market makers) who post bid and ask prices. Dealers maintain an inventory of securities. Quote-driven markets are thus sometimes called **dealer markets**, **price-driven markets**, or **over-the-counter markets**. Most securities other than stocks trade in quote-driven markets. Trading often takes place electronically.

Order-Driven Markets

In **order-driven markets**, orders are executed using trading rules, which are necessary because traders are usually anonymous. Exchanges and automated trading systems are examples of order-driven markets. Two sets of rules are used in these markets: order matching rules and trade pricing rules.

Order matching rules establish an *order precedence hierarchy*. **Price priority** is one criteria, where the trades given highest priority are those at the highest bid (buy) and lowest ask (sell). If orders are at the same prices, a **secondary precedence rule** gives priority to non-hidden orders and earliest arriving orders. These rules encourage traders to price their trades aggressively, display their entire orders, and trade earlier, thereby improving liquidity.

After orders are created using order matching rules, **trade pricing rules** are used to determine the price. Under the *uniform pricing rule*, all orders trade at the same price, which is the price that results in the highest volume of trading. The *discriminatory pricing rule* uses the limit price of the order that arrived first as the trade price.

In an electronic crossing network, the typical trader is an institution. Orders are batched together and crossed (matched) at fixed points in time during the day at the average of the bid and ask quotes from the exchange where the stock primarily trades. This pricing rule is referred to as the *derivative pricing rule* because it is derived from the security's main market. The price is not determined by orders in the crossing network.

Brokered Markets

In **brokered markets**, brokers find the counterparty in order to execute a trade. This service is especially valuable when the trader has a security that is unique or illiquid. Examples are large blocks of stock, real estate, and artwork. Dealers typically do not carry an inventory of these assets and there are too few trades for these assets to trade in order-driven markets.

Market Information

A market is said to be **pre-trade transparent** if investors can obtain pre-trade information regarding quotes and orders. A market is **post-trade transparent** if investors can obtain post-trade information regarding completed trade prices and sizes.

Buy-side traders value transparency because it allows them to better understand security values and trading costs. Dealers, on the other hand, prefer opaque markets because this

provides them with an informational advantage over traders who trade less frequently in the security. Transactions costs and bid-ask spreads are larger in opaque markets.

LOS 36.k: Describe characteristics of a well-functioning financial system.

A well-functioning financial system allows entities to achieve their purposes. More specifically, **complete markets** fulfill the following:

- Investors can save for the future at fair rates of return.
- Creditworthy borrowers can obtain funds.
- Hedgers can manage their risks.
- Traders can obtain the currencies, commodities, and other assets they need.

If a market can perform these functions at low trading costs (including commissions, bid-ask spreads, and price impacts), it is said to be **operationally efficient**. If security prices reflect all the information associated with fundamental value in a timely fashion, then the financial system is **informationally efficient**. A well-functioning financial system has complete markets that are operationally and informationally efficient, with prices that reflect fundamental values.

A well-functioning financial system has financial intermediaries that:

- Organize trading venues, including exchanges, brokerages, and alternative trading systems.
- Supply liquidity.
- Securitize assets so that borrowers can obtain funds inexpensively.
- Manage banks that use depositor capital to fund borrowers.
- Manage insurance firms that pool unrelated risks.
- Manage investment advisory services that assist investors with asset management inexpensively.
- Provide clearinghouses that settle trades.
- Manage depositories that provide for asset safety.

The benefits of a well-functioning financial system are tremendous. Savers can fund entrepreneurs who need capital to fund new companies. Company risks can be shared so that risky companies can be funded. These benefits are enhanced because the transactions can occur among strangers, widening the opportunities for capital formation and risk sharing in the economy.

Furthermore, in informationally efficiently markets, capital is allocated to its most productive use. That is, they are **allocationally efficient**. Informational efficiency is brought about by traders who bid prices up and down in response to new information that changes estimates of securities' fundamental values. If markets are operationally efficient, security prices will be more informationally efficient because low trading costs encourage trading based on new information. The existence of accounting standards and financial reporting requirements also reduces the costs of obtaining information and increases security values.

LOS 36.l: Describe objectives of market regulation.

Without market regulation, many problems could persist in financial markets:

- *Fraud and theft*: In complex financial markets, the potential for theft and fraud increases because investment managers and others can take advantage of unsophisticated investors. Furthermore, if returns are often random, it is difficult for investors to determine if their agents (e.g., investment managers and brokers) are performing well.
- *Insider trading*: If investors believe traders with inside information will exploit them, they will exit the market and liquidity will be reduced.
- *Costly information*: If obtaining information is relatively expensive, markets will not be as informationally efficient and investors will not invest as much.
- *Defaults*: Parties might not honor their obligations in markets.

To solve these problems, market regulation should:

- Protect unsophisticated investors so that trust in the markets is preserved.
- Require minimum standards of competency and make it easier for investors to evaluate performance. The CFA Program and the Global Investment Performance Standards are part of this effort.
- Prevent insiders from exploiting other investors.
- Require common financial reporting requirements (e.g., those of the International Accounting Standards Board) so that information gathering is less expensive.
- Require minimum levels of capital so that market participants will be able to honor their long-term commitments. This is especially important for insurance companies and pension funds that individuals depend on for their financial future. With capital at stake, market participants have more incentive to be careful about the risks they take.

Regulation can be provided by governments as well as industry groups. For example, most exchanges, clearinghouses, and dealer trade organizations are self-regulating organizations (SROs), meaning that they regulate their members. Governments sometimes delegate regulatory authority to SROs.

When they fail to address the problems mentioned previously, financial markets do not function well. Liquidity declines, firms shun risky projects, new ideas go unfunded, and economic growth slows.



MODULE QUIZ 36.3

1. A stock is selling at \$50. An investor's valuation model estimates its intrinsic value to be \$40. Based on her estimate, she would *most likely* place:
 - A. a short-sale order.
 - B. a stop order to buy.
 - C. a market order to buy.
2. Which of the following limit buy orders would be the *most likely* to go unexecuted?
 - A. A marketable order.
 - B. An order behind the market.
 - C. An order making a new market.
3. New issues of securities are transactions in:
 - A. the primary market.

- B. the secondary market.
 - C. the seasoned market.
4. In which of the following types of markets do stocks trade any time the market is open?
- A. Exchange markets.
 - B. Call markets.
 - C. Continuous markets.
5. A market is said to be informationally efficient if it features:
- A. market prices that reflect all available information about the value of the securities traded.
 - B. timely and accurate information about current supply and demand conditions.
 - C. many buyers and sellers that are willing to trade at prices above and below the prevailing market price.
6. Which of the following would *least likely* be an objective of market regulation?
- A. Reduce burdensome accounting standards.
 - B. Make it easier for investors to evaluate performance.
 - C. Prevent investors from using inside information in securities trading.

KEY CONCEPTS

LOS 36.a

The three main functions of the financial system are to:

1. Allow entities to save, borrow, issue equity capital, manage risks, exchange assets, and utilize information.
2. Determine the return that equates aggregate savings and borrowing.
3. Allocate capital efficiently.

LOS 36.b

Assets and markets can be classified as:

- Financial assets (e.g., securities, currencies, derivatives) versus real assets (e.g., real estate, equipment).
- Debt securities versus equity securities.
- Public securities that trade on exchanges or through dealers versus private securities.
- Physical derivative contracts (e.g., on grains or metals) versus financial derivative contracts (e.g., on bonds or equity indexes).
- Spot versus future delivery markets.
- Primary markets (issuance of new securities) versus secondary markets (trading of previously issued securities).
- Money markets (short-term debt instruments) versus capital markets (longer-term debt instruments and equities).
- Traditional investment markets (bonds, stocks) versus alternative investment markets (e.g., real estate, hedge funds, fine art).

LOS 36.c

The major types of assets are securities, currencies, contracts, commodities, and real assets.

Securities include fixed income (e.g., bonds, notes, commercial paper), equity (common stock, preferred stock, warrants), and pooled investment vehicles (mutual funds, exchange-traded funds, hedge funds, asset-backed securities).

Contracts include futures, forwards, options, swaps, and insurance contracts.

Commodities include agricultural products, industrial and precious metals, and energy products and are traded in spot, forward, and futures markets.

Most national currencies are traded in spot markets and some are also traded in forward and futures markets.

LOS 36.d

Financial intermediaries perform the following roles:

- Brokers, exchanges, and alternative trading systems connect buyers and sellers of the same security at the same location and time. They provide a centralized location for trading.
- Dealers match buyers and sellers of the same security at different points in time.
- Arbitrageurs connect buyers and sellers of the same security at the same time but in different venues. They also connect buyers and sellers of non-identical securities of similar risk.
- Securitizers and depository institutions package assets into a diversified pool and sell interests in it. Investors obtain greater liquidity and choose their desired risk level.
- Insurance companies create a diversified pool of risks and manage the risk inherent in providing insurance.
- Clearinghouses reduce counterparty risk and promote market integrity.

LOS 36.e

A long position in an asset represents current or future ownership. A long position benefits when the asset increases in value.

A short position represents an agreement to sell or deliver an asset or results from borrowing an asset and selling it (i.e., a short sale). A short position benefits when the asset decreases in value.

When an investor buys a security by borrowing from a broker, the investor is said to buy on margin and has a leveraged position. The risk of investing borrowed funds is referred to as financial leverage. More leverage results in greater risk.

LOS 36.f

The leverage ratio is the value of the asset divided by the value of the equity position. Higher leverage ratios indicate greater risk.

The return on a margin transaction is the increase in the value of the position after deducting selling commissions and interest charges, divided by the amount of funds initially invested, including purchase commissions.

The maintenance margin is the minimum percentage of equity that a margin investor is required to maintain in his account. If the investor's equity falls below the maintenance margin, the investor will receive a margin call. The stock price that will result in a margin call is:

$$\text{margin call price} = P_0 \left(\frac{1 - \text{initial margin}}{1 - \text{maintenance margin}} \right)$$

where:

P_0 = initial purchase price

LOS 36.g

Execution instructions specify how to trade. Market orders and limit orders are examples of execution instructions.

Validity instructions specify when an order can be filled. Day orders, good til canceled orders, and stop orders are examples of validity instructions.

Clearing instructions specify how to settle a trade.

LOS 36.h

A market order is an order to execute the trade immediately at the best possible price. A market order is appropriate when the trader wants to execute a transaction quickly. The disadvantage of a market order is that it may execute at an unfavorable price.

A limit order is an order to trade at the best possible price, subject to the price satisfying the limit condition. A limit order avoids price execution uncertainty. The disadvantage of a limit order is that it may not be filled. A buy (sell) order with a limit of \$18 will only be executed if the security can be bought (sold) at a price of \$18 or less (more).

LOS 36.i

New issues of securities are sold in primary capital markets. Secondary financial markets are where securities trade after their initial issuance.

In an underwritten offering, the investment bank guarantees that the issue will be sold at a price that is negotiated between the issuer and bank. In a best efforts offering, the bank acts only as a broker.

In a private placement, a firm sells securities directly to qualified investors, without the disclosures of a public offering.

A liquid secondary market makes it easier for firms to raise external capital in the primary market, which results in a lower cost of capital for firms.

LOS 36.j

There are three main categories of securities markets:

1. Quote-driven markets: Investors trade with dealers that maintain inventories of securities, currencies, or contracts.
2. Order-driven markets: Order-matching and trade-pricing rules are used to match the orders of buyers and sellers.
3. Brokered markets: Brokers locate a counterparty to take the other side of a buy or sell order.

In call markets, securities are only traded at specific times. In continuous markets, trades occur at any time the market is open.

LOS 36.k

A well-functioning financial system has the following characteristics:

- Complete markets: Savers receive a return, borrowers can obtain capital, hedgers can manage risks, and traders can acquire needed assets.
- Operational efficiency: Trading costs are low.
- Informational efficiency: Prices reflect fundamental information quickly.
- Allocational efficiency: Capital is directed to its highest valued use.

LOS 36.1

The objectives of market regulation are to:

- Protect unsophisticated investors.
- Establish minimum standards of competency.
- Help investors to evaluate performance.
- Prevent insiders from exploiting other investors.
- Promote common financial reporting requirements so that information gathering is less expensive.
- Require minimum levels of capital so that market participants will be able to honor their commitments and be more careful about their risks.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 36.1, 36.2

- C** Bonds are financial assets. Real assets are physical things such as a commodity or a factory. Buying a bond from a dealer is a secondary market transaction. A primary market transaction is an issuance of securities by an entity that is raising funds. (Module 36.1, LOS 36.b)
- C** Daniel holds a derivative contract that has a value determined by another financial contract; in this case, the long-term bond. (Module 36.1, LOS 36.c)
- B** This situation best describes a dealer. A dealer buys an asset for its inventory in the hopes of reselling it later at a higher price. Brokers stand between buyers and sellers of the same security at the same location and time. Arbitrageurs trade in the same security simultaneously in different markets. (Module 36.1, LOS 36.d)
- A** Buying a put is most similar to a short position in the underlying asset because the put increases in value if the underlying asset value decreases. The writer of a put and the holder of a call have a long exposure to the underlying asset because their positions increase in value if the underlying asset value increases. (Module 36.2, LOS 36.e)
- B** The total purchase price is $1,000 \times \$50 = \$50,000$. The investor must post initial margin of $40\% \times \$50,000 = \$20,000$. The remaining $\$30,000$ is borrowed. The commission on the purchase is $1,000 \times \$0.01 = \10 . Thus, the initial equity investment is $\$20,010$.
In one year, the sales price is $1,000 \times \$56 = \$56,000$. Dividends received are $1,000 \times \$0.30 = \300 . Interest paid is $\$30,000 \times 3\% = \900 . The commission on the sale is $1,000 \times \$0.01 = \10 . Thus, the ending value is $\$56,000 - \$30,000 + \$300 - \$900 - \$10 = \$25,390$.

The return on the equity investment is $\$25,390 / \$20,010 - 1 = 26.89\%$. (Module 36.2, LOS 36.f)

Module Quiz 36.3

1. **A** If the investor believes the stock is overvalued in the market, the investor should place a short-sale order, which would be profitable if the stock moves toward her value estimate. (LOS 36.g, 36.h)
2. **B** A behind-the-market limit order would be least likely executed. In the case of a buy, the limit buy order price is below the best bid. It will likely not execute until security prices decline. A marketable buy order is the most likely to trade because it is close to the best ask price. In an order that is making a new market or inside the market, the limit buy order price is between the best bid and ask. (LOS 36.h)
3. **A** The primary market refers to the market for newly issued securities. (LOS 36.i)
4. **C** Continuous markets are defined as markets where stocks can trade any time the market is open. Some exchange markets are call markets where orders are accumulated and executed at specific times. (LOS 36.j)
5. **A** Informational efficiency means the prevailing price reflects all available information about the value of the asset, and the price reacts quickly to new information. (LOS 36.k)
6. **A** Market regulation should require financial reporting standards so that information gathering is less expensive and the informational efficiency of the markets is enhanced. (LOS 36.l)

READING 37

SECURITY MARKET INDEXES

EXAM FOCUS

Security market indexes are used to measure the performance of markets and investment managers. Understand the construction, calculation, and weaknesses of price-weighted, market capitalization-weighted, and equal-weighted indexes. Be familiar with the various security indexes and their potential weaknesses.

MODULE 37.1: INDEX WEIGHTING METHODS



LOS 37.a: Describe a security market index.

Video covering
this content is
available online.

A **security market index** is used to represent the performance of an asset class, security market, or segment of a market. They are usually created as portfolios of individual securities, which are referred to as the **constituent securities** of the index. An index has a numerical value that is calculated from the market prices (actual when available, or estimated) of its constituent securities at a point in time. An index return is the percentage change in the index's value over a period of time.

LOS 37.b: Calculate and interpret the value, price return, and total return of an index.

An index return may be calculated using a **price index** or a **return index**. A price index uses only the prices of the constituent securities in the return calculation. A rate of return that is calculated based on a price index is referred to as a **price return**.

A return index includes both prices and income from the constituent securities. A rate of return that is calculated based on a return index is called a **total return**. If the assets in an index produce interim cash flows such as dividends or interest payments, the total return will be greater than the price return.

Once returns are calculated for each period, they then can be compounded together to arrive at the return for the measurement period:

$$R_P = (1 + R_{S1})(1 + R_{S2})(1 + R_{S3})(1 + R_{S4}) \dots (1 + R_{Sk}) - 1$$

where:

R_P = portfolio return during the measurement period

k = total number of subperiods

R_{Sk} = portfolio return during the subperiod k

For example, if the returns for the first two periods were 0.50% and 1.04%, they would be geometrically linked to produce 1.55%:

$$R_P = (1 + R_{S1})(1 + R_{S2}) - 1 = (1.005)(1.0104) - 1 = 0.0155 \text{ or } 1.55\%$$

If the starting index value is 100, its value after two periods would be $100 \times 1.0155 = 101.55$.

LOS 37.c: Describe the choices and issues in index construction and management.

Index providers must make several decisions:

- What is the *target market* the index is intended to measure?
- Which securities from the target market should be included?
- How should the securities be weighted in the index?
- How often should the index be rebalanced?
- When should the selection and weighting of securities be re-examined?

The target market may be defined very broadly (e.g., stocks in the United States) or narrowly (e.g., small-cap value stocks in the United States). It may also be defined by geographic region or by economic sector (e.g., cyclical stocks). The constituent stocks in the index could be all the stocks in that market or just a representative sample. The selection process may be determined by an objective rule or subjectively by a committee.

LOS 37.d: Compare the different weighting methods used in index construction.

Weighting schemes for stock indexes include price weighting, equal weighting, market capitalization weighting, float-adjusted market capitalization weighting, and fundamental weighting.

A **price-weighted index** is simply an arithmetic average of the prices of the securities included in the index. The divisor of a price-weighted index is adjusted for stock splits and changes in the composition of the index when securities are added or deleted, such that the index value is unaffected by such changes.

The advantage of a price-weighted index is that its computation is simple. One disadvantage is that a given percentage change in the price of a higher priced stock has a greater impact on the index's value than does an equal percentage change in the price of a lower priced stock. Put another way, higher priced stocks have more weight in the calculation of a price-weighted index. Additionally, a stock's weight in the index going forward changes if the firm splits its stock, repurchases stock, or issues stock dividends, as all of these actions will affect the price of the stock and therefore its weight in the index. A portfolio that has an equal number of shares in each of the constituent stocks will have price returns (ignoring dividends) that will match the returns of a price-weighted index.

Two major price-weighted indexes are the Dow Jones Industrial Average (DJIA) and the Nikkei Dow Jones Stock Average. The DJIA is a price-weighted index based on 30 U.S. stocks. The Nikkei

Dow is constructed from the prices of 225 stocks that trade in the first section of the Tokyo Stock Exchange.

An **equal-weighted index** is calculated as the arithmetic average return of the index stocks and, for a given time period, would be matched by the returns on a portfolio that had equal dollar amounts invested in each index stock. As with a price-weighted index, an advantage of an equal-weighted index is its simplicity.

One complication with an equal-weighted index return is that a matching portfolio would have to be adjusted periodically (rebalanced) as prices change so that the values of all security positions are made equal each period. The portfolio rebalancing required to match the performance of an equal-weighted index creates high transactions costs that would decrease portfolio returns.

Another concern with an equal-weighted index is that the weights placed on the returns of the securities of smaller capitalization firms are greater than their proportions of the overall market value of the index stocks. Conversely, the weights on the returns of large capitalization firms in the index are smaller than their proportions of the overall market value of the index stocks.

The Value Line Composite Average and the Financial Times Ordinary Share Index are well-known examples of equal-weighted indexes.

A **market capitalization-weighted index** (or **value-weighted index**) has weights based on the market capitalization of each index stock (current stock price multiplied by the number of shares outstanding) as a proportion of the total market capitalization of all the stocks in the index. A market capitalization-weighted index return can be matched with a portfolio in which the value of each security position in the portfolio is the same proportion of the total portfolio value as the proportion of that security's market capitalization to the total market capitalization of all of the securities included in the index. This weighting method more closely represents changes in aggregate investor wealth than price weighting. Because the weight of an index stock is based on its market capitalization, a market capitalization-weighted index does not need to be adjusted when a stock splits or pays a stock dividend.

An alternative to using a firm's market capitalization to calculate its weight in an index is to use its **market float**. A firm's market float is the total value of the shares that are actually available to the investing public and excludes the value of shares held by controlling stockholders because they are unlikely to sell their shares. For example, the float for Microsoft would exclude shares owned by Bill Gates and Paul Allen (the founders) and those of certain other large shareholders as well. The market float is often calculated excluding those shares held by corporations or governments as well. Sometimes the market float calculation excludes shares that are not available to foreign buyers and is then referred to as the **free float**. The reason for this is to better match the index weights of stocks to their proportions of the total value of all the shares of index stocks that are actually available to investors.

A **float-adjusted market capitalization-weighted index** is constructed like a market capitalization-weighted index. The weights, however, are based on the proportionate value of each firm's shares that are available to investors to the total market value of the shares of index stocks that are available to investors. Firms with relatively large percentages of their shares

held by controlling stockholders will have less weight than they have in an unadjusted market-capitalization index.

The advantage of market capitalization-weighted indexes of either type is that index security weights represent proportions of total market value. The primary disadvantage of value-weighted indexes is that the relative impact of a stock's return on the index increases as its price rises and decreases as its price falls. This means that stocks that are possibly overvalued are given disproportionately high weights in the index and stocks that are possibly undervalued are given disproportionately low weights. Holding a portfolio that tracks a value-weighted index is, therefore, similar to following a momentum strategy, under which the most successful stocks are given the greatest weights and poor performing stocks are underweighted.

The Standard and Poor's 500 (S&P 500) Index Composite is an example of a market capitalization-weighted index.

An index that uses **fundamental weighting** uses weights based on firm fundamentals, such as earnings, dividends, or cash flow. In contrast to market capitalization index weights, these weights are unaffected by the share prices of the index stocks (although related to them over the long term). Fundamental weights can be based on a single measure or some combination of fundamental measures.

An advantage of a fundamental-weighted index is that it avoids the bias of market capitalization-weighted indexes toward the performance of the shares of overvalued firms and away from the performance of the shares of undervalued firms. A fundamental-weighted index will actually have a value tilt, overweighting firms with high value-based metrics such as book-to-market ratios or earnings yields. Note that a firm with a high earnings yield (total earnings to total market value) relative to other index firms will by construction have a higher weight in an earnings-weighted index because, among index stocks, its earnings are high relative to its market value.

LOS 37.e: Calculate and analyze the value and return of an index given its weighting method.

Price Weighting

A price-weighted index adds the market prices of each stock in the index and divides this total by the number of stocks in the index. The divisor, however, must be adjusted for stock splits and other changes in the index portfolio to maintain the continuity of the series over time.

$$\text{price-weighted index} = \frac{\text{sum of stock prices}}{\text{number of stocks in index adjusted for splits}}$$

EXAMPLE: Price-weighted index

Given the information for the three stocks presented in the following figure, calculate a price-weighted index return over a one-month period.

Index Firm Data

	Share Price December 31, 20X6	Share Price January 31, 20X7
Stock X	\$10	\$20
Stock Y	\$20	\$15
Stock Z	\$60	\$40

Answer:

The price-weighted index is $(10 + 20 + 60) / 3 = 30$ as of December 31 and $(20 + 15 + 40) / 3 = 25$ as of January 31. Hence, the price-weighted 1-month percentage return is:

$$\frac{25}{30} - 1 = -16.7\%$$

EXAMPLE: Adjusting a price-weighted index for stock splits

At the market close on day 1, Stock A has a price of \$10, Stock B has a price of \$20, and Stock C has a price of \$90. The value of a price-weighted index of these three stocks is $(10 + 20 + 90) / 3 = 40$ at the close of trading. If Stock C splits 2-for-1, effective on day 2, what is the new denominator for the index?

Answer:

The effect of the split on the price of Stock C, in the absence of any change from the price at the end of day 1, would be to reduce it to $\$90 / 2 = \45 . The index denominator will be adjusted so that the index value would remain at 40 if there were no changes in the stock prices other than to adjust for the split. The new denominator, d , must satisfy $(10 + 20 + 45) / d = 40$ and equals 1.875.

The returns on a price-weighted index could be matched by purchasing an equal number of shares of each stock represented in the index. Because the index is price weighted, a percentage change in a high-priced stock will have a relatively greater effect on the index than the same percentage change in a low-priced stock.

Market Capitalization Weighting

A market capitalization-weighted index is calculated by summing the total value (current stock price multiplied by the number of shares outstanding) of all the stocks in the index. This sum is then divided by a similar sum calculated during the selected base period. The ratio is then multiplied by the index's base value (typically 100).

For example, if the total market values of the index portfolio on December 31 and January 31 are \$80 million and \$95 million, respectively, the index value at the end of January is:

$$\begin{aligned}\text{current index value} &= \frac{\text{current total market value of index stocks}}{\text{base year total market value of index stocks}} \\ &\quad \times \text{base year index value}\end{aligned}$$

$$\text{current index value} = \frac{\$95 \text{ million}}{\$80 \text{ million}} \times 100 = 118.75$$

Thus, the market capitalization-weighted index percentage return is:

$$(118.75 / 100) - 1 = 18.75\%$$

The following example of price-weighting versus market value-weighting shows how these two indexes are calculated and how they differ.

EXAMPLE: Price-weighted vs. market capitalization-weighted indexes

Consider the three firms described in the following table. Compare the effects on a price-weighted index and a market capitalization-weighted index if Stock A doubles in price or if Stock C doubles in price. Assume the period shown in the table is the base period for the market capitalization-weighted index and that its base value is 100.

Index Firm Data

Company	Number of Shares Outstanding (000s)	Stock Price	Capitalization (000s)
A	100	\$100	\$10,000
B	1,000	\$10	\$10,000
C	20,000	\$1	\$20,000

Answer:

The price-weighted index equals:

$$\frac{100 + 10 + 1}{3} = 37$$

If Stock A doubles in price to \$200, the price-weighted index value is:

$$\frac{200 + 10 + 1}{3} = 70.33$$

If Stock C doubles in price to \$2, the price-weighted index value is:

$$\frac{100 + 10 + 2}{3} = 37.33$$

If Stock A doubles in value, the index goes up 33.33 points, while if Stock C doubles in value, the index only goes up 0.33 points. Changes in the value of the firm with the highest stock price have a disproportionately large influence on a price-weighted index.

For a market capitalization-weighted index, the base period market capitalization is $(100,000 \times \$100) + (1,000,000 \times \$10) + (20,000,000 \times \$1) = \$40,000,000$.

If Stock A doubles in price to \$200, the index goes to:

$$\frac{100,000 \times \$200 + 1,000,000 \times \$10 + 20,000,000 \times \$1}{\$40,000,000} \times 100 = 125$$

If Stock C doubles in price to \$2, the index goes to:

$$\frac{100,000 \times \$100 + 1,000,000 \times \$10 + 20,000,000 \times \$2}{\$40,000,000} \times 100 = 150$$

In the market capitalization-weighted index, the returns on Stock C have the greatest influence on the index return because Stock C's market capitalization is larger than that of Stock A or Stock B.

Equal Weighting

An equal-weighted index places an equal weight on the returns of all index stocks, regardless of their prices or market values. A \$2 change in the price of a \$20 stock has the same effect on the

index as a \$30 change in the price of a \$300 stock regardless of the size of the company. The return of an equal-weighted index over a given period is often calculated as a simple average of the returns of the index stocks.

EXAMPLE: Equally weighted index

Calculate the equal-weighted index value for the three stocks described in the following table, assuming an initial index value of 131.

Equal-Weighted Index Data

Stock	Initial Price	Current Price	Price Change
A	\$12	\$15	+25.0%
B	\$52	\$48	-7.7%
C	\$38	\$45	+18.4%

Answer:

$$\text{change in index} = \frac{25\% - 7.7\% + 18.4\%}{3} = 11.9\%$$

$$\text{new index value} = 131(1 + 0.119) = 146.59$$

Note that for a total return index, period returns would include any dividends paid over the period.



MODULE QUIZ 37.1

- Choices that must be made when constructing a security market index *least likely* include whether to:
 - use a nominal or interval scale.
 - measure the performance of an entire market or market segment.
 - weight the securities equally or by some firm-specific characteristic.

Use the information in the following table to answer Questions 2 through 4.

	As of January 1		As of December 31	
	Share Price	Number of Shares Outstanding (thousands)	Share Price	Number of Shares Outstanding (thousands)
Stock A	\$22	1,500	\$28	1,500
Stock B	\$40	10,000	\$50	10,000
Stock C	\$34	3,000	\$30	3,000

- The 1-year return on a price-weighted index of these three stocks is *closest* to:
 - 12.5%.
 - 13.5%.
 - 18.0%.
- The 1-year return on an equal-weighted index of these three stocks is *closest* to:
 - 12.0%.
 - 12.5%.
 - 13.5%.
- The 1-year return on a market capitalization-weighted index of these stocks is *closest* to:
 - 12.0%.
 - 12.5%.
 - 13.5%.
- The 1-year return on a market capitalization-weighted index of these stocks is *closest* to:

- A. 12.5%.
 - B. 13.5%.
 - C. 18.0%.
5. Market float of a stock is *best* described as its:
- A. total outstanding shares.
 - B. shares that are available to domestic investors.
 - C. outstanding shares, excluding those held by controlling shareholders.
6. For which of the following indexes will rebalancing occur *most* frequently?
- A. A price-weighted index.
 - B. An equal-weighted index.
 - C. A market capitalization-weighted index.

MODULE 37.2: USES AND TYPES OF INDEXES



LOS 37.f: Describe rebalancing and reconstitution of an index.

Video covering
this content is
available online.

Rebalancing refers to adjusting the weights of securities in a portfolio to their target weights after price changes have affected the weights. For index calculations, rebalancing to target weights on the index securities is done on a periodic basis, usually quarterly. Because the weights in price- and value-weighted indexes (portfolios) are adjusted to their correct values by changes in prices, rebalancing is an issue primarily for equal-weighted indexes. As noted previously, the weights on security returns in an (initially) equal-weighted portfolio are not equal as securities prices change over time. Therefore, rebalancing the portfolio at the end of each period used to calculate index returns is necessary for the portfolio return to match the index return.

Index reconstitution refers to periodically adding and deleting securities that make up an index. Securities are deleted if they no longer meet the index criteria and are replaced by other securities that do. Indexes are reconstituted to reflect corporate events such as bankruptcy or delisting of index firms and are at the subjective judgment of a committee.

When a security is added to an index, its price tends to rise as portfolio managers seeking to track that index in a portfolio buy the security. The prices of deleted securities tend to fall as portfolio managers sell them. Note that additions and deletions also require that the weights on the returns of other index stocks be adjusted to conform to the desired weighting scheme.

LOS 37.g: Describe uses of security market indexes.

Security market indexes have several uses:

- *Reflection of market sentiment.* Indexes provide a representative market return and thus reflect investor confidence. Although the Dow Jones Industrial Average is a popular index, it reflects the performance of only 30 stocks and thus may not be a good measure of sentiment with regard to the broader market.
- *Benchmark of manager performance.* An index can be used to evaluate the performance of an active manager. Because portfolio performance depends to a large degree on its chosen style,

the benchmark should be consistent with the manager's investment approach and style to assess the manager's skill accurately. The index stocks should be those that the manager will actually choose from. For example, a value manager should be compared against a value index, not a broad market index, because portfolio securities will be selected from among value stocks.

- *Measure of market return and risk.* In asset allocation, estimates of the expected return and standard deviation of returns for various asset classes are based on historical returns for an index of securities representing that asset class.
 - *Measure of beta and risk-adjusted return.* The use of the capital asset pricing model (CAPM) to determine a stock's expected return requires an estimate of its beta and the return on the market. Index portfolio returns are used as a proxy for the returns on the market portfolio, both in estimating a stock's beta, and then again in calculating its expected return based on its systematic (beta) risk. Expected returns can then be compared to actual stock returns to determine systematic risk-adjusted returns.
 - *Model portfolio for index funds.* Investors who wish to invest passively can invest in an index fund, which seeks to replicate the performance of a market index. There are index mutual funds and index exchange-traded funds, as well as private portfolios that are structured to match the return of an index.
-

LOS 37.h: Describe types of equity indexes.

Investors can use a variety of equity market indexes. These equity indexes can be classified as follows:

- *Broad market index.* Provides a measure of a market's overall performance and usually contains more than 90% of the market's total value. For example, the Wilshire 5000 Index contains more than 6,000 equity securities and is, therefore, a good representation of the overall performance of the U.S. equity market.
- *Multi-market index.* Typically constructed from the indexes of markets in several countries and is used to measure the equity returns of a geographic region (e.g., Latin America indexes), markets based on their stage of economic development (e.g., emerging markets indexes), or the entire world (e.g., MSCI World Index).
- *Multi-market index with fundamental weighting.* Uses market capitalization-weighting for the country indexes but then weights the country index returns in the global index by a fundamental factor (e.g., GDP). This prevents a country with previously high stock returns from being overweighted in a multi-market index.
- *Sector index.* Measures the returns for an industry sector such as health care, financial, or consumer goods firms. Investors can use these indexes in cyclical analysis because some sectors do better than others in various phases of the business cycle. Sector indexes can be for a particular country or global. These indexes are used to evaluate portfolio managers and to construct index portfolios.
- *Style index.* Measures the returns to market capitalization and value or growth strategies. Some indexes reflect a combination of the two (e.g., small-cap value fund). Because there is no widely accepted definition of large-cap, mid-cap, or small-cap stocks, different indexes use

different definitions. These definitions may be specified values of market capitalization or relative definitions, such as defining large-cap stocks as the largest 500 firms in a given market. In constructing value stock and growth stock indexes, price-to-earnings ratios or dividend yields are often used to identify value and growth stocks. Over time, stocks can migrate from one classification to another. For example, a successful small-cap company might grow to become a mid-cap or large-cap company. This causes style indexes to typically have higher turnover of constituent firms than broad market indexes.

LOS 37.i: Compare types of security market indexes.

The following table summarizes some of the noteworthy characteristics of various global indexes. Notice from the table that most security market indexes are market capitalization-weighted and often adjusted for the float (securities actually available for purchase). The number of securities in many of these indexes can vary.

Index	Reflects	Number of Constituent Securities	Weighting Method	Notes
Dow Jones Industrial Average	Large U.S. stocks	30	Price	Stocks are chosen by Wall Street Journal editors
Nikkei Stock Average	Large Japanese stocks	225	Modified price	Price weighted and adjusted for high-priced shares
TOPIX	All stocks on the Tokyo Stock Exchange First Section	Variable	Market capitalization, adjusted for float	Has a large number of small illiquid stocks making it hard to replicate. Contains 93% of the market cap of Japanese equities
MSCI All Country World Index	Stocks in 23 developed and 24 emerging markets	Variable	Market capitalization, adjusted for float	Available in both U.S. dollars and local currency
S&P Developed Ex-U.S. BMI Energy Sector Index	Global energy stocks outside the United States	Variable	Market capitalization, adjusted for float	Is the model portfolio for an ETF
Barclays Capital Global Aggregate Bond Index	Global investment-grade bonds	Variable	Market capitalization	Formerly compiled by Lehman Brothers
Markit iBoxx Euro High-Yield Bond Indexes	Below investment-grade bonds	Variable	Market capitalization	Represents liquid portion of market and rebalanced monthly
FTSE EPRA/NAREIT Global Real Estate Index	Global real estate	Variable	Market capitalization, adjusted for float	Represents publicly traded REITs
HFRX Global Hedge Fund Index	Global hedge funds	Variable	Asset weighted	Contains a variety of hedge fund strategies and is weighted based on the amount invested in each hedge fund
HFRX Equal Weighted Strategies EUR Index	Global hedge funds	Variable	Equal weighted	Contains same strategy funds as HFRX Global Hedge Fund Index and is equal weighted
Morningstar Style Indexes	U.S. stocks grouped by value/growth and market cap	Variable	Market capitalization, adjusted for float	Nine categories classified by combinations of three cap categories and three value/growth categories

LOS 37.j: Describe types of fixed-income indexes.

Fixed-income securities vary widely with respect to their coupon rates, ratings, maturities, and embedded options such as convertibility to common stock. Consequently, a wide variety of fixed-income indexes is available. Like equity indexes, fixed-income indexes are created for various sectors, geographic regions, and levels of country economic development. They can also be constructed based on type of issuer or collateral, coupon, maturity, default risk, or inflation protection. Broad market indexes, sector indexes, style indexes, and other specialized indexes are available.

Investors should be aware of several issues with the construction of fixed-income indexes:

- *Large universe of securities.* The fixed-income security universe is much broader than the universe of stocks. Fixed-income securities are issued not just by firms, but also by

governments and government agencies. Each of these entities may also issue various types of fixed-income securities. Also, unlike stocks, bonds mature and must be replaced in fixed-income indexes. As a result, turnover is high in fixed-income indexes.

- *Dealer markets and infrequent trading.* Fixed-income securities are primarily traded by dealers, so index providers must depend on dealers for recent prices. Because fixed-income securities are typically illiquid, a lack of recent trades may require index providers to estimate the value of index securities from recent prices of securities with similar characteristics.

The large number of fixed-income securities results in large differences in the number of index securities among fixed-income indexes. Illiquidity, transaction costs, and high turnover of constituent securities make it both difficult and expensive for fixed-income portfolio managers to replicate a fixed-income index.

LOS 37.k: Describe indexes representing alternative investments.

Alternative assets are of interest to investors because of their potential diversification benefits. Three of the most widely held alternative assets are commodities, real estate, and hedge funds.

Commodity indexes represent futures contracts on commodities such as grains, livestock, metals, and energy. Examples include the Thomson Reuters/Core Commodity CRB Index (previously the Commodity Research Bureau Index) and the S&P GSCI (previously the Goldman Sachs Commodity Index).

The issues in commodity indexes relevant for investors are as follows:

- *Weighting method.* Commodity index providers use a variety of weighting schemes. Some use equal weighting, others weight commodities by their global production values, and others use fixed weights that the index provider determines. As a result, different indexes have significantly different commodity exposures and risk and return characteristics. For example, one index may have a large exposure to the prices of energy commodities while another has a large exposure to the prices of agricultural products.
- *Futures vs. actual.* Commodity indexes are based on the prices of commodity futures contracts, not the spot prices of commodities. Commodity futures contracts reflect the risk-free rate of return, changes in futures prices, and the roll yield. Furthermore, the contracts mature and must be replaced over time by other contracts. For these reasons, the return on commodity futures differs from the returns on a long position in the commodity itself.

Real estate indexes can be constructed using returns based on appraisals of properties, repeat property sales, or the performance of Real Estate Investment Trusts (REITs). REITs are similar to closed-end mutual funds in that they invest in properties or mortgages and then issue ownership interests in the pool of assets to investors. While real properties are quite illiquid, REIT shares trade like any common shares and many offer very good liquidity to investors. FTSE International produces a family of REIT indexes.

Hedge funds pool investor money and invest in nontraditional assets, using leverage (borrowed money or derivative contracts) and both long and short positions. Most **hedge fund indexes** equally weight the returns of the hedge funds included in the index.

Hedge funds are largely unregulated and are not required to report their performance to index providers. Consequently, some funds will report to one index but not another. The performance of different indexes can thus vary substantially.

Furthermore, it is often the case that those funds that report are the funds that have been successful, as the poorly performing funds do not want to publicize their performance. Funds that have reported in the past but have recently had poor returns may stop reporting their performance. The result is an upward bias in index returns, with hedge funds appearing to be better investments than they actually are.



PROFESSOR'S NOTE

Commodities, real estate, and hedge funds are discussed further in Alternative Investments.



1. The publisher of an index that includes 50 corporate bonds removes from the index three bonds that are nearing maturity and one whose issuer has defaulted and selects four actively traded bonds to replace them in the index. This bond index is said to have been:
 - A. redefined.
 - B. rebalanced.
 - C. reconstituted.
2. Which of the following would *most likely* represent an inappropriate use of an index?
 - A. As a reflection of market sentiment.
 - B. Comparing a small-cap manager against a broad market.
 - C. Using the CAPM to determine the expected return and beta.
3. An index of 200 mid-cap growth stocks is *best* described as:
 - A. a style index.
 - B. a sector index.
 - C. a broad market index.
4. Which of the following is *least accurate* regarding fixed-income indexes?
 - A. Replicating the return on a fixed-income security index is difficult for investors.
 - B. There is a great deal of heterogeneity in the composition of fixed-income security indexes.
 - C. Due to the large universe of fixed-income security issues, data for fixed-income securities are relatively easy to obtain.
5. Which of the following indexes of alternative investments is *most likely* to be calculated from derivatives prices?
 - A. Real estate index.
 - B. Commodity index.
 - C. Hedge fund index.
6. Most of the widely used global security indexes are:
 - A. price weighted.
 - B. equal weighted.
 - C. market capitalization weighted.

KEY CONCEPTS

LOS 37.a

A security market index represents the performance of an asset class, security market, or segment of a market. The performance of the market or segment over a period of time is represented by the percentage change in (i.e., the return on) the value of the index.

LOS 37.b

A price index uses only the prices of the constituent securities in the return calculation. The rate of return is called a price return.

A total return index uses both the price of and the income from the index securities in the return calculation.

LOS 37.c

Decisions that index providers must make when constructing and managing indexes include:

- The target market the index will measure.
- Which securities from the target market to include.
- The appropriate weighting method.
- How frequently to rebalance the index to its target weights.
- How frequently to re-examine the selection and weighting of securities.

LOS 37.d

A price-weighted index is the arithmetic mean of the prices of the index securities. The divisor, which is initially equal to the number of securities in the index, must be adjusted for stock splits and changes in the composition of the index over time.

An equal-weighted index assigns the same weight to each of its constituent securities.

A market capitalization-weighted index gives each constituent security a weight equal to its proportion of the total market value of all securities in the index. Market capitalization can be adjusted for a security's market float or free float to reflect the fact that not all outstanding shares are available for purchase.

A fundamental-weighted index uses weights that are independent of security prices, such as company earnings, revenue, assets, or cash flow.

LOS 37.e

$$\text{Price-weighted index} = \frac{\text{sum of stock prices}}{\text{number of stocks in index adjusted for splits}}$$

$$\text{Market capitalization-weighted index} =$$

$$\frac{\text{current total market value of index stocks}}{\text{base year total market value of index stocks}} \times \text{base year index value}$$

$$\text{Equal-weighted index} = (1 + \text{average percentage change in index stocks}) \times \text{initial index value}$$

LOS 37.f

Index providers periodically rebalance the weights of the constituent securities. This is most important for equal-weighted indexes.

Reconstitution refers to changing the securities that are included in an index. This is necessary when securities mature or when they no longer have the required characteristics to be included.

LOS 37.g

Indexes are used for the following purposes:

- Reflection of market sentiment.
- Benchmark of manager performance.
- Measure of market return.
- Measure of beta and excess return.
- Model portfolio for index funds.

LOS 37.h

Broad market equity indexes represent the majority of stocks in a market.

Multi-market equity indexes contain the indexes of several countries. Multi-market equity indexes with fundamental weighting use market capitalization weighting for the securities within a country's market but then weight the countries within the global index by a fundamental factor.

Sector indexes measure the returns for a sector (e.g., health care) and are useful because some sectors do better than others in certain business cycle phases. These indexes are used to evaluate portfolio managers and as models for sector investment funds.

Style indexes measure the returns to market capitalization and value or growth strategies. Stocks tend to migrate among classifications, which causes style indexes to have higher constituent turnover than broad market indexes.

LOS 37.i

Security market indexes available from commercial providers represent a variety of asset classes and reflect target markets that can be classified by:

- Geographic location, such as country, regional, or global indexes.
- Sector or industry, such as indexes of energy producers.
- Level of economic development, such as emerging market indexes.
- Fundamental factors, such as indexes of value stocks or growth stocks.

LOS 37.j

Fixed-income indexes can be classified by issuer, collateral, coupon, maturity, credit risk (e.g., investment grade versus high-yield), and inflation protection. They can be delineated as broad market, sector, style, or other specialized indexes. Indexes exist for various sectors, regions, and levels of development.

The fixed-income security universe is much broader than the equity universe, and fixed-income indexes have higher turnover. Index providers must depend on dealers for fixed-income security prices, and the securities are often illiquid. Fixed-income security indexes vary widely in their numbers of constituent securities and can be difficult and expensive to replicate.

LOS 37.k

Indexes have been developed to represent markets for alternative assets such as commodities, real estate, and hedge funds.

Issues in creating commodity indexes include the weighting method (different indexes can have vastly different commodity weights and resulting risk and return) and the fact that commodity

indexes are based on the performance of commodity futures contracts, not the actual commodities, which can result in different performance for a commodity index versus the actual commodity.

Real estate indexes include appraisal indexes, repeat property sales indexes, and indexes of real estate investment trusts.

Because hedge funds report their performance to index providers voluntarily, the performance of different hedge fund indexes can vary substantially and index returns have an upward bias.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 37.1

1. **A** To be useful, a security market index must have a numerical value. Selecting the target market and determining the weighting method are among the choices that must be made when constructing a securities index. (LOS 37.a, 37.c)

2. **A** $\frac{22 + 40 + 34}{3} = 32, \frac{28 + 50 + 30}{3} = 36, \frac{36}{32} - 1 = 0.125 = 12.5\%$
(LOS 37.b, 37.d, 37.e)

3. **C** $\left[\left(\frac{28}{22} - 1 \right) + \left(\frac{50}{40} - 1 \right) + \left(\frac{30}{34} - 1 \right) \right] \left(\frac{1}{3} \right) = 0.135 = 13.5\%$
(LOS 37.b, 37.d, 37.e)

4. **C** Total portfolio value January 1:

$$22(1,500) + 40(10,000) + 34(3,000) = \$535,000$$

Total portfolio value December 31:

$$28(1,500) + 50(10,000) + 30(3,000) = \$632,000$$

$$\frac{632}{535} - 1 = 0.1813 \approx 18\%$$

From a base value of 100, the December 31 index value would be

$$\frac{632}{535} \times 100 = 118.13.$$

(LOS 37.b, 37.d, 37.e)

5. **C** Market float represents shares available to the investing public and excludes shares held by controlling shareholders. Free float is a narrower measure that also excludes shares that are not available to foreign investors. (LOS 37.d)

6. **B** An equal-weighted index will be rebalanced most frequently because as stock prices change, their representation in the index needs to be adjusted. Price-weighted and market capitalization-weighted indexes do not usually need rebalancing. (LOS 37.d)

Module Quiz 37.2

1. **C** Changing the constituent securities of an index is referred to as reconstituting the index. Rebalancing refers to adjusting the index weights to their target levels. (LOS 37.f)

2. **B** Comparing a small-cap manager against a broad market would be an inappropriate use of an index. A benchmark should be consistent with the manager's investment approach and style. A manager's performance will depend to a large degree on its chosen style. (LOS 37.g)

3. **A** An index for mid-cap growth stocks is best described as a style index. Sector indexes typically measure returns for a specific industry or sector of the economy (e.g., utilities or financial services firms). (LOS 37.h)
4. **C** Fixed-income securities are largely traded by dealers and trade infrequently. Data are therefore difficult to obtain. (LOS 37.j)
5. **B** Commodity indexes are typically calculated from prices of commodity futures contracts. (LOS 37.k)
6. **C** Most global security indexes are market capitalization-weighted with a float adjustment to reflect the amount of shares available to investors. (LOS 37.i)

READING 38

MARKET EFFICIENCY

EXAM FOCUS

The informational efficiency of market prices is a very important concept to a portfolio manager. When markets are truly efficient, careful analysis and security selection using publicly available information will not lead to positive risk-adjusted returns on average. For the exam, you must understand the three forms of market efficiency and know the evidence from tests of each form of market efficiency. Focus your attention on the implications of this evidence about the value of technical and fundamental analysis and about the role of portfolio managers in the investment process. Finally, be familiar with market anomalies listed and the perspective provided by behavioral finance.

MODULE 38.1: MARKET EFFICIENCY



LOS 38.a: Describe market efficiency and related concepts, including their importance to investment practitioners.

Video covering
this content is
available online.

An **informationally efficient capital market** is one in which the current price of a security fully, quickly, and rationally reflects all available information about that security. This is really a statistical concept. An academic might say, “Given all available information, current securities prices are unbiased estimates of their values, so that the expected return on any security is just the equilibrium return necessary to compensate investors for the risk (uncertainty) regarding its future cash flows.” This concept is often put more intuitively as, “You can’t beat the market.”

In a perfectly efficient market, investors should use a **passive investment** strategy (i.e., buying a broad market index of stocks and holding it) because **active investment** strategies will underperform due to transactions costs and management fees. However, to the extent that market prices are inefficient, active investment strategies can generate positive risk-adjusted returns.

One method of measuring a market’s efficiency is to determine the time it takes for trading activity to cause information to be reflected in security prices (i.e., the lag from the time information is disseminated to the time prices reflect the value implications of that information). In some very efficient markets, such as foreign currency markets, this lag can be as short as a minute. If there is a significant lag, informed traders can use the information to potentially generate positive risk-adjusted returns.

Note that market prices should not be affected by the release of information that is well anticipated. Only new information (information that is unexpected and changes expectations)

should move prices. The announcement that a firm's earnings were up 45% over the last quarter may be good news if the expected increase was 20%. On the other hand, this may be bad news if a 70% increase was anticipated or no news at all if market participants correctly anticipated quarterly earnings.

LOS 38.b: Contrast market value and intrinsic value.

The **market value** of an asset is its current price. The **intrinsic value** or **fundamental value** of an asset is the value that a rational investor with full knowledge about the asset's characteristics would willingly pay. For example, a bond investor would fully know and understand a bond's coupon, maturity, default risk, liquidity, and other characteristics and would use these to estimate its intrinsic value.

In markets that are highly efficient, investors can typically expect market values to reflect intrinsic values. If markets are not completely efficient, active managers will buy assets for which they think intrinsic values are greater than market values and sell assets for which they think intrinsic values are less than market values.

Intrinsic values cannot be known with certainty and are estimated by investors who will have differing estimates of an asset's intrinsic value. The more complex an asset, the more difficult it is to estimate its intrinsic value. Furthermore, intrinsic value is constantly changing as new (unexpected) information becomes available.

LOS 38.c: Explain factors that affect a market's efficiency.

Markets are generally neither perfectly efficient nor completely inefficient. The degree of informational efficiency varies across countries, time, and market types. The following factors affect the degree of market efficiency.

Number of market participants. The larger the number of investors, analysts, and traders who follow an asset market, the more efficient the market. The number of participants can vary through time and across countries. For example, some countries prevent foreigners from trading in their markets, reducing market efficiency.

Availability of information. The more information is available to investors, the more efficient the market. In large, developed markets such as the New York Stock Exchange, information is plentiful and markets are quite efficient. In emerging markets, the availability of information is lower, and consequently, market prices are relatively less efficient. Some assets, such as bonds, currencies, swaps, forwards, mortgages, and money market securities that trade in over-the-counter (OTC) markets, may have less available information.

Access to information should not favor one party over another. Therefore, regulations such as the U.S. Securities and Exchange Commission's Regulation FD (fair disclosure) require that firms disclose the same information to the public that they disclose to stock analysts. Traders with material inside information about a firm are prohibited from trading on that information.

Impediments to trading. **Arbitrage** refers to buying an asset in one market and simultaneously selling it at a higher price in another market. This buying and selling of assets will continue until the prices in the two markets are equal. Impediments to arbitrage, such as high transactions costs or lack of information, will limit arbitrage activity and allow some price inefficiencies (i.e., mispricing of assets) to persist.

Short selling improves market efficiency. The sales pressure from short selling prevents assets from becoming overvalued. Restrictions on short selling, such as an inability to borrow stock cheaply, can reduce market efficiency.

Transaction and information costs. To the extent that the costs of information, analysis, and trading are greater than the potential profit from trading misvalued securities, market prices will be inefficient. It is generally accepted that markets are efficient if, after deducting costs, there are no risk-adjusted returns to be made from trading based on publicly available information.

LOS 38.d: Contrast weak-form, semi-strong-form, and strong-form market efficiency.

Professor Eugene Fama originally developed the concept of market efficiency and identified three forms of market efficiency. The difference among them is that each is based on a different set of information.

1. **Weak-form market efficiency.** The weak form of the efficient markets hypothesis (EMH) states that current security prices *fully reflect all currently available security market data*. Thus, past price and volume (market) information will have no predictive power about the future direction of security prices because price changes will be independent from one period to the next. In a weak-form efficient market, an investor cannot achieve positive risk-adjusted returns on average by using technical analysis.
2. **Semi-strong-form market efficiency.** The semi-strong form of the EMH holds that security prices rapidly adjust without bias to the arrival of all new public information. As such, current security prices *fully reflect all publicly available information*. The semi-strong form says security prices include all past security market information and nonmarket information available to the public. The implication is that an investor cannot achieve positive risk-adjusted returns on average by using fundamental analysis.
3. **Strong-form market efficiency.** The strong form of the EMH states that security prices *fully reflect all information from both public and private sources*. The strong form includes all types of information: past security market information, public, and private (inside) information. This means that no group of investors has monopolistic access to information relevant to the formation of prices, and none should be able to consistently achieve positive abnormal returns.

Given the prohibition on insider trading in most markets, it would be unrealistic to expect markets to reflect all private information. The evidence supports the view that markets are not strong-form efficient.

PROFESSOR'S NOTE



As a base level knowledge of the EMH, you should know that the weak form is based on past security market information; the semi-strong form is based on all public information (including market information); and the strong form is based on both public information and inside or private information.

LOS 38.e: Explain the implications of each form of market efficiency for fundamental analysis, technical analysis, and the choice between active and passive portfolio management.

Abnormal profit (or risk-adjusted returns) calculations are often used to test market efficiency. To calculate abnormal profits, the expected return for a trading strategy is calculated given its risk, using a model of expected returns such as the CAPM or a multifactor model. If returns are, on average, greater than equilibrium expected returns, we can reject the hypothesis of efficient prices with respect to the information on which the strategy is based.

The results of tests of the various forms of market efficiency have implications about the value of technical analysis, fundamental analysis, and portfolio management in general.

Technical analysis seeks to earn positive risk-adjusted returns by using historical price and volume (trading) data. Tests of weak-form market efficiency have examined whether technical analysis produces abnormal profits. Generally, the evidence indicates that technical analysis does not produce abnormal profits, so we cannot reject the hypothesis that markets are weak-form efficient. However, technical analysis has been shown to have success in emerging markets, and there are so many possible technical analysis trading strategies that they cannot all be tested. As noted previously, the success of any technical analysis strategy should be evaluated considering the costs of information, analysis, and trading.

Fundamental analysis is based on public information such as earnings, dividends, and various accounting ratios and estimates. The semi-strong form of market efficiency suggests that all public information is already reflected in stock prices. As a result, investors should not be able to earn abnormal profits by trading on this information.

One method of testing the semi-strong form is an **event study**. Event studies examine abnormal returns before and after the release of new information that affects a firm's intrinsic value, such as earnings announcements or dividend changes. The null hypothesis is that investors should not be able to earn positive abnormal returns on average by trading based on firm events because prices will rapidly reflect news about a firm's prospects. The evidence in developed markets indicates that markets are generally semi-strong form efficient. However, there is evidence of semi-strong form inefficiency in some emerging markets.

The evidence that developed markets are generally semi-strong form efficient raises questions about the usefulness of fundamental analysis. It must be fundamental analysis, however, that results in informationally efficient market prices. Fundamental analysis can also be of use to those exceptionally skilled investors who can generate abnormal profits through its use and to those who act rapidly before new information is reflected in prices.



PROFESSOR'S NOTE

Markets can be weak-form efficient without being semi-strong or strong-form efficient. If markets are semi-strong form efficient, they must be weak-form efficient because public information includes market information, but semi-strong form efficient markets need not be strong-form efficient.

Active vs. Passive Portfolio Management

If markets are semi-strong form efficient, investors should invest passively (i.e., invest in an index portfolio that replicates the returns on a market index). Indeed, the evidence shows that most mutual fund managers cannot outperform a passive index strategy over time.

If so, what is the role of a portfolio manager? Even if markets are efficient, portfolio managers can add value by establishing and implementing portfolio risk and return objectives and by assisting clients with portfolio diversification, asset allocation, and tax management.

LOS 38.f: Describe market anomalies.

An anomaly is something that deviates from the common rule. Tests of the EMH are frequently called *anomaly studies*, so in the efficient markets literature, a **market anomaly** is something that would lead us to reject the hypothesis of market efficiency.

Just by chance, some variables will be related to abnormal returns over a given period, although in fact these relationships are unlikely to persist over time. Thus, analysts using historical data can find patterns in security returns that appear to violate market efficiency but are unlikely to recur in the future. If the analyst uses a 5% significance level and examines the relationship between stock returns and 40 variables, two of the variables are expected to show a statistically significant relationship with stock returns by random chance. Recall that the significance level of a hypothesis test is the probability that the null hypothesis (efficiency here) will be rejected purely by chance, even when it is true. Investigating data until a statistically significant relation is found is referred to as **data snooping** or **data mining**. Note that 1,000 analysts, each testing different hypotheses on the same data set, could produce the same results as a single researcher who performed 1,000 hypothesis tests.

To avoid data snooping bias, analysts should first ask if there is an economic basis for the relationships they find between certain variables and stock returns and then test the discovered relationships with a large sample of data to determine if the relationships are persistent and present in various subperiods.

Anomalies in Time-Series Data

Calendar anomalies. The **January effect** or **turn-of-the-year effect** is the finding that during the first five days of January, stock returns, especially for small firms, are significantly higher than they are the rest of the year. In an efficient market, traders would exploit this profit opportunity in January, and in so doing, eliminate it.

Possible explanations for the January effect are **tax-loss selling**, as investors sell losing positions in December to realize losses for tax purposes and then repurchase stocks in January, pushing their prices up, and **window dressing**, as portfolio managers sell risky stocks in December to remove them from their year-end statements and repurchase them in January.

Evidence indicates that each of these explains only a portion of the January effect. However, after adjustments are made for risk, the January effect does not appear to persist over time.

Other calendar anomalies that were found at one time but no longer appear to persist are the *turn-of-the-month effect* (stock returns are higher in the days surrounding month end), the *day-of-the-week effect* (average Monday returns are negative), the *weekend effect* (positive Friday returns are followed by negative Monday returns), and the *holiday effect* (pre-holiday returns are higher).

Overreaction and momentum anomalies. The **overreaction effect** refers to the finding that firms with poor stock returns over the previous three or five years (losers) have better subsequent returns than firms that had high stock returns over the prior period. This pattern has been attributed to investor overreaction to both unexpected good news and unexpected bad news. This pattern is also present for bonds and in some international markets. **Momentum effects** have also been found where high short-term returns are followed by continued high returns. This pattern is present in some international markets as well.

Both the overreaction and momentum effects violate the weak form of market efficiency because they provide evidence of a profitable strategy based only on market data. Some researchers argue that the evidence of overreaction to new information is due to the nature of the statistical tests used and that evidence of momentum effects in securities prices reflects rational investor behavior.

Anomalies in Cross-Sectional Data

The **size effect** refers to initial findings that small-cap stocks outperform large-cap stocks. This effect could not be confirmed in later studies, suggesting that either investors had traded on, and thereby eliminated, this anomaly or that the initial finding was simply a random result for the time period examined.

The **value effect** refers to the finding that **value stocks** [those with lower price-to-earnings (P/E), lower market-to-book (M/B), and higher dividend yields] have outperformed **growth stocks** (those with higher P/E, higher M/B, and lower dividend yields). This violates the semi-strong form of market efficiency because the information necessary to classify stocks as value or growth is publicly available. However, some researchers attribute the value effect to greater risk of value stocks that is not captured in the risk adjustment procedure used in the studies.

Other Anomalies

Closed-end investment funds. The shares of **closed-end investment funds** trade at prices that sometimes deviate from the **net asset value** (NAV) of the fund shares, often trading at large discounts to NAV. Such large discounts are an anomaly because, by arbitrage, the value of the pool of assets should be the same as the market price for closed-end shares. Various explanations have been put forth to explain this anomaly, including management fees, taxes on future capital gains, and share illiquidity. None of these explanations fully explains the pricing discrepancy. However, transaction costs would eliminate any profits from exploiting the unexplained portion of closed-end fund discounts.

Earnings announcements. An **earnings surprise** is that portion of announced earnings that was not expected by the market. Positive earnings surprises (earnings higher than expected) precede periods of positive risk-adjusted post-announcement stock returns, and negative

surprises lead to predictable negative risk-adjusted returns. The anomaly is that the adjustment process does not occur entirely on the announcement day. Investors could exploit this anomaly by buying positive earnings surprise firms and selling negative earnings surprise firms. Some researchers argue that evidence of predictable abnormal returns after earnings surprises is a result of estimating risk-adjusted returns incorrectly in the tests and that transaction costs would eliminate any abnormal profits from attempting to exploit this returns anomaly.

Initial public offerings. IPOs are typically underpriced, with the offer price below the market price once trading begins. However, the long-term performance of IPO shares as a group is below average. This suggests that investors overreact, in that they are too optimistic about a firm's prospects on the offer day. Some believe this is not an anomaly, but rather a result of the statistical methodologies used to estimate abnormal returns.

Economic fundamentals. Research has found that stock returns are related to known economic fundamentals such as dividend yields, stock volatility, and interest rates. However, we would expect stock returns to be related to economic fundamentals in efficient markets. The relationship between stock returns and dividend yields is also not consistent over all time periods.

Implications for Investors

The majority of the evidence suggests that reported anomalies are not violations of market efficiency but are due to the methodologies used in the tests of market efficiency. Furthermore, both underreaction and overreaction have been found in the markets, meaning that prices are efficient on average. Other explanations for the evidence of anomalies are that they are transient relations, too small to profit from, or simply reflect returns to risk that the researchers have failed to account for.

The bottom line for investors is that portfolio management based on previously identified anomalies will likely be unprofitable. Investment management based solely on anomalies has no sound economic basis.

LOS 38.g: Describe behavioral finance and its potential relevance to understanding market anomalies.

Behavioral finance examines the actual decision-making processes of investors. Many observers have concluded that investors are not the rational utility-maximizing decision makers with complete information that traditional finance assumes they are. Investors appear to exhibit bias in their decision making, base decisions on the actions of others, and not evaluate risk in the way traditional models assume they do.

Various types of investor irrationality have been proposed as explanations for reported pricing anomalies. Whether widespread investor irrationality is the underlying cause of reported returns anomalies is an open question. Market efficiency does not require an assumption that every investor acts rationally in accordance with traditional finance theory. Semi-strong form market efficiency requires that investors cannot earn positive abnormal returns on average (beat the market) using public information. The evidence on market efficiency certainly suggests that this is the case. Evidence that some investors exhibit bias, or other deviations from

perfect rationality, in their investment decision making does not necessarily mean that market prices themselves are irrational, at least not in ways that lead to violations of market efficiency.

Observed investor behaviors and biases that are considered evidence of irrational behavior include:

- **Loss aversion**, which refers to the tendency of investors to be more risk averse when faced with potential losses than they are when faced with potential gains. Put another way, investors dislike a loss more than they like a gain of an equal amount.
- **Investor overconfidence**, which is a tendency of investors to overestimate their abilities to analyze security information and identify differences between securities' market prices and intrinsic values.
- **Herding**, which is a tendency of investors to act in concert on the same side of the market, acting not on private analysis, but mimicking the investment actions of other investors.

An **information cascade** results when investors mimic the decisions of others. The idea is that uninformed or less-informed traders watch the actions of informed traders and follow their investment actions. If those who act first are more knowledgeable investors, others following their actions may, in fact, be part of the process of incorporating new information into securities prices and actually move market prices toward their intrinsic values, improving informational efficiency.

Behavioral finance can explain how securities' market prices can deviate from rational prices and be biased estimates of intrinsic value. If investor rationality is viewed as a prerequisite for market efficiency, then markets are not efficient. If market efficiency only requires that investors cannot consistently earn abnormal risk-adjusted returns, then research supports the belief that markets are efficient.



MODULE QUIZ 38.1

1. In an informationally efficient capital market:
 - A. active managers can generate abnormal profits.
 - B. security prices quickly reflect new information.
 - C. investors react to all information releases rapidly.
2. The intrinsic value of an asset:
 - A. changes through time as new information is released.
 - B. is the price at which the asset can be bought or sold at a given point in time.
 - C. can be easily determined with a financial calculator, given investor risk preferences.
3. In terms of market efficiency, short selling *most likely*:
 - A. leads to excess volatility, which reduces market efficiency.
 - B. promotes market efficiency by making assets less likely to become overvalued.
 - C. has little effect on market efficiency because short sellers face the risk of unlimited losses.
4. The weak-form EMH asserts that stock prices fully reflect which of the following types of information?
 - A. Market only.
 - B. Market and public.
 - C. Public and private.
5. Research has revealed that the performance of professional money managers tends to be:
 - A. equal to the performance of a passive investment strategy.
 - B. inferior to the performance of a passive investment strategy.

- C. superior to the performance of a passive investment strategy.
6. Which of the following *best* describes the majority of the evidence regarding anomalies in stock returns?
- Weak-form market efficiency holds, but semi-strong form efficiency does not.
 - Neither weak-form nor semi-strong form market efficiency holds.
 - Reported anomalies are not violations of market efficiency but are the result of research methodologies.
7. Investors who exhibit loss aversion *most likely*:
- have symmetric risk preferences.
 - are highly risk averse.
 - dislike losses more than they like equal gains.

KEY CONCEPTS

LOS 38.a

In an informationally efficient capital market, security prices reflect all available information fully, quickly, and rationally. The more efficient a market is, the quicker its reaction will be to new information. Only unexpected information should elicit a response from traders.

If the market is fully efficient, active investment strategies cannot earn positive risk-adjusted returns consistently, and investors should therefore use a passive strategy.

LOS 38.b

An asset's market value is the price at which it can currently be bought or sold.

An asset's intrinsic value is the price that investors with full knowledge of the asset's characteristics would place on the asset.

LOS 38.c

Large numbers of market participants and greater information availability tend to make markets more efficient.

Impediments to arbitrage and short selling and high costs of trading and gathering information tend to make markets less efficient.

LOS 38.d

The weak form of the efficient markets hypothesis (EMH) states that security prices fully reflect all past price and volume information.

The semi-strong form of the EMH states that security prices fully reflect all publicly available information.

The strong form of the EMH states that security prices fully reflect all public and private information.

LOS 38.e

If markets are weak-form efficient, technical analysis does not consistently result in abnormal profits.

If markets are semi-strong form efficient, fundamental analysis does not consistently result in abnormal profits. However, fundamental analysis is necessary if market prices are to be semi-strong form efficient.

If markets are strong-form efficient, active investment management does not consistently result in abnormal profits.

Even if markets are strong-form efficient, portfolio managers can add value by establishing and implementing portfolio risk and return objectives and assisting with portfolio diversification, asset allocation, and tax minimization.

LOS 38.f

A market anomaly is something that deviates from the efficient market hypothesis. Most evidence suggests anomalies are not violations of market efficiency but are due to the methodologies used in anomaly research, such as data mining or failing to adjust adequately for risk.

Anomalies that have been identified in time-series data include calendar anomalies such as the January effect (small firm stock returns are higher at the beginning of January), overreaction anomalies (stock returns subsequently reverse), and momentum anomalies (high short-term returns are followed by continued high returns).

Anomalies that have been identified in cross-sectional data include a size effect (small-cap stocks outperform large-cap stocks) and a value effect (value stocks outperform growth stocks).

Other identified anomalies involve closed-end investment funds selling at a discount to NAV, slow adjustments to earnings surprises, investor overreaction to and long-term underperformance of IPOs, and a relationship between stock returns and prior economic fundamentals.

LOS 38.g

Behavioral finance examines whether investors behave rationally, how investor behavior affects financial markets, and how cognitive biases may result in anomalies. Behavioral finance describes investor irrationality but does not necessarily refute market efficiency as long as investors cannot consistently earn abnormal risk-adjusted returns.

ANSWER KEY FOR MODULE QUIZ

Module Quiz 38.1

1. **B** In informationally efficient capital markets, new information is quickly reflected in security prices. Investors react only to unexpected information releases because information releases that are expected will already be reflected in securities prices. Active strategies will underperform in an efficient market because they have greater transactions and management costs than passive strategies and will not consistently create positive abnormal returns after adjusting for risk. (LOS 38.a)
2. **A** Intrinsic value changes as new information arrives in the marketplace. It cannot be known with certainty and can only be estimated. The price of an asset at a given point in time is its market value, which will differ from its intrinsic value if markets are not fully efficient. (LOS 38.b)
3. **B** Short selling promotes market efficiency because the sales pressure from short selling can reduce the prices of assets that have become overvalued. (LOS 38.c)

4. **A** Weak-form EMH states that stock prices fully reflect all market (i.e., price and volume) information. (LOS 38.d)
5. **B** Tests indicate that mutual fund performance has been inferior to that of a passive index strategy. (LOS 38.e)
6. **C** The majority of evidence is that anomalies are not violations of market efficiency but are due to the research methodologies used. Portfolio management based on anomalies will likely be unprofitable after transactions costs are considered. (LOS 38.f)
7. **C** Loss aversion refers to the tendency of investors to be more risk averse when faced with potential losses and less risk averse when faced with potential gains. That is, they dislike losses more than they like gains of an equal amount. Their risk preferences are asymmetric. (LOS 38.g)

READING 39

OVERVIEW OF EQUITY SECURITIES

EXAM FOCUS

Equities have higher returns than bonds and bills, but also higher risk. Know the characteristics of common and preferred equity types, as well as the methods of investing in foreign stock. Understand the difference between the book value of equity and market value of equity and what this difference represents.

MODULE 39.1: TYPES OF EQUITY INVESTMENTS



LOS 39.a: Describe characteristics of types of equity securities.

Video covering this content is available online.

Common shares are the most common form of equity and represent an ownership interest. Common shareholders have a residual claim (after the claims of debtholders and preferred stockholders) on firm assets if the firm is liquidated and govern the corporation through voting rights. Firms are under no obligation to pay dividends on common equity; the firm determines what dividend will be paid periodically. Common stockholders are able to vote for the board of directors, on merger decisions, and on the selection of auditors. If they are unable to attend the annual meeting, shareholders can vote by **proxy** (having someone else vote as they direct them, on their behalf).

In a **statutory voting** system, each share held is assigned one vote in the election of each member of the board of directors. Under **cumulative voting**, shareholders can allocate their votes to one or more candidates as they choose. For example, consider a situation where a shareholder has 100 shares and three directors will be elected. Under statutory voting, the shareholder can vote 100 shares for his director choice in each election. Under cumulative voting, the shareholder has 300 votes, which can be cast for a single candidate or spread across multiple candidates. The three receiving the greatest number of votes are elected. Cumulative voting makes it possible for a minority shareholder to have more proportional representation on the board. The way the math works, a holder of 30% of the firm's shares could choose three of ten directors with cumulative voting but could elect no directors under statutory voting.

Preference shares (or **preferred stock**) have features of both common stock and debt. As with common stock, preferred stock dividends are not a contractual obligation, and the shares usually do not mature. Like debt, preferred shares typically make fixed periodic payments to investors and do not usually have voting rights. Preference shares may be callable, giving the firm the right to repurchase the shares at a pre-specified call price. They may also be putable,

giving the shareholder the right to sell the preference shares back to the issuer at a specified price.

Cumulative preference shares are usually promised fixed dividends, and any dividends that are not paid must be made up before common shareholders can receive dividends. The dividends of **non-cumulative preference shares** do not accumulate over time when they are not paid, but dividends for any period must be paid before common shareholders can receive dividends.

Preferred shares have a stated par value and pay a percentage dividend based on the par value of the shares. An \$80 par value preferred with a 10% dividend pays a dividend of \$8 per year. Investors in **participating preference shares** receive extra dividends if firm profits exceed a predetermined level and may receive a value greater than the par value of the preferred stock if the firm is liquidated. **Non-participating preference shares** have a claim equal to par value in the event of liquidation and do not share in firm profits. Smaller and riskier firms whose investors may be concerned about the firm's future often issue participating preferred stock so investors can share in the upside potential of the firm.

Convertible preference shares can be exchanged for common stock at a conversion ratio determined when the shares are originally issued. It has the following advantages:

- The preferred dividend is higher than a common dividend.
- If the firm is profitable, the investor can share in the profits by converting his shares into common stock.
- The conversion option becomes more valuable when the common stock price increases.
- Preferred shares have less risk than common shares because the dividend is stable and they have priority over common stock in receiving dividends and in the event of liquidation of the firm.

Because of their upside potential, convertible preferred shares are often used to finance risky venture capital and private equity firms. The conversion feature compensates investors for the additional risk they take when investing in such firms.

LOS 39.b: Describe differences in voting rights and other ownership characteristics among different equity classes.

A firm may have different classes of common stock (e.g., "Class A" and "Class B" shares). One class may have greater voting power and seniority if the firm's assets are liquidated. The classes may also be treated differently with respect to dividends, stock splits, and other transactions with shareholders. Information on the ownership and voting rights of different classes of equity shares can be found in the company's filings with securities regulators, such as the Securities and Exchange Commission in the United States.

LOS 39.c: Compare and contrast public and private equity securities.

The discussion so far has centered on equity that is publicly traded. **Private equity** is usually issued to institutional investors via private placements. Private equity markets are smaller than public markets but are growing rapidly.

Compared to public equity, private equity has the following characteristics:

- Less liquidity because no public market for the shares exists.
- Share price is negotiated between the firm and its investors, not determined in a market.
- More limited firm financial disclosure because there is no government or exchange requirement to do so.
- Lower reporting costs because of less onerous reporting requirements.
- Potentially weaker corporate governance because of reduced reporting requirements and less public scrutiny.
- Greater ability to focus on long-term prospects because there is no public pressure for short-term results.
- Potentially greater return for investors once the firm goes public.

The three main types of private equity investments are venture capital, leveraged buyouts, and private investments in public equity.

Venture capital refers to the capital provided to firms early in their life cycles to fund their development and growth. Venture capital financing at various stages of a firm's development is referred to as *seed* or *start-up*, *early stage*, or *mezzanine* financing. Investors can be family, friends, wealthy individuals, or private equity funds. Venture capital investments are illiquid and investors often have to commit funds for three to ten years before they can cash out (exit) their investment. Investors hope to profit when they can sell their shares after (or as part of) an initial public offering or to an established firm.

In a **leveraged buyout (LBO)**, investors buy all of a firm's equity using debt financing (leverage). If the buyers are the firm's current management, the LBO is referred to as a **management buyout (MBO)**. Firms in LBOs usually have cash flow that is adequate to service the issued debt or have undervalued assets that can be sold to pay down the debt over time.

In a **private investment in public equity (PIPE)**, a public firm that needs capital quickly sells private equity to investors. The firm may have growth opportunities, be in distress, or have large amounts of debt. The investors can often buy the stock at a sizeable discount to its market price.



MODULE QUIZ 39.1

1. The advantage of participating preferred shares versus non-participating preferred shares is that participating preferred shares can:
 - A. obtain voting rights.
 - B. receive extra dividends.
 - C. be converted into common stock.
2. Which of the following *best* describes the benefit of cumulative share voting?
 - A. It provides significant minority shareholders with proportional representation on the board.
 - B. It prevents minority shareholders from exercising excessive control.
 - C. If cumulative dividends are not paid, preferred shareholders are given voting rights.
3. Compared to public equity, which of the following is *least likely* to characterize private equity?

- A. Lower reporting costs.
- B. Potentially weaker corporate governance.
- C. Lower returns because of its less liquid market.

MODULE 39.2: FOREIGN EQUITIES AND EQUITY RISK



Video covering
this content is
available online.

LOS 39.d: Describe methods for investing in non-domestic equity securities.

When capital flows freely across borders, markets are said to be *integrated*. The world's financial markets have become more integrated over time, especially as a result of improved communications and trading technologies. However, barriers to global capital flows still exist. Some countries restrict foreign ownership of their domestic stocks, primarily to prevent foreign control of domestic companies and to reduce the variability of capital flows in and out of their countries.

An increasing number of countries have dropped foreign capital restrictions. Studies have shown that reducing capital barriers improves equity market performance. Furthermore, companies are increasingly turning to foreign investors for capital by listing their stocks on foreign stock exchanges or by encouraging foreign ownership of shares.

From the firm's perspective, listing on foreign stock exchanges increases publicity for the firm's products and the liquidity of the firm's shares. Foreign listing also increases firm transparency due to the stricter disclosure requirements of many foreign markets.

Direct investing in the securities of foreign companies simply refers to buying a foreign firm's securities in foreign markets. Some obstacles to direct foreign investment are that:

- The investment and return are denominated in a foreign currency.
- The foreign stock exchange may be illiquid.
- The reporting requirements of foreign stock exchanges may be less strict, impeding analysis.
- Investors must be familiar with the regulations and procedures of each market in which they invest.

Other methods for investing in foreign companies are provided by global depository receipts (GDRs), American depository receipts (ADRs), global registered shares (GRSs), and baskets of listed depository receipts (BLDRs).

Depository receipts (DRs) represent ownership in a foreign firm and are traded in the markets of other countries in local market currencies. A bank deposits shares of the foreign firm and then issues receipts representing ownership of a specific number of the foreign shares. The **depository bank** acts as a custodian and manages dividends, stock splits, and other events. Although the investor does not have to convert to the foreign currency, the value of the DR is affected by exchange rate changes, as well as firm fundamentals, economic events, and any other factors that affect the value of any stock.

If the firm is involved with the issue, the depository receipt is a **sponsored DR**; otherwise, it is an **unsponsored DR**. A sponsored DR provides the investor voting rights and is usually subject to greater disclosure requirements. In an unsponsored DR, the depository bank retains the voting rights.

Global depository receipts (GDRs) are issued outside the United States and the issuer's home country. Most GDRs are traded on the London and Luxembourg exchanges. Although not listed on U.S. exchanges, they are usually denominated in U.S. dollars and can be sold to U.S. institutional investors. GDRs are not subject to the capital flow restrictions imposed by governments and thus offer the firm and the investor greater opportunities for foreign investment. The firm usually chooses to list the GDR in a market where many investors are familiar with the firm.

American depository receipts (ADRs) are denominated in U.S. dollars and trade in the United States. The security on which the ADR is based is the **American depository share (ADS)**, which trades in the firm's domestic market. Some ADRs allow firms to raise capital in the United States or use the shares to acquire other firms. Most require U.S. Securities and Exchange Commission (SEC) registration, but some are privately placed (Rule 144A or Regulation S receipts).

The four types of ADRs, with different levels of trading availability and firm requirements, are summarized in Figure 39.1.

Figure 39.1: Types of ADRs

	Level I	Level II	Level III	Rule 144A
Trading location	Over-the-counter (OTC)	NYSE, Nasdaq, and AMEX	NYSE, Nasdaq, and AMEX	Private
SEC registration required	Yes	Yes	Yes	No
Ability to raise capital in United States	No	No	Yes	Yes
Firm listing expenses	Low	High	High	Low

Global registered shares (GRS) are traded in different currencies on stock exchanges around the world.

A basket of listed depository receipts (BLDR) is an exchange-traded fund (ETF) that is a collection of DRs. ETF shares trade in markets just like common stocks.

LOS 39.e: Compare the risk and return characteristics of different types of equity securities.

The returns on equity investments consist of price changes, dividend payments, and, in the case of equities denominated in a foreign currency, gains or losses from changes in exchange rates. A Japanese investor who invests in euro-denominated shares will have greater yen-based returns if the euro appreciates relative to the yen.

Gains from dividends and the reinvestment of dividends have been an important part of equity investors' long-term returns. For example, \$1 invested in U.S. stocks in 1900 would have been

worth \$834 in real terms in 2011 with dividends reinvested but only \$8.10 with price appreciation alone. Over the same time period, the terminal wealth for bonds and bills would have been \$9.30 and \$2.80, respectively.¹

The risk of equity securities is most commonly measured as the standard deviation of returns. Preferred stock is less risky than common stock because preferred stock pays a known, fixed dividend to investors that is a large part of the return, whereas common dividends are variable and can vary with earnings. Also, preferred stockholders receive their distributions before common shareholders and have a claim in liquidation equal to the par value of their shares that has priority over the claims of common stock owners. Because it is less risky, preferred stock has a lower average return than common stock.

Cumulative preferred shares have less risk than non-cumulative preferred shares because they retain the right to receive any missed dividends before any common stock dividends can be paid.

For both common and preferred shares, putable shares are less risky and callable shares are more risky compared to shares with neither option. Putable shares are less risky because if the market price drops, the investor can put the shares back to the firm at a fixed price (assuming the firm has the capital to honor the put). Because of this feature, putable shares usually pay a lower dividend yield than non-putable shares.

Callable shares are the most risky because if the market price rises, the firm can call the shares, limiting the upside potential of the shares. Callable shares, therefore, usually have higher dividend yields than non-callable shares.

LOS 39.f: Explain the role of equity securities in the financing of a company's assets.

Equity capital is used for the purchase of long-term assets, equipment, research and development, and expansion into new businesses or geographic areas. Equity securities provide the firm with “currency” that can be used to buy other companies or that can be offered to employees as incentive compensation. Having publicly traded equity securities provides liquidity, which may be especially important to firms that need to meet regulatory requirements, capital adequacy ratios, and liquidity ratios.

LOS 39.g: Contrast the market value and book value of equity securities.

The primary goal of firm management is to increase the book value of the firm's equity and thereby increase the market value of its equity. The **book value of equity** is the value of the firm's assets on the balance sheet minus its liabilities. It increases when the firm has positive net income and retained earnings that flow into the equity account. When management makes decisions that increase income and retained earnings, they increase the book value of equity.

The **market value of equity** is the total value of a firm's outstanding equity shares based on market prices and reflects the expectations of investors about the firm's future performance. Investors use their perceptions of the firm's risk and the amounts and timing of future cash

flows to determine the market value of equity. The market value and book value of equity are seldom equal. Although management may be maximizing the book value of equity, this may not be reflected in the market value of equity because book value does not reflect investor expectations about future firm performance.

LOS 39.h: Compare a company's cost of equity, its (accounting) return on equity, and investors' required rates of return.

A key ratio used to determine management efficiency is the **accounting return on equity**, usually referred to simply as the **return on equity (ROE)**. ROE is calculated as net income available to common (net income minus preferred dividends) divided by the average book value of common equity over the period:

$$ROE_t = \frac{NI_t}{\text{average } BV_t} = \frac{NI_t}{(BV_t + BV_{t-1})/2}$$

Alternatively, ROE is often calculated using only beginning-of-year book value of equity (i.e., book value of equity for end of year $t - 1$):

$$ROE_t = \frac{NI_t}{BV_{t-1}}$$

The first method is more appropriate when it is the industry convention or when book value is volatile. The latter method is more appropriate when examining ROE for a number of years or when book value is stable.

Higher ROE is generally viewed as a positive for a firm, but the reason for an increase should be examined. For example, if book value is decreasing more rapidly than net income, ROE will increase. This is not, however, a positive for the firm. A firm can also issue debt to repurchase equity, thereby decreasing the book value of equity. This would increase the ROE but also make the firm's shares riskier due to the increased financial leverage (debt).



PROFESSOR'S NOTE

The DuPont formula discussed in the reading on Financial Analysis Techniques can help the analyst determine the reasons for changes in ROE.

The book value of equity reflects a firm's financial decisions and operating results since its inception, whereas the market value of equity reflects the market's consensus view of a firm's future performance. The **price-to-book ratio** (also called the **market-to-book ratio**) is the market value of a firm's equity divided by the book value of its equity. The more optimistic investors are about the firm's future growth, the greater its price-to-book ratio. The price-to-book ratio is used as a measure of relative value. Often, firms with low price-to-book ratios are considered *value stocks*, while firms with high price-to-book ratios are considered *growth stocks*.

EXAMPLE: ROE, market, and book value of equity calculations

Given the following data for O'Grady Industries, calculate the return on average equity for 20X9 and the total market value of equity, the book value per share, and the price-to-book

ratio at the end of 20X9.

Fiscal Year-End Dec. 31	20X9	20X8
Total stockholder's equity	18,503	17,143
Net income available to common	3,526	3,056
Stock price	\$16.80	\$15.30
Shares outstanding	3,710	2,790

Answer:

The return on average equity for 20X9 is:

$$ROE_t = \frac{NI_t}{\text{average } BV_t} = \frac{NI_t}{(BV_t + BV_{t-1})/2}$$

$$= \frac{\$3,526}{(\$18,503 + \$17,143)/2} = 19.78\%$$

The total market value of the firm's equity at the end of 20X9 is:

$$\$16.80 \times 3,710 = \$62,328$$

The book value per share at the end of 20X9 is:

$$= \frac{\$18,503}{3,710} = \$4.99$$

The price-to-book ratio at the end of 20X9 is:

$$= \frac{\$16.80}{\$4.99} = 3.37$$

Investors' Required Return and the Cost of Equity

A firm's **cost of equity** is the expected equilibrium total return (including dividends) on its shares in the market. It is usually estimated in practice using a dividend discount model or the capital asset pricing model. At any point in time, a decrease in share price will increase the expected return on the shares and an increase in share price will decrease expected returns, other things equal. Because the intrinsic value of a firm's shares is the discounted present value of its future cash flows, an increase (decrease) in the required return used to discount future cash flows will decrease (increase) intrinsic value.

Investors also estimate the expected market returns on equity shares and compare this to the minimum return they will accept for bearing the risk inherent in a particular stock.

If an investor estimates the expected return on a stock to be greater than her minimum required rate of return on the shares, given their risk, then the shares are an attractive investment. Investors can have different required rates of return for a given risk, different estimates of a firm's future cash flows, and different estimates of the risk of a firm's equity shares. A firm's cost of equity can be interpreted as the minimum rate of return required by investors (in the aggregate) to compensate them for the risk of the firm's equity shares.



MODULE QUIZ 39.2

1. Global depository receipts are most often denominated in:
 - A. the currency of the country where they trade and issued outside the United States.

- B. U.S. dollars and issued in the United States.
 - C. U.S. dollars and issued outside the United States.
2. Which of the following types of preferred shares has the *most* risk for investors?
- A. Putable shares.
 - B. Callable shares.
 - C. Non-putable, non-callable shares.
3. Which of the following *best* describes the book value of equity?
- A. Management should attempt to maximize book value of equity.
 - B. Book value of equity decreases when retained earnings increase.
 - C. Book value of equity reflects investors' perceptions of the firm's future.
4. Which of the following causes of an increase in return on equity is *most likely* a positive sign for a firm's equity investors?
- A. A firm issues debt to repurchase equity.
 - B. Net income is increasing at a faster rate than book value of equity.
 - C. Net income is decreasing at a slower rate than book value of equity.

KEY CONCEPTS

LOS 39.a

Common shareholders have a residual claim on firm assets and govern the corporation through voting rights. Common shares have variable dividends which the firm is under no legal obligation to pay.

Preferred stock typically does not mature, does not have voting rights, and has dividends that are fixed in amount but are not a contractual obligation of the firm.

Cumulative preferred shares require any dividends that were missed in the past (dividends in arrears) to be paid before common shareholders receive any dividends. Participating preferred shares receive extra dividends if firm profits exceed a pre-specified level and a value greater than the par value if the firm is liquidated. Convertible preferred stock can be converted to common stock at a pre-specified conversion ratio.

Callable shares allow the firm the right to repurchase the shares at a pre-specified price. Putable shares give the shareholder the right to sell the shares back to the firm at a pre-specified price.

LOS 39.b

Some companies' equity shares are divided into different classes, such as Class A and Class B shares. Different classes of common equity may have different voting rights and priority in liquidation.

LOS 39.c

Compared to publicly traded firms, private equity firms have lower reporting costs, greater ability to focus on long-term prospects, and potentially greater return for investors once the firm goes public. However, private equity investments are illiquid, firm financial disclosure may be limited, and corporate governance may be weaker.

LOS 39.d

Investors who buy foreign stock directly on a foreign stock exchange receive a return denominated in a foreign currency, must abide by the foreign stock exchange's regulations and procedures, and may be faced with less liquidity and less transparency than is available in the

investor's domestic markets. Investors can often avoid these disadvantages by purchasing depository receipts for the foreign stock that trade on their domestic exchange.

Global depository receipts are issued outside the United States and outside the issuer's home country. American depository receipts are denominated in U.S. dollars and are traded on U.S. exchanges.

Global registered shares are common shares of a firm that trade in different currencies on stock exchanges throughout the world.

Baskets of listed depository receipts are exchange-traded funds that invest in depository receipts.

LOS 39.e

Equity investor returns consist of dividends, capital gains or losses from changes in share prices, and any foreign exchange gains or losses on shares traded in a foreign currency. Compounding of reinvested dividends has been an important part of an equity investor's long-term return.

Preferred stock is less risky than common stock because preferred stock pays a known, fixed dividend to investors; preferred stockholders must receive dividends before common stock dividends can be paid; and preferred stockholders have a claim equal to par value if the firm is liquidated. Putable shares are the least risky and callable shares are the most risky. Cumulative preferred shares are less risky than non-cumulative preferred shares, as any dividends missed must be paid before a common stock dividend can be paid.

LOS 39.f

Equity securities provide funds to the firm to buy productive assets, to buy other companies, or to offer to employees as compensation. Equity securities provide liquidity that may be important when the firm must raise additional funds.

LOS 39.g

The book value of equity is the difference between the financial statement value of the firm's assets and liabilities. Positive retained earnings increase the book value of equity. Book values reflect the firm's past operating and financing choices.

The market value of equity is the share price multiplied by the number of shares outstanding. Market value reflects investors' expectations about the timing, amount, and risk of the firm's future cash flows.

LOS 39.h

The accounting return on equity (ROE) is calculated as the firm's net income divided by the book value of common equity. ROE measures whether management is generating a return on common equity but is affected by the firm's accounting methods.

The firm's cost of equity is the minimum rate of return that investors in the firm's equity require. Investors' required rates of return are reflected in the market prices of the firm's shares.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 39.1

1. **B** Participating preferred shares can receive extra dividends if firm profits exceed a pre-specified level and a value greater than the par value if the firm is liquidated. (LOS 39.a)
2. **A** Cumulative voting allows minority shareholders to gain representation on the board because they can use all of their votes for specific board members. (LOS 39.b)
3. **C** Private equity has less liquidity because no public market for it exists. The lower liquidity of private equity would increase required returns. (LOS 39.c)

Module Quiz 39.2

1. **C** Global Depository Receipts are not listed on U.S. exchanges and are most often denominated in U.S. dollars. They are not issued in the United States. (LOS 39.d)
2. **B** Callable shares are the most risky because if the market price rises, the firm can call in the shares, limiting the investor's potential gains. Putable shares are the least risky because if the market price drops, the investor can put the shares back to the firm at a predetermined price. The risk of non-putable, non-callable shares falls in between. (LOS 39.e)
3. **A** The primary goal of firm management is to increase the book value of equity. It increases when retained earnings are positive. The *market* value of equity reflects the collective expectations of investors about the firm's future performance. (LOS 39.g)
4. **B** Net income increasing at a faster rate than book value of equity generally would be a positive sign. If a firm issues debt to repurchase equity, this decreases the book value of equity and increases the ROE. However, now the firm becomes riskier due to the increased debt. Net income decreasing at a slower rate than book value of equity would increase ROE, but decreasing net income is not a positive sign. (LOS 39.h)

¹ Ryan C. Fuhrmann, CFA, and Asjeet S. Lamba, CFA, *Overview of Equity Securities*, CFA Program Level I 2020 Curriculum, Volume 5 (CFA Institute, 2019).

READING 40

INTRODUCTION TO INDUSTRY AND COMPANY ANALYSIS

EXAM FOCUS

This topic provides a great deal of material on industry analysis. Understand the effects of business cycles and the stage of an industry's life cycle. Porter's five forces and two competitive strategies are very important to know. Beyond these, make sure that you know the effects on price competition and profitability of the items considered in industry analysis and of the various firm characteristics discussed.

MODULE 40.1: INDUSTRY ANALYSIS



LOS 40.a: Explain uses of industry analysis and the relation of industry analysis to company analysis.

Video covering this content is available online.

Industry analysis is important for company analysis because it provides a framework for understanding the firm. Analysts will often focus on a group of specific industries so that they can better understand the business conditions the firms in those industries face.

Understanding a firm's business environment can provide insight about the firm's potential growth, competition, and risks. For a credit analyst, industry conditions can provide important information about whether a firm will be able to meet its obligations during the next recession.

In an active management strategy, industry analysis can identify industries that are undervalued or overvalued in order to weight them appropriately. Some investors engage in **industry rotation**, which is overweighting or underweighting industries based on the current phase of the business cycle. A firm's industry has been found to be as important as its home country in determining its performance.

In performance attribution analysis, the sources of portfolio return are determined relative to a benchmark. The industry representation within a portfolio is often a significant component of attribution analysis.

LOS 40.b: Compare methods by which companies can be grouped.

One way to group companies into an industry is by the *products and services* they offer. For example, the firms that produce automobiles constitute the auto industry. A **sector** is a group of

similar industries. Hospitals, doctors, pharmaceutical firms, and other industries are included in the health care sector. Systems that are grouped by products and services usually use a firm's **principal business activity** (the largest source of sales or earnings) to classify firms. Examples of these systems are discussed in the following and include the Global Industry Classification Standard (GICS), Russell Global Sectors (RGS), and Industry Classification Benchmark.

Firms can also be classified by their *sensitivity to business cycles*. This system has two main classifications: cyclical and non-cyclical firms.

Statistical methods, such as cluster analysis, can also be used. This method groups firms that historically have had highly correlated returns. The groups (i.e., industries) formed will then have lower returns correlations between groups.

This method has several limitations:

- Historical correlations may not be the same as future correlations.
 - The groupings of firms may differ over time and across countries.
 - The grouping of firms is sometimes non-intuitive.
 - The method is susceptible to statistical error (i.e., firms can be grouped by a relationship that occurs by chance, or not grouped together when they should be).
-

LOS 40.c: Explain the factors that affect the sensitivity of a company to the business cycle and the uses and limitations of industry and company descriptors such as "growth," "defensive," and "cyclical."

A **cyclical firm** is one whose earnings are highly dependent on the stage of the business cycle. These firms have high earnings volatility and high operating leverage. Their products are often expensive, non-necessities whose purchase can be delayed until the economy improves.

Examples of cyclical industries include basic materials and processing, consumer discretionary, energy, financial services, industrial and producer durables, and technology.

In contrast, a **non-cyclical firm** produces goods and services for which demand is relatively stable over the business cycle. Examples of non-cyclical industries include health care, utilities, telecommunications, and consumer staples.

Sectors can also be classified by their sensitivity to the phase of the business cycle. Cyclical sector examples include energy, financials, technology, materials, and consumer discretionary. Non-cyclical sector examples include health care, utilities, and consumer staples.

Non-cyclical industries can be further separated into defensive (stable) or growth industries. **Defensive industries** are those that are least affected by the stage of the business cycle and include utilities, consumer staples (such as food producers), and basic services (such as drug stores). **Growth industries** have demand so strong they are largely unaffected by the stage of the business cycle.

Descriptors such as "growth," "defensive," and "cyclical" should be used with caution. The term **growth cyclical** is used to describe firms with strong long-term growth potential that have revenue that is quite sensitive to economic cycles. Cyclical industries, which are supposed to be

dependent on the business cycle, often include growth firms that are less dependent on the business cycle. Non-cyclical industries can be affected by severe recessions, as was the case in the 2008–09 downturn. Defensive industries may not always be safe investments. For example, grocery stores are classified as defensive, but they are subject to intense price competition that reduces earnings. Defensive industries may also contain some truly defensive and some growth firms. Because business cycle phases differ across countries and regions, two cyclical firms operating in different countries may be simultaneously experiencing different cyclical effects on earnings growth.

LOS 40.d: Describe current industry classification systems, and identify how a company should be classified, given a description of its activities and the classification system.

Classifying firms by industry provides a method of examining trends and firm valuations. It also allows analysts to compare firms in different countries on a similar basis. The following are the industry classification systems currently available to investors.

Commercial Classifications

Several index providers classify firms. Some use three levels of classification, while others use four levels. The providers generally use firm fundamentals such as revenue to classify firms. Although the nomenclature differs among providers, the broadest category is generally the sector level, followed by industry and sub-industry.

Commercial industry classifications include the Global Industry Classification Standard developed by Standard & Poor's and MSCI Barra, Russell Global Sectors, and the Industry Classification Benchmark developed by Dow Jones and FTSE.

Sectors and firm compositions representative of those used by commercial providers are as follows.

Basic materials and processing firms produce:

- Building materials.
- Chemicals.
- Paper and forest products.
- Containers and packaging.
- Metals, minerals, and mining.

Consumer discretionary firms are cyclical and sell goods and services in industries such as:

- Automotive.
- Apparel.
- Hotels and restaurants.

Consumer staples firms are less cyclical and sell goods and services in industries such as:

- Food.
- Beverage.
- Tobacco.

- Personal care products.

Energy firms are involved in:

- Energy exploration.
- Refining.
- Production.
- Energy equipment.
- Energy services.

Financial services firms include firms involved in:

- Banking.
- Insurance.
- Real estate financing.
- Asset management.
- Brokerage.

Health care includes:

- Pharmaceuticals.
- Biotech.
- Medical devices.
- Health care equipment.
- Medical supplies.
- Health care services.

Industrial and producer durables firms produce capital goods for commercial services industries including:

- Heavy machinery and equipment.
- Aerospace.
- Defense.
- Transportation.
- Commercial services and supplies.

Real estate firms are involved in the development, management, and operation of real properties, including:

- Real estate investment trusts (REITs).
- Real estate services firms.

Technology firms sell or produce:

- Computers.
- Software.
- Semiconductors.
- Communications equipment.

- Internet services.
- Electronic entertainment.
- Consulting and services.

Telecommunications firms include wired and wireless service providers. *Utilities* includes electric, gas, and water utilities. Some industry classification providers include telecommunication and utilities in the same group, while others separate them.

To classify a firm accurately, an analyst should have detailed knowledge about the firm and the delineation of industry classifications.

Government Classifications

Several government bodies also provide industry classification of firms. They frequently do so to organize the economic data they publish. A main thrust of their systems is to make comparisons of industries consistent across time and country. The main systems are similar to each other.

- *International Standard Industrial Classification of All Economic Activities* (ISIC) was produced by the United Nations in 1948 to increase global comparability of data.
- *Statistical Classification of Economic Activities in the European Community* is similar to the ISIC but is designed for Europe.
- *Australian and New Zealand Standard Industrial Classification* was jointly developed by those countries.
- *North American Industry Classification System* (NAICS) was jointly developed by the United States, Canada, and Mexico.

The methodologies that government providers use in their compilation of industry groups differ from those used by commercial providers. Most governments do not identify individual firms in a group, so an analyst cannot know the groups' exact composition. Commercial providers identify the constituent firms. Government systems are updated less frequently; for example, the NAICS is updated every five years. Governments do not distinguish between small and large firms, for-profit and not-for-profit organizations, or private and public firms. Commercial providers only include for-profit and public firms and can delineate by the size of the firm.

An analyst should not assume that two firms in the same narrowest industry classification can be compared with each other for fundamental analysis and valuation. Instead, the analyst should construct *peer groups*.

LOS 40.e: Explain how a company's industry classification can be used to identify a potential “peer group” for equity valuation.

A **peer group** is a set of similar companies an analyst will use for valuation comparisons. More specifically, a peer group will consist of companies with similar business activities, demand drivers, cost structure drivers, and availability of capital.

To form a peer group, an analyst will often start by identifying companies in the same industry classification, using the commercial classification providers previously described. Usually, the

analyst will use other information to verify that the firms in an industry are indeed peers. An analyst might include a company in more than one peer group.

The following are steps an analyst would use to form a peer group:

- Use commercial classification providers to determine which firms are in the same industry.
 - Examine firms' annual reports to see if they identify key competitors.
 - Examine competitors' annual reports to see if other competitors are named.
 - Use industry trade publications to identify competitors.
 - Confirm that comparable firms have similar sources of sales and earnings, have similar sources of demand, and are in similar geographic markets.
 - Adjust financial statements of non-financial companies for any financing subsidiary data they include.
-

LOS 40.f: Describe the elements that need to be covered in a thorough industry analysis.

A thorough industry analysis should include the following elements:

- Evaluate the relationships between macroeconomic variables and industry trends using information from industry groups, firms in the industry, competitors, suppliers, and customers.
 - Estimate industry variables using different approaches and scenarios.
 - Compare with other analysts' forecasts of industry variables to confirm the validity of the analysis and potentially find industries that are misvalued as a result of consensus forecasts.
 - Determine the relative valuation of different industries.
 - Compare the valuations of industries across time to determine the volatility of their performance over the long run and during different phases of the business cycle. This is useful for long-term investing as well as short-term industry rotation based on the current economic environment.
 - Analyze industry prospects based on **strategic groups**, which are groups of firms that are distinct from the rest of the industry due to the delivery or complexity of their products or barriers to entry. For example, full-service hotels are a distinct market segment within the hotel industry.
 - Classify industries by **life-cycle stage**, whether it is embryonic, growth, shakeout, mature, or declining.
 - Position the industry on the **experience curve**, which shows the cost per unit relative to output. The curve declines because of increases in productivity and economies of scale, especially in industries with high fixed costs.
 - Consider the forces that affect industries, which include demographic, macroeconomic, governmental, social, and technological influences.
 - Examine the forces that determine competition within an industry.
-

LOS 40.g: Describe the principles of strategic analysis of an industry.

Industries differ markedly in profitability because of differences in economic fundamentals, industry structure, and degree of competition. In some industries, competition is intense and few firms earn economic profits. **Economic profits**, the return on invested capital minus its cost, are greater than 20% in some industries and negative in others. The degree of economic profits depends in part on pricing power (elasticity of demand for the firm's products). An analyst should understand that industry conditions and profits can change dramatically over time, so industry analysis should be forward-looking.

One component of an analyst's industry analysis should be **strategic analysis**, which examines how an industry's competitive environment influences a firm's strategy. The analysis framework developed by Michael Porter¹ delineates five forces that determine industry competition.

1. *Rivalry among existing competitors.* Rivalry increases when many firms of relatively equal size compete within an industry. Slow growth leads to competition as firms fight for market share, and high fixed costs lead to price decreases as firms try to operate at full capacity. For example, the high fixed costs in the auto industry from capital investments and labor contracts force firms to produce a large number of vehicles that they can only sell at low margins. Industries with products that are undifferentiated or have barriers (are costly) to exit tend to have high levels of competition.
2. *Threat of entry.* Industries that have significant barriers to entry (e.g., large capital outlays for facilities) will find it easier to maintain premium pricing. It is costly to enter the steel or oil production industries. Those industries have large barriers to entry and thus less competition from newcomers. An analyst should identify factors that discourage new entrants, such as economies of scale.
3. *Threat of substitutes.* Substitute products limit the profit potential of an industry because they limit the prices firms can charge by increasing the elasticity of demand. Commodity-like products have high levels of competition and low profit margins. The more differentiated the products are within an industry, the less price competition there will be. For example, in the pharmaceutical industry, patents protect a producer from competition in the markets for patented drugs.
4. *Power of buyers.* Buyers' ability to bargain for lower prices or higher quality influences industry profitability. Bargaining by governments and ever-larger health care providers have put downward pressure even on patented drugs.
5. *Power of suppliers.* Suppliers' ability to raise prices or limit supply influences industry profitability. Suppliers are more powerful if there are just a few of them and their products are scarce. For example, Microsoft is one of the few suppliers of operating system software and thus has pricing power.

The first two forces deserve further attention because almost all firms must be concerned about the threat of new entrants and competition that would erode profits. Studying these forces also helps the analyst better understand the subject firm's competitors and prospects. The following summary describes how these two factors influence the competitive environment in an industry:

- Higher barriers to entry reduce competition.

- Greater concentration (a small number of firms control a large part of the market) reduces competition, whereas market fragmentation (a large number of firms, each with a small market share) increases competition.
- Unused capacity in an industry, especially if prolonged, results in intense price competition. For example, underutilized capacity in the auto industry has resulted in very competitive pricing.
- Stability in market share reduces competition. For example, loyalty of a firm's customers tends to stabilize market share and profits.
- More price sensitivity in customer buying decisions results in greater competition.
- Greater maturity of an industry results in slowing growth.



MODULE QUIZ 40.1

1. Industry classification systems from commercial index providers typically classify firms by:
 - A. statistical methods.
 - B. products and services.
 - C. business cycle sensitivity.
2. Firms and industries are *most appropriately* classified as cyclical or non-cyclical based on:
 - A. their stock price fluctuations relative to the market.
 - B. the sensitivity of their earnings to the business cycle.
 - C. the volatility of their earnings relative to a peer group.
3. An analyst should *most likely* include two firms in the same peer group for analysis if the firms:
 - A. are both grouped in the same industry classification.
 - B. are similar in size, industry life-cycle stage, and cyclicalities.
 - C. derive their revenue and earnings from similar business activities.
4. The industry experience curve shows the cost per unit relative to:
 - A. output.
 - B. age of firms.
 - C. industry life-cycle stage.
5. Which of the following is *least likely* an element of an industry strategic analysis?
 - A. Market correlations.
 - B. Demographic influences.
 - C. Influence of industry capacity on pricing.
6. Two of the five competitive forces in the Porter framework are:
 - A. threat of entry and barriers to exit.
 - B. power of suppliers and threat of substitutes.
 - C. rivalry among competitors and power of regulators.

MODULE 40.2: PRICING POWER AND COMPANY ANALYSIS



Video covering this content is available online.

LOS 40.h: Explain the effects of barriers to entry, industry concentration, industry capacity, and market share stability on pricing power and price competition.

Barriers to Entry

High barriers to entry benefit existing industry firms because they prevent new competitors from competing for market share and reducing the existing firms' return on capital. In industries with low barriers to entry, firms have little pricing power and competition reduces existing firms' return on capital. To assess the ease of entry, the analyst should determine how easily a new entrant to the industry could obtain the capital, intellectual property, and customer base needed to be successful. One method of determining the ease of entry is to examine the composition of the industry over time. If the same firms dominate the industry today as ten years ago, entry is probably difficult.

High barriers to entry do not necessarily mean firm pricing power is high. Industries with high barriers to entry may have strong price competition among existing firms. This is more likely when the products sold are undifferentiated and commodity-like or when high barriers to exit result in overcapacity. For example, an automobile factory may have a low value in an alternative use, making firm owners less likely to exit the industry. They continue to operate even when losing money, hoping to turn things around, which can result in industry overcapacity and intense price competition.

Low barriers to entry do not ensure success for new entrants. Barriers to entry may change over time, and so might the competitive environment.

Industry Concentration

High industry concentration does not guarantee pricing power.

- Absolute market share may not matter as much as a firm's market share relative to its competitors. A firm may have a 50% market share, but if a single competitor has the other 50%, their 50% share would not result in a great degree of pricing power. Return on capital is limited by intense competition between the two firms.
- Conversely, a firm that has a 10% market share when no competitor has more than 2% may have a good degree of pricing power and high return on capital.
- If industry products are undifferentiated and commodity-like, then consumers will switch to the lowest-priced producer. The more importance consumers place on price, the greater the competition in an industry. Greater competition leads to lower return on capital.
- Industries with greater product differentiation in regard to features, reliability, and service after the sale will have greater pricing power. Return on capital can be higher for firms that can better differentiate their products.
- If the industry is capital intensive, and therefore costly to enter or exit, overcapacity can result in intense price competition.

Tobacco, alcohol, and confections are examples of highly concentrated industries in which firms' pricing power is relatively strong. Automobiles, aircraft, and oil refining are examples of highly concentrated industries with relatively weak pricing power.

Although industry concentration does not guarantee pricing power, a fragmented market does usually result in strong competition. When there are many industry members, firms cannot coordinate pricing, firms will act independently, and because each member has such a small market share, any incremental increase in market share may make a price decrease profitable.

Industry Capacity

Industry capacity has a clear impact on pricing power. Undercapacity, a situation in which demand exceeds supply at current prices, results in pricing power and higher return on capital. Overcapacity, with supply greater than demand at current prices, will result in downward pressure on price and lower return on capital.

An analyst should be familiar with the industry's current capacity and its planned investment in additional capacity. Capacity is fixed in the short run and variable in the long run. In other words, given enough time, producers will build enough factories and raise enough capital to meet demand at a price close to minimum average cost. However, producers may overshoot the optimal industry capacity, especially in cyclical markets. For example, producers may start to order new equipment during an economic expansion to increase capacity. By the time they bring the additional production on to the market, the economy may be in a recession with decreased demand. A diligent analyst can look for signs that the planned capacity increases of all producers (who may not take into account the capacity increases of other firms) sum to more output than industry demand will support.

Capacity is not necessarily physical. For example, an increase in demand for insurance can be more easily and quickly met than an increase in demand for a product requiring physical capacity, such as electricity or refined petroleum products.

If capacity is physical and specialized, overcapacity can exist for an extended period if producers expand too much over the course of a business cycle. Specialized physical capacity may have a low liquidation value and be costly to reallocate to a different product. Non-physical capacity (e.g., financial capital) can be reallocated more quickly to new industries than physical capacity.

Market Share Stability

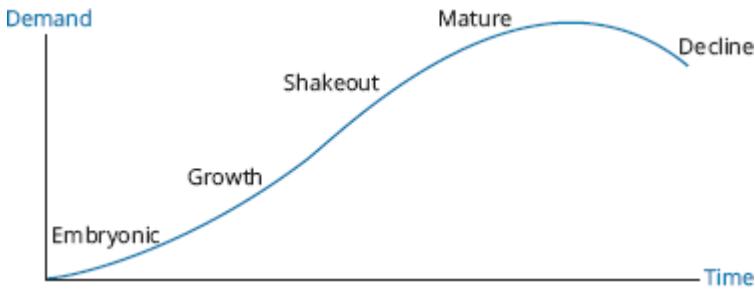
An analyst should examine whether firms' market shares in an industry have been stable over time. Market shares that are highly variable likely indicate a highly competitive industry in which firms have little pricing power. More stable market shares likely indicate less intense competition in the industry.

Factors that affect market share stability include barriers to entry, introductions of new products and innovations, and the **switching costs** that customers face when changing from one firm's products to another. Switching costs, such as the time and expense of learning to use a competitor's product, tend to be higher for specialized or differentiated products. High switching costs contribute to market share stability and pricing power.

LOS 40.i: Describe industry life-cycle models, classify an industry as to life-cycle stage, and describe limitations of the life-cycle concept in forecasting industry performance.

Industry life cycle analysis should be a component of an analyst's strategic analysis. An industry's stage in the cycle has an impact on industry competition, growth, and profits. An industry's stage will change over time, so the analyst must monitor the industry on an ongoing basis. The five phases of the industry life-cycle model are illustrated in Figure 40.1.

Figure 40.1: Stages of the Industry Life Cycle



In the **embryonic stage**, the industry has just started. The characteristics of this stage are as follows:

- *Slow growth*: customers are unfamiliar with the product.
- *High prices*: the volume necessary for economies of scale has not been reached.
- *Large investment required*: to develop the product.
- *High risk of failure*: most embryonic firms fail.

In the **growth stage**, industry growth is rapid. The characteristics of this stage are as follows:

- *Rapid growth*: new consumers discover the product.
- *Limited competitive pressures*: the threat of new firms coming into the market peaks during the growth phase, but rapid growth allows firms to grow without competing on price.
- *Falling prices*: economies of scale are reached and distribution channels increase.
- *Increasing profitability*: due to economies of scale.

In the **shakeout stage**, industry growth and profitability are slowing due to strong competition. The characteristics of this stage are as follows:

- *Growth has slowed*: demand reaches saturation level with few new customers to be found.
- *Intense competition*: industry growth has slowed, so firm growth must come at the expense of competitors.
- *Increasing industry overcapacity*: firm investment exceeds increases in demand.
- *Declining profitability*: due to overcapacity.
- *Increased cost cutting*: firms restructure to survive and attempt to build brand loyalty.
- *Increased failures*: weaker firms liquidate or are acquired.

In the **mature stage**, there is little industry growth and firms begin to consolidate. The characteristics of this stage are as follows:

- *Slow growth*: market is saturated and demand is only for replacement.
- *Consolidation*: market evolves to an oligopoly.
- *High barriers to entry*: surviving firms have brand loyalty and low cost structures.
- *Stable pricing*: firms try to avoid price wars, although periodic price wars may occur during recessions.
- *Superior firms gain market share*: the firms with better products may grow faster than the industry average.

In the **decline stage**, industry growth is negative. The characteristics of this stage are as follows:

- *Negative growth*: due to development of substitute products, societal changes, or global competition.
- *Declining prices*: competition is intense and there are price wars due to overcapacity.
- *Consolidation*: failing firms exit or merge.

An analyst should determine whether a firm is “acting its age” or stage of industry development. Growth firms should be reinvesting in operations in an attempt to increase product offerings, increase economies of scale, and build brand loyalty. They are not yet worried about cost efficiency. They should not pay out cash flows to investors but save them for internal growth. On the other hand, mature firms focus on cost efficiency because demand is largely from replacement. They find few opportunities to introduce new products. These firms should typically pay out cash to investors as dividends or stock repurchases because cash flows are strong but internal growth is limited. An analyst should be concerned about firms that do not act their stage, such as a mature firm that is investing in low-return projects for the sake of increasing firm size.

Although life-cycle analysis is a useful tool, industries do not always conform to its framework. Life-cycle stages may not be as long or short as anticipated, or they might be skipped altogether. An industry’s product may become obsolete quickly due to technological change, government regulation, societal change, or demographics. Life-cycle analysis is likely most useful during stable periods, not during periods of upheaval when conditions are changing rapidly. Furthermore, some firms will experience growth and profits that are dissimilar to others in their industries due to competitive advantages or disadvantages.

LOS 40.j: Describe macroeconomic, technological, demographic, governmental, social, and environmental influences on industry growth, profitability, and risk.

The external influences on industry growth, profitability, and risk should be a component of an analyst’s strategic analysis. These external factors include macroeconomic, technological, demographic, governmental, and social influences.

Macroeconomic factors can be cyclical or structural (longer-term) trends, most notably economic output as measured by GDP or some other measure. Interest rates affect financing costs for firms and individuals, as well as financial institution profitability. Credit availability affects consumer and business expenditures and funding. Inflation affects costs, prices, interest rates, and business and consumer confidence. An example of a structural economic factor is the education level of the work force. More education can increase workers’ productivity and real wages, which in turn can increase their demand for consumer goods.

Technology can change an industry dramatically through the introduction of new or improved products. Computer hardware is an example of an industry that has undergone dramatic transformation. Radical improvements in circuitry were assisted by transformations in other industries, including the computer software and telecommunications industries. Another

example of an industry that has been changed by technology is photography, which has largely moved from film to digital media.

Demographic factors include age distribution and population size, as well as other changes in the composition of the population. As a large segment of the population reaches their twenties, residential construction, furniture, and related industries see increased demand. An aging of the overall population can mean significant growth for the health care industry and developers of retirement communities. For example, the aging of the post-World War II Baby Boomers is an example of demographics that will increase demand in these industries.

Governments have an important and widespread effect on businesses through various channels, including taxes and regulation. The level of tax rates certainly affects industries, but analysts should also be aware of the differential taxation applied to some goods. For example, tobacco is heavily taxed in the United States. Specific regulations apply to many industries. Entry into the health care industry, for example, is controlled by governments that license doctors and other providers. Governments can also empower self-regulatory organizations, such as stock exchanges that regulate their members. Some industries, such as the U.S. defense industry, depend heavily on government purchases of goods and services.

Social influences relate to how people work, play, spend their money, and conduct their lives; these factors can have a large impact on industries. For example, when women entered the U.S. workforce, the restaurant industry benefitted because there was less cooking at home. Child care, women's clothing, and other industries were also dramatically affected.

Environmental influences, such as climate change and concerns about environmental sustainability of some industries, continue to grow in importance as a factor in industry growth and profitability. Three primary considerations are as follows:

1. Consumer perceptions about industries and specific companies can have a significant impact on growth and profitability.
2. Increased government regulation of some industries will produce profound changes in how they operate and the drivers of demand for products in those industries.
3. Changing weather patterns and shortages of water and energy will affect various industries, such as agriculture and utilities.

Some examples of industries and businesses affected by environmental factors are the following:

- Perceptions that raising cattle may be cruel to animals, produces significant greenhouse gases, and require a disproportionately large amount of resources is changing that industry. Other businesses, notably producers of plant-based meat substitutes, have experienced tremendous growth.
- Other types of agriculture will likely be significantly affected by environmental concerns and resource shortages—and quite possibly by the near-term effects of climate change.
- Airlines have been criticized for their large contribution to carbon emissions.
- The automobile industry is changing rapidly, with government regulation increasing and some outright bans on internal combustion vehicles in favor of electric-powered vehicles on the horizon. There is an associated perception that oil producers will find themselves facing a much smaller industry.

- Relatedly, the electric vehicle segment of the auto and truck industry is growing rapidly, along with battery manufacture, the construction of vehicle charging stations, and the alternative energy industry in general.
 - Energy producers may face increasing regulation, especially related to fracking and the release of natural gas into the atmosphere.
 - New businesses (and possibly industries) focused on combating climate change, improving sustainability, and reducing the negative effects of economic activity on the environment will likely experience rapid growth in the coming decades.
-

LOS 40.k: Compare characteristics of representative industries from the various economic sectors.

To illustrate the long list of factors to be considered in industry analysis, we use the following strategic analysis of the candy/confections industry.

- *Major firms:* Cadbury, Hershey, Mars, and Nestle.
 - *Barriers to entry and success:* Very high. Low capital and technological barriers, but consumers have strong brand loyalty.
 - *Industry concentration:* Very concentrated. Largest four firms dominate global market share.
 - *Influence of industry capacity on pricing:* None. Pricing is determined by strength of brand, not production capacity.
 - *Industry stability:* Very stable. Market share changes slowly.
 - *Life cycle:* Very mature. Growth is driven by population changes.
 - *Competition:* Low. Lack of unbranded candy makers in market reduces competition. Consumer decision is based on brand awareness, not price.
 - *Demographic influences:* Not applicable.
 - *Government influence:* Low. Industry is largely unregulated, but regulation arising from concerns about obesity is possible.
 - *Social influence:* Not applicable.
 - *Technological influence:* Very low. Limited impact from technology.
 - *Business cycle sensitivity:* Non-cyclical and defensive. Demand for candy is very stable.
-

LOS 40.l: Describe the elements that should be covered in a thorough company analysis.

Having gained understanding of an industry's external environment, an analyst can then focus on **company analysis**. This involves analyzing the firm's financial condition, products and services, and **competitive strategy**. Competitive strategy is how a firm responds to the opportunities and threats of the external environment. The strategy may be defensive or offensive.

Porter has identified two important competitive strategies that can be employed by firms within an industry: a **cost leadership (low-cost) strategy** or a **product or service**

differentiation strategy. According to Porter, a firm must decide to focus on one of these two areas to compete effectively.

In a *low-cost strategy*, the firm seeks to have the lowest costs of production in its industry, offer the lowest prices, and generate enough volume to make a superior return. The strategy can be used defensively to protect market share or offensively to gain market share. If industry competition is intense, pricing can be aggressive or even predatory. In **predatory pricing**, the firm hopes to drive out competitors and later increase prices. Although there are often laws prohibiting predatory pricing, it can be hard to prove if the firm's costs are not easily traced to a particular product. A low-cost strategy firm should have managerial incentives that are geared toward improving operating efficiency.

In a *differentiation strategy*, the firm's products and services should be distinctive in terms of type, quality, or delivery. For success, the firm's cost of differentiation must be less than the price premium buyers place on product differentiation. The price premium should also be sustainable over time. Successful differentiators will have outstanding marketing research teams and creative personnel.

A company analysis should include the following elements:

- Firm overview, including information on operations, governance, and strengths and weaknesses.
- Industry characteristics.
- Product demand.
- Product costs.
- Pricing environment.
- Financial ratios, with comparisons to other firms and over time.
- Projected financial statements and firm valuation.

A firm's return on equity (ROE) should be part of the financial analysis. The ROE is a function of profitability, total asset turnover, and financial leverage (debt).

Analysts often use **spreadsheet modeling** to analyze and forecast company fundamentals. The problem with this method is that the models' complexity can make their conclusions seem precise. However, estimation is performed with error that can compound over time. As a check on a spreadsheet model's output, an analyst should consider which factors are likely to be different going forward and how this will affect the firm. Analysts should also be able to explain the assumptions of a spreadsheet model.



MODULE QUIZ 40.2

1. Greater pricing power is *most likely* to result from greater:
 - A. unused capacity.
 - B. market concentration.
 - C. volatility in market share.
2. Which of the following statements *best* describes the relationship between pricing power and ease of entry and exit? Greater ease of entry:
 - A. and greater ease of exit decrease pricing power.
 - B. and greater ease of exit increase pricing power.
 - C. decreases pricing power and greater ease of exit increases pricing power.

3. Industry overcapacity and increased cost cutting characterize which stage of the industry life cycle?
 - A. Growth.
 - B. Shakeout.
 - C. Maturity.
4. In which of these characteristics is the oil producing industry *most likely* similar to the home building industry?
 - A. Industry concentration.
 - B. Demographic influences.
 - C. Business cycle sensitivity.
5. Which of the following is *least likely* a significant external influence on industry growth?
 - A. Social influences.
 - B. Macroeconomic factors.
 - C. Supplier bargaining power.
6. Which of the following *best* describes a low-cost competitive strategy?
 - A. Volume sold is typically modest.
 - B. Managerial incentives promote operational efficiency.
 - C. Success depends heavily on creative marketing and product development.

KEY CONCEPTS

LOS 40.a

Industry analysis is necessary for understanding a company's business environment before engaging in analysis of the company. The industry environment can provide information about the firm's potential growth, competition, risks, appropriate debt levels, and credit risk.

Industry valuation can be used in an active management strategy to determine which industries to overweight or underweight in a portfolio.

Industry representation is often a component in a performance attribution analysis of a portfolio's return.

LOS 40.b

Firms can be grouped into industries according to their products and services or business cycle sensitivity, or through statistical methods that group firms with high historical correlation in returns.

LOS 40.c

A cyclical firm has earnings that are highly dependent on the business cycle. A non-cyclical firm has earnings that are less dependent on the business cycle. Industries can also be classified as cyclical or non-cyclical. Non-cyclical industries or firms can be classified as defensive (demand for the product tends not to fluctuate with the business cycle) or growth (demand is so strong that it is largely unaffected by the business cycle).

Limitations of descriptors such as growth, defensive, and cyclical include the facts that cyclical industries often include growth firms; even non-cyclical industries can be affected by severe recessions; defensive industries are not always safe investments; business cycle timing differs across countries and regions; and the classification of firms is somewhat arbitrary.

LOS 40.d

Industry classification systems from commercial providers include the Global Industry Classification Standard (Standard & Poor's and MSCI Barra), Russell Global Sectors, and the Industry Classification Benchmark (Dow Jones and FTSE).

Industry classification systems developed by government agencies include the International Standard Industrial Classification (ISIC), the North American Industry Classification System (NAICS), and systems designed for the European Union and Australia/New Zealand.

LOS 40.e

A peer group should consist of companies with similar business activities, demand drivers, cost structure drivers, and availability of capital. To form a peer group, the analyst will often start by identifying companies in the same industry, but the analyst should use other information to verify that the firms in an industry are comparable.

LOS 40.f

A thorough industry analysis should:

- Evaluate the relationships between macroeconomic variables and industry trends.
- Estimate industry variables using different approaches and scenarios.
- Check estimates against those from other analysts.
- Compare the valuation for different industries.
- Compare the valuation for industries across time to determine risk and rotation strategies.
- Analyze industry prospects based on strategic groups.
- Classify industries by their life-cycle stage.
- Position the industry on the experience curve.
- Consider demographic, macroeconomic, governmental, social, and technological influences.
- Examine the forces that determine industry competition.

LOS 40.g

Strategic analysis of an industry involves analyzing the competitive forces that determine the possibility of economic profits.

Porter's five forces that determine industry competition are:

1. Rivalry among existing competitors.
2. Threat of entry.
3. Threat of substitutes.
4. Power of buyers.
5. Power of suppliers.

LOS 40.h

High barriers to entry prevent new competitors from taking away market share, but they do not guarantee pricing power or high return on capital, especially if the products are undifferentiated or barriers to exit result in overcapacity. Barriers to entry may change over time.

While market fragmentation usually results in strong competition and low return on capital, high industry concentration may not guarantee pricing power. If industry products are

undifferentiated, consumers will switch to the cheapest producer. Overcapacity may result in price wars.

Capacity is fixed in the short run and variable in the long run. Undercapacity typically results in pricing power. Producers may overinvest in new capacity, especially in cyclical industries or if the capacity is physical and specialized. Non-physical capacity comes into production and can be reallocated more quickly than physical capacity.

Highly variable market shares indicate a highly competitive industry. Stable market shares suggest less intense competition. High switching costs contribute to market share stability.

LOS 40.i

Phases of the industry life-cycle model are the embryonic, growth, shakeout, maturity, and decline stages.

- Embryonic stage: Slow growth; high prices; large investment required; high risk of failure.
- Growth stage: Rapid growth; little competition; falling prices; increasing profitability.
- Shakeout stage: Slowing growth; intense competition; industry overcapacity; declining profitability; cost cutting; increased failures.
- Mature stage: Slow growth; consolidation; high barriers to entry; stable pricing; superior firms gain market share.
- Decline stage: Negative growth; declining prices; consolidation.

A limitation of life-cycle analysis is that life-cycle stages may not be as long or short as anticipated or might be skipped altogether due to technological change, government regulation, societal change, or demographics. Firms in the same life-cycle stage will experience dissimilar growth and profits due to their competitive positions.

LOS 40.j

Macroeconomic influences on industries include long-term trends in factors such as GDP growth, interest rates, and inflation, as well as structural factors such as the education level of the workforce.

Demographic influences include the size and age distribution of the population.

Government factors include tax rates, regulations, empowerment of self-regulatory organizations, and government purchases of goods and services.

Social influences relate to how people interact and conduct their lives.

Technology can dramatically change an industry through the introduction of new or improved products.

Environmental influences continue to grow in importance as a factor in industry growth and profitability.

LOS 40.k

The elements of an industry strategic analysis are the major firms, barriers to entry, industry concentration, influence of industry capacity on pricing, industry stability, life cycle, competition, demographic influences, government influence, social influence, technological influence, and whether the industry is growth, defensive, or cyclical.

LOS 40.1

Company analysis should include an overview of the firm, industry characteristics, and analysis of product demand, product costs, the pricing environment, the firm's financial ratios, and projected financial statements and firm valuation. The analysis should describe the company's competitive strategy.

Companies can employ a cost leadership (low-cost) strategy or a product or service differentiation strategy. A cost leadership firm seeks to have the lowest costs of production in its industry, offer the lowest prices, and generate enough volume to make a superior return. A differentiating firm's products and services should be distinctive in terms of type, quality, or delivery.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 40.1

1. **B** The classification systems provided by S&P/MSCI Barra, Russell, and Dow Jones/FTSE classify firms according to the product or service they produce. (LOS 40.d)
2. **B** For industry analysis, cyclical firms and industries are those with earnings that are highly dependent on the business cycle, while non-cyclical firms and industries are those with earnings that are relatively less sensitive to the business cycle. (LOS 40.c)
3. **C** Firms should be included in a peer group if their business activities are comparable. An analyst may begin with available industry classifications when forming peer groups but should refine them based on factors including the firms' sources of demand and earnings and the geographic markets in which they operate. (LOS 40.e)
4. **A** The experience curve shows the cost per unit relative to output. Unit cost declines at higher output volume because of increases in productivity and economies of scale, especially in industries with high fixed costs. (LOS 40.f)
5. **A** Elements of an industry strategic analysis include the major firms, barriers to entry/success, industry concentration, influence of industry capacity on pricing, industry stability, life cycle, competition, demographic influences, government influence, social influence, technological influence, and whether the industry is growth, defensive, or cyclical. (LOS 40.g)
6. **B** Porter's five forces are rivalry among existing competitors, threat of entry, threat of substitutes, bargaining power of buyers, and bargaining power of suppliers. (LOS 40.g)

Module Quiz 40.2

1. **B** Greater concentration (a small number of firms control a large part of the market) typically reduces competition and results in greater pricing power. Greater unused capacity in an industry, especially if chronic, results in greater price competition and less pricing power. Greater stability in market share is typically associated with greater pricing power. (LOS 40.h)

2. **C** In industries with greater ease of entry, firms have little pricing power because new competitors can take away market share. High costs of exiting result in overcapacity and likely price wars. Greater ease of exit (i.e., low costs of exit) increases pricing power. (LOS 40.h)
 3. **B** The shakeout stage is characterized by slowed growth, intense competition, industry overcapacity, increased cost cutting, declining profitability, and increased failures. (LOS 40.i)
 4. **C** Oil production and home building are both highly cyclical industries. Oil production is dominated by a small number of large global firms, while home construction is characterized by a large number of relatively smaller firms. Demographics have more influence on housing construction (e.g., the rate of new household formation) than on oil production. (LOS 40.k)
 5. **C** Supplier bargaining power is best characterized as a force internal to the industry. External influences on industry growth, profitability, and risk include macroeconomic, technological, demographic, governmental, and social influences. (LOS 40.j)
 6. **B** Firms that use a low-cost strategy should have managerial incentives suitable to create efficient operations. In a low-cost strategy, the firm seeks to generate high enough sales volume to make a superior return. Marketing and product development are key elements of a differentiation strategy. (LOS 40.l)
-

¹ Michael Porter, "The Five Competitive Forces That Shape Strategy," *Harvard Business Review*, Volume 86, No. 1: pp. 78–93.

READING 41

EQUITY VALUATION: CONCEPTS AND BASIC TOOLS

EXAM FOCUS

This reading discusses the use of discounted cash flow models, price multiples, and asset-based models for stock valuation. Know when the various models are appropriate, how to apply them, and their advantages and disadvantages. This topic is foundational material for all three levels of the CFA exams. Be sure you understand these fundamental concepts.

MODULE 41.1: DIVIDENDS, SPLITS, AND REPURCHASES



Video covering this content is available online.

LOS 41.a: Evaluate whether a security, given its current market price and a value estimate, is overvalued, fairly valued, or undervalued by the market.

Recall from the reading on Market Efficiency that **intrinsic value** or **fundamental value** is defined as the rational value investors would place on the asset if they had full knowledge of the asset's characteristics. Analysts use valuation models to estimate the intrinsic values of stocks and compare them to the stocks' market prices to determine whether individual stocks are overvalued, undervalued, or fairly valued. In doing valuation analysis for stocks, analysts are assuming that some stocks' prices deviate significantly from their intrinsic values.

To the extent that market prices deviate from intrinsic values, analysts who can estimate a stock's intrinsic value better than the market can earn abnormal profits if the stock's market price moves toward its intrinsic value over time. There are several things to consider, however, in deciding whether to invest based on differences between market prices and estimated intrinsic values.

1. The larger the percentage difference between market prices and estimated values, the more likely the investor is to take a position based on the estimate of intrinsic value. Small differences between market prices and estimates of intrinsic values are to be expected.
2. The more confident the investor is about the appropriateness of the valuation model used, the more likely the investor is to take an investment position in a stock that is identified as overvalued or undervalued.
3. The more confident the investor is about the estimated inputs used in the valuation model, the more likely the investor is to take an investment position in a stock that is identified as

overvalued or undervalued. Analysts must also consider the sensitivity of a model value to each of its inputs in deciding whether to act on a difference between model values and market prices. If a decrease of one-half percent in the long-term growth rate used in the valuation model would produce an estimated value equal to the market price, an analyst would have to be quite sure of the model's growth estimate to take a position in the stock based on its estimated value.

4. Even if we assume that market prices sometimes deviate from intrinsic values, market prices must be treated as fairly reliable indications of intrinsic value. Investors must consider why a stock is mispriced in the market. Investors may be more confident about estimates of value that differ from market prices when few analysts follow a particular security.
 5. Finally, to take a position in a stock identified as mispriced in the market, an investor must believe that the market price will actually move toward (and certainly not away from) its estimated intrinsic value and that it will do so to a significant extent within the investment time horizon.
-

LOS 41.b: Describe major categories of equity valuation models.

Analysts use a variety of models to estimate the value of equities. Usually, an analyst will use more than one model with several different sets of inputs to determine a range of possible stock values.

In **discounted cash flow models** (or **present value models**), a stock's value is estimated as the present value of cash distributed to shareholders (*dividend discount models*) or the present value of cash available to shareholders after the firm meets its necessary capital expenditures and working capital expenses (*free cash flow to equity models*).

There are two basic types of **multiplier models** (or **market multiple models**) that can be used to estimate intrinsic values. In the first type, the ratio of stock price to such fundamentals as earnings, sales, book value, or cash flow per share is used to determine if a stock is fairly valued. For example, the price to earnings (P/E) ratio is frequently used by analysts.

The second type of multiplier model is based on the ratio of **enterprise value** to either earnings before interest, taxes, depreciation, and amortization (EBITDA) or revenue. Enterprise value is the market value of all a firm's outstanding securities minus cash and short-term investments. Common stock value can be estimated by subtracting the value of liabilities and preferred stock from an estimate of enterprise value.

In **asset-based models**, the intrinsic value of common stock is estimated as total asset value minus liabilities and preferred stock. Analysts typically adjust the book values of the firm's assets and liabilities to their fair values when estimating the market value of its equity with an asset-based model.

LOS 41.c: Describe regular cash dividends, extra dividends, stock dividends, stock splits, reverse stock splits, and share repurchases.

Cash dividends, as the name implies, are payments made to shareholders in cash. They may be regularly scheduled dividends or one-time special dividends. **Regular dividends** occur when a company pays out a portion of profits on a consistent schedule (e.g., quarterly). A long-term record of stable or increasing dividends is widely viewed by investors as a sign of a company's financial stability. **Special dividends** are used when favorable circumstances allow the firm to make a one-time cash payment to shareholders, in addition to any regular dividends the firm pays. Many cyclical firms (e.g., automakers) will use a special dividend to share profits with shareholders when times are good but maintain the flexibility to conserve cash when profits are poor. Other names for special dividends include *extra dividends* and *irregular dividends*.

Stock dividends are dividends paid out in new shares of stock rather than cash. In this case, there will be more shares outstanding, but each one will be worth less. Total shareholders' equity remains unchanged. Stock dividends are commonly expressed as a percentage. A 20% stock dividend means every shareholder gets 20% more stock.

Stock splits divide each existing share into multiple shares, creating more shares. There are now more shares, but the price of each share will drop correspondingly to the number of shares created, so there is no change in the owner's wealth. Splits are expressed as a ratio. In a 3-for-1 stock split, each old share is split into three new shares. Stock splits are currently more common than stock dividends.

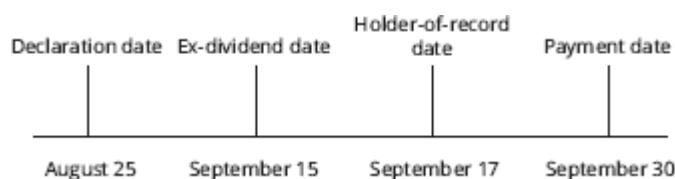
Reverse stock splits are the opposite of stock splits. After a reverse split, there are fewer shares outstanding but there is a higher stock price. Because these factors offset one another, shareholder wealth is unchanged.

A **share repurchase** is a transaction in which a company buys outstanding shares of its own common stock. Share repurchases are an alternative to cash dividends as a way of distributing cash to shareholders, and they have the same effect on shareholders' wealth as cash dividends of the same size. A company might repurchase shares to support their price or to signal that management believes the shares are undervalued. Share repurchases may also be used to offset an increase in outstanding shares from the exercise of employee stock options. In countries that tax capital gains at lower rates than dividends, shareholders may prefer share repurchases to dividend payments as a way to distribute cash to shareholders.

LOS 41.d: Describe dividend payment chronology.

The dates relevant to dividend payments are shown in Figure 41.1.

Figure 41.1: Dividend Payment Chronology



Declaration date. The date the board of directors approves payment of a dividend, specifying the per-share dividend amount, the date shareholders must own the stock to receive the dividend (record date), and the date the dividend payment will be made (payment date).

Ex-dividend date. The first day on which a share purchaser will not receive the next dividend. The ex-dividend date is one or two business days before the holder-of-record date, depending on the settlement period for stock purchases. If you buy the share on or after the ex-dividend date, you will not receive the dividend.

Holder-of-record date (record date). The date on which all owners of shares become entitled to receive the dividend payment on their shares.

Payment date. The date dividend checks are mailed to, or payment is made electronically to, holders of record.

On the ex-dividend date, the share price will decrease from the previous day's closing price by approximately the amount of the dividend, in the absence of other factors affecting the stock price. Consider shares that are trading at \$25 on the day prior to the ex-dividend date and will pay a \$1 dividend. Purchasing a share on the day prior to the ex-dividend date will give the owner a share of stock and the \$1 dividend on the payment date. Purchasing a share on the ex-dividend date will entitle the owner only to the share; the dividend payment will go to the seller.



MODULE QUIZ 41.1

1. An analyst estimates a value of \$45 for a stock with a market price of \$50. The analyst is *most likely* to conclude that a stock is overvalued if:
 - A. few analysts follow the stock and the analyst has less confidence in his model inputs.
 - B. few analysts follow the stock and the analyst is confident in his model inputs.
 - C. many analysts follow the stock and the analyst is confident in his model inputs.
2. A valuation model based on free cash flow to equity is *most likely* to be:
 - A. a multiplier model.
 - B. an asset-based model.
 - C. a present value model.
3. A company is evaluating the likely effects on its share price of declaring a 50% stock dividend or a 3-for-2 stock split. Other things equal, which of these will result in a lower share price?
 - A. 3-for-2 stock split.
 - B. 50% stock dividend.
 - C. Both should have the same effect.
4. The first date on which the purchaser of a stock will not receive a dividend that has been declared is:
 - A. the declaration date.
 - B. the ex-dividend date.
 - C. the holder-of-record date.

MODULE 41.2: DIVIDEND DISCOUNT MODELS



LOS 41.e: Explain the rationale for using present value models to value equity and describe the dividend discount and free-cash-flow-to-equity models.

Video covering this content is available online.

The **dividend discount model (DDM)** is based on the rationale that the intrinsic value of stock is the present value of its future dividends.

The most general form of the model is as follows:

$$V_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1 + k_e)^t}$$

where:

V_0 = current stock value

D_t = dividend at time t

k_e = required rate of return on common equity

One-year holding period DDM. For a holding period of one year, the value of the stock today is the present value of any dividends during the year plus the present value of the expected price of the stock at the end of the year (referred to as its **terminal value**).

The one-year holding period DDM is simply:

$$\text{value} = \frac{\text{dividend to be received}}{(1 + k_e)} + \frac{\text{year-end price}}{(1 + k_e)}$$

EXAMPLE: One-period DDM valuation

Calculate the value of a stock that paid a \$1 dividend last year, if next year's dividend will be 5% higher and the stock will sell for \$13.45 at year-end. The required return is 13.2%.

Answer:

The next dividend is the current dividend increased by the estimated growth rate. In this case, we have:

$$D_1 = D_0 \times (1 + \text{dividend growth rate}) = \$1.00 \times (1 + 0.05) = \$1.05$$

The present value of the expected future cash flows is:

$$\text{dividend: } \frac{\$1.05}{1.132} = \$0.93$$

$$\text{year-end price: } \frac{\$13.45}{1.132} = \$11.88$$

The current value based on the investor's expectations is:

$$\text{stock value} = \$0.93 + \$11.88 = \$12.81$$

Multiple-year holding period DDM. With a multiple-year holding period, we simply sum the present values of the estimated dividends over the holding period and the estimated terminal value.

For a two-year holding period, we have:

$$\text{value} = \frac{D_1}{(1 + k_e)} + \frac{D_2}{(1 + k_e)^2} + \frac{P_2}{(1 + k_e)^2}$$



PROFESSOR'S NOTE

It is useful to think of the subscript t on dividends (D_t) and prices (P_t) as the end of period t . For example, in the preceding equation, P_2 is the price at the end of Year 2.

Think of it as the selling price of a share, immediately after D_2 is received.

EXAMPLE: Multiple-period DDM valuation

A stock recently paid a dividend of \$1.50 which is expected to grow at 8% per year. The required rate of return of 12%. Calculate the value of this stock assuming that it will be priced at \$51.00 three years from now.

Answer:

Find the PV of the future dividends:

$$D_1 = \$1.50(1.08) = \$1.62$$

$$D_2 = \$1.50(1.08)^2 = \$1.75$$

$$D_3 = \$1.50(1.08)^3 = \$1.89$$

$$\text{PV of dividends} = \$1.62 / 1.12 + \$1.75 / (1.12)^2 + \$1.89 / (1.12)^3 = \$4.19$$

Find the PV of the future price:

$$\$51.00 / (1.12)^3 = \$36.30$$

Add the present values. The current value based on the investor's expectations is $\$4.19 + \$36.30 = \$40.49$.

The most general form of the DDM uses an infinite holding period because a corporation has an indefinite life. In an infinite-period DDM model, the present value of all expected future dividends is calculated and there is no explicit terminal value for the stock. In practice, as we will see, a terminal value can be calculated at a time in the future after which the growth rate of dividends is expected to be constant.

Free cash flow to equity (FCFE) is often used in discounted cash flow models instead of dividends because it represents the potential amount of cash that could be paid out to common shareholders. That is, FCFE reflects the firm's capacity to pay dividends. FCFE is also useful for firms that do not currently pay dividends.

FCFE is defined as the cash remaining after a firm meets all of its debt obligations and provides for the capital expenditures necessary to maintain existing assets and to purchase the new assets needed to support the assumed growth of the firm. In other words, it is the cash available to the firm's equity holders after a firm meets all of its other obligations. FCFE for a period is often calculated as:

$$\text{FCFE} = \text{net income} + \text{depreciation} - \text{increase in working capital} - \text{fixed capital investment (FCInv)} - \text{debt principal repayments} + \text{new debt issues}$$

FCFE can also be calculated as:

$$\text{FCFE} = \text{cash flow from operations} - \text{FCInv} + \text{net borrowing}$$

In the second formula, **net borrowing** is the increase in debt during the period (i.e., amount borrowed minus amount repaid) and is assumed to be available to shareholders. Fixed capital

investment must be subtracted because the firm must invest in assets to sustain itself. FCFE is projected for future periods using the firm's financial statements.

Restating the general form of the DDM in terms of FCFE, we have:

$$V_0 = \sum_{t=1}^{\infty} \frac{FCFE_t}{(1 + k_e)^t}$$

Estimating the Required Return for Equity

The capital asset pricing model (CAPM) provides an estimate of the required rate of return (k_i) for security i as a function of its systematic risk (β_i), the risk-free rate (R_f), and the expected return on the market [$E(R_{mkt})$] as:

$$k_i = R_f + \beta_i [E(R_{mkt}) - R_f]$$

There is some controversy over whether the CAPM is the best model to calculate the required return on equity. Also, different analysts will likely use different inputs, so there is no single number that is correct.



PROFESSOR'S NOTE

The CAPM is discussed in detail in Portfolio Management.

Recall from the Corporate Issuers reading on the cost of capital that for firms with publicly traded debt, analysts often estimate the required return on the firm's common equity by adding a risk premium to the firm's current bond yield. If the firm does not have publicly traded debt, an analyst can add a larger risk premium to a government bond yield.

LOS 41.g: Calculate the intrinsic value of a non-callable, non-convertible preferred stock.



PROFESSOR'S NOTE

At the end of this reading we will address the LOS that concerns advantages and disadvantages of each category of valuation model.

Preferred stock pays a dividend that is usually fixed and usually has an indefinite maturity. When the dividend is fixed and the stream of dividends is infinite, the infinite period dividend discount model reduces to a simple ratio:

$$\text{preferred stock value} = \frac{D_p}{(1 + k_p)^1} + \frac{D_p}{(1 + k_p)^2} + \dots + \frac{D_p}{(1 + k_p)^x} = \frac{D_p}{k_p}$$

EXAMPLE: Preferred stock valuation

A company's \$100 par preferred stock pays a \$5.00 annual dividend and has a required return of 8%. Calculate the value of the preferred stock.

Answer:

Value of the preferred stock: $D_p / k_p = \$5.00 / 0.08 = \62.50

LOS 41.h: Calculate and interpret the intrinsic value of an equity security based on the Gordon (constant) growth dividend discount model or a two-stage dividend discount model, as appropriate.

The **Gordon growth model** (or **constant growth model**) assumes the annual growth rate of dividends, g_c , is constant. Hence, next period's dividend, D_1 , is $D_0(1 + g_c)$, the second year's dividend, D_2 , is $D_0(1 + g_c)^2$, and so on. The extended equation using this assumption gives the present value of the expected future dividends (V_0) as:

$$V_0 = \frac{D_0(1 + g_c)}{(1 + k_e)} + \frac{D_0(1 + g_c)^2}{(1 + k_e)^2} + \frac{D_0(1 + g_c)^3}{(1 + k_e)^3} + \dots + \frac{D_0(1 + g_c)^\infty}{(1 + k_e)^\infty}$$

When the growth rate of dividends is constant, this equation simplifies to the Gordon (constant) growth model:

$$V_0 = \frac{D_0(1 + g_c)}{k_e - g_c} = \frac{D_1}{k_e - g_c}$$



PROFESSOR'S NOTE

In much of the finance literature, you will see this model referred to as the constant growth DDM, infinite period DDM, or the Gordon growth model. Whatever you call it, memorize D_1 over (k minus g). Note that our valuation model for preferred stock is the same as the constant growth model with no growth ($g = 0$).

The assumptions of the Gordon growth model are:

- Dividends are the appropriate measure of shareholder wealth.
- The constant dividend growth rate, g_c , and required return on stock, k_e , are never expected to change.
- k_e must be greater than g_c . If not, the math will not work.

If any one of these assumptions is not met, the model is not appropriate.

EXAMPLE: Gordon growth model valuation

Calculate the value of a stock that paid a \$1.50 dividend last year, if dividends are expected to grow at 8% forever and the required return on equity is 12%.

Answer:

Determine D_1 : $D_0(1 + g_c) = \$1.50(1.08) = \1.62

Calculate the stock's value = $D_1 / (k_e - g_c)$
= $\$1.62 / (0.12 - 0.08)$
= $\$40.50$



PROFESSOR'S NOTE

When doing stock valuation problems on the exam, watch for words like "forever," "infinitely," "indefinitely," "for the foreseeable future," and so on. This will tell you that

the Gordon growth model should be used. Also watch for words like “just paid” or “recently paid.” These will refer to the last dividend, D_0 . Words like “will pay” or “is expected to pay” refer to D_1 .

This example demonstrates that the stock’s value is determined by the relationship between the investor’s required rate of return on equity, k_e , and the projected growth rate of dividends, g_c :

- As the difference between k_e and g_c widens, the value of the stock falls.
- As the difference narrows, the value of the stock rises.
- Small changes in the difference between k_e and g_c can cause large changes in the stock’s value.

Because the estimated stock value is very sensitive to the denominator, an analyst should calculate several different value estimates using a range of required returns and growth rates.

An analyst can also use the Gordon growth model to determine how much of the estimated stock value is due to dividend growth. To do this, assume the growth rate is zero and calculate a value. Then, subtract this value from the stock value estimated using a positive growth rate.

EXAMPLE: Amount of estimated stock value due to dividend growth

Using the data from the previous example, calculate how much of the estimated stock value is due to dividend growth.

Answer:

The estimated stock value with a growth rate of zero is:

$$V_0 = D / k = \$1.50 / 0.12 = \$12.50$$

The amount of the estimated stock value due to estimated dividend growth is:

$$\$40.50 - \$12.50 = \$28.00$$

Estimating the Growth Rate in Dividends

To estimate the growth rate in dividends, the analyst can use three methods:

1. Use the historical growth in dividends for the firm.
2. Use the median industry dividend growth rate.
3. Estimate the sustainable growth rate.

The **sustainable growth rate** is the rate at which equity, earnings, and dividends can continue to grow indefinitely assuming that ROE is constant, the dividend payout ratio is constant, and no new equity is sold.

$$\text{sustainable growth} = (1 - \text{dividend payout ratio}) \times \text{ROE}$$

The quantity $(1 - \text{dividend payout ratio})$ is also referred to as the **retention rate**, the proportion of net income that is not paid out as dividends and goes to retained earnings, thus increasing equity.

EXAMPLE: Sustainable growth rate

Green, Inc., is expected to pay dividends equal to 25% of earnings. Green's ROE is 21%. Calculate and interpret its sustainable growth rate.

Answer:

$$g = (1 - 0.25) \times 21\% = 15.75\%$$

With long-run economic growth typically in the single digits, it is unlikely that a firm could sustain 15.75% growth forever. The analyst should also examine the growth rate for the industry and the firm's historical growth rate to determine whether the estimate is reasonable.

Some firms do not currently pay dividends but are expected to begin paying dividends at some point in the future. A firm may not currently pay a dividend because it is in financial distress and cannot afford to pay out cash or because the return the firm can earn by reinvesting cash is greater than what stockholders could expect to earn by investing dividends elsewhere.

For these firms, an analyst must estimate the amount and timing of the first dividend in order to use the Gordon growth model. Because these parameters are highly uncertain, the analyst should check the estimate from the Gordon growth model against estimates made using other models.

EXAMPLE: A firm with no current dividend

A firm currently pays no dividend but is expected to pay a dividend at the end of Year 4. Year 4 earnings are expected to be \$1.64, and the firm will maintain a payout ratio of 50%. Assuming a constant growth rate of 5% and a required rate of return of 10%, estimate the current value of this stock.

Answer:

The first step is to find the value of the stock at the end of Year 3. Remember, P_3 is the present value of dividends in Years 4 through infinity, calculated at the end of Year 3, one period *before* the first dividend is paid.

Calculate D_4 , the estimate of the dividend that will be paid at the end of Year 4:

$$D_4 = (\text{dividend payout ratio})(E_4) = (0.5)(1.64) = \$0.82$$

Apply the constant growth model to estimate V_3 :

$$V_3 = D_4 / (k_e - g_c) = \$0.82 / (0.10 - 0.05) = \$16.40$$

The second step is to calculate the current value, V_0 :

$$V_0 = 16.40 / 1.1^3 = \$12.32$$

Multistage Dividend Growth Models

A firm may temporarily experience a growth rate that exceeds the required rate of return on the firm's equity, but no firm can maintain this relationship indefinitely. A firm with an extremely high growth rate will attract competition, and its growth rate will eventually fall. We must assume the firm will return to a more sustainable rate of growth at some point in the future in order to calculate the present value of expected future dividends.

One way to value a dividend-paying firm that is experiencing temporarily high growth is to add the present values of dividends expected during the high-growth period to the present value of the constant growth value of the firm at the end of the high-growth period. This is referred to as the **multistage dividend discount model**.

$$\text{value} = \frac{D_1}{(1+k_e)} + \frac{D_2}{(1+k_e)^2} + \dots + \frac{D_n}{(1+k_e)^n} + \frac{P_n}{(1+k_e)^n}$$

where $P_n = \frac{D_{n+1}}{k_e - g_c}$ is the terminal stock value, assuming that dividends at $t = n + 1$ and beyond grow at a constant rate of g_c .

Steps in using the multistage model:

- Determine the discount rate, k_e .
- Project the size and duration of the high initial dividend growth rate, g^* .
- Estimate dividends during the high-growth period.
- Estimate the constant growth rate at the end of the high-growth period, g_c .
- Estimate the first dividend that will grow at the constant rate.
- Use the constant growth value to calculate the stock value at the end of the high-growth period.
- Add the PVs of all dividends to the PV of the terminal value of the stock.

EXAMPLE: Multistage growth

Consider a stock with dividends that are expected to grow at 15% per year for two years, after which they are expected to grow at 5% per year, indefinitely. The last dividend paid was \$1.00, and $k_e = 11\%$. Calculate the value of this stock using the multistage growth model.

Answer:

Calculate the dividends over the high-growth period:

$$D_1 = D_0(1 + g^*) = 1.00(1.15) = \$1.15$$

$$D_2 = D_1(1 + g^*) = 1.15(1.15) = 1.15^2 = \$1.32$$

Although we increase D_1 by the high growth rate of 15% to get D_2 , D_2 will grow at the constant growth rate of 5% for the foreseeable future. This property of D_2 allows us to use the constant growth model formula with D_2 to get P_1 , a time = 1 value for all the (infinite) dividends expected from time = 2 onward.

$$P_1 = \frac{D_2}{k_e - g_c} = \frac{1.32}{0.11 - 0.05} = 22.00$$

Finally, we can sum the present values of dividend 1 and of P_1 to get the present value of all the expected future dividends during both the high- and constant growth periods:

$$\frac{1.15 + 22.00}{1.11} = \$20.86$$



PROFESSOR'S NOTE

Many finance textbooks solve multiple stage growth problems like this one by using the first dividend that has grown at the constant long-term rate to calculate the terminal value, one period after the dividend we have used. Except for rounding, this results in the same current stock value. In fact, the constant growth model can be employed using any dividend during the assumed constant growth period.

A common mistake with multistage growth problems is to calculate the future value, P_1 in this example, and then to either forget to discount it back to the present or to discount over the number of periods until the constant growth dividend is paid (two in this example) rather than using the correct number of periods for discounting the constant growth value (one period in the example). Don't make these mistakes because question writers like to present these common errors as answer choices.

LOS 41.i: Identify characteristics of companies for which the constant growth or a multistage dividend discount model is appropriate.

The Gordon growth model uses a single constant growth rate of dividends and is most appropriate for valuing stable and mature, non-cyclical, dividend-paying firms.

For dividend-paying firms with dividends that are expected to grow rapidly, slowly, or erratically over some period, followed by constant dividend growth, some form of the multistage growth model should be employed. The important points are that dividends must be estimable and must grow at a constant rate after some initial period so that the constant growth model can be used to determine the terminal value of the stock. Thus, we can apply multistage dividend growth models to a firm with high current growth that will drop to a stable rate in the future or to a firm that is temporarily losing market share and growing slowly or getting smaller, as long as its growth is expected to stabilize to a constant rate at some point in the future.

One variant of a multistage growth model assumes that the firm has three stages of dividend growth, not just two. These three stages can be categorized as growth, transition, and maturity. A 3-stage model would be suitable for firms with an initial high growth rate, followed by a lower growth rate during a second, transitional period, followed by the constant growth rate in the long run, such as a young firm still in the high growth phase.

When a firm does not pay dividends, estimates of dividend payments some years in the future are highly speculative. In this case, and in any case where future dividends cannot be estimated with much confidence, valuation based on FCFE is appropriate as long as growth rates of earnings can be estimated. In other cases, valuation based on price multiples may be more appropriate.

MODULE QUIZ 41.2



1. The constant growth model requires which of the following?
 - A. $g < k$.
 - B. $g > k$.
 - C. $g \neq k$.
2. What would an investor be willing to pay for a share of preferred stock that pays an annual \$7 dividend if the required return is 7.75%?
 - A. \$77.50.
 - B. \$87.50.
 - C. \$90.32.
3. An analyst estimates that a stock will pay a \$2 dividend next year and that it will sell for \$40 at year-end. If the required rate of return is 15%, what is the value of the stock?
 - A. \$33.54.
 - B. \$36.52.
 - C. \$43.95.
4. What is the intrinsic value of a company's stock if dividends are expected to grow at 5%, the most recent dividend was \$1, and investors' required rate of return for this stock is 10%?
 - A. \$20.00.
 - B. \$21.00.
 - C. \$22.05.
5. Assume that a stock is expected to pay dividends at the end of Year 1 and Year 2 of \$1.25 and \$1.56, respectively. Dividends are expected to grow at a 5% rate thereafter. Assuming that k_e is 11%, the value of the stock is *closest* to:
 - A. \$22.30.
 - B. \$23.42.
 - C. \$24.55.
6. An analyst feels that Brown Company's earnings and dividends will grow at 25% for two years, after which growth will fall to a constant rate of 6%. If the projected discount rate is 10%, and Brown's most recently paid dividend was \$1, the value of Brown's stock using the multistage dividend discount model is *closest* to:
 - A. \$31.25.
 - B. \$33.54.
 - C. \$36.65.
7. Which of the following firms would *most likely* be appropriately valued using the constant growth DDM?
 - A. An auto manufacturer.
 - B. A producer of bread and snack foods.
 - C. A biotechnology firm in existence for two years.

MODULE 41.3: RELATIVE VALUATION MEASURES



LOS 41.j: Explain the rationale for using price multiples to value equity, how the price to earnings multiple relates to fundamentals, and the use of multiples based on comparables.

Video covering this content is available online.

Because the dividend discount model is very sensitive to its inputs, many investors rely on other methods. In a **price multiple** approach, an analyst compares a stock's price multiple to a benchmark value based on an index, industry group of firms, or a peer group of firms within an

industry. Common price multiples used for valuation include price-to-earnings, price-to-cash flow, price-to-sales, and price-to-book value ratios.

Price multiples are widely used by analysts and readily available in numerous media outlets. Price multiples are easily calculated and can be used in time series and cross-sectional comparisons. Many of these ratios have been shown to be useful for predicting stock returns, with low multiples associated with higher future returns.

A critique of price multiples is that they reflect only the past because historical (trailing) data are often used in the denominator. For this reason, many practitioners use forward (leading or prospective) values in the denominator (sales, book value, earnings, etc.). The use of projected values can result in much different ratios. An analyst should be sure to use price multiple calculations consistently across firms.

When we compare a price multiple, such as P/E, for a firm to those of other firms based on market prices, we are using **price multiples based on comparables**. By contrast, **price multiples based on fundamentals** tell us what a multiple should be based on some valuation model and therefore are not dependent on the current market prices of other companies to establish value.

LOS 41.k: Calculate and interpret the following multiples: price to earnings, price to an estimate of operating cash flow, price to sales, and price to book value.

Price multiples used for valuation include:

- **Price-earnings (P/E) ratio:** The P/E ratio is a firm's stock price divided by earnings per share and is widely used by analysts and cited in the press.
- **Price-sales (P/S) ratio:** The P/S ratio is a firm's stock price divided by sales per share.
- **Price-book value (P/B) ratio:** The P/B ratio is a firm's stock price divided by book value of equity per share.
- **Price-cash flow (P/CF) ratio:** The P/CF ratio is a firm's stock price divided by cash flow per share, where cash flow may be defined as operating cash flow or free cash flow.

Other multiples can be used that are industry specific. For example, in the cable television industry, stock market capitalization is compared to the number of subscribers.

Multiples Based on Fundamentals

To understand fundamental price multiples, consider the Gordon growth valuation model:

$$P_0 = \frac{D_1}{k - g}$$

If we divide both sides of the equation by next year's projected earnings, E_1 , we get

$$\frac{P_0}{E_1} = \frac{D_1/E_1}{k - g}$$

which is the leading P/E for this stock if it is valued in the market according to the constant growth DDM.

This P/E based on fundamentals is also referred to as a **justified P/E**. It is “justified” because, assuming we have the correct inputs for D_1 , E_1 , k_e , and g , the previous equation will provide a P/E ratio that is based on the present value of the future cash flows. We refer to this as a *leading P/E ratio* because it is based on expected earnings next period, not on actual earnings for the previous period, which would produce a lagging or *trailing P/E ratio*.

One advantage of this approach is that it makes clear how the firm’s P/E ratio should be related to its fundamentals. It illustrates that the P/E ratio is a function of:

- D_1 / E_1 = expected dividend payout ratio.
- k = required rate of return on the stock.
- g = expected constant growth rate of dividends.

EXAMPLE: P/E based on fundamentals

A firm has an expected dividend payout ratio of 30%, a required rate of return of 13%, and an expected dividend growth rate of 6%. Calculate the firm’s fundamental (justified) leading P/E ratio.

Answer:

$$\text{expected P/E ratio: } 0.3 / (0.13 - 0.06) = 4.3$$

The justified P/E ratio serves as a benchmark for the price at which the stock should trade. In the previous example, if the firm’s actual P/E ratio (based on the market price and expected earnings) was 8, the stock would be considered overvalued. If the firm’s market P/E ratio was 2, the stock would be considered undervalued.

P/E ratios based on fundamentals are very sensitive to the inputs (especially the denominator, $k - g$), so the analyst should use several different sets of inputs to indicate a range for the justified P/E.

Because we started with the equation for the constant growth DDM, the P/E ratio calculated in this way is the P/E ratio consistent with the constant growth DDM. We can see from the formula that, *other things equal*, the P/E ratio we have defined here will increase with (1) a higher dividend payout rate, (2) a higher growth rate, or (3) a lower required rate of return. So, if the subject firm has a higher dividend payout ratio, higher growth rate, and lower required return than its peers, a higher P/E ratio may be justified.

In practice, other things are not equal. An increase in the dividend payout ratio, for example, will reduce the firm’s sustainable growth rate. While higher dividends will increase firm value, a lower growth rate will decrease firm value. This relationship is referred to as the **dividend displacement of earnings**. The net effect on firm value of increasing the dividend payout ratio is ambiguous. As intuition would suggest, firms cannot continually increase their P/Es or market values by increasing the dividend payout ratio. Otherwise, all firms would have 100% payout ratios.



PROFESSOR’S NOTE

Watch for the wording “other things equal” or “other variables unchanged” in any exam questions about the effect of changing one variable.

EXAMPLE: Fundamental P/E ratio comparison

Holt Industries makes decorative items. The following figures are for Holt and its industry.

Holt Industries	Industry Average
Dividend payout ratio	25%
Sales growth	7.5%
Total debt to equity	113%

Which of these factors suggest a higher fundamental P/E ratio for Holt?

Answer:

- The higher dividend payout ratio supports Holt having a higher P/E ratio than the industry.
- Higher growth in sales suggests that Holt will be able to increase dividends at a faster rate, which supports Holt having a higher P/E ratio than the industry.
- The higher level of debt, however, indicates that Holt has higher risk and a higher required return on equity, which supports Holt having a lower P/E ratio than the industry.

Multiples Based on Comparables

Valuation based on price multiple comparables (or comps) involves using a price multiple to evaluate whether an asset is valued properly relative to a benchmark. Common benchmarks include the stock’s historical average (a time series comparison) or similar stocks and industry averages (a cross-sectional comparison). Comparing firms within an industry is useful for analysts who are familiar with a particular industry. Price multiples are readily calculated and provided by many media outlets.

The economic principle guiding this method is the **law of one price**, which asserts that two identical assets should sell at the same price, or in this case, two comparable assets should have approximately the same multiple.

The analyst should be sure that any comparables used really are comparable. Price multiples may not be comparable across firms if the firms are different sizes, are in different industries, or will grow at different rates. Furthermore, using P/E ratios for cyclical firms is complicated due to their sensitivity to economic conditions. In this case, the P/S ratio may be favored over the P/E ratio because the sales are less volatile than earnings due to both operating and financial leverage.

The disadvantages of using price multiples based on comparables are (1) a stock may appear overvalued by the comparable method but undervalued by the fundamental method, or vice versa; (2) different accounting methods can result in price multiples that are not comparable across firms, especially internationally; and (3) price multiples for cyclical firms may be greatly affected by economic conditions at a given point in time.

EXAMPLE: Valuation using comparables

The following figures are for Renee's Bakery. All figures except the stock price are in millions.

Fiscal Year-End	20X3	20X2	20X1
Total stockholder's equity	\$55.60	\$54.10	\$52.60
Net revenues	\$77.30	\$73.60	\$70.80
Net income	\$3.20	\$1.10	\$0.40
Net cash flow from operations	\$17.90	\$15.20	\$12.20
Stock price	\$11.40	\$14.40	\$12.05
Shares outstanding	4.476	3.994	3.823

Calculate Renee's lagging P/E, P/CF, P/S, and P/B ratios. Judge whether the firm is undervalued or overvalued using the following relevant industry averages for 20X3 and the firm's historical trend.

Lagging Industry Ratios	20X3
Price-to-earnings	8.6
Price-to-cash flow	4.6
Price-to-sales	1.4
Price-to-book value	3.6

Answer:

To calculate the lagging price multiples, first divide the relevant financial statement items by the number of shares to get per-share amounts. Then, divide the stock price by this figure.

For example, for the P/S ratio for 20X3, divide net revenue (net sales) by the number of shares:

$$\frac{\text{sales}}{\text{number of shares}} = \frac{\$77.30}{4.476} = 17.270$$

Then, divide the stock price by sales per share:

$$\frac{P}{S} = \frac{\$11.40}{17.3} = 0.7$$

Using the net income for earnings, the net cash flow from operations for the cash flow, and stockholder's equity for book value, the ratios for Renee's Bakery are:

	20X3	20X2	20X1
P/E	15.9	52.3	115.2
P/CF	2.9	3.8	3.8
P/S	0.7	0.8	0.7
P/B	0.9	1.1	0.9

Comparing Renee's Bakery's ratios to the industry averages for 20X3, the price multiples are lower in all cases except for the P/E multiple. This cross-sectional evidence suggests that Renee's Bakery is undervalued.

The P/E ratio merits further investigation. Renee's Bakery may have a higher P/E because its earnings are depressed by high depreciation, interest expense, or taxes. Calculating the price-EBITDA ratio would provide an alternative measure that is unaffected by these expenses.

On a time series basis, the ratios are trending downward. This indicates that Renee's Bakery may be currently undervalued relative to its past valuations. We could also calculate average price multiples for the ratios over 20X1–20X3 as a benchmark for the current values:

Company average P/E 20X1–20X3	61.1
Company average P/CF 20X1–20X3	3.5
Company average P/S 20X1–20X3	0.7
Company average P/B 20X1–20X3	1.0

The current P/E, P/CF, and P/B ratios are lower than their 3-year averages. This indicates that Renee's Bakery may be currently undervalued. It also may be the case, however, that P/E ratios for the market as a whole have been decreasing over the period due to systematic factors.

LOS 41.I: Describe enterprise value multiples and their use in estimating equity value.

Enterprise value (EV) measures total company value. EV can be viewed as what it would cost to acquire the firm:

$$EV = \text{market value of common and preferred stock} + \text{market value of debt} - \text{cash and short-term investments}$$

Cash and short-term investments are subtracted because an acquirer's cost for a firm would be decreased by the amount of the target's liquid assets. Although an acquirer assumes the firm's debt, it also receives the firm's cash and short-term investments. Enterprise value is appropriate when an analyst wants to compare the values of firms that have significant differences in capital structure.

EBITDA (earnings before interest, taxes, depreciation, and amortization are subtracted) is probably the most frequently used denominator for EV multiples; operating income can also be used. Because the numerator represents total company value, it should be compared to earnings of both debt and equity owners. An advantage of using EBITDA instead of net income is that EBITDA is usually positive even when earnings are not. When net income is negative, value multiples based on earnings are meaningless. A disadvantage of using EBITDA is that it often includes non-cash revenues and expenses.

A potential problem with using enterprise value is that the market value of a firm's debt is often not available. In this case, the analyst can use the market values of similar bonds or can use their book values. Book value, however, may not be a good estimate of market value if firm and market conditions have changed significantly since the bonds were issued.

EXAMPLE: Calculating EV/EBITDA multiples

Daniel, Inc., is a manufacturer of small refrigerators and other appliances. The following figures are from Daniel's most recent financial statements except for the market value of long-term debt, which has been estimated from financial market data.

Stock price	\$40.00
Shares outstanding	200,000
Market value of long-term debt	\$600,000
Book value of long-term debt	\$900,000
Book value of total debt	\$2,100,000
Cash and marketable securities	\$250,000
EBITDA	\$1,000,000

Calculate the EV/EBITDA multiple.

Answer:

First, we must estimate the market value of the firm's short-term debt and liabilities. To do so, subtract the book value of long-term debt from the book value of total debt: \$2,100,000 – \$900,000 = \$1,200,000. This is the book value of the firm's short-term debt. We can assume the market value of these short-term items is close to their book value. (As we will see in the Fixed Income topic area, the market values of debt instruments approach their face values as they get close to maturity.)

Add the market value of long-term debt to get the market value of total debt: \$600,000 + \$1,200,000 = \$1,800,000.

The market value of equity is the stock price multiplied by the number of shares: \$40 × 200,000 = \$8,000,000.

The enterprise value of the firm is the sum of debt and equity minus cash: \$1,800,000 + \$8,000,000 – \$250,000 = \$9,550,000.

$$\text{EV/EBITDA} = \$9,550,000 / \$1,000,000 \approx 9.6.$$

If the competitor or industry average EV/EBITDA is above 9.6, Daniel is relatively undervalued. If the competitor or industry average EV/EBITDA is below 9.6, Daniel is relatively overvalued.

LOS 41.m: Describe asset-based valuation models and their use in estimating equity value.

Our third category of valuation model is **asset-based models**, which are based on the idea that equity value is the market or fair value of assets minus the market or fair value of liabilities. Because market values of firm assets are usually difficult to obtain, the analyst typically starts with the balance sheet to determine the values of assets and liabilities. In most cases, market values are not equal to book values. Possible approaches to valuing assets are to value them at their depreciated values, inflation-adjusted depreciated values, or estimated replacement values.

Applying asset-based models is especially problematic for a firm that has a large amount of intangible assets, on or off the balance sheet. The effect of the loss of the current owners' talents and customer relationships on forward earnings may be quite difficult to measure. Analysts often consider asset-based model values as floor or minimum values when significant intangibles, such as business reputation, are involved. An analyst should consider supplementing

an asset-based valuation with a more forward-looking valuation, such as one from a discounted cash flow model.

Asset-based model valuations are most reliable when the firm has primarily tangible short-term assets, assets with ready market values (e.g., financial or natural resource firms), or when the firm will cease to operate and is being liquidated. Asset-based models are often used to value private companies but may be increasingly useful for public firms as they move toward fair value reporting on the balance sheet.

EXAMPLE: Using an asset-based model for a public firm

Williams Optical is a publicly traded firm. An analyst estimates that the market value of net fixed assets is 120% of book value. Liability and short-term asset market values are assumed to equal their book values. The firm has 2,000 shares outstanding.

Using the selected financial results in the table, calculate the value of the firm's net assets on a per-share basis.

Cash	\$10,000
Accounts receivable	\$20,000
Inventories	\$50,000
Net fixed assets	<u>\$120,000</u>
Total assets	\$200,000
Accounts payable	\$5,000
Notes payable	\$30,000
Term loans	\$45,000
Common stockholder equity	<u>\$120,000</u>
Total liabilities and equity	\$200,000

Answer:

Estimate the market value of assets, adjusting the fixed assets for the analyst's estimates of their market values:

$$\$10,000 + \$20,000 + \$50,000 + \$120,000(1.20) = \$224,000$$

Determine the market value of liabilities:

$$\$5,000 + 30,000 + \$45,000 = \$80,000$$

Calculate the adjusted equity value:

$$\$224,000 - \$80,000 = \$144,000$$

Calculate the adjusted equity value per share:

$$\$144,000 / 2,000 = \$72$$

LOS 41.f: Explain advantages and disadvantages of each category of valuation model.

Advantages of discounted cash flow models:

- They are based on the fundamental concept of discounted present value and are well grounded in finance theory.
- They are widely accepted in the analyst community.

Disadvantages of discounted cash flow models:

- Their inputs must be estimated.
- Value estimates are very sensitive to input values.

Advantages of comparable valuation using price multiples:

- Evidence that some price multiples are useful for predicting stock returns.
- Price multiples are widely used by analysts.
- Price multiples are readily available.
- They can be used in time series and cross-sectional comparisons.
- EV/EBITDA multiples are useful when comparing firm values independent of capital structure or when earnings are negative and the P/E ratio cannot be used.

Disadvantages of comparable valuation using price multiples:

- Lagging price multiples reflect the past.
- Price multiples may not be comparable across firms if the firms have different size, products, and growth.
- Price multiples for cyclical firms may be greatly affected by economic conditions at a given point in time.
- A stock may appear overvalued by the comparable method but undervalued by a fundamental method or vice versa.
- Different accounting methods can result in price multiples that are not comparable across firms, especially internationally.
- A negative denominator in a price multiple results in a meaningless ratio. The P/E ratio is especially susceptible to this problem.

Advantages of price multiple valuations based on fundamentals:

- They are based on theoretically sound valuation models.
- They correspond to widely accepted value metrics.

Disadvantage of price multiple valuations based on fundamentals:

- Price multiples based on fundamentals will be very sensitive to the inputs (especially the k – g denominator).

Advantages of asset-based models:

- They can provide floor values.
- They are most reliable when the firm has primarily tangible short-term assets, assets with ready market values, or when the firm is being liquidated.
- They are increasingly useful for valuing public firms that report fair values.

Disadvantages of asset-based models:

- Market values are often difficult to obtain.

- Market values are usually different than book values.
- They are inaccurate when a firm has a high proportion of intangible assets or future cash flows not reflected in asset values.
- Assets can be difficult to value during periods of hyperinflation.



MODULE QUIZ 41.3

1. Which of the following is *least likely* a rationale for using price multiples?
 - A. Price multiples are easily calculated.
 - B. The fundamental P/E ratio is insensitive to its inputs.
 - C. The use of forward values in the divisor provides an incorporation of the future.
2. A firm has an expected dividend payout ratio of 60% and an expected future growth rate of 7%. What should the firm's fundamental price-to-earnings (P/E) ratio be if the required rate of return on stocks of this type is 15%?
 - A. 5.0×.
 - B. 7.5×.
 - C. 10.0×.
3. Enterprise value is defined as the market value of equity plus:
 - A. the face value of debt minus cash and short-term investments.
 - B. the market value of debt minus cash and short-term investments.
 - C. cash and short-term investments minus the market value of debt.
4. Which of the following firms would *most appropriately* be valued using an asset-based model?
 - A. An energy exploration firm in financial distress that owns drilling rights for offshore areas.
 - B. A paper firm located in a country that is experiencing high inflation.
 - C. A software firm that invests heavily in research and development and frequently introduces new products.
5. Which type of valuation model is viewed as having the disadvantage of producing results that may not be comparable across firms?
 - A. Asset-based models.
 - B. Price multiple models.
 - C. Discounted cash flow models.

KEY CONCEPTS

LOS 41.a

An asset is fairly valued if the market price is equal to its estimated intrinsic value, undervalued if the market price is less than its estimated value, and overvalued if the market price is greater than the estimated value.

For security valuation to be profitable, the security must be mispriced now and price must converge to intrinsic value over the investment horizon.

Securities that are followed by many investors are more likely to be fairly valued than securities that are neglected by analysts.

LOS 41.b

Discounted cash flow models estimate the present value of cash distributed to shareholders (dividend discount models) or the present value of cash available to shareholders after meeting capital expenditures and working capital expenses (free cash flow to equity models).

Multiplier models compare the stock price to earnings, sales, book value, or cash flow. Alternatively, enterprise value is compared to sales or EBITDA.

Asset-based models define a stock's value as the firm's total asset value minus liabilities and preferred stock, on a per-share basis.

LOS 41.c

Regular cash dividends are paid at set intervals. A special dividend is a one-time cash payment to shareholders.

Stock dividends are additional shares of stock. Stock splits divide each existing share into multiple shares. In either case, the value of each share will decrease because the total value of outstanding shares is unchanged. The portion of the company owned by each shareholder is also unchanged.

In a reverse stock split, the number of shares owned by each shareholder is decreased, so total shares outstanding are decreased and the value of a single share is increased.

A share repurchase is a purchase by the company of its outstanding shares. Share repurchases are an alternative to cash dividends as a way to distribute cash to shareholders.

LOS 41.d

Dividend payment chronology:

- Declaration date: The date the board of directors approves payment of the dividend.
- Ex-dividend date: The first day a share of stock trades without the dividend, one or two business days before the holder-of-record date. On the ex-dividend date, the value of each share decreases by the amount of the dividend.
- Holder-of-record date: The date on which share owners who will receive the dividend are identified.
- Payment date. The date the dividend checks are sent to, or payment is transferred to, shareholders.

LOS 41.e

The dividend discount model is based on the rationale that a corporation has an indefinite life, and a stock's value is the present value of its future cash dividends. The most general form of the model is:

$$V_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1 + k_e)^t}$$

Free cash flow to equity (FCFE) can be used instead of dividends. FCFE is the cash remaining after a firm meets all of its debt obligations and provides for necessary capital expenditures. FCFE reflects the firm's capacity for dividends and is useful for firms that currently do not pay a dividend. By using FCFE, an analyst does not need to project the amount and timing of future dividends.

LOS 41.f

Advantages of discounted cash flow models:

- Easy to calculate.

- Widely accepted in the analyst community.
- FCFE model is useful for firms that currently do not pay a dividend.
- Gordon growth model is useful for stable, mature, noncyclical firms.
- Multistage models can be used for firms with nonconstant growth.

Disadvantages of discounted cash flow models:

- Inputs must be forecast.
- Estimates are very sensitive to inputs.
- For the Gordon growth model specifically:
 - Very sensitive to the $k - g$ denominator.
 - Required return on equity must be greater than the growth rate.
 - Required return on equity and growth rate must remain constant.
 - Firm must pay dividends.

Advantages of price multiples:

- Often useful for predicting stock returns.
- Widely used by analysts.
- Easily calculated and readily available.
- Can be used in time series and cross-sectional comparisons.
- EV/EBITDA multiples are useful when comparing firm values independent of capital structure or when earnings are negative and the P/E ratio cannot be used.

Disadvantages of price multiples:

- P/E ratio based on fundamentals will be very sensitive to the inputs.
- May not be comparable across firms, especially internationally.
- Multiples for cyclical firms may be greatly affected by economic conditions. P/E ratio may be especially inappropriate. (The P/S multiple may be more appropriate for cyclical firms.)
- A stock may appear overvalued by the comparable method but undervalued by the fundamental method or vice versa.
- Negative denominator results in a meaningless ratio; the P/E ratio is especially susceptible to this problem.
- A potential problem with EV/EBITDA multiples is that the market value of a firm's debt is often not available.

Advantages of asset-based models:

- Can provide floor values.
- Most reliable when the firm has mostly tangible short-term assets, assets with a ready market value, or when the firm is being liquidated.
- May be increasingly useful for valuing public firms if they report fair values.

Disadvantages of asset-based models:

- Market values of assets can be difficult to obtain and are usually different than book values.

- Inaccurate when a firm has a large amount of intangible assets or future cash flows not reflected in asset value.
- Asset values can be difficult to value during periods of hyperinflation.

LOS 41.g

Preferred stock typically pays a fixed dividend and does not mature. It is valued as:

$$\text{preferred stock value} = \frac{D_p}{k_p}$$

LOS 41.h

The Gordon growth model assumes the growth rate in dividends is constant:

$$V_0 = \frac{D_1}{k_e - g_c}$$

The sustainable growth rate is the rate at which earnings and dividends can continue to grow indefinitely:

$$g = b \times \text{ROE}$$

where:

b = earnings retention rate = 1 – dividend payout rate

ROE = return on equity

A firm with high growth over some number of periods followed by a constant growth rate of dividends forever can be valued using a multistage model:

$$\text{value} = \frac{D_1}{(1 + k_e)} + \frac{D_2}{(1 + k_e)^2} + \dots + \frac{D_n}{(1 + k_e)^n} + \frac{P_n}{(1 + k_e)^n}$$

where:

$$P_n = \frac{D_{n+1}}{k_e - g_c}$$

g_c = constant growth rate of dividends

n = number of periods of supernormal growth

LOS 41.i

The constant growth model is most appropriate for firms that pay dividends that grow at a constant rate, such as stable and mature firms or noncyclical firms such as utilities and food producers in mature markets.

A 2-stage DDM would be most appropriate for a firm with high current growth that will drop to a stable rate in the future, an older firm that is experiencing a temporary high growth phase, or an older firm with a market share that is decreasing but expected to stabilize.

A 3-stage model would be appropriate for a young firm still in a high growth phase.

LOS 41.j

The P/E ratio based on fundamentals is calculated as:

$$\frac{P_0}{E_1} = \frac{D_1/E_1}{k - g}$$

If the subject firm has a higher dividend payout ratio, higher growth rate, and lower required return than its peers, it may be justified in having a higher P/E ratio.

Price multiples are widely used by analysts, are easily calculated and readily available, and can be used in time series and cross-sectional comparisons.

LOS 41.k

The price-earnings (P/E) ratio is a firm's stock price divided by earnings per share.

The price-sales (P/S) ratio is a firm's stock price divided by sales per share.

The price-book value (P/B) ratio is a firm's stock price divided by book value per share.

The price-cash flow (P/CF) ratio is a firm's stock price divided by cash flow per share. Cash flow may be defined as operating cash flow or free cash flow.

LOS 41.l

Enterprise value (EV) measures total company value:

$$EV = \text{market value of common and preferred stock} + \text{market value of debt} - \text{cash and short-term investments}$$

EBITDA is frequently used as the denominator in EV multiples because EV represents total company value, and EBITDA represents earnings available to all investors.

LOS 41.m

Asset-based models value equity as the market or fair value of assets minus liabilities. These models are most appropriate when a firm's assets are largely tangible and have fair values that can be established easily.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 41.1

1. **B** If the analyst is more confident of his input values, he is more likely to conclude that the security is overvalued. The market price is more likely to be correct for a security followed by many analysts and less likely correct when few analysts follow the security. (LOS 41.a)
2. **C** One example of a present value model is valuation based on the present value of future cash flows available to equity holders. (LOS 41.b)
3. **C** Both a 50% stock dividend and a 3-for-2 stock split will increase the number of shares by 50%, while neither will affect value of the company. Therefore, the decrease in the share price should be the same in either case. (LOS 41.c)
4. **B** The chronology of a dividend payout is declaration date, ex-dividend date, holder-of-record date, and payment date. The ex-dividend date is the cutoff date for receiving the dividend: stocks purchased on or after the ex-dividend date will not receive the dividend. (LOS 41.d)

Module Quiz 41.2

1. **A** For the constant growth model, the constant growth rate (g) must be less than the required rate of return (k). (LOS 41.e)
2. **C** The share value is $7.0 / 0.0775 = \$90.32$. (LOS 41.g)

3. **B** $(\$40 + \$2) / 1.15 = \$36.52$. (LOS 41.h)
4. **B** Using the constant growth model, $\$1(1.05) / (0.10 - 0.05) = \21.00 . (LOS 41.h)
5. **C** $(\$1.25 / 1.11) + [1.56 / (0.11 - 0.05)] / 1.11 = \24.55 . (LOS 41.h)
6. **C** $\$1(1.25) / 1.1 + [\$1(1.25)^2 / (0.10 - 0.06)] / 1.1 = \36.65 . (LOS 41.h)
7. **B** The constant growth DDM assumes that the dividend growth rate is constant. The most likely choice here is the bread and snack producer. Auto manufacturers are more likely to be cyclical than to experience constant growth. A biotechnology firm in existence for two years is unlikely to pay a dividend, and if it does, dividend growth is unlikely to be constant. (LOS 41.i)

Module Quiz 41.3

1. **B** The fundamental P/E ratio is sensitive to its inputs. It uses the DDM as its framework, and the denominator $k - g$ in both has a large impact on the calculated P/E or stock value. (LOS 41.j)
2. **B** Using the earnings multiplier model, $0.6 / (0.15 - 0.07) = 7.5\times$. (LOS 41.k)
3. **B** Enterprise value is market value of equity plus market value of debt minus cash and short-term investments. (LOS 41.l)
4. **A** The energy exploration firm would be most appropriately valued using an asset-based model. Its near-term cash flows are likely negative, so a forward-looking model is of limited use. Furthermore, it has valuable assets in the form of drilling rights that likely have a readily determined market value. The paper firm would likely not be appropriately valued using an asset-based model because high inflation makes the values of a firm's assets more difficult to estimate. An asset-based model would not be appropriate to value the software firm because the firm's value largely consists of internally developed intangible assets. (LOS 41.m)
5. **B** Results that may not be comparable across firms are considered a disadvantage of valuation models based on price multiples. (LOS 41.f)

TOPIC QUIZ: EQUITY INVESTMENTS

You have now finished the Equity Investments topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Equity Investments. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

FORMULAS

$$NPV = CF_0 + \frac{CF_1}{(1+k)^1} + \frac{CF_2}{(1+k)^2} + \dots + \frac{CF_n}{(1+k)^n} = \sum_{t=0}^n \frac{CF_t}{(1+k)^t}$$

$$IRR: 0 = CF_0 + \frac{CF_1}{(1+IRR)^1} + \frac{CF_2}{(1+IRR)^2} + \dots + \frac{CF_n}{(1+IRR)^n} = \sum_{t=0}^n \frac{CF_t}{(1+IRR)^t}$$

$$WACC = (w_d)[k_d(1-t)] + (w_{ps})(k_{ps}) + (w_{ce})(k_{ce})$$

$$\text{after-tax cost of debt} = k_d(1-t)$$

$$\text{cost of preferred stock} = k_{ps} = D_{ps} / P$$

cost of common equity:

$$k_{ce} = R_f + \beta[E(R_m) - R_f]$$

$$k_{ce} = \text{bond yield} + \text{risk premium}$$

$$\text{adjusted beta} = 2/3 \times \text{unadjusted beta} + 1/3$$

unlevered asset beta: target beta:

$$\beta_{\text{ASSET}} = \beta_{\text{EQUITY}} \left\{ \frac{1}{1 + [(1-t)\frac{D}{E}]} \right\} \quad \beta_{\text{TARGET}} = \beta_{\text{ASSET}} \left\{ 1 + [(1-t)\frac{D}{E}] \right\}$$

$$\text{degree of operating leverage} = \frac{Q(P-V)}{Q(P-V)-F} = \frac{\% \Delta EBIT}{\% \Delta \text{sales}}$$

$$\text{degree of financial leverage} = \frac{EBIT}{EBIT-I} = \frac{\% \Delta EPS}{\% \Delta EBIT}$$

$$\text{degree of total leverage} = DOL \times DFL = \frac{\% \Delta EPS}{\% \Delta \text{sales}}$$

$$\text{breakeven quantity of sales} = \frac{\text{fixed operating costs} + \text{fixed financing costs}}{\text{price} - \text{variable cost per unit}}$$

$$\text{operating breakeven quantity of sales} = \frac{\text{fixed operating costs}}{\text{price} - \text{variable cost per unit}}$$

$$\text{current ratio} = \frac{\text{current assets}}{\text{current liabilities}}$$

$$\text{quick ratio} = \frac{\text{cash} + \text{short-term marketable securities} + \text{receivables}}{\text{current liabilities}}$$

$$\text{receivables turnover} = \frac{\text{credit sales}}{\text{average receivables}}$$

$$\text{number of days of receivables} = \frac{365}{\text{receivables turnover}} = \frac{\text{average receivables}}{\text{average day's credit sales}}$$

$$\text{inventory turnover} = \frac{\text{cost of goods sold}}{\text{average inventory}}$$

$$\text{number of days of inventory} = \frac{365}{\text{inventory turnover}} = \frac{\text{average inventory}}{\text{average day's COGS}}$$

$$\text{payables turnover ratio} = \frac{\text{purchases}}{\text{average trade payables}}$$

$$\text{number of days of payables} = \frac{365}{\text{payables turnover ratio}} = \frac{\text{average payables}}{\text{average day's purchases}}$$

operating cycle = average days of inventory + average days of receivables

$$\text{cash conversion cycle} = \left(\frac{\text{average days of receivables}}{\text{of receivables}} \right) + \left(\frac{\text{average days of inventory}}{\text{of inventory}} \right) - \left(\frac{\text{average days of payables}}{\text{of payables}} \right)$$

$$\text{margin call price} = P_0 \left(\frac{1 - \text{initial margin}}{1 - \text{maintenance margin}} \right)$$

$$\text{price-weighted index} = \frac{\text{sum of stock prices}}{\text{number of stocks in index adjusted for splits}}$$

$$\text{market cap-weighted index} = \frac{\sum [(\text{price}_{\text{today}})(\text{number of shares outstanding})]}{\sum [(\text{price}_{\text{base year}})(\text{number of shares outstanding})]} \times \text{base year index value}$$

$$\text{preferred stock valuation model: } P_0 = \frac{D_p}{k_p}$$

$$\text{one-period stock valuation model: } P_0 = \frac{D_1}{1 + k_e} + \frac{P_1}{1 + k_e}$$

$$\text{infinite period model: } P_0 = \frac{D_1}{k_e - g} = \frac{D_0 \times (1 + g)}{k_e - g}$$

multistage model:

$$P_0 = \frac{D_1}{(1 + k_e)} + \frac{D_2}{(1 + k_e)^2} + \dots + \frac{D_n}{(1 + k_e)^n} + \frac{P_n}{(1 + k_e)^n}$$

where:

$P_n = \frac{D_{n+1}}{k_e - g_c}$, and D_{n+1} is a dividend that will grow at the constant rate of g_c forever

$$\text{earnings multiplier: } \frac{P_0}{E_1} = \frac{\frac{D_1}{E_1}}{k - g}$$

expected growth rate: $g = (\text{retention rate})(\text{ROE})$

$$\text{trailing P/E} = \frac{\text{market price per share}}{\text{EPS over previous 12 months}}$$

$$\text{leading P/E} = \frac{\text{market price per share}}{\text{forecast EPS over next 12 months}}$$

$$\text{P/B ratio} = \frac{\text{market value of equity}}{\text{book value of equity}} = \frac{\text{market price per share}}{\text{book value per share}}$$

where:

$$\begin{aligned} \text{book value of equity} &= \text{common shareholders' equity} \\ &= (\text{total assets} - \text{total liabilities}) - \text{preferred stock} \end{aligned}$$

$$\text{P/S ratio} = \frac{\text{market value of equity}}{\text{total sales}} = \frac{\text{market price per share}}{\text{sales per share}}$$

$$\text{P/CF ratio} = \frac{\text{market value of equity}}{\text{cash flow}} = \frac{\text{market price per share}}{\text{cash flow per share}}$$

$$\begin{aligned} \text{enterprise value} &= \text{market value of common and preferred stock} \\ &+ \text{market value of debt} \\ &- \text{cash and short-term investments} \end{aligned}$$

INDEX

A

- abandonment options, 40
- abnormal profit, 158
- accelerated book build, 121
- accounting return on equity, 173
- accounts payable, 47
- accounts receivable, 47
- acid-test ratio, 51
- activist shareholders, 13
- adverse selection, 113
- affiliate marketing, 25
- after-tax cost of debt, 60
- agency costs of debt, 80
- agency costs of equity, 83
- aggregators, 26
- allocational efficiency, 125
- all-or-nothing orders, 119
- alternative trading systems (ATS), 111
- American depository receipts (ADRs), 171
- American depository share (ADS), 171
- annual general meeting, 12
- arbitrage, 113, 157
- ask price, 118
- ask size, 118
- asset-backed securities, 109
- asset-based models, 207, 225
- asset beta, 66
- asset-light model, 26
- asymmetric information, 83
- auction pricing, 24
- audit committee, 14
- average days' sales outstanding, 52

B

B2B (business to business), 24
B2C (business to consumer), 24
backup line of credit, 49
basket of listed depository receipts (BLDR), 172
behavioral finance, 162
best efforts, 122
bid-ask spread, 118
bid price, 118
bid size, 118
blank check companies, 4
blanket lien, 48
block brokers, 111
board of directors, 2, 10, 14
bond indenture, 13
book runner, 121
book value of equity, 173
borrowed capital, 2
breakeven quantity of sales, 96
broker-dealers, 112
broketed markets, 124
brokers, 111
bundling, 24
business growth investments, 32
business maintenance investments, 31
business model, 23
business risk, 27, 92
business structures, 1

C

calendar anomalies, 160
call markets, 123
call money rate, 115
call option, 110
cannibalization, 33
capital allocation process, 32
capital asset pricing model (CAPM), 63
capital components, 59

capital investment risk, 27
capital rationing, 34
capital structure, 73
capital structure irrelevance, 76
cash conversion cycle, 53
cash dividends, 207
cash flow management, 50
channel strategy, 24
civil law system, 15
clearinghouses, 113
closed-end funds, 161
collateral, 13
commercial paper, 49
committed (regular) line of credit, 48
commodities, 110
commodity indexes, 146
common equity, 49
common-law system, 15
common shares, 167
common stock, 108
company analysis, 195
compensation committee, 14
competitive risks, 27
competitive strategy, 195
complete markets, 124
component cost of capital, 60
constant growth model, 212
constituent securities, 135
continuous markets, 123
contracts, 109
contribution margin, 96
conventional cash flow pattern, 33
convertible preference shares, 168
corporate governance, 12
corporation, 2
cost-based pricing, 24
cost leadership (low-cost) strategy, 195
cost of debt capital, 61
cost of equity capital, 63, 175
cost of preferred stock, 62
costs of asymmetric information, 83

costs of financial distress, 80
counterparty risk, 114
covenants, 13
credit default swaps, 110
creditor committees, 13
creditors, 10
crowdsourcing, 26
cumulative preference shares, 168
cumulative voting, 13, 167
currencies, 109
current ratio, 51
custodians, 114
cyclical firm, 182

D

data mining, 159
data snooping, 159
day orders, 120
dealer markets, 123
debt ratings, 75
debt securities, 107
declaration date, 208
decline stage, 192
defensive industries, 183
degree of financial leverage (DFL), 93
degree of operating leverage (DOL), 92
degree of total leverage (DTL), 94
depository bank, 171
depository institutions, 112
depository receipts (DRs), 171
derivative contracts, 107
direct investing, 171
direct listing, 4
direct sales, 24
disadvantages of the IRR method, 37
discounted cash flow models, 206
dividend discount model (DDM), 209
dividend displacement of earnings, 220
dividend reinvestment plan, 122

double taxation, 3
drags on liquidity, 50
dynamic pricing, 24

E

earnings surprise, 161
economic profits, 187
efficient market hypothesis (EMH), 157
electronic communication networks (ECNs), 111
embryonic stage, 192
employees, 10
engagement/active ownership, 18
enterprise value, 206, 223
environmental influences, 194
equal-weighted index, 137
equilibrium interest rate, 107
equity securities, 107, 108
ESG investing, 17
ESG risk, 28
event study, 158
exchanges, 111
exchange-traded funds (ETFs), 109
exchange-traded notes (ETNs), 109
ex-dividend date, 208
execution risk, 27
expansion options, 40
expansion projects, 32
experience curve, 187
externalities, 33
extraordinary general meetings, 13

F

factoring, 48
freemium pricing, 24
fill-or-kill order, 120
financial assets, 107
financial derivative contracts, 107

financial distress, 80
financial intermediaries, 111
financial leverage, 28, 115
financial plan, 23
financial risk, 28, 92
fixed-income securities, 108
float-adjusted market capitalization-weighted index, 138
flotation costs, 67
forward contract, 109
fractional ownership, 25
franchising, 25
free cash flow to equity (FCFE), 211
free float, 138
full integration, 18
fundamental analysis, 158
fundamental options, 40
fundamental value, 156, 205
fundamental weighting, 138
futures contract, 110

G

general partnership, 2
global depository receipts (GDRs), 171
global registered shares (GRS), 172
going concern projects, 31
good-on-close orders, 120
good-on-open orders, 120
good til canceled orders, 120
Gordon growth model, 212
governance committee, 14
green bonds, 18
green finance, 18
growth cyclical, 183
growth industries, 183
growth stage, 192
growth stocks, 160

H

hedge fund indexes, 147
hedge funds, 109
herding, 162
hidden revenue, 25
holder-of-record date, 208
hostile takeover, 13
hurdle rate, 36
hybrid business models, 26

I

immediate-or-cancel orders, 120
incremental cash flows, 33
independent projects, 33
indications of interest, 121
industry life cycle, 191
industry rotation, 181
informational efficiency, 125
informationally efficient capital market, 155
information asymmetry, 11
information cascade, 162
initial margin requirement, 115
initial public offerings (IPOs), 3, 121, 161
insurance companies, 113
insurance contract, 110
internal rate of return (IRR), 36
intrinsic value, 156, 205
inventory, 48
inventory turnover, 52
investment banks, 111
investment committee, 14
investor overconfidence, 162

J

January effect, 160
justified P/E, 220

K

key advantage of IRR, 37
key advantage of NPV, 37

L

lean start-ups, 26
leverage, 91
leveraged buyout (LBO), 4, 169
leveraged position, 115
leverage ratio, 115
licensing, 25
licensing agreements, 25
life-cycle stage, 187
limited company, 2
limited liability, 2
limited partners, 2
limited partnership, 2
lines of credit, 48
liquidity ratios, 51
long position, 114
long-term debt, 49
loss aversion, 162

M

macro risk, 27
maintenance margin requirement, 116
majority voting, 13
management buyout (MBO), 4
marginal cost of capital (MCC), 59
margin call, 117
margin loan, 115
marketable securities, 48
market anomaly, 159
market capitalization-weighted index, 137
market float, 137
market multiple models, 206

market-on-close orders, 120
marketplace businesses, 25
market-to-book ratio, 174
market value, 156
market value of equity, 173
market value-weighted index, 140
matrix pricing, 62
mature stage, 192
MM I with taxes, 79
MM Proposition I (no taxes), 76
momentum effects, 160
money markets, 107
multilateral trading facilities (MTFs), 111
multiplier models, 206
mutual funds, 109
mutually exclusive projects, 34

N

negative screening, 17
net agency cost of equity, 83
net asset value (NAV), 161
net borrowing, 211
net operating cycle, 53
net operating profit after tax (NOPAT), 39
net present value (NPV), 35
network effects, 26
nominations committee, 14
non-bank lenders, 49
non-cumulative preference shares, 168
non-cyclical firm, 182
non-participating preference shares, 168
not-for-profit corporation, 2
number of days of payables, 52
number of days of receivables, 52

O

offer price, 118

omnichannel strategy, 24
operating breakeven quantity of sales, 98
operating cash flow, 47
operating cycle, 53
operating leverage, 28
operating risk, 92
operational efficiency, 125
opportunity costs, 33
optimal capital structure, 80
optional products, 24
option contract, 110
order-driven markets, 123
order matching rules, 124
overdraft lines of credit, 48
overreaction effect, 160
over-the-counter markets, 123
ownership capital, 2

P

participating preference shares, 168
partnership agreement, 2
passive investment strategy, 155
payables payment period, 52
payables turnover ratio, 52
pay-in-advance models, 26
payment date, 208
payments-in-lieu, 115
pecking order theory, 84
peer group, 186
penetration pricing, 24
physical derivative contracts, 107
pooled investments, 109
positive screening, 17
post-trade transparent, 124
predatory pricing, 196
preference shares, 168
preferred stock, 108, 168, 212
present value models, 206
pre-trade transparent, 124

price-book value (P/B) ratio, 174, 219
price-cash flow (P/CF) ratio, 219
price discrimination, 24
price-driven markets, 123
price-earnings (P/E) ratio, 219
price index, 135
price multiple approach, 218
price multiples based on comparables, 219
price multiples based on fundamentals, 219
price priority, 124
price return, 135
price-sales (P/S) ratio, 219
price-setting options, 40
price versus market value weighting, 140
price-weighted index, 136, 139
primary capital markets, 121
primary dealers, 112
primary market, 107
principal-agent conflict, 10
principal business activity, 182
private corporation, 3
private debt, 49
private investment in public equity (PIPE), 169
private label manufacturers, 25
private placement, 3, 122
private placement memorandum (PPR), 3
private securities, 107
probability of financial distress, 80
production-flexibility options, 40
product market risk, 27
product or service differentiation strategy, 195
project sequencing, 42
proxy, 12, 167
proxy fight, 13
public corporation, 3
public debt, 49
public securities, 107
pulls on liquidity, 51
put option, 110

Q

- quick ratio, 51
- quote-driven markets, 123

R

- razors-and-blades, 24
- real assets, 110
- real estate indexes, 147
- real options, 40
- rebalancing, 143
- receivables turnover, 51
- reconstitution, 143
- regular dividends, 207
- regulatory/compliance projects, 31
- related party transactions, 11
- relative/best-in-class, 17
- remuneration committee, 14
- residual claim, 4
- residual losses, 84
- responsible investing, 17
- return index, 135
- return on capital, 39
- return on equity (ROE), 173
- return on invested capital (ROIC), 39
- reverse stock splits, 207
- revolving line of credit, 48
- rights offering, 122
- risk-adjusted returns, 158
- risk committee, 14

S

- sales risk, 92
- seasoned offerings, 121
- secondary financial markets, 121
- secondary issues, 121
- secondary market, 107

secondary precedence rule, 124
secondary sources of liquidity, 50
secured (asset-backed) loans, 48
securities, 108
security market index, 135
semi-strong form market efficiency, 157
senior managers, 10
shakeout stage, 192
shareholders, 10
shareholder theory, 9
share repurchase, 207
shelf registration, 122
short position, 114
short rebate rate, 115
short sale, 115
size effect, 160
socially responsible investing, 17
sole proprietorship, 1
special dividends, 207
special purpose acquisition company (SPAC), 4
special resolutions, 13
sponsored depository receipt, 171
spot markets, 107
spreadsheet modeling, 196
stakeholder management, 12
stakeholder theory, 9
standing limit orders, 119
static trade-off theory, 80
statutory voting, 167
stock dividends, 207
stock splits, 207
stop-buy, 120
stop loss orders, 120
stop orders, 120
stop-sell order, 120
strategic analysis, 187
strategic groups, 187
strong-form market efficiency, 157
subscription model, 25
sunk costs, 33
suppliers, 10

supply chain, 25
sustainable growth rate, 214
sustainable investing, 17
swap contract, 110
switching costs, 27, 191

T

target capital structure, 82
tax-loss selling, 160
tax shield provided by debt, 79
technical analysis, 158
tender offer, 13
terminal value, 209
thematic investing, 18
tiered pricing, 24
timing options, 40
trade pricing rules, 124
traditional investment markets, 108
tranche, 109
turn-of-the-year effect, 160

U

uncommitted line of credit, 48
unconventional cash flow pattern, 33
underwritten offering, 122
unlevered beta, 66
unsponsored depository receipt, 171
unweighted index bias, 141

V

value-added resellers, 25
value-based pricing, 24
value chain, 25
value effect, 160
value of a levered firm, 79

value of an unlevered firm, 79
value proposition, 25
value stocks, 160
value-weighted index, 137
venture capital, 169

W

warrants, 108
weak-form market efficiency, 157
web-based lender, 49
weighted average cost of capital (WACC), 59
window dressing, 160
working capital, 51

Notes



2023 CFA® Exam Prep

SchweserNotes™

Fixed Income, Derivatives, and
Alternative Investments



LEVEL I BOOK 4

KAPLAN SCHWEISER

Book 4: Fixed Income, Derivatives, and Alternative Investments

SchweserNotes™ 2023

Level I CFA®



SCHWESERNOTES™ 2023 LEVEL I CFA® BOOK 4: FIXED INCOME, DERIVATIVES, AND ALTERNATIVE INVESTMENTS

©2022 Kaplan, Inc. All rights reserved.

Published in 2022 by Kaplan, Inc.

Printed in the United States of America.

ISBN: 978-1-0788-2591-7

These materials may not be copied without written permission from the author. The unauthorized duplication of these notes is a violation of global copyright laws and the CFA Institute Code of Ethics. Your assistance in pursuing potential violators of this law is greatly appreciated.

Required CFA Institute disclaimer: Kaplan Schweser is a CFA Institute Prep Provider. Only CFA Institute Prep Providers are permitted to make use of CFA Institute copyrighted materials which are the building blocks of the exam. We are also required to update our materials every year and this is validated by CFA Institute.

CFA Institute does not endorse, promote, review or warrant the accuracy or quality of the product and services offered by Kaplan Schweser. CFA Institute®, CFA® and “Chartered Financial Analyst®” are trademarks owned by CFA Institute.

Certain materials contained within this text are the copyrighted property of CFA Institute. The following is the copyright disclosure for these materials: “Copyright, 2022, CFA Institute. Reproduced and republished from 2023 Learning Outcome Statements, Level I, II, and III questions from CFA® Program Materials, CFA Institute Standards of Professional Conduct, and CFA Institute’s Global Investment Performance Standards with permission from CFA Institute. All Rights Reserved.”

Disclaimer: The SchweserNotes should be used in conjunction with the original readings as set forth by CFA Institute in their 2023 Level I CFA Study Guide. The information contained in these Notes covers topics contained in the readings referenced by CFA Institute and is believed to be accurate. However, their accuracy cannot be guaranteed nor is any warranty conveyed as to your ultimate exam success. The authors of the referenced readings have not endorsed or sponsored these Notes.

CONTENTS

Learning Outcome Statements (LOS)

FIXED INCOME

READING 42

Fixed-Income Securities: Defining Elements

Exam Focus

Module 42.1: Bond Indentures, Regulation, and Taxation

Module 42.2: Bond Cash Flows and Contingencies

Key Concepts

Answer Key for Module Quizzes

READING 43

Fixed-Income Markets: Issuance, Trading, and Funding

Exam Focus

Module 43.1: Types of Bonds and Issuers

Module 43.2: Corporate Debt and Funding Alternatives

Key Concepts

Answer Key for Module Quizzes

READING 44

Introduction to Fixed-Income Valuation

Exam Focus

Module 44.1: Bond Valuation and Yield to Maturity

Module 44.2: Spot Rates and Accrued Interest

Module 44.3: Yield Measures

Module 44.4: Yield Curves

Module 44.5: Yield Spreads

Key Concepts

Answer Key for Module Quizzes

READING 45

Introduction to Asset-Backed Securities

Exam Focus

Module 45.1: Structure of Mortgage-Backed Securities

Module 45.2: Prepayment Risk and Non-Mortgage-Backed ABS

Key Concepts

Answer Key for Module Quizzes

READING 46

Understanding Fixed-Income Risk and Return
Exam Focus
Module 46.1: Sources of Returns, Duration
Module 46.2: Interest Rate Risk and Money Duration
Module 46.3: Convexity and Yield Volatility
Key Concepts
Answer Key for Module Quizzes

READING 47

Fundamentals of Credit Analysis
Exam Focus
Module 47.1: Credit Risk and Bond Ratings
Module 47.2: Evaluating Credit Quality
Key Concepts
Answer Key for Module Quizzes

Topic Quiz: Fixed Income

DERIVATIVES

READING 48

Derivative Instrument and Derivative Market Features
Module 48.1: Derivatives Markets
Key Concepts
Answer Key for Module Quizzes

READING 49

Forward Commitment and Contingent Claim Features and Instruments
Module 49.1: Forwards and Futures
Module 49.2: Swaps and Options
Key Concepts
Answer Key for Module Quizzes

READING 50

Derivative Benefits, Risks, and Issuer and Investor Uses
Module 50.1: Uses, Benefits, and Risks of Derivatives
Key Concepts
Answer Key for Module Quizzes

READING 51

Arbitrage, Replication, and the Cost of Carry in Pricing Derivatives
Module 51.1: Arbitrage, Replication, and Carrying Costs
Key Concepts
Answer Key for Module Quizzes

READING 52

Pricing and Valuation of Forward Contracts and for an Underlying with Varying Maturities

Module 52.1: Forward Contract Valuation

Key Concepts

Answer Key for Module Quizzes

READING 53

Pricing and Valuation of Futures Contracts

Module 53.1: Futures Valuation

Key Concepts

Answer Key for Module Quizzes

READING 54

Pricing and Valuation of Interest Rates and Other Swaps

Module 54.1: Swap Valuation

Key Concepts

Answer Key for Module Quizzes

READING 55

Pricing and Valuation of Options

Module 55.1: Option Valuation

Key Concepts

Answer Key for Module Quizzes

READING 56

Option Replication Using Put–Call Parity

Module 56.1: Put–Call Parity

Key Concepts

Answer Key for Module Quizzes

READING 57

Valuing a Derivative Using a One-Period Binomial Model

Module 57.1: Binomial Model for Option Values

Key Concepts

Answer Key for Module Quizzes

Topic Quiz: Derivatives

ALTERNATIVE INVESTMENTS

READING 58

Categories, Characteristics, and Compensation Structures of Alternative Investments

Module 58.1: Alternative Investment Structures

Key Concepts

Answer Key for Module Quizzes

READING 59

Performance Calculation and Appraisal of Alternative Investments

Module 59.1: Performance Appraisal and Return Calculations

Key Concepts

Answer Key for Module Quizzes

READING 60

Private Capital, Real Estate, Infrastructure, Natural Resources, and Hedge Funds

Module 60.1: Private Equity and Private Debt

Module 60.2: Real Estate, Infrastructure, and Natural Resources

Module 60.3: Hedge Funds

Key Concepts

Answer Key for Module Quizzes

Topic Quiz: Alternative Investments

Appendix

Formulas

Index

LEARNING OUTCOME STATEMENTS (LOS)

42. Fixed-Income Securities: Defining Elements

The candidate should be able to:

- a. describe basic features of a fixed-income security.
- b. describe content of a bond indenture.
- c. compare affirmative and negative covenants and identify examples of each.
- d. describe how legal, regulatory, and tax considerations affect the issuance and trading of fixed-income securities.
- e. describe how cash flows of fixed-income securities are structured.
- f. describe contingency provisions affecting the timing and/or nature of cash flows of fixed-income securities and whether such provisions benefit the borrower or the lender.

43. Fixed-Income Markets: Issuance, Trading, and Funding

The candidate should be able to:

- a. describe classifications of global fixed-income markets.
- b. describe the use of interbank offered rates as reference rates in floating-rate debt.
- c. describe mechanisms available for issuing bonds in primary markets.
- d. describe secondary markets for bonds.
- e. describe securities issued by sovereign governments.
- f. describe securities issued by non-sovereign governments, quasi-government entities, and supranational agencies.
- g. describe types of debt issued by corporations.
- h. describe structured financial instruments.
- i. describe short-term funding alternatives available to banks.
- j. describe repurchase agreements (repos) and the risks associated with them.

44. Introduction to Fixed-Income Valuation

The candidate should be able to:

- a. calculate a bond's price given a market discount rate.
- b. identify the relationships among a bond's price, coupon rate, maturity, and market discount rate (yield-to-maturity).
- c. define spot rates and calculate the price of a bond using spot rates.
- d. describe and calculate the flat price, accrued interest, and the full price of a bond.
- e. describe matrix pricing.
- f. calculate annual yield on a bond for varying compounding periods in a year.
- g. calculate and interpret yield measures for fixed-rate bonds and floating-rate notes.
- h. calculate and interpret yield measures for money market instruments.
- i. define and compare the spot curve, yield curve on coupon bonds, par curve, and forward curve.
- j. define forward rates and calculate spot rates from forward rates, forward rates from spot rates, and the price of a bond using forward rates.
- k. compare, calculate, and interpret yield spread measures.

45. Introduction to Asset-Backed Securities

The candidate should be able to:

- a. explain benefits of securitization for economies and financial markets.

- b. describe securitization, including the parties involved in the process and the roles they play.
- c. describe typical structures of securitizations, including credit tranching and time tranching.
- d. describe types and characteristics of residential mortgage loans that are typically securitized.
- e. describe types and characteristics of residential mortgage-backed securities, including mortgage pass-through securities and collateralized mortgage obligations, and explain the cash flows and risks for each type.
- f. define prepayment risk and describe the prepayment risk of mortgage-backed securities.
- g. describe characteristics and risks of commercial mortgage-backed securities.
- h. describe types and characteristics of non-mortgage asset-backed securities, including the cash flows and risks of each type.
- i. describe collateralized debt obligations, including their cash flows and risks.
- j. describe characteristics and risks of covered bonds and how they differ from other asset-backed securities.

46. Understanding Fixed-Income Risk and Return

The candidate should be able to:

- a. calculate and interpret the sources of return from investing in a fixed-rate bond.
- b. define, calculate, and interpret Macaulay, modified, and effective durations.
- c. explain why effective duration is the most appropriate measure of interest rate risk for bonds with embedded options.
- d. define key rate duration and describe the use of key rate durations in measuring the sensitivity of bonds to changes in the shape of the benchmark yield curve.
- e. explain how a bond's maturity, coupon, and yield level affect its interest rate risk.
- f. calculate the duration of a portfolio and explain the limitations of portfolio duration.
- g. calculate and interpret the money duration of a bond and price value of a basis point (PVBP).
- h. calculate and interpret approximate convexity and compare approximate and effective convexity.
- i. calculate the percentage price change of a bond for a specified change in yield, given the bond's approximate duration and convexity.
- j. describe how the term structure of yield volatility affects the interest rate risk of a bond.
- k. describe the relationships among a bond's holding period return, its duration, and the investment horizon.
- l. explain how changes in credit spread and liquidity affect yield-to-maturity of a bond and how duration and convexity can be used to estimate the price effect of the changes.
- m. describe the difference between empirical duration and analytical duration.

47. Fundamentals of Credit Analysis

The candidate should be able to:

- a. describe credit risk and credit-related risks affecting corporate bonds.
- b. describe default probability and loss severity as components of credit risk.
- c. describe seniority rankings of corporate debt and explain the potential violation of the priority of claims in a bankruptcy proceeding.
- d. compare and contrast corporate issuer credit ratings and issue credit ratings and describe the rating agency practice of "notching".

- e. explain risks in relying on ratings from credit rating agencies.
- f. explain the four Cs (Capacity, Collateral, Covenants, and Character) of traditional credit analysis.
- g. calculate and interpret financial ratios used in credit analysis.
- h. evaluate the credit quality of a corporate bond issuer and a bond of that issuer, given key financial ratios of the issuer and the industry.
- i. describe macroeconomic, market, and issuer-specific factors that influence the level and volatility of yield spreads.
- j. explain special considerations when evaluating the credit of high-yield, sovereign, and non-sovereign government debt issuers and issues.

48. Derivative Instrument and Derivative Market Features

The candidate should be able to:

- a. define a derivative and describe basic features of a derivative instrument.
- b. describe the basic features of derivative markets, and contrast over-the-counter and exchange-traded derivative markets.

49. Forward Commitment and Contingent Claim Features and Instruments

The candidate should be able to:

- a. define forward contracts, futures contracts, swaps, options (calls and puts), and credit derivatives and compare their basic characteristics.
- b. determine the value at expiration and profit from a long or a short position in a call or put option.
- c. contrast forward commitments with contingent claims.

50. Derivative Benefits, Risks, and Issuer and Investor Uses

The candidate should be able to:

- a. describe benefits and risks of derivative instruments.
- b. compare the use of derivatives among issuers and investors.

51. Arbitrage, Replication, and the Cost of Carry in Pricing Derivatives

The candidate should be able to:

- a. explain how the concepts of arbitrage and replication are used in pricing derivatives.
- b. explain the difference between the spot and expected future price of an underlying and the cost of carry associated with holding the underlying asset.

52. Pricing and Valuation of Forward Contracts and for an Underlying with Varying Maturities

The candidate should be able to:

- a. explain how the value and price of a forward contract are determined at initiation, during the life of the contract, and at expiration.
- b. explain how forward rates are determined for an underlying with a term structure and describe their uses.

53. Pricing and Valuation of Futures Contracts

The candidate should be able to:

- a. compare the value and price of forward and futures contracts.
- b. explain why forward and futures prices differ.

54. Pricing and Valuation of Interest Rates and Other Swaps

The candidate should be able to:

- a. describe how swap contracts are similar to but different from a series of forward contracts.
- b. contrast the value and price of swaps.

55. Pricing and Valuation of Options

The candidate should be able to:

- a. explain the exercise value, moneyness, and time value of an option.
- b. contrast the use of arbitrage and replication concepts in pricing forward commitments and contingent claims.
- c. identify the factors that determine the value of an option and describe how each factor affects the value of an option.

56. Option Replication Using Put–Call Parity

The candidate should be able to:

- a. explain put–call parity for European options.
- b. explain put–call *forward* parity for European options.

57. Valuing a Derivative Using a One-Period Binomial Model

The candidate should be able to:

- a. explain how to value a derivative using a one-period binomial model.
- b. describe the concept of risk neutrality in derivatives pricing.

58. Categories, Characteristics, and Compensation Structures of Alternative Investments

The candidate should be able to:

- a. describe types and categories of alternative investments.
- b. describe characteristics of direct investment, co-investment, and fund investment methods for alternative investments.
- c. describe investment and compensation structures commonly used in alternative investments.

59. Performance Calculation and Appraisal of Alternative Investments

The candidate should be able to:

- a. describe issues in performance appraisal of alternative investments.
- b. calculate and interpret returns of alternative investments both before and after fees.

60. Private Capital, Real Estate, Infrastructure, Natural Resources, and Hedge Funds

The candidate should be able to:

- a. explain investment characteristics of private equity.
- b. explain investment characteristics of private debt.
- c. explain investment characteristics of real estate.
- d. explain investment characteristics of infrastructure.
- e. explain investment characteristics of natural resources.
- f. explain investment characteristics of hedge funds.

READING 42

FIXED-INCOME SECURITIES: DEFINING ELEMENTS

EXAM FOCUS

Here your focus should be on learning the basic characteristics of debt securities and as much of the bond terminology as you can remember. Key items are the coupon structure of bonds and options embedded in bonds: call options, put options, and conversion (to common stock) options.

MODULE 42.1: BOND INDENTURES, REGULATION, AND TAXATION



Video covering this content is available online.

There are two important points about fixed-income securities that we will develop further along in the Fixed Income readings but may be helpful as you study this reading.

- The most common type of fixed-income security is a bond that promises to make a series of interest payments in fixed amounts and to repay the principal amount at maturity. When market interest rates (i.e., yields on bonds) *increase*, the value of such bonds *decreases* because the present value of a bond's promised cash flows decreases when a higher discount rate is used.
- Bonds are rated based on their relative probability of default (failure to make promised payments). Because investors prefer bonds with lower probability of default, bonds with lower credit quality must offer investors higher yields to compensate for the greater probability of default. Other things equal, a decrease in a bond's rating (an increased probability of default) will decrease the price of the bond, thus increasing its yield.

LOS 42.a: Describe basic features of a fixed-income security.

The features of a fixed-income security include specification of:

- The issuer of the bond.
- The maturity date of the bond.
- The par value (principal value to be repaid).
- Coupon rate and frequency.
- Currency in which payments will be made.

Issuers of Bonds

There are several types of entities that issue bonds when they borrow money, including:

- **Corporations.** Often corporate bonds are divided into those issued by financial companies and those issued by nonfinancial companies.
- **Sovereign national governments.** A prime example is U.S. Treasury bonds, but many countries issue sovereign bonds.
- **Non-sovereign governments.** Issued by government entities that are not national governments, such as the state of California or the city of Toronto.
- **Quasi-government entities.** Not a direct obligation of a country's government or central bank. An example is the Federal National Mortgage Association (Fannie Mae).
- **Supranational entities.** Issued by organizations that operate globally such as the World Bank, the European Investment Bank, and the International Monetary Fund (IMF).
- **Special purpose entities.** These are corporations set up to purchase financial assets and issue *asset-backed securities*, which are bonds backed by the cash flows from those assets.

Bond Maturity

The maturity date of a bond is the date on which the principal is to be repaid. Once a bond has been issued, the time remaining until maturity is referred to as the **term to maturity** or **tenor** of a bond.

When bonds are issued, their terms to maturity range from one day to 30 years or more. Both Disney and Coca-Cola have issued bonds with original maturities of 100 years. Bonds that have no maturity date are called **perpetual bonds**. They make periodic interest payments but do not promise to repay the principal amount.

Bonds with original maturities of one year or less are referred to as **money market securities**. Bonds with original maturities of more than one year are referred to as **capital market securities**.

Par Value

The **par value** of a bond is the **principal amount** that will be repaid at maturity. The par value is also referred to as the *face value*, *maturity value*, *redemption value*, or *principal value* of a bond. Bonds can have a par value of any amount, and their prices are quoted as a percentage of par. A bond with a par value of \$1,000 quoted at 98 is selling for \$980.

A bond that is selling for more than its par value is said to be trading at a **premium** to par; a bond that is selling at less than its par value is said to be trading at a **discount** to par; and a bond that is selling for exactly its par value is said to be trading **at par**.

Coupon Payments

The coupon rate on a bond is the annual percentage of its par value that will be paid to bondholders. Some bonds make coupon interest payments annually, while others make semiannual, quarterly, or monthly payments. A \$1,000 par value semiannual-pay bond with a 5% coupon would pay 2.5% of \$1,000, or \$25, every six months. A bond with a fixed coupon rate is called a **plain vanilla bond** or a **conventional bond**.

Some bonds pay no interest prior to maturity and are called **zero-coupon bonds** or **pure discount bonds**. *Pure discount* refers to the fact that these bonds are sold at a discount to their par value and the interest is all paid at maturity when bondholders receive the par value. A 10-year, \$1,000, zero-coupon bond yielding 7% would sell at about \$500 initially and pay \$1,000 at maturity. We discuss various other coupon structures later in this reading.

Currencies

Bonds are issued in many currencies. Sometimes borrowers from countries with volatile currencies issue bonds denominated in euros or U.S. dollars to make them more attractive to a wide range of investors. A **dual-currency bond** makes coupon interest payments in one currency and the principal repayment at maturity in another currency. A **currency option bond** gives bondholders a choice of which of two currencies they would like to receive their payments in.

LOS 42.b: Describe content of a bond indenture.

LOS 42.c: Compare affirmative and negative covenants and identify examples of each.

The legal contract between the bond issuer (borrower) and bondholders (lenders) is called a **trust deed**, and in the United States and Canada, it is also often referred to as the **bond indenture**. The indenture defines the obligations of and restrictions on the borrower and forms the basis for all future transactions between the bondholder and the issuer.

The provisions in the bond indenture are known as *covenants* and include both *negative covenants* (prohibitions on the borrower) and *affirmative covenants* (actions the borrower promises to perform).

Negative covenants include restrictions on asset sales (the company can't sell assets that have been pledged as collateral), negative pledge of collateral (the company can't claim that the same assets back several debt issues simultaneously), and restrictions on additional borrowings (the company can't borrow additional money unless certain financial conditions are met).

Negative covenants serve to protect the interests of bondholders and prevent the issuing firm from taking actions that would increase the risk of default. At the same time, the covenants must not be so restrictive that they prevent the firm from taking advantage of opportunities that arise or responding appropriately to changing business circumstances.

Affirmative covenants do not typically restrict the operating decisions of the issuer. Common affirmative covenants are to make timely interest and principal payments to bondholders, to insure and maintain assets, and to comply with applicable laws and regulations.

Two examples of affirmative covenants are cross-default and pari passu provisions. A **cross-default** clause states that if the issuer defaults on any other debt obligation, they will also be considered in default on this bond. A **pari passu** clause states that this bond will have the same priority of claims as the issuer's other senior debt issues.

LOS 42.d: Describe how legal, regulatory, and tax considerations affect the issuance and trading of fixed-income securities.

Bonds are subject to different legal and regulatory requirements depending on where they are issued and traded. Bonds issued by a firm domiciled in a country and also traded in that country's currency are referred to as **domestic bonds**. Bonds issued by a firm incorporated in a foreign country that trade on the **national bond market** of another country in that country's currency are referred to as **foreign bonds**. Examples include bonds issued by foreign firms that trade in China and are denominated in yuan, which are called *panda bonds*, and bonds issued by firms incorporated outside the United States that trade in the United States and are denominated in U.S. dollars, which are called *Yankee bonds*.

Eurobonds are issued outside the jurisdiction of any one country and denominated in a currency different from the currency of the countries in which they are sold. They are subject to less regulation than domestic bonds in most jurisdictions and were initially introduced to avoid U.S. regulations. Eurobonds should not be confused with bonds denominated in euros or thought to originate in Europe, although they can be both. Eurobonds got the "euro" name because they were first introduced in Europe, and most are still traded by firms in European capitals. A bond issued by a Chinese firm that is denominated in yen and traded in markets outside Japan would fit the definition of a eurobond. Eurobonds that trade in the national bond market of a country other than the country that issues the currency the bond is denominated in, and in the eurobond market, are referred to as **global bonds**.

Eurobonds are referred to by the currency they are denominated in. Eurodollar bonds are denominated in U.S. dollars, and euroyen bonds are denominated in yen. At one time, the majority of eurobonds were issued in **bearer** form. Ownership of bearer bonds is evidenced simply by possessing the bonds, whereas ownership of **registered bonds** is recorded. Bearer bonds may be more attractive than registered bonds to those seeking to avoid taxes. As with most other bonds, eurobonds are now issued in registered form.

Other legal and regulatory issues addressed in a trust deed include:

- Legal information about the entity issuing the bond.
- Any assets (collateral) pledged to support repayment of the bond.
- Any additional features that increase the probability of repayment (credit enhancements).
- Covenants describing any actions the firm must take and any actions the firm is prohibited from taking.

Issuing entities

Bonds are issued by several types of legal entities, and bondholders must be aware of which entity has actually promised to make the interest and principal payments. Sovereign bonds are most often issued by the treasury of the issuing country.

Corporate bonds may be issued by a well-known corporation such as Microsoft, by a subsidiary of a company, or by a holding company that is the overall owner of several operating companies. Bondholders must pay attention to the specific entity issuing the bonds because the credit quality can differ among related entities.

Sometimes an entity is created solely for the purpose of owning specific assets and issuing bonds to provide the funds to purchase the assets. These entities are referred to as **special purpose entities (SPEs)** in the United States and special purpose vehicles (SPVs) in Europe.

Bonds issued by these entities are called **securitized bonds**. As an example, a firm could sell loans it has made to customers to an SPE that issues bonds to purchase the loans. The interest and principal payments on the loans are then used to make the interest and principal payments on the bonds.

Often, an SPE can issue bonds at a lower interest rate than bonds issued by the originating corporation. This is because the assets supporting the bonds are owned by the SPE and are used to make the payments to holders of the securitized bonds even if the company itself runs into financial trouble. For this reason, SPEs are called **bankruptcy remote vehicles** or entities.

Sources of repayment

Sovereign bonds are typically repaid by the tax receipts of the issuing country. Bonds issued by non-sovereign government entities are repaid by either general taxes, revenues of a specific project (e.g., an airport), or by special taxes or fees dedicated to bond repayment (e.g., a water district or sewer district).

Corporate bonds are generally repaid from cash generated by the firm's operations. As noted previously, securitized bonds are repaid from the cash flows of the financial assets owned by the SPE.

Collateral and credit enhancements

Unsecured bonds represent a claim to the overall assets and cash flows of the issuer. **Secured bonds** are backed by a claim to specific assets of a corporation, which reduces their risk of default and, consequently, the yield that investors require on the bonds. Assets pledged to support a bond issue (or any loan) are referred to as **collateral**.

Because they are backed by collateral, secured bonds are *senior* to unsecured bonds. Among unsecured bonds, two different issues may have different priority in the event of bankruptcy or liquidation of the issuing entity. The claim of senior unsecured debt is below (after) that of secured debt but ahead of *subordinated*, or junior, debt.

Sometimes secured debt is referred to by the type of collateral pledged. **Equipment trust certificates** are debt securities backed by equipment such as railroad cars and oil drilling rigs. **Collateral trust bonds** are backed by financial assets, such as stocks and (other) bonds. Be aware that while the term **debentures** refers to unsecured debt in the United States and elsewhere, in Great Britain and some other countries the term refers to bonds collateralized by specific assets.

The most common type of securitized bond is a **mortgage-backed security (MBS)**. The underlying assets are a pool of mortgages, and the interest and principal payments from the mortgages are used to pay the interest and principal on the MBS.

In some countries, especially European countries, financial companies issue **covered bonds**. Covered bonds are similar to asset-backed securities, but the underlying assets (the cover pool), although segregated, remain on the balance sheet of the issuing corporation (i.e., no SPE is created).

Credit enhancement can be either internal (built into the structure of a bond issue) or external (provided by a third party). One method of internal credit enhancement is *overcollateralization*,

in which the collateral pledged has a value greater than the par value of the debt issued. One limitation of this method of credit enhancement is that the additional collateral is also the underlying assets, so when asset defaults are high, the value of the excess collateral declines in value.

Two other methods of internal credit enhancement are a *cash reserve fund* and an *excess spread account*. A cash reserve fund is cash set aside to make up for credit losses on the underlying assets. With an excess spread account, the yield promised on the bonds issued is less than the promised yield on the assets supporting the ABS. This gives some protection if the yield on the financial assets is less than anticipated. If the assets perform as anticipated, the excess cash flow from the collateral can be used to retire (pay off the principal on) some of the outstanding bonds.

Another method of internal credit enhancement is to divide a bond issue into *tranches* (French for *slices*) with different seniority of claims. Any losses due to poor performance of the assets supporting a securitized bond are first absorbed by the bonds with the lowest seniority, then the bonds with the next-lowest priority of claims. The most senior tranches in this structure can receive very high credit ratings because the probability is very low that losses will be so large that they cannot be absorbed by the subordinated tranches. The subordinated tranches must have higher yields to compensate investors for the additional risk of default. This is sometimes referred to as *waterfall structure* because available funds first go to the most senior tranche of bonds, then to the next-highest priority bonds, and so forth.

External credit enhancements include surety bonds, bank guarantees, and letters of credit from financial institutions. *Surety bonds* are issued by insurance companies and are a promise to make up any shortfall in the cash available to service the debt. *Bank guarantees* serve the same function. A *letter of credit* is a promise to lend money to the issuing entity if it does not have enough cash to make the promised payments on the covered debt. While all three of these external credit enhancements increase the credit quality of debt issues and decrease their yields, deterioration of the credit quality of the guarantor will also reduce the credit quality of the covered issue.

Taxation of Bond Income

Most often, the interest income paid to bondholders is taxed as ordinary income at the same rate as wage and salary income. The interest income from bonds issued by municipal governments in the United States, however, is most often exempt from national income tax and often from any state income tax in the state of issue.

When a bondholder sells a coupon bond prior to maturity, it may be at a gain or a loss relative to its purchase price. Such gains and losses are considered capital gains income (rather than ordinary taxable income). Capital gains are often taxed at a lower rate than ordinary income. Capital gains on the sale of an asset that has been owned for more than some minimum amount of time may be classified as *long-term* capital gains and taxed at an even lower rate.

Pure-discount bonds and other bonds sold at significant discounts to par when issued are termed **original issue discount (OID)** bonds. Because the gains over an OID bond's tenor as the price moves towards par value are really interest income, these bonds can generate a tax liability even when no cash interest payment has been made. In many tax jurisdictions, a portion of the discount from par at issuance is treated as taxable interest income each year. This tax

treatment also allows that the tax basis of the OID bonds is increased each year by the amount of interest income recognized, so there is no additional capital gains tax liability at maturity.

Some tax jurisdictions provide a symmetric treatment for bonds issued at a premium to par, allowing part of the premium to be used to reduce the taxable portion of coupon interest payments.



MODULE QUIZ 42.1

1. A dual-currency bond pays coupon interest in a currency:
 - A. of the bondholder's choice.
 - B. other than the home currency of the issuer.
 - C. other than the currency in which it repays principal.
2. A bond's indenture:
 - A. contains its covenants.
 - B. is the same as a debenture.
 - C. relates only to its interest and principal payments.
3. A clause in a bond indenture that requires the borrower to perform a certain action is *most accurately* described as:
 - A. a trust deed.
 - B. a negative covenant.
 - C. an affirmative covenant.
4. An investor buys a pure-discount bond, holds it to maturity, and receives its par value. For tax purposes, the increase in the bond's value is *most likely* to be treated as:
 - A. a capital gain.
 - B. interest income.
 - C. tax-exempt income.

MODULE 42.2: BOND CASH FLOWS AND CONTINGENCIES



Video covering
this content is
available online.

LOS 42.e: Describe how cash flows of fixed-income securities are structured.

A typical bond has a **bullet** structure. Periodic interest payments (coupon payments) are made over the life of the bond, and the principal value is paid with the final interest payment at maturity. The interest payments are referred to as the bond's **coupons**. When the final payment includes a lump sum in addition to the final period's interest, it is referred to as a **balloon payment**.

Consider a \$1,000 face value 5-year bond with an annual coupon rate of 5%. With a bullet structure, the bond's promised payments at the end of each year would be as follows.

Year	1	2	3	4	5
PMT	\$50	\$50	\$50	\$50	\$1,050
Principal remaining	\$1,000	\$1,000	\$1,000	\$1,000	\$0

A loan structure in which the periodic payments include both interest and some repayment of principal (the amount borrowed) is called an **amortizing loan**. If a bond (loan) is **fully amortizing**, this means the principal is fully paid off when the last periodic payment is made. Typically, automobile loans and home loans are fully amortizing loans. If the 5-year, 5% bond in

the previous table had a fully amortizing structure rather than a bullet structure, the payments and remaining principal balance at each year-end would be as follows (final payment reflects rounding of previous payments).

Year	1	2	3	4	5
PMT	\$230.97	\$230.97	\$230.97	\$230.97	\$230.98
Principal remaining	\$819.03	\$629.01	\$429.49	\$219.99	\$0

A bond can also be structured to be **partially amortizing** so that there is a balloon payment at bond maturity, just as with a bullet structure. However, unlike a bullet structure, the final payment includes just the remaining unamortized principal amount rather than the full principal amount. In the following table, the final payment includes \$200 to repay the remaining principal outstanding.

Year	1	2	3	4	5
PMT	\$194.78	\$194.78	\$194.78	\$194.78	\$394.78
Principal remaining	\$855.22	\$703.20	\$543.58	\$375.98	\$0

Sinking fund provisions provide for the repayment of principal through a series of payments over the life of the issue. For example, a 20-year issue with a face amount of \$300 million may require that the issuer retire \$20 million of the principal every year beginning in the sixth year.

Details of sinking fund provisions vary. There may be a period during which no sinking fund redemptions are made. The amount of bonds redeemed according to the sinking fund provision could decline each year or increase each year.

The price at which bonds are redeemed under a sinking fund provision is typically par but can be different from par. If the market price is less than the sinking fund redemption price, the issuer can satisfy the sinking fund provision by buying bonds in the open market with a par value equal to the amount of bonds that must be redeemed. This would be the case if interest rates had risen since issuance so that the bonds were trading below the sinking fund redemption price.

Sinking fund provisions offer both advantages and disadvantages to bondholders. On the plus side, bonds with a sinking fund provision have less credit risk because the periodic redemptions reduce the total amount of principal to be repaid at maturity. The presence of a sinking fund, however, can be a disadvantage to bondholders when interest rates fall.

This disadvantage to bondholders can be seen by considering the case where interest rates have fallen since bond issuance, so the bonds are trading at a price above the sinking fund redemption price. In this case, the bond trustee will select outstanding bonds for redemption randomly. A bondholder would suffer a loss if her bonds were selected to be redeemed at a price below the current market price. This means the bonds have more *reinvestment risk* because bondholders who have their bonds redeemed can only reinvest the funds at the new, lower yield (assuming they buy bonds of similar risk).



PROFESSOR'S NOTE

The concept of reinvestment risk is developed more in subsequent readings. It can be defined as the uncertainty about the interest to be earned on cash flows from a bond

that are reinvested in other debt securities. In the case of a bond with a sinking fund, the greater probability of receiving the principal repayment prior to maturity increases the expected cash flows during the bond's life and, therefore, the uncertainty about interest income on reinvested funds.

There are several coupon structures besides a fixed-coupon structure, and we summarize the most important ones here.

Floating-Rate Notes

Some bonds pay periodic interest that depends on a current market rate of interest. These bonds are called **floating-rate notes (FRN)** or **floaters**. The market rate of interest is called the **market reference rate (MRR)**, and an FRN promises to pay the reference rate plus some interest margin. This added margin is typically expressed in **basis points**, which are hundredths of 1%. A 120 basis point margin is equivalent to 1.2%.

As an example, consider a floating-rate note that pays the London Interbank Offered Rate (LIBOR) plus a margin of 0.75% (75 basis points) annually. If 1-year LIBOR is 2.3% at the beginning of the year, the bond will pay $2.3\% + 0.75\% = 3.05\%$ of its par value at the end of the year. The new 1-year rate at that time will determine the rate of interest paid at the end of the next year. Most floaters pay quarterly and are based on a quarterly (90-day) reference rate.

A floating-rate note may have a **cap**, which benefits the issuer by placing a limit on how high the coupon rate can rise. Often, FRNs with caps also have a **floor**, which benefits the bondholder by placing a minimum on the coupon rate (regardless of how low the reference rate falls). An **inverse floater** has a coupon rate that increases when the reference rate decreases and decreases when the reference rate increases.

Other Coupon Structures

Step-up coupon bonds are structured so that the coupon rate increases over time according to a predetermined schedule. Typically, step-up coupon bonds have a *call feature* that allows the firm to redeem the bond issue at a set price at each step-up date. If the new higher coupon rate is greater than what the market yield would be at the call price, the firm will call the bonds and retire them. This means if market yields rise, a bondholder may, in turn, get a higher coupon rate because the bonds are less likely to be called on the step-up date.

Yields could increase because an issuer's credit rating has fallen, in which case the higher step-up coupon rate simply compensates investors for greater credit risk. Aside from this, we can view step-up coupon bonds as having some protection against increases in market interest rates to the extent they are offset by increases in bond coupon rates.

A **credit-linked coupon bond** carries a provision stating that the coupon rate will go up by a certain amount if the credit rating of the issuer falls and go down if the credit rating of the issuer improves. While this offers some protection against a credit downgrade of the issuer, the higher required coupon payments may make the financial situation of the issuer worse and possibly increase the probability of default.

A **payment-in-kind (PIK) bond** allows the issuer to make the coupon payments by increasing the principal amount of the outstanding bonds, essentially paying bond interest with more bonds. Firms that issue PIK bonds typically do so because they anticipate that firm cash flows

may be less than required to service the debt, often because of high levels of debt financing (leverage). These bonds typically have higher yields because of a lower perceived credit quality from cash flow shortfalls or simply because of the high leverage of the issuing firm.

With a **deferred coupon bond**, also called a **split coupon bond**, regular coupon payments do not begin until a period of time after issuance. These are issued by firms that anticipate cash flows will increase in the future to allow them to make coupon interest payments.

Deferred coupon bonds may be appropriate financing for a firm financing a large project that will not be completed and generating revenue for some period of time after bond issuance. Deferred coupon bonds may offer bondholders tax advantages in some jurisdictions. Zero-coupon bonds can be considered a type of deferred coupon bond.

An **index-linked bond** has coupon payments and/or a principal value that is based on a commodity index, an equity index, or some other published index number. **Inflation-linked bonds** (also called **linkers**) are the most common type of index-linked bonds. Their payments are based on the change in an inflation index, such as the Consumer Price Index (CPI) in the United States. Indexed bonds that will not pay less than their original par value at maturity, even when the index has decreased, are termed **principal protected bonds**.

The different structures of inflation-indexed bonds include the following:

- **Indexed-annuity bonds.** Fully amortizing bonds with the periodic payments directly adjusted for inflation or deflation.
- **Indexed zero-coupon bonds.** The payment at maturity is adjusted for inflation.
- **Interest-indexed bonds.** The coupon rate is adjusted for inflation while the principal value remains unchanged.
- **Capital-indexed bonds.** This is the most common structure. An example is U.S. Treasury Inflation Protected Securities (TIPS). The coupon rate remains constant, and the principal value of the bonds is increased by the rate of inflation (or decreased by deflation).

To better understand the structure of capital-indexed bonds, consider a bond with a par value of \$1,000 at issuance, a 3% annual coupon rate paid semiannually, and a provision that the principal value will be adjusted for inflation (or deflation). If six months after issuance the reported inflation has been 1% over the period, the principal value of the bonds is increased by 1% from \$1,000 to \$1,010, and the six-month coupon of 1.5% is calculated as 1.5% of the new (adjusted) principal value of \$1,010 (i.e., $1,010 \times 1.5\% = \$15.15$).

With this structure we can view the coupon rate of 3% as a real rate of interest. Unexpected inflation will not decrease the purchasing power of the coupon interest payments, and the principal value paid at maturity will have approximately the same purchasing power as the \$1,000 par value did at bond issuance.

LOS 42.f: Describe contingency provisions affecting the timing and/or nature of cash flows of fixed-income securities and whether such provisions benefit the borrower or the lender.

A **contingency provision** in a contract describes an action that may be taken if an event (the contingency) actually occurs. Contingency provisions in bond indentures are referred to as

embedded options, embedded in the sense that they are an integral part of the bond contract and are not a separate security. Some embedded options are exercisable at the option of the issuer of the bond and, therefore, are valuable to the issuer; others are exercisable at the option of the purchaser of the bond and, thus, have value to the bondholder.

Bonds that do not have contingency provisions are referred to as **straight** or **option-free** bonds.

A **call option** gives the *issuer* the right to redeem all or part of a bond issue at a specific price (call price) if they choose to. As an example of a call provision, consider a 6% 20-year bond issued at par on June 1, 2012, for which the indenture includes the following *call schedule*:

- The bonds can be redeemed by the issuer at 102% of par after June 1, 2017.
- The bonds can be redeemed by the issuer at 101% of par after June 1, 2020.
- The bonds can be redeemed by the issuer at 100% of par after June 1, 2022.

For the 5-year period from the issue date until June 2017, the bond is not callable. We say the bond has five years of *call protection*, or that the bond is *call protected* for five years. This 5-year period is also referred to as a *lockout period*, a *cushion*, or a *deferment period*.

June 1, 2017, is referred to as the *first call date*, and the *call price* is 102 (102% of par value) between that date and June 2020. The amount by which the call price is above par is referred to as the *call premium*. The call premium at the first call date in this example is 2%, or \$20 per \$1,000 bond. The call price declines to 101 (101% of par) after June 1, 2020. After June 1, 2022, the bond is callable at par, and that date is referred to as the *first par call date*.

For a bond that is currently callable, the call price puts an upper limit on the value of the bond in the market.

A call option has value to the issuer because it gives the issuer the right to redeem the bond and issue a new bond (borrow) if the market yield on the bond declines. This could occur either because interest rates in general have decreased or because the credit quality of the bond has increased (default risk has decreased).

Consider a situation where the market yield on the previously discussed 6% 20-year bond has declined from 6% at issuance to 4% on June 1, 2017 (the first call date). If the bond did not have a call option, it would trade at approximately \$1,224. With a call price of 102, the issuer can redeem the bonds at \$1,020 each and borrow that amount at the current market yield of 4%, reducing the annual interest payment from \$60 per bond to \$40.80.



PROFESSOR'S NOTE

This is analogous to refinancing a home mortgage when mortgage rates fall in order to reduce the monthly payments.

The issuer will only choose to exercise the call option when it is to their advantage to do so. That is, they can reduce their interest expense by calling the bond and issuing new bonds at a lower yield. Bond buyers are disadvantaged by the call provision and have more reinvestment risk because their bonds will only be called (redeemed prior to maturity) when the proceeds can be reinvested only at a lower yield. For this reason, a callable bond must offer a higher yield (sell at a lower price) than an otherwise identical noncallable bond. The difference in price

between a callable bond and an otherwise identical noncallable bond is equal to the value of the call option to the issuer.

There are three *styles of exercise* for callable bonds:

1. American style—the bonds can be called anytime after the first call date.
2. European style—the bonds can only be called on the call date specified.
3. Bermuda style—the bonds can be called on specified dates after the first call date, often on coupon payment dates.

Note that these are only style names and are not indicative of where the bonds are issued.

To avoid the higher interest rates required on callable bonds but still preserve the option to redeem bonds early when corporate or operating events require it, issuers introduced bonds with **make-whole** call provisions. With a make-whole bond, the call price is not fixed but includes a lump-sum payment based on the present value of the future coupons the bondholder will not receive if the bond is called early.

With a make-whole call provision, the calculated call price is unlikely to be lower than the market value of the bond. Therefore the issuer is unlikely to call the bond except when corporate circumstances, such as an acquisition or restructuring, require it. The make-whole provision does not put an upper limit on bond values when interest rates fall as does a regular call provision. The make-whole provision actually penalizes the issuer for calling the bond. The net effect is that the bond can be called if necessary, but it can also be issued at a lower yield than a bond with a traditional call provision.

Putable Bonds

A **put option** gives the *bondholder* the right to sell the bond back to the issuing company at a prespecified price, typically par. Bondholders are likely to exercise such a put option when the fair value of the bond is less than the put price because interest rates have risen or the credit quality of the issuer has fallen. Exercise styles used are similar to those we enumerated for callable bonds.

Unlike a call option, a put option has value to the bondholder because the choice of whether to exercise the option is the bondholder's. For this reason, a putable bond will sell at a higher price (offer a lower yield) compared to an otherwise identical option-free bond.

Convertible Bonds

Convertible bonds, typically issued with maturities of 5–10 years, give bondholders the option to exchange the bond for a specific number of shares of the issuing corporation's common stock. This gives bondholders the opportunity to profit from increases in the value of the common shares. Regardless of the price of the common shares, the value of a convertible bond will be at least equal to its bond value without the conversion option. Because the conversion option is valuable to bondholders, convertible bonds can be issued with lower yields compared to otherwise identical straight bonds.

Essentially, the owner of a convertible bond has the downside protection (compared to equity shares) of a bond, but at a reduced yield, and the upside opportunity of equity shares. For this reason convertible bonds are often referred to as a *hybrid security*—part debt and part equity.

To issuers, the advantages of issuing convertible bonds are a lower yield (interest cost) compared to straight bonds and the fact that debt financing is converted to equity financing when the bonds are converted to common shares. Some terms related to convertible bonds are:

- **Conversion price.** The price per share at which the bond (at its par value) may be converted to common stock.
- **Conversion ratio.** Equal to the par value of the bond divided by the conversion price. If a bond with a \$1,000 par value has a conversion price of \$40, its *conversion ratio* is $1,000 / 40 = 25$ shares per bond.
- **Conversion value.** This is the market value of the shares that would be received upon conversion. A bond with a conversion ratio of 25 shares when the current market price of a common share is \$50 would have a conversion value of $25 \times 50 = \$1,250$.

Even if the share price increases to a level where the conversion value is significantly above the bond's par value, bondholders might not convert the bonds to common stock until they must because the interest yield on the bonds is higher than the dividend yield on the common shares received through conversion. For this reason, many convertible bonds have a call provision. Because the call price will be less than the conversion value of the shares, by exercising their call provision, the issuers can force bondholders to exercise their conversion option when the conversion value is significantly above the par value of the bonds.

Warrants

An alternative way to give bondholders an opportunity for additional returns when the firm's common shares increase in value is to include **warrants** with straight bonds when they are issued. Warrants give their holders the right to buy the firm's common shares at a given price over a given period of time. As an example, warrants that give their holders the right to buy shares for \$40 will provide profits if the common shares increase in value above \$40 prior to expiration of the warrants. For a young firm, issuing debt can be difficult because the downside (probability of firm failure) is significant, and the upside is limited to the promised debt payments. Including warrants, which are sometimes referred to as a "sweetener," makes the debt more attractive to investors because it adds potential upside profits if the common shares increase in value.

Contingent Convertible Bonds

Contingent convertible bonds (referred to as *CoCos*) are bonds that convert from debt to common equity automatically if a specific event occurs. This type of bond has been issued by some European banks. Banks must maintain specific levels of equity financing. If a bank's equity falls below the required level, they must somehow raise more equity financing to comply with regulations. CoCos are often structured so that if the bank's equity capital falls below a given level, they are automatically converted to common stock. This has the effect of decreasing the bank's debt liabilities and increasing its equity capital at the same time, which helps the bank to meet its minimum equity requirement.



MODULE QUIZ 42.2

1. A 10-year bond pays no interest for three years, then pays \$229.25, followed by payments of \$35 semiannually for seven years, and an additional \$1,000 at maturity. This bond is:
 - A step-up bond.
 - A zero-coupon bond.

- C. a deferred-coupon bond.
2. Which of the following statements is *most accurate* with regard to floating-rate issues that have caps and floors?
- A cap is an advantage to the bondholder, while a floor is an advantage to the issuer.
 - A floor is an advantage to the bondholder, while a cap is an advantage to the issuer.
 - A floor is an advantage to both the issuer and the bondholder, while a cap is a disadvantage to both the issuer and the bondholder.
3. Which of the following *most accurately* describes the maximum price for a currently callable bond?
- Its par value.
 - The call price.
 - The present value of its par value.

KEY CONCEPTS

LOS 42.a

Basic features of a fixed income security include the issuer, maturity date, par value, coupon rate, coupon frequency, and currency.

- Issuers include corporations, governments, quasi-government entities, and supranational entities.
- Bonds with original maturities of one year or less are money market securities. Bonds with original maturities of more than one year are capital market securities.
- Par value is the principal amount that will be repaid to bondholders at maturity. Bonds are trading at a premium if their market price is greater than par value or trading at a discount if their price is less than par value.
- Coupon rate is the percentage of par value that is paid annually as interest. Coupon frequency may be annual, semiannual, quarterly, or monthly. Zero-coupon bonds pay no coupon interest and are pure discount securities.
- Bonds may be issued in a single currency, dual currencies (one currency for interest and another for principal), or with a bondholder's choice of currency.

LOS 42.b

A bond indenture or trust deed is a contract between a bond issuer and the bondholders, which defines the bond's features and the issuer's obligations. An indenture specifies the entity issuing the bond, the source of funds for repayment, assets pledged as collateral, credit enhancements, and any covenants with which the issuer must comply.

LOS 42.c

Covenants are provisions of a bond indenture that protect the bondholders' interests. Negative covenants are restrictions on a bond issuer's operating decisions, such as prohibiting the issuer from issuing additional debt or selling the assets pledged as collateral. Affirmative covenants are administrative actions the issuer must perform, such as making the interest and principal payments on time.

LOS 42.d

Legal and regulatory matters that affect fixed income securities include the places where they are issued and traded, the issuing entities, sources of repayment, and collateral and credit

enhancements.

- Domestic bonds trade in the issuer's home country and currency. Foreign bonds are from foreign issuers but denominated in the currency of the country where they trade.
Eurobonds are issued outside the jurisdiction of any single country and denominated in a currency other than that of the countries in which they trade.
- Issuing entities may be a government or agency; a corporation, holding company, or subsidiary; or a special purpose entity.
- The source of repayment for sovereign bonds is the country's taxing authority. For non-sovereign government bonds, the sources may be taxing authority or revenues from a project. Corporate bonds are repaid with funds from the firm's operations. Securitized bonds are repaid with cash flows from a pool of financial assets.
- Bonds are secured if they are backed by specific collateral or unsecured if they represent an overall claim against the issuer's cash flows and assets.
- Credit enhancement may be internal (overcollateralization, excess spread, tranches with different priority of claims) or external (surety bonds, bank guarantees, letters of credit).

Interest income is typically taxed at the same rate as ordinary income, while gains or losses from selling a bond are taxed at the capital gains tax rate. However, the increase in value toward par of original issue discount bonds is considered interest income. In the United States, interest income from municipal bonds is usually tax-exempt at the national level and in the issuer's state.

LOS 42.e

A bond with a bullet structure pays coupon interest periodically and repays the entire principal value at maturity.

A bond with an amortizing structure repays part of its principal at each payment date. A fully amortizing structure makes equal payments throughout the bond's life. A partially amortizing structure has a balloon payment at maturity, which repays the remaining principal as a lump sum.

A sinking fund provision requires the issuer to retire a portion of a bond issue at specified times during the bonds' life.

Floating-rate notes have coupon rates that adjust based on a reference rate such as LIBOR.

Other coupon structures include step-up coupon notes, credit-linked coupon bonds, payment-in-kind bonds, deferred coupon bonds, and index-linked bonds.

LOS 42.f

Embedded options benefit the party who has the right to exercise them. Call options benefit the issuer, while put options and conversion options benefit the bondholder.

Call options allow the issuer to redeem bonds at a specified call price.

Put options allow the bondholder to sell bonds back to the issuer at a specified put price.

Conversion options allow the bondholder to exchange bonds for a specified number of shares of the issuer's common stock.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 42.1

1. **C** Dual-currency bonds pay coupon interest in one currency and principal in a different currency. These currencies may or may not include the home currency of the issuer. A currency option bond allows the bondholder to choose a currency in which to be paid. (LOS 42.a)
2. **A** An indenture is the contract between the company and its bondholders and contains the bond's covenants. (LOS 42.b)
3. **C** Affirmative covenants require the borrower to perform certain actions. Negative covenants restrict the borrower from performing certain actions. Trust deed is another name for a bond indenture. (LOS 42.c)
4. **B** Tax authorities typically treat the increase in value of a pure-discount bond toward par as interest income to the bondholder. In many jurisdictions this interest income is taxed periodically during the life of the bond even though the bondholder does not receive any cash until maturity. (LOS 42.d)

Module Quiz 42.2

1. **C** This pattern describes a deferred-coupon bond. The first payment of \$229.25 is the value of the accrued coupon payments for the first three years. (LOS 42.e)
2. **B** A cap is a maximum on the coupon rate and is advantageous to the issuer. A floor is a minimum on the coupon rate and is, therefore, advantageous to the bondholder. (LOS 42.e)
3. **B** Whenever the price of the bond increases above the strike price stipulated on the call option, it will be optimal for the issuer to call the bond. Theoretically, the price of a currently callable bond should never rise above its call price. (LOS 42.f)

READING 43

FIXED-INCOME MARKETS: ISSUANCE, TRADING, AND FUNDING

EXAM FOCUS

This reading introduces many terms and definitions. Focus on different types of issuers, features of the various debt security structures, and why different sources of funds have different interest costs. Understand well the differences between fixed-rate and floating-rate debt and how rates are determined on floating-rate debt and for repurchase agreements.

MODULE 43.1: TYPES OF BONDS AND ISSUERS



LOS 43.a: Describe classifications of global fixed-income markets.

Video covering this content is available online.

Global bond markets can be classified by several bond characteristics, including type of issuer, credit quality, maturity, coupon, currency, geography, indexing, and taxable status.

Type of issuer. Common classifications are households, nonfinancial corporations, governments, and financial institutions. In developed markets, the largest issuers by total value of bonds outstanding in global markets are financial corporations and governments. In emerging markets, nonfinancial corporations are the largest issuers.

Credit quality. Standard & Poor's (S&P), Moody's, and Fitch all provide credit ratings on bonds. For S&P and Fitch, the highest bond ratings are AAA, AA, A, and BBB, and are considered *investment grade bonds*. The equivalent ratings by Moody's are Aaa through Baa3. Bonds BB+ or lower (Ba1 or lower) are termed high-yield, speculative, or "junk" bonds. Some institutions are prohibited from investing in bonds of less than investment grade.

Original maturities. Securities with original maturities of one year or less are classified as **money market securities**. Examples include U.S. Treasury bills, commercial paper (issued by corporations), and negotiable certificates of deposit, or CDs (issued by banks). Securities with original maturities greater than one year are referred to as **capital market securities**.

Coupon structure. Bonds are classified as either floating-rate or fixed-rate bonds, depending on whether their coupon interest payments are stated in the bond indenture or depend on the level of a short-term *market reference rate* determined over the life of the bond. Purchasing floating-rate debt is attractive to some institutions that have variable-rate sources of funds (liabilities), such as banks. This allows these institutions to avoid the balance sheet effects of interest rate increases that would increase the cost of funds but leave the interest income at a fixed rate. The

value of fixed-rate bonds (assets) held would fall in the value, while the value of their liabilities would be much less affected.

Currency denomination. A bond's price and returns are determined by the interest rates in the bond's currency. The majority of bonds issued are denominated in either U.S. dollars or euros.

Geography. Bonds may be classified by the markets in which they are issued. Recall the discussion in the previous reading of domestic (or national) bond markets, foreign bonds, and eurobonds, and the differences among them. Bond markets may also be classified as **developed markets** or **emerging markets**. Emerging markets are countries whose capital markets are less well-established than those in developed markets. Emerging market bonds are typically viewed as riskier than developed market bonds and therefore have higher yields. In most emerging markets, and some developed markets, publicly traded debt securities are chiefly issued by governments. Investors who want exposure to private sector debt in these markets can obtain it indirectly by investing in financial institutions that lend to private sector borrowers.

Indexing. As discussed previously, the cash flows on some bonds are based on an index (**index-linked bonds**). Bonds with cash flows determined by inflation rates are referred to as inflation-indexed or inflation-linked bonds. Inflation-linked bonds are issued primarily by governments but also by some corporations of high credit quality.

Tax status. In various countries, some issuers may issue bonds that are exempt from income taxes. In the United States, these bonds can be issued by municipalities and are called **municipal bonds**, or **munis**.

LOS 43.b: Describe the use of interbank offered rates as reference rates in floating-rate debt.

Until recently, the most widely used reference rate for floating-rate bonds was the London Interbank Offered Rate (LIBOR). However, the fact that LIBOR is not based on actual transactions, and has been subject to manipulation by bankers reporting their expected interbank lending rates, has led to an effort to replace LIBOR with market-determined rates. It has been agreed that by the end of 2021, banks will no longer be required to report the estimated rates that are used to determine LIBOR. Thus, alternatives to LIBOR must be found for each of the various currencies involved. In the United States, the new rate will likely be the secured overnight financing rate (SOFR), which is based on the actual rates of repurchase (repo) transactions and reported daily by the Federal Reserve.

For floating-rate bonds, the market reference rate must match the frequency with which the coupon rate on the bond is reset. For example, a bond with a coupon rate that is reset twice each year would use a six-month MRR.

LOS 43.c: Describe mechanisms available for issuing bonds in primary markets.

Sales of newly issued bonds are referred to as **primary market** transactions. Newly issued bonds can be registered with securities regulators for sale to the public, a **public offering**, or sold only to qualified investors, a **private placement**.

A public offering of bonds in the primary market is typically done with the help of an investment bank. The investment bank has expertise in the various steps of a public offering, including:

- Determining funding needs.
- Structuring the debt security.
- Creating the bond indenture.
- Naming a bond trustee (a trust company or bank trust department).
- Registering the issue with securities regulators.
- Assessing demand and pricing the bonds given market conditions.
- Selling the bonds.

Bonds can be sold through an **underwritten offering** or a **best efforts offering**. In an underwritten offering, the entire bond issue is purchased from the issuing firm by the investment bank, termed the underwriter in this case. While smaller bond issues may be sold by a single investment bank, for larger issues, the *lead underwriter* heads a **syndicate** of investment banks who collectively establish the pricing of the issue and are responsible for selling the bonds to dealers, who in turn sell them to investors. The syndicate takes the risk that the bonds will not all be sold.

A new bond issue is publicized and dealers indicate their interest in buying the bonds, which provides information about appropriate pricing. Some bonds are traded on a *when issued* basis in what is called the **grey market**. Such trading prior to the offering date of the bonds provides additional information about the demand for and market clearing price (yield) for the new bond issue.

In a *best efforts* offering, the investment banks sell the bonds on a commission basis. Unlike an underwritten offering, the investment banks do not commit to purchase the whole issue (i.e., underwrite the issue).

Some bonds, especially government bonds, are sold through an auction.



PROFESSOR'S NOTE

Recall that auction procedures were explained in detail in the prerequisite readings for Economics.

U.S. Treasury securities are sold through single price auctions with the majority of purchases made by **primary dealers** that participate in purchases and sales of bonds with the Federal Reserve Bank of New York to facilitate the open market operations of the Fed. Individuals can purchase U.S. Treasury securities through the periodic auctions as well, but are a small part of the total.

In a **shelf registration**, a bond issue is registered with securities regulators in its aggregate value with a master prospectus. Bonds can then be issued over time when the issuer needs to raise funds. Because individual offerings under a shelf registration require less disclosure than a separate registration of a bond issue, only financially sound companies are granted this option. In some countries, bonds registered under a shelf registration can be sold only to qualified investors.

LOS 43.d: Describe secondary markets for bonds.

Secondary markets refer to the trading of previously issued bonds. While some government bonds and corporate bonds are traded on exchanges, the great majority of bond trading in the secondary market is made in the dealer, or over-the-counter, market. Dealers post bid (purchase) prices and ask or offer (selling) prices for various bond issues. The difference between the bid and ask prices is the dealer's spread. The average spread is often between 10 and 12 basis points but varies across individual bonds according to their liquidity and may be more than 50 basis points for an illiquid issue.¹

Bond trades are cleared through a clearing system, just as equities trades are. Settlement (the exchange of bonds for cash) for government bonds is either the day of the trade (cash settlement) or the next business day ($T + 1$). Corporate bonds typically settle on $T + 2$ or $T + 3$, although in some markets it is longer.

One example of a secondary market transaction in bonds is a **tender offer**, in which an issuer offers to repurchase some of its outstanding bonds at a specified price. Typically, a tender offer involves bonds that are trading at a discount. For example, if a corporate bond is trading for 90% of par value, the company might offer to repurchase part of the issue for a higher price (say 93% of par value). This has advantages for both the issuer and the bondholders. The bondholders can receive a higher price for their bonds than they can currently obtain in the market, and the issuer can pay less than face value to retire the bonds.

LOS 43.e: Describe securities issued by sovereign governments.

National governments or their treasuries issue bonds backed by the taxing power of the government that are referred to as **sovereign bonds**. Bonds issued in the currency of the issuing government carry high credit ratings and are considered to be essentially free of default risk. Both a sovereign's ability to collect taxes and its ability to print the currency support these high credit ratings.

Sovereign nations also issue bonds denominated in currencies different from their own. Credit ratings are often higher for a sovereign's local currency bonds than for example, its euro or U.S. dollar-denominated bonds. This is because the national government cannot print the developed market currency and the developed market currency value of local currency tax collections is dependent on the exchange rate between the two currencies.

Trading is most active and prices most informative for the most recently issued government securities of a particular maturity. These issues are referred to as **on-the-run** bonds and also as **benchmark** bonds because the yields of other bonds are determined relative to the "benchmark" yields of sovereign bonds of similar maturities.

Sovereign governments issue fixed-rate, floating-rate, and inflation-indexed bonds.

LOS 43.f: Describe securities issued by non-sovereign governments, quasi-government entities, and supranational agencies.

Non-sovereign government bonds are issued by states, provinces, counties, and sometimes by entities created to fund and provide services such as for the construction of hospitals, airports, and other municipal services. Payments on the bonds may be supported by the revenues of a specific project, from general tax revenues, or from special taxes or fees dedicated to the repayment of project debt.

Non-sovereign bonds are typically of high credit quality, but sovereign bonds typically trade with lower yields (higher prices) because their credit risk is perceived to be less than that of non-sovereign bonds.



PROFESSOR'S NOTE

We will examine the credit quality of sovereign and non-sovereign government bonds in our reading on Fundamentals of Credit Analysis.

Agency or quasi-government bonds are issued by entities created by national governments for specific purposes such as financing small businesses or providing mortgage financing. In the United States, bonds are issued by government-sponsored enterprises (GSEs), such as the Federal National Mortgage Association and the Tennessee Valley Authority.

Some quasi-government bonds are backed by the national government, which gives them high credit quality. Even those not backed by the national government typically have high credit quality although their yields are marginally higher than those of sovereign bonds.

Supranational bonds are issued by supranational agencies, also known as *multilateral agencies*. Examples are the World Bank, the IMF, and the Asian Development Bank. Bonds issued by supranational agencies typically have high credit quality and can be very liquid, especially large issues of well-known entities.



MODULE QUIZ 43.1

1. LIBOR rates are determined:
 - A. by countries' central banks.
 - B. by money market regulators.
 - C. in the interbank lending market.
2. In which type of primary market transaction does an investment bank sell bonds on a commission basis?
 - A. Single-price auction.
 - B. Best-efforts offering.
 - C. Underwritten offering.
3. Secondary market bond transactions *most likely* take place:
 - A. in dealer markets.
 - B. in brokered markets.
 - C. on organized exchanges.
4. Sovereign bonds are described as on-the-run when they:
 - A. are the most recent issue in a specific maturity.
 - B. have increased substantially in price since they were issued.
 - C. receive greater-than-expected demand from auction bidders.
5. Bonds issued by the World Bank would *most likely* be:
 - A. quasi-government bonds.
 - B. global bonds.
 - C. supranational bonds.

MODULE 43.2: CORPORATE DEBT AND FUNDING ALTERNATIVES



Video covering this content is available online.

LOS 43.g: Describe types of debt issued by corporations.

Bank Debt

Most corporations fund their businesses to some extent with bank loans. These are typically variable-rate loans. When the loan involves only one bank, it is referred to as a **bilateral loan**. In contrast, when a loan is funded by several banks, it is referred to as a **syndicated loan** and the group of banks is the syndicate. There is a secondary market in syndicated loan interests that are also securitized, creating bonds that are sold to investors.

Commercial Paper

For larger creditworthy corporations, funding costs can be reduced by issuing short-term debt securities referred to as **commercial paper**. For these firms, the interest cost of commercial paper is less than the interest on a bank loan. Commercial paper yields more than short-term sovereign debt because it has, on average, more credit risk and less liquidity.

Firms use commercial paper to fund working capital and as a temporary source of funds prior to issuing longer-term debt. Debt that is temporary until permanent financing can be secured is referred to as **bridge financing**.

Commercial paper is a short-term, unsecured debt instrument. In the United States, commercial paper is issued with maturities of 270 days or less, because debt securities with maturities of 270 days or less are exempt from SEC registration. Eurocommercial paper (ECP) is issued in several countries with maturities as long as 364 days. Commercial paper is issued with maturities as short as one day (overnight paper), with most issues maturing in about 90 days.

Commercial paper is often reissued or *rolled over* when it matures. The risk that a company will not be able to sell new commercial paper to replace maturing paper is termed *rollover risk*. The two important circumstances in which a company will face rollover difficulties are (1) there is a deterioration in a company's actual or perceived ability to repay the debt at maturity, which will significantly increase the required yield on the paper or lead to less-than-full subscription to a new issue, and (2) significant systemic financial distress, as was experienced in the 2008 financial crisis, that may "freeze" debt markets so that very little commercial paper can be sold at all.

In order to get an acceptable credit rating from the ratings services on their commercial paper, corporations maintain **backup lines of credit** with banks. These are sometimes referred to as *liquidity enhancement* or *backup liquidity lines*. The bank agrees to provide the funds when the paper matures, if needed, except in the case of a *material adverse change* (i.e., when the company's financial situation has deteriorated significantly).

Similar to U.S. T-bills, commercial paper in the United States is typically issued as a pure discount security, making a single payment equal to the face value at maturity. Prices are quoted as a percentage discount from face value. In contrast, ECP rates may be quoted as either a discount yield or an *add-on yield*, that is, the percentage interest paid at maturity in addition to

the par value of the commercial paper. As an example, consider 240-day commercial paper with a holding period yield of 1.35%. If it is quoted with a discount yield, it will be issued at $100 / 1.0135 = 98.668$ and pay 100 at maturity. If it is quoted with an add-on yield, it will be issued at 100 and pay 101.35 at maturity.

Corporate Bonds

In the previous reading, we discussed several features of corporate bonds. **Corporate bonds** are issued with various coupon structures and with both fixed-rate and floating-rate coupon payments. They may be secured by collateral or unsecured and may have call, put, or conversion provisions.

We also discussed a sinking fund provision as a way to reduce the credit risk of a bond by redeeming part of the bond issue periodically over a bond's life. An alternative to a sinking fund provision is to issue a **serial bond issue**. With a serial bond issue, bonds are issued with several maturity dates so that a portion of the issue is redeemed periodically. An important difference between a serial bond issue and an issue with a sinking fund is that with a serial bond issue, investors know at issuance when specific bonds will be redeemed. A bond issue that does not have a serial maturity structure is said to have a **term maturity structure** with all the bonds maturing on the same date.

In general, corporate bonds are referred to as short-term if they are issued with maturities of up to 5 years, medium-term when issued with maturities from 5 to 12 years, and long-term when maturities exceed 12 years.

Corporations issue debt securities called **medium-term notes (MTNs)**, which are not necessarily medium-term in maturity. MTNs are issued in various maturities, ranging from nine months to periods as long as 100 years. Issuers provide *maturity ranges* (e.g., 18 months to two years) for MTNs they wish to sell and provide yield quotes for those ranges. Investors interested in purchasing the notes make an offer to the issuer's agent, specifying the face value and an exact maturity within one of the ranges offered. The agent then confirms the issuer's willingness to sell those MTNs and effects the transaction.

MTNs can have fixed- or floating-rate coupons, but longer-maturity MTNs are typically fixed-rate bonds. Most MTNs, other than long-term MTNs, are issued by financial corporations and most buyers are financial institutions. MTNs can be structured to meet an institution's specifications. While custom bond issues have less liquidity, they provide slightly higher yields compared to an issuer's publicly traded bonds.

LOS 43.h: Describe structured financial instruments.

Structured financial instruments are securities designed to change the risk profile of an underlying debt security, often by combining a debt security with a derivative. Sometimes structured financial instruments redistribute risk. Examples of this type of structured instruments are asset-backed securities and collateralized debt obligations. Both of these types of structured securities are discussed in some detail in our review of asset-backed securities.

Here, we describe several other types of structured instruments with which candidates should be familiar.

1. Yield enhancement instruments

A **credit-linked note (CLN)** has regular coupon payments, but its redemption value depends on whether a specific credit event occurs. If the credit event (e.g., a credit rating downgrade or default of a reference asset) does not occur, the CLN will be redeemed at its par value. If the credit event occurs, the CLN will make a lower redemption payment. Thus, the realized yield on a CLN will be lower if the credit event occurs. Purchasing a CLN can be viewed as buying a note and simultaneously selling a credit default swap (CDS), a derivative security. The buyer of a CDS makes periodic payments to the seller, who will make a payment to the buyer if a specified credit event occurs. The yield on a CLN is higher than it would be on the note alone, without the credit link. This extra yield compensates the buyer of the note (seller of the CDS) for taking on the credit risk of the reference asset, which is why we classify CLNs as a yield enhancement instrument.

2. Capital protected instruments

A capital protected instrument offers a guarantee of a minimum value at maturity as well as some potential upside gain. An example is a security that promises to pay \$1,000 at maturity plus a percentage of any gains on a specified stock index over the life of the security. Such a security could be created by combining a zero-coupon bond selling for \$950 that matures at \$1,000 in 1 year, with a 1-year call option on the reference stock index with a cost of \$50. The total cost of the security is \$1,000, and the minimum payoff at maturity (if the call option expires with a value of zero) is \$1,000. If the call option has a positive value at maturity, the total payment at maturity is greater than \$1,000. A structured financial instrument that promises the \$1,000 payment at maturity under this structure is called a **guarantee certificate**, because the guaranteed payoff is equal to the initial cost of the structured security. Capital protected instruments that promise a payment at maturity less than the initial cost of the instrument offer less-than-full protection, but greater potential for upside gains because more calls can be purchased.

3. Participation instruments

A participation instrument has payments that are based on the value of an underlying instrument, often a reference interest rate or equity index. Participation instruments do not offer capital protection. One example of a participation instrument is a floating-rate note. With a floating-rate note, the coupon payments are based on the value of a short-term interest rate, such as 90-day LIBOR (the reference rate). When the reference rate increases, the coupon payment increases. Because the coupon payments move with the reference rates on floating-rate securities, their market values remain relatively stable, even when interest rates change.

Participation is often based on the performance of an equity price, an equity index value, or the price of another asset. Fixed-income portfolio managers who are only permitted to invest in “debt” securities can use participation instruments to gain exposure to returns on an equity index or asset price.

4. Leveraged instruments

An **inverse floater** is an example of a leveraged instrument. An inverse floater has coupon payments that increase when a reference rate decreases and decrease when a reference rate increases, the opposite of coupon payments on a floating-rate note. A simple structure

might promise to pay a coupon rate, C , equal to a specific rate minus a reference rate, for example, $C = 6\% - 180\text{-day LIBOR}$. When 180-day LIBOR increases, the coupon rate on the inverse floater decreases.

Inverse floaters can also be structured with leverage so that the change in the coupon rate is some multiple of the change in the reference rate. As an example, consider a note with $C = 6\% - (1.2 \times 90\text{-day LIBOR})$ so that the coupon payment rate changes by 1.2 times the change in the reference rate. Such a floater is termed a **leveraged inverse floater**. When the multiplier on the reference rate is less than one, such as $7\% - (0.5 \times 180\text{-day LIBOR})$, the instrument is termed a **deleveraged inverse floater**. In either case, a minimum or floor rate for the coupon rate, often 0%, is specified for the inverse floater.

LOS 43.i: Describe short-term funding alternatives available to banks.

Customer deposits (retail deposits) are a short-term funding source for banks. Checking accounts provide transactions services and immediate availability of funds but typically pay no interest. Money market mutual funds and savings accounts provide less liquidity or less transactions services, or both, and pay periodic interest.

In addition to funds from retail accounts, banks offer interest-bearing **certificates of deposit (CDs)** that mature on specific dates and are offered in a range of short-term maturities. Nonnegotiable CDs cannot be sold and withdrawal of funds often incurs a significant penalty.

Negotiable certificates of deposit can be sold. At the wholesale level, large denomination (typically more than \$1 million) negotiable CDs are an important funding source for banks. They typically have maturities of one year or less and are traded in domestic bond markets as well as in the eurobond market.

Another source of short-term funding for banks is to borrow excess reserves from other banks in the **central bank funds market**. Banks in most countries must maintain a portion of their funds as reserves on deposit with the central bank. At any point in time, some banks may have more than the required amount of reserves on deposit, while others require more reserve deposits. In the market for central bank funds, banks with excess reserves lend them to other banks for periods of one day (overnight funds) and for longer periods up to a year (term funds). **Central bank funds rates** refer to rates for these transactions, which are strongly influenced by the effect of the central bank's open market operations on the money supply and availability of short-term funds.

In the United States, the central bank funds rate is called the Fed funds rate and this rate influences the interest rates of many short-term debt securities.

Other than reserves on deposit with the central bank, funds that are loaned by one bank to another are referred to as **interbank funds**. Interbank funds are loaned between banks for periods of one day to a year. These loans are unsecured and, as with many debt markets, liquidity may decrease severely during times of systemic financial distress.

LOS 43.j: Describe repurchase agreements (repos) and the risks associated with them.

A **repurchase (repo) agreement** is an arrangement by which one party sells a security to a counterparty with a commitment to buy it back at a later date at a specified (higher) price. The *repurchase price* is greater than the selling price and accounts for the interest charged by the buyer, who is, in effect, lending funds to the seller with the security as collateral. The interest rate implied by the two prices is called the *repo rate*, which is the annualized percentage difference between the two prices. A repurchase agreement for one day is called an *overnight repo* and an agreement covering a longer period is called a *term repo*. The interest cost of a repo is customarily less than the rate on bank loans or other short-term borrowing.

As an example, consider a firm that enters into a repo agreement to sell a 4%, 12-year bond with a par value of \$1 million and a market value of \$970,000 for \$940,000 and to repurchase it 90 days later (the **repo date**) for \$947,050.

The implicit interest rate for the 90-day loan period is $947,050 / 940,000 - 1 = 0.75\%$ and the *repo rate* would be expressed as the equivalent annual rate.

The percentage difference between the market value and the amount loaned is called the **repo margin** or the **haircut**. In our example, it is $940,000 / 970,000 - 1 = -3.1\%$. This margin protects the lender in the event that the value of the security decreases over the term of the repo agreement.

The repo rate is:

- Higher, the longer the repo term.
- Lower, the higher the credit quality of the collateral security.
- Lower when the collateral security is delivered to the lender.
- Higher when the interest rates for alternative sources of funds are higher.

The repo margin is influenced by similar factors. The repo margin is:

- Higher, the longer the repo term.
- Lower, the higher the credit quality of the collateral security.
- Lower, the higher the credit quality of the borrower.
- Lower when the collateral security is in high demand or low supply.

The reason the supply and demand conditions for the collateral security affects pricing is that some lenders want to own a specific bond or type of bond as collateral. For a bond that is high demand, lenders must compete for bonds by offering lower repo lending rates.

Viewed from the standpoint of a bond dealer, a **reverse repo agreement** refers to taking the opposite side of a repurchase transaction, lending funds by buying the collateral security rather than selling the collateral security to borrow funds.



MODULE QUIZ 43.2

1. With which of the following features of a corporate bond issue does an investor *most likely* face the risk of redemption prior to maturity?
 - A. Serial bonds.
 - B. Sinking fund.
 - C. Term maturity structure.
2. A financial instrument is structured such that cash flows to the security holder increase if a specified reference rate increases. This structured financial instrument is *best* described as:

- A. a participation instrument.
 - B. a capital protected instrument.
 - C. a yield enhancement instrument.
3. Smith Bank lends Johnson Bank excess reserves on deposit with the central bank for a period of three months. Is this transaction said to occur in the interbank market?
- A. Yes.
 - B. No, because the interbank market refers to loans for more than one year.
 - C. No, because the interbank market does not include reserves at the central bank.
4. In a repurchase agreement, the percentage difference between the repurchase price and the amount borrowed is *most accurately* described as:
- A. the haircut.
 - B. the repo rate.
 - C. the repo margin.

KEY CONCEPTS

LOS 43.a

Global bond markets can be classified by the following:

- Type of issuer: Households, nonfinancial corporations, governments, financial institutions.
- Credit quality: Investment grade, noninvestment grade.
- Original maturity: Money market (one year or less), capital market (more than one year).
- Coupon: Fixed rate, floating rate.
- Currency and geography: Domestic, foreign, global, eurobond markets; developed, emerging markets.
- Other classifications: Indexing, taxable status.

LOS 43.b

Interbank lending rates, such as London Interbank Offered Rate (LIBOR), are frequently used as reference rates for floating-rate debt. An appropriate reference rate is one that matches a floating-rate note's currency and frequency of rate resets, such as six-month U.S. dollar LIBOR for a semiannual floating-rate note issued in U.S. dollars.

LOS 43.c

Bonds may be issued in the primary market through a public offering or a private placement.

A public offering using an investment bank may be underwritten, with the investment bank or syndicate purchasing the entire issue and selling the bonds to dealers; or on a best-efforts basis, in which the investment bank sells the bonds on commission. Public offerings may also take place through auctions, which is the method commonly used to issue government debt.

A private placement is the sale of an entire issue to a qualified investor or group of investors, which are typically large institutions.

LOS 43.d

Bonds that have been issued previously trade in secondary markets. While some bonds trade on exchanges, most are traded in dealer markets. Spreads between bid and ask prices are narrower for liquid issues and wider for less liquid issues.

Trade settlement is typically T + 2 or T + 3 for corporate bonds and either cash settlement or T + 1 for government bonds.

LOS 43.e

Sovereign bonds are issued by national governments and backed by their taxing power. Sovereign bonds may be denominated in the local currency or a foreign currency.

LOS 43.f

Non-sovereign government bonds are issued by governments below the national level, such as provinces or cities, and may be backed by taxing authority or revenues from a specific project.

Agency or quasi-government bonds are issued by government sponsored entities and may be explicitly or implicitly backed by the government.

Supranational bonds are issued by multilateral agencies that operate across national borders.

LOS 43.g

Debt issued by corporations includes bank debt, commercial paper, corporate bonds, and medium-term notes.

Bank debt includes bilateral loans from a single bank and syndicated loans from multiple banks.

Commercial paper is a money market instrument issued by corporations of high credit quality.

Corporate bonds may have a term maturity structure (all bonds in an issue mature at the same time) or a serial maturity structure (bonds in an issue mature on a predetermined schedule) and may have a sinking fund provision.

Medium-term notes are corporate issues that can be structured to meet the requirements of investors.

LOS 43.h

Structured financial instruments include asset-backed securities and collateralized debt securities as well as the following types:

- Yield enhancement instruments include credit linked notes, which are redeemed at an amount less than par value if a specified credit event occurs on a reference asset, or at par if it does not occur. The buyer receives a higher yield for bearing the credit risk of the reference asset.
- Capital protected instruments offer a guaranteed payment, which may be equal to the purchase price of the instrument, along with participation in any increase in the value of an equity, an index, or other asset.
- Participation instruments are debt securities with payments that depend on the returns on an asset or index, or depend on a reference interest rate. One example is a floating rate bond, which makes coupon payments that change with a short-term reference rate, such as LIBOR. Other participation instruments make coupon payments based on the returns on an index of equity securities or on some other asset.
- An inverse floater is a leveraged instrument that has a coupon rate that varies inversely with a specified reference interest rate, for example, $6\% - (L \times 180\text{-day LIBOR})$. L is the leverage of the inverse floater. An inverse floater with $L > 1$, so that the coupon rate

changes by more than the reference rate, is termed a leveraged inverse floater. An inverse floater with $L < 1$ is a deleveraged floater.

LOS 43.i

Short-term funding alternatives available to banks include:

- Customer deposits, including checking accounts, savings accounts, and money market mutual funds.
- Negotiable CDs, which may be sold in the wholesale market.
- Central bank funds market. Banks may buy or sell excess reserves deposited with their central bank.
- Interbank funds. Banks make unsecured loans to one another for periods up to a year.

LOS 43.j

A repurchase agreement is a form of short-term collateralized borrowing in which one party sells a security to another party and agrees to buy it back at a predetermined future date and price. The repo rate is the implicit interest rate of a repurchase agreement. The repo margin, or haircut, is the difference between the amount borrowed and the value of the security.

Repurchase agreements are an important source of short-term financing for bond dealers. If a bond dealer is lending funds instead of borrowing, the agreement is known as a reverse repo.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 43.1

1. **C** LIBOR rates are determined in the market for interbank lending. (LOS 43.b)
2. **B** In a best-efforts offering, the investment bank or banks do not underwrite (i.e., purchase all of) a bond issue, but rather sell the bonds on a commission basis. Bonds sold by auction are offered directly to buyers by the issuer (typically a government). (LOS 43.c)
3. **A** The secondary market for bonds is primarily a dealer market in which dealers post bid and ask prices. (LOS 43.d)
4. **A** Sovereign bonds are described as *on-the-run* or *benchmark* when they represent the most recent issue in a specific maturity. (LOS 43.e)
5. **C** Bonds issued by the World Bank, which is a multilateral agency operating globally, are termed *supranational bonds*. (LOS 43.f)

Module Quiz 43.2

1. **B** With a sinking fund, the issuer must redeem part of the issue prior to maturity, but the specific bonds to be redeemed are not known. Serial bonds are issued with a schedule of maturities and each bond has a known maturity date. In an issue with a term maturity structure, all the bonds are scheduled to mature on the same date. (LOS 43.g)
2. **A** Floating-rate notes are an example of a participation instrument. (LOS 43.h)

3. **C** The interbank market refers to short-term borrowing and lending among banks of funds other than those on deposit at a central bank. Loans of reserves on deposit with a central bank are said to occur in the central bank funds market. (LOS 43.i)
 4. **B** The repo rate is the percentage difference between the repurchase price and the amount borrowed. The repo margin or haircut is the percentage difference between the amount borrowed and the value of the collateral. (LOS 43.j)
-

¹ Fixed Income Markets: Issuance, Trading, and Funding, Choudhry, M.; Mann, S.; and Whitmer, L.; in CFA Program 2023 Level I Curriculum, Volume 4 (CFA Institute, 2022).

READING 44

INTRODUCTION TO FIXED-INCOME VALUATION

EXAM FOCUS

The concepts introduced here are very important for understanding the factors that determine the value of debt securities and various yield measures. The relationships between yield to maturity, spot rates, and forward rates are core material and come up in many contexts throughout the CFA curriculum. Yield spread measures also have many applications. Note that while several of the required learning outcomes have the command word “calculate” in them, a good understanding of the underlying concepts is just as important for exam success on this material.

MODULE 44.1: BOND VALUATION AND YIELD TO MATURITY



Video covering this content is available online.

LOS 44.a: Calculate a bond's price given a market discount rate.

Calculating the Value of an Annual Coupon Bond

The value of a coupon bond can be calculated by summing the present values of all of the bond's promised cash flows. The market discount rate appropriate for discounting a bond's cash flows is called the bond's **yield-to-maturity (YTM)**. If we know a bond's yield-to-maturity, we can calculate its value, and if we know its value (market price), we can calculate its yield-to-maturity.

Consider a newly issued 10-year, \$1,000 par value, 10% coupon, annual-pay bond. The coupon payments will be \$100 at the end of each year the \$1,000 par value will be paid at the end of year 10. First, let's value this bond assuming the appropriate discount rate is 10%. The present value of the bond's cash flows discounted at 10% is:

$$\frac{100}{1.1} + \frac{100}{1.1^2} + \frac{100}{1.1^3} + \dots + \frac{100}{1.1^9} + \frac{1,100}{1.1^{10}} = 1,000$$

The calculator solution is:

$$N = 10; PMT = 100; FV = 1,000; I/Y = 10; CPT \rightarrow PV = -1,000$$

where:

N = number of years

PMT = the *annual* coupon payment

I/Y = the *annual* discount rate

FV = the par value or selling price at the end of an assumed holding period



PROFESSOR'S NOTE

Take note of a couple of points here. The discount rate is entered as a whole number in percent, 10, not 0.10. The 10 coupon payments of \$100 each are taken care of in the N = 10 and PMT = 100 entries. The principal repayment is in the FV = 1,000 entry. Lastly, note that the PV is negative; it will be the opposite sign to the sign of PMT and FV. The calculator is just “thinking” that to receive the payments and future value (to own the bond), you must pay the present value of the bond today (you must buy the bond). That’s why the PV amount is negative; it is a cash outflow to a bond buyer.

Now let's value that same bond with a discount rate of 8%:

$$\frac{100}{1.08} + \frac{100}{1.08^2} + \frac{100}{1.08^3} + \dots + \frac{100}{1.08^9} + \frac{1,100}{1.08^{10}} = 1,134.20$$

The calculator solution is:

$$N = 10; PMT = 100; FV = 1,000; I/Y = 8; CPT \rightarrow PV = -1,134.20$$

If the market discount rate for this bond were 8%, it would sell at a premium of \$134.20 above its par value. **When bond yields decrease, the present value of a bond's payments, its market value, increases.**

If we discount the bond's cash flows at 12%, the present value of the bond is:

$$\frac{100}{1.12} + \frac{100}{1.12^2} + \frac{100}{1.12^3} + \dots + \frac{100}{1.12^9} + \frac{1,100}{1.12^{10}} = 887.00$$

The calculator solution is:

$$N = 10; PMT = 100; FV = 1,000; I/Y = 12; CPT \rightarrow PV = -887$$

If the market discount rate for this bond were 12%, it would sell at a discount of \$113 to its par value. **When bond yields increase, the present value of a bond's payments, its market value, decreases.**



PROFESSOR'S NOTE

It's worth noting here that a 2% decrease in yield-to-maturity increases the bond's value by more than a 2% increase in yield decreases the bond's value. This illustrates that the bond's price-yield relationship is convex, as we will explain in more detail in a later reading.

Calculating the value of a bond with semiannual coupon payments. Let's calculate the value of the same bond with semiannual payments.

Rather than \$100 per year, the security will pay \$50 every six months. With an annual YTM of 8%, we need to discount the coupon payments at 4% per period which results in a present value of:

$$\frac{50}{1.04} + \frac{50}{1.04^2} + \frac{50}{1.04^3} + \dots + \frac{50}{1.04^{19}} + \frac{1,050}{1.04^{20}} = 1,135.90$$

The calculator solution is:

$$N = 20; PMT = 50; FV = 1,000; I/Y = 4; CPT \rightarrow PV = -1,135.90$$

The value of a zero-coupon bond is simply the present value of the maturity payment. With a discount rate of 3% per period, a 5-period zero-coupon bond with a par value of \$1,000 has a value of:

$$\frac{1,000}{1.03^5} = \$862.61$$

LOS 44.b: Identify the relationships among a bond's price, coupon rate, maturity, and market discount rate (yield-to-maturity).

So far we have used a bond's cash flows and an assumed discount rate to calculate the value of the bond. We can also calculate the market discount rate given a bond's price in the market, because there is an inverse relationship between price and yield. For a 3-year, 8% annual coupon bond that is priced at 90.393, the market discount rate is:

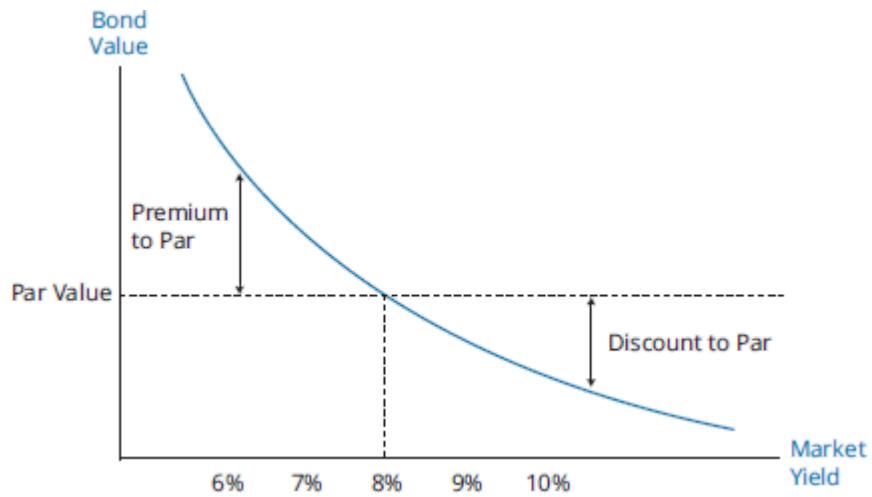
$$N = 3; PMT = 8; FV = 100; PV = -90.393; CPT \rightarrow I/Y = 12\%$$

We can summarize the relationships between price and yield as follows:

1. At a point in time, a decrease (increase) in a bond's YTM will increase (decrease) its price.
2. If a bond's coupon rate is greater than its YTM, its price will be at a premium to par value.
If a bond's coupon rate is less than its YTM, its price will be at a discount to par value.
3. The percentage decrease in value when the YTM increases by a given amount is smaller than the increase in value when the YTM decreases by the same amount (the price-yield relationship is convex).
4. Other things equal, the price of a bond with a lower coupon rate is more sensitive to a change in yield than is the price of a bond with a higher coupon rate.
5. Other things equal, the price of a bond with a longer maturity is more sensitive to a change in yield than is the price of a bond with a shorter maturity.

Figure 44.1 illustrates the convex relationship between a bond's price and its yield-to-maturity:

Figure 44.1: Market Yield vs. Bond Value for an 8% Coupon Bond



Relationship Between Price and Maturity

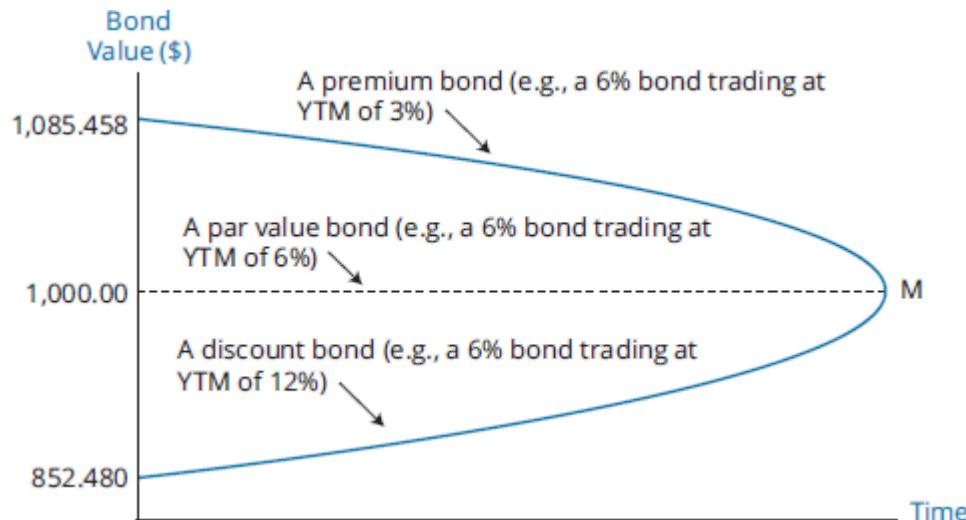
Prior to maturity, a bond can be selling at a significant discount or premium to par value. However, regardless of its required yield, the price will converge to par value as maturity approaches. Consider a bond with \$1,000 par value and a 3-year life paying 6% semiannual coupons. The bond values corresponding to required yields of 3%, 6%, and 12% as the bond approaches maturity are presented in Figure 44.2.

Figure 44.2: Bond Values and the Passage of Time

Time to Maturity (in years)	YTM = 3%	YTM = 6%	YTM = 12%
3.0	\$1,085.46	\$1,000.00	\$852.48
2.5	1,071.74	1,000.00	873.63
2.0	1,057.82	1,000.00	896.05
1.5	1,043.68	1,000.00	919.81
1.0	1,029.34	1,000.00	945.00
0.5	1,014.78	1,000.00	971.69
0.0	1,000.00	1,000.00	1,000.00

The change in value associated with the passage of time for the three bonds represented in Figure 44.2 is presented graphically in Figure 44.3. This convergence to par value at maturity is known as the **constant-yield price trajectory** because it shows how the bond's price would change as time passes if its yield-to-maturity remained constant.

Figure 44.3: Premium, Par, and Discount Bonds



MODULE QUIZ 44.1

1. A 20-year, 10% annual-pay bond has a par value of \$1,000. What is the price of the bond if it has a yield-to-maturity of 15%?
 - A. \$685.14.
 - B. \$687.03.
 - C. \$828.39.
2. An analyst observes a 5-year, 10% semiannual-pay bond. The face amount is £1,000. The analyst believes that the yield-to-maturity on a semiannual bond basis should be 15%. Based on this yield estimate, the price of this bond would be:
 - A. £828.40.
 - B. £1,189.53.
 - C. £1,193.04.
3. An analyst observes a 20-year, 8% option-free bond with semiannual coupons. The required yield-to-maturity on a semiannual bond basis was 8%, but suddenly it decreased to 7.25%. As a result, the price of this bond:
 - A. increased.
 - B. decreased.
 - C. stayed the same.
4. A \$1,000, 5%, 20-year annual-pay bond has a YTM of 6.5%. If the YTM remains unchanged, how much will the bond value increase over the next three years?
 - A. \$13.62.
 - B. \$13.78.
 - C. \$13.96.

MODULE 44.2: SPOT RATES AND ACCRUED INTEREST



Video covering this content is available online.

LOS 44.c: Define spot rates and calculate the price of a bond using spot rates.

The yield-to-maturity is calculated as if the discount rate for every bond cash flow is the same. In reality, discount rates depend on the time period in which the bond payment will be made. **Spot rates** are the market discount rates for a single payment to be received in the future. The discount rates for zero-coupon bonds are spot rates and we sometimes refer to spot rates as *zero-coupon rates* or simply *zero rates*.

In order to price a bond with spot rates, we sum the present values of the bond's payments, each discounted at the spot rate for the number of periods before it will be paid. The general equation for calculating a bond's value using spot rates (S_i) is:

$$\frac{CPN_1}{1 + S_1} + \frac{CPN_2}{(1 + S_2)^2} + \dots + \frac{CPN_N + FV_N}{(1 + S_N)^N} = PV$$

EXAMPLE: Valuing a bond using spot rates

Given the following spot rates, calculate the value of a 3-year, 5% annual-coupon bond.

Spot rates

1-year: 3%

2-year: 4%

3-year: 5%

Answer:

$$\frac{50}{1.03} + \frac{50}{(1.04)^2} + \frac{1,050}{(1.05)^3} = 48.54 + 46.23 + 907.03 = \$1,001.80$$

This price, calculated using spot rates, is sometimes called the *no-arbitrage price* of a bond because if a bond is priced differently there will be a profit opportunity from arbitrage among bonds.

Because the bond value is slightly greater than its par value, we know its YTM is slightly less than its coupon rate of 5%. Using the price of 1,001.80, we can calculate the YTM for this bond as:

$$N = 3; PMT = 50; FV = 1,000; PV = -1,001.80; CPT \rightarrow I/Y = 4.93\%$$

LOS 44.d: Describe and calculate the flat price, accrued interest, and the full price of a bond.

The coupon bond values we have calculated so far are calculated on the date a coupon is paid, as the present value of the remaining coupons. For most bond trades, the settlement date, which is when cash is exchanged for the bond, will fall between coupon payment dates. As time passes (and future coupon payment dates get closer), the value of the bond will increase.

The value of a bond between coupon dates can be calculated, using its current YTM, as the value of the bond on its last coupon date (PV) times $(1 + YTM / \# \text{ of coupon periods per year})^{t/T}$, where t is the number of days since the last coupon payment, and T is the number of days in the coupon period. For a given settlement date, this value is referred to as the **full price** of the bond.

Let's work an example for a specific bond:

EXAMPLE: Calculating the full price of a bond

A 5% bond makes coupon payments on June 15 and December 15 and is trading with a YTM of 4%. The bond is purchased and will settle on August 21 when there will be four coupons remaining until maturity. Calculate the full price of the bond using actual days.

Step 1: Calculate the value of the bond on the last coupon date (coupons are semiannual, so we use $4 / 2 = 2\%$ for the periodic discount rate):

$$N = 4; PMT = 25; FV = 1,000; I/Y = 2; CPT \rightarrow PV = -1,019.04$$

Step 2: Adjust for the number of days since the last coupon payment:

Days between June 15 and December 15 = 183 days.

Days between June 15 and settlement on August 21 = 67 days.

$$\text{Full price} = 1,019.04 \times (1.02)^{67/183} = 1,026.46.$$

The accrued interest since the last payment date can be calculated as the coupon payment times the portion of the coupon period that has passed between the last coupon payment date and the settlement date of the transaction. For the bond in the previous example, the accrued interest on the settlement date of August 21 is:

$$\$25 (67 / 183) = \$9.15$$

The full price (invoice price) minus the accrued interest is referred to as the **flat price** of the bond.

$$\text{flat price} = \text{full price} - \text{accrued interest}$$

So for the bond in our example, the flat price = $1,026.46 - 9.15 = 1,017.31$.

The flat price of the bond is also referred to as the bond's **clean price** or **quoted price**, and the full price is also referred to as the **dirty price**.

Note that the flat price is not the present value of the bond on its last coupon payment date, $1,017.31 < 1,019.04$.

So far, in calculating accrued interest, we used the actual number of days between coupon payments and the actual number of days between the last coupon date and the settlement date. This actual/actual method is used most often with government bonds. The 30/360 method is most often used for corporate bonds. This method assumes that there are 30 days in each month and 360 days in a year.

EXAMPLE: Accrued interest

An investor buys a \$1,000 par value, 4% annual-pay bond that pays its coupons on May 15. The investor's buy order settles on August 10. Calculate the accrued interest that is owed to the bond seller, using the 30/360 method and the actual/actual method.

Answer:

The annual coupon payment is $4\% \times \$1,000 = \40 .

Using the 30/360 method, interest is accrued for $30 - 15 = 15$ days in May; 30 days each in June and July; and 10 days in August, or $15 + 30 + 30 + 10 = 85$ days.

$$\text{accrued interest (30/360 method)} = \frac{85}{360} \times \$40 = \$9.44$$

Using the actual/actual method, interest is accrued for $31 - 15 = 16$ days in May; 30 days in June; 31 days in July; and 10 days in August, or $16 + 30 + 31 + 10 = 87$ days.

$$\text{accrued interest (actual/actual method)} = \frac{87}{365} \times \$40 = \$9.53$$

LOS 44.e: Describe matrix pricing.

Matrix pricing is a method of estimating the required yield-to-maturity (or price) of bonds that are currently not traded or infrequently traded. The procedure is to use the YTMs of traded bonds that have credit quality very close to that of a nontraded or infrequently traded bond and are similar in maturity and coupon, to estimate the required YTM.

EXAMPLE: Pricing an illiquid bond

Rob Phelps, CFA, is estimating the value of a nontraded 4% annual-pay, A+ rated bond that has three years remaining until maturity. He has obtained the following yields-to-maturity on similar corporate bonds:

- A+ rated, 2-year annual-pay, YTM = 4.3%
- A+ rated, 5-year annual-pay, YTM = 5.1%
- A+ rated, 5-year annual-pay, YTM = 5.3%

Estimate the value of the nontraded bond.

Answer:

Step 1: Take the average YTM of the 5-year bonds: $(5.1 + 5.3) / 2 = 5.2\%$.

Step 2: Interpolate the 3-year YTM based on the 2-year and average 5-year YTMs:

$$\begin{aligned} & 4.3\% + (5.2\% - 4.3\%) \times [(3 \text{ years} - 2 \text{ years}) / (5 \text{ years} - 2 \text{ years})] \\ & = 4.6\% \end{aligned}$$

Step 3: Price the nontraded bond with a YTM of 4.6%:

$$N = 3; PMT = 40; FV = 1,000; I/Y = 4.6; CPT \rightarrow PV = -983.54$$

The estimated value is \$983.54 per \$1,000 par value.

In using the averages in the preceding example, we have used simple *linear interpolation*. Because the maturity of the nontraded bond is three years, we estimate the YTM on the 3-year bond as the yield on the 2-year bond, plus one-third of the difference between the YTM of the 2-year bond and the average YTM of the 5-year bonds. Note that the difference in maturity between the 2-year bond and the 3-year bond is one year and the difference between the maturities of the 2-year and 5-year bonds is three years.

A variation of matrix pricing used for pricing new bond issues focuses on the spreads between bond yields and the yields of a benchmark bond of similar maturity that is essentially default risk free. Often the yields on Treasury bonds are used as benchmark yields for U.S. dollar-denominated corporate bonds. When estimating the YTM for the new issue bond, the appropriate spread to the yield of a Treasury bond of the same maturity is estimated and added to the yield of the benchmark issue.

EXAMPLE: Estimating the spread for a new 6-year, A rated bond issue

Consider the following market yields:

$$5\text{-year, U.S. Treasury bond, YTM } 1.48\%$$

5-year, A rated corporate bond, YTM 2.64%

7-year, U.S. Treasury bond, YTM 2.15%

7-year, A rated corporate bond, YTM 3.55%

6-year U.S. Treasury bond, YTM 1.74%

Estimate the required yield on a newly issued 6-year, A rated corporate bond.

Answer:

1. Calculate the spreads to the benchmark (Treasury) yields.

Spread on the 5-year corporate bond is $2.64 - 1.48 = 1.16\%$.

Spread on the 7-year corporate bond is $3.55 - 2.15 = 1.40\%$.

2. Calculate the average spread because the 6-year bond is the midpoint of five and seven years:

$$\text{average spread} = (1.16 + 1.40) / 2 = 1.28\%$$

3. Add the average spread to the YTM of the 6-year Treasury (benchmark) bond.

$1.74 + 1.28 = 3.02\%$, which is our estimate of the YTM on the newly issued 6-year, A rated bond.



MODULE QUIZ 44.2

1. If spot rates are 3.2% for one year, 3.4% for two years, and 3.5% for three years, the price of a \$100,000 face value, 3-year, annual-pay bond with a coupon rate of 4% is *closest* to:
 - A. \$101,420.
 - B. \$101,790.
 - C. \$108,230.

2. An investor paid a full price of \$1,059.04 each for 100 bonds. The purchase was between coupon dates, and accrued interest was \$23.54 per bond. What is each bond's flat price?
 - A. \$1,000.00.
 - B. \$1,035.50.
 - C. \$1,082.58.

3. Cathy Moran, CFA, is estimating a value for an infrequently traded bond with six years to maturity, an annual coupon of 7%, and a single-B credit rating. Moran obtains yields-to-maturity for more liquid bonds with the same credit rating:
 - 5% coupon, eight years to maturity, yielding 7.20%.
 - 6.5% coupon, five years to maturity, yielding 6.40%.

The infrequently traded bond is *most likely* trading at:

- A. par value.
- B. a discount to par value.
- C. a premium to par value.

MODULE 44.3: YIELD MEASURES



LOS 44.f: Calculate annual yield on a bond for varying compounding periods in a year.

Video covering this content is available online.

LOS 44.g: Calculate and interpret yield measures for fixed-rate bonds and floating-rate notes.

Given a bond's price in the market, we can say that the YTM is the discount rate that makes the present value of a bond's cash flows equal to its price. For a 5-year, annual pay 7% bond that is priced in the market at \$1,020.78, the YTM will satisfy the following equation:

$$\frac{70}{1 + \text{YTM}} + \frac{70}{(1 + \text{YTM})^2} + \frac{70}{(1 + \text{YTM})^3} + \frac{70}{(1 + \text{YTM})^4} + \frac{1,070}{(1 + \text{YTM})^5}$$

$$= 1,020.78$$

We can calculate the YTM (discount rate) that satisfies this equality as:

$$N = 5; \text{PMT} = 70; \text{FV} = 1,000; \text{PV} = -1,020.78; \text{CPT} \rightarrow I/Y = 6.5\%$$

By convention, the YTM on a semiannual coupon bond is expressed as two times the semiannual discount rate. For a 5-year, semiannual pay 7% coupon bond, we can calculate the semiannual discount rate as $\text{YTM}/2$ and then double it to get the YTM expressed as an annual yield:

$$\frac{35}{1 + \text{YTM}/2} + \frac{35}{(1 + \text{YTM}/2)^2} + \frac{35}{(1 + \text{YTM}/2)^3} + \dots + \frac{35}{(1 + \text{YTM}/2)^9} + \frac{1,035}{(1 + \text{YTM}/2)^{10}}$$

$$= 1,020.78$$

$$N = 10; \text{PMT} = 35; \text{FV} = 1,000; \text{PV} = -1,020.78; \text{CPT} \rightarrow I/Y = 3.253\%$$

The YTM is $3.253 \times 2 = 6.506\%$.

Yield Measures for Fixed-Rate Bonds

The number of bond coupon payments per year is referred to as the **periodicity** of a bond. A bond with a periodicity of 2 will have its yield to maturity quoted on a **semiannual bond basis**. For a given coupon rate, the greater the periodicity, the more compounding periods, and the greater the annual yield.



PROFESSOR'S NOTE

This is analogous to the relationship among a stated annual rate, the number of compounding periods per year, and the effective annual yield explained in Quantitative Methods.

In general, the annual (effective) yield for bond with its YTM stated for a periodicity of n , and n compounding periods per year, is:

$$\text{annual yield} = \left(1 + \frac{\text{YTM}}{n}\right)^n - 1$$

EXAMPLE: Effective annual yields

What is the annual yield for a bond with a stated YTM of 10%:

1. When the periodicity of the bond is 2 (pays semiannually)?
2. When the periodicity of the bond is 4 (pays quarterly)?

Answer:

$$1. \text{annual yield} = \left(1 + \frac{0.10}{2}\right)^2 - 1 = 1.05^2 - 1 = 0.1025 = 10.25\%$$

$$2. \text{annual yield} = \left(1 + \frac{0.10}{4}\right)^4 - 1 = 1.025^4 - 1 = 0.1038 = 10.38\%$$

It may be necessary to adjust the quoted yield on a bond to make it comparable with the yield on a bond with a different periodicity. This is illustrated in the following example.

EXAMPLE: Adjusting yields for periodicity

An Atlas Corporation bond is quoted with a YTM of 4% on a semiannual bond basis. What yields should be used to compare it with a quarterly-pay bond and an annual-pay bond?

Answer:

The first thing to note is that 4% on a semiannual bond basis is an effective yield of 2% per 6-month period. To compare this with the yield on an annual-pay bond, which is an effective annual yield, we need to calculate the effective annual yield on the semiannual coupon bond, which is $1.02^2 - 1 = 4.04\%$.

For the annual YTM on the quarterly-pay bond, we need to calculate the effective quarterly yield and multiply by four. The quarterly yield (yield per quarter) that is equivalent to a yield of 2% per six months is $1.02^{1/2} - 1 = 0.995\%$. The quoted annual rate for the equivalent yield on a quarterly bond basis is $4 \times 0.995 = 3.98\%$.

Note that we have shown that the effective annual yields are the same for:

- An annual coupon bond with a yield of 4.04% on an annual basis (periodicity of one).
- A semiannual coupon bond with a yield of 4.0% on a semiannual basis (periodicity of two).
- A quarterly coupon bond with a yield of 3.98% on quarterly basis (periodicity of four).

Bond yields calculated using the stated coupon payment dates are referred to as following the **street convention**. Because some coupon dates will fall on weekends and holidays, coupon payments will actually be made the next business day. The yield calculated using these actual coupon payment dates is referred to as the **true yield**. Some coupon payments will be made later when holidays and weekends are taken into account, so true yields will be slightly lower than street convention yields, if only by a few basis points.

When calculating spreads between government bond yields and the yield on a corporate bond, the corporate bond yield is often restated to its yield on actual/actual basis to match the day count convention used on government bonds (rather than the 30/360 day count convention used for calculating corporate bond yields).

Current yield (also called **income yield** or **running yield**) is simple to calculate, but offers limited information. This measure looks at just one source of return: *a bond's annual interest income*—it does not consider capital gains/losses or reinvestment income. The formula for the current yield is:

$$\text{current yield} = \frac{\text{annual cash coupon payment}}{\text{bond price}}$$

EXAMPLE: Computing current yield

Consider a 20-year, \$1,000 par value, 6% *semiannual-pay* bond that is currently trading at a flat price of \$802.07. Calculate the current yield.

Answer:

The *annual* cash coupon payments total:

$$\begin{aligned}\text{annual cash coupon payment} &= \text{par value} \times \text{stated coupon rate} \\ &= \$1,000 \times 0.06 = \$60\end{aligned}$$

Because the bond is trading at \$802.07, the current yield is:

$$\text{current yield} = \frac{60}{802.07} = 0.0748, \text{ or } 7.48\%$$

Note that current yield is based on *annual* coupon interest so that it is the same for a semiannual-pay and annual-pay bond with the same coupon rate and price.

The current yield does not account for gains or losses as the bond's price moves toward its par value over time. A bond's **simple yield** takes a discount or premium into account by assuming that any discount or premium declines evenly over the remaining years to maturity. The sum of the annual coupon payment plus (minus) the straight-line amortization of a discount (premium) is divided by the flat price to get the simple yield.

EXAMPLE: Computing simple yield

A 3-year, 8% coupon, semiannual-pay bond is priced at 90.165. Calculate the simple yield.

Answer:

The discount from par value is $100 - 90.165 = 9.835$. Annual straight-line amortization of the discount is $9.835 / 3 = 3.278$.

$$\text{simple yield} = \frac{8 + 3.278}{90.165} = 12.51\%$$

For a callable bond, an investor's yield will depend on whether and when the bond is called. The **yield-to-call** can be calculated for each possible call date and price. The lowest of yield-to-maturity and the various yields-to-call is termed the **yield-to-worst**. The following example illustrates these calculations.

EXAMPLE: Yield-to-call and yield-to-worst

Consider a 10-year, semiannual-pay 6% bond trading at 102 on January 1, 2014. The bond is callable according to the following schedule:

Callable at 102 on or after January 1, 2019.

Callable at 100 on or after January 1, 2022.

Calculate the bond's YTM, yield-to-first call, yield-to-first par call, and yield-to-worst.

Answer:

The *yield-to-maturity* on the bond is calculated as:

$$N = 20; PMT = 30; FV = 1,000; PV = -1,020; CPT \rightarrow I/Y = 2.867\%$$

$$2 \times 2.867 = 5.734\% = \text{YTM}$$

To calculate the *yield-to-first call*, we calculate the yield-to-maturity using the number of semiannual periods until the first call date (10) for *N* and the call price (1,020) for *FV*:

$$N = 10; PMT = 30; FV = 1,020; PV = -1,020; CPT \rightarrow I/Y = 2.941\%$$

$$2 \times 2.941 = 5.882\% = \text{yield-to-first call}$$

To calculate the *yield-to-first par call* (second call date), we calculate the yield-to-maturity using the number of semiannual periods until the first par call date (16) for *N* and the call price (1,000) for *FV*:

$$N = 16; PMT = 30; FV = 1,000; PV = -1,020; CPT \rightarrow I/Y = 2.843\%$$

$$2 \times 2.843 = 5.686\% = \text{yield-to-first par call}$$

The lowest yield, 5.686%, is realized if the bond is called at par on January 1, 2022, so the *yield-to-worst* is 5.686%.

The **option-adjusted yield** is calculated by adding the value of the call option to the bond's current flat price. The value of a callable bond is equal to the value of the bond if it did not have the call option, minus the value of the call option (because the issuer *owns* the call option).

The option-adjusted yield will be less than the yield-to-maturity for a callable bond because callable bonds have higher yields to compensate bondholders for the issuer's call option. The option-adjusted yield can be used to compare the yields of bonds with various embedded options to each other and to similar option-free bonds.

Floating-Rate Note Yields

The values of floating rate notes (FRNs) are more stable than those of fixed-rate debt of similar maturity because the coupon interest rates are reset periodically based on a reference rate.

Recall that the coupon rate on a floating-rate note is the reference rate plus or minus a margin based on the credit risk of the bond relative to the credit risk of the reference rate instrument. The coupon rate for the next period is set using the current reference rate for the reset period, and the payment at the end of the period is based on this rate. For this reason, we say that interest is paid *in arrears*.

If an FRN is issued by a company that has more (less) credit risk than the financial institutions from which the MRR is derived, a margin is added to (subtracted from) the MRR. The liquidity of an FRN and its tax treatment can also affect the margin.

We call the margin used to calculate the bond coupon payments the **quoted margin** and we call the margin required to return the FRN to its par value the **required margin** (also called the **discount margin**). When the credit quality of an FRN is unchanged, the quoted margin is equal to the required margin and the FRN returns to its par value at each reset date when the next coupon payment is reset to the current market rate (plus or minus the appropriate margin).

If the credit quality of the issuer decreases, the quoted margin will be less than the required margin and the FRN will sell at a discount to its par value. If credit quality has improved, the quoted margin will be greater than the required margin and the FRN will sell at a premium to its par value.

A somewhat simplified way of calculating the value of an FRN on a reset date is to use the current reference rate plus the quoted margin to estimate the future cash flows for the FRN and to discount these future cash flows at the reference rate plus the required (discount) margin. More complex models produce better estimates of value.

EXAMPLE: Valuation of a floating-rate note

A \$100,000 floating rate note is based on a 180-day MRR with a quoted margin of 120 basis points. On a reset date with 5 years remaining to maturity, the 180-day MRR is quoted as 3.0% (annualized) and the required rate of return (based on the issuer's current credit rating) is 4.5% (annualized). What is the market value of the floating rate note?

Answer:

The current annualized coupon rate on the note is $3.0\% + 1.2\% = 4.2\%$, so the next semiannual coupon payment will be $4.2\% / 2 = 2.1\%$ of face value. The required return in the market (discount margin) as an effective 180-day discount rate is $4.5\% / 2 = 2.25\%$.

Using a face value of 100%, 10 coupon payments of 2.1%, and a discount rate per period of 2.25%, we can calculate the present value of the floating rate note as:

$$N = 10; I/Y = 2.25\%; FV = 100; PMT = 2.1; CPT PV = 98.67$$

The current value of the note is 98.67% of its face value, or \$98,670.

LOS 44.h: Calculate and interpret yield measures for money market instruments.

For money market securities, debt securities maturing in a year or less, yields are quoted based on various conventions. Some yield quotes are based on a 360-day year while others are based on a 365-day year. Some yield quotes are add-on yields and others are discount yields. Add-on yields are simply the interest to be earned on the amount paid or deposited. Discount yields are annualized current discounts from the face (maturity) values of money market securities.

Bank CDs and market reference rates are typically quoted as annualized add-on rates. U.S. Treasury bills are quoted as their annualized discount from face value, based on a 360-day year.

The relation between a security's yield quoted as an annualized add-on yield based on a 365-day year and its holding period yield (HPY) is:

$$\text{Quoted add-on yield} = \text{HPY} \times 365/\text{days to maturity}$$

Consider a 100-day bank CD with an add-on yield (annualized) of 1.5%, based on a 365-day year. We can calculate the HPY of the CD as the quoted yield of $1.5\% \times 100/365 = 0.41\%$. The purchase of a \$1,000 CD would provide a payment of \$1,004.10 in 100 days.

The relation between a quoted discount and the actual discount based on a 360-day year is:

$$\text{Quoted discount yield} = \text{discount on the security} \times 360/\text{days to maturity}$$

Consider a 180-day U.S. T-bill quoted at 2.2% (annualized) discount yield based on a 360-day year. The actual discount from face value on the T-bill is $180/360 \times 2.2\% = 1.1\%$. A \$1,000 T-bill would be priced at $(1 - 0.011) \times 1,000 = \989 . The HPY of the T-bill is $1,000/989 - 1 = 1.11\%$, slightly higher than its discount from face value of 1.1%.

An analyst should be able to convert the yield of a security calculated on one basis to its yield on another basis. Such adjustments allow us to compare the yields of two money market securities for which quoted yields are calculated differently. The following provides some examples of converting a yield to one based on a different convention.

EXAMPLE: Money market yields

1. A \$1,000 90-day T-bill is priced with an annualized discount of 1.2%. Calculate its market price and its annualized add-on yield based on a 365-day year.
2. A \$1 million negotiable CD with 120 days to maturity is quoted with an add-on yield of 1.4% based on a 365-day year. Calculate the payment at maturity for this CD and its bond equivalent yield.
3. A bank deposit for 100 days is quoted with an add-on yield of 1.5% based on a 360-day year. Calculate the bond equivalent yield and the yield on a semiannual bond basis.

Answer:

1. The discount from face value is $1.2\% \times 90 / 360 \times 1,000 = \3 so the current price is $1,000 - 3 = \$997$. The equivalent add-on yield for 90 days is $3 / 997 = 0.3009\%$. The annualized add-on yield based on a 365-day year is $365 / 90 \times 0.3009 = 1.2203\%$. This add-on yield based on a 365-day year is referred to as the **bond equivalent yield** for a money market security.
2. The add-on interest for the 120-day period is $120 / 365 \times 1.4\% = 0.4603\%$. At maturity, the CD will pay $\$1 \text{ million} \times (1 + 0.004603) = \$1,004,603$. The quoted yield on the CD is the bond equivalent yield because it is an add-on yield annualized based on a 365-day year.
3. Because the yield of 1.5% is an annualized yield calculated based on a 360-day year, the bond equivalent yield, which is based on a 365-day year, is:

$$(365 / 360) \times 1.5\% = 1.5208\%$$

We may want to compare the yield on a money market security to the YTM of a semiannual-pay bond. The method is to convert the money market security's holding period return to an effective semiannual yield, and then double it.

Because the yield of 1.5% is calculated as the add-on yield for 100 days times $100 / 360$, the 100-day holding period return is $1.5\% \times 100 / 360 = 0.4167\%$. The effective annual yield is $1.004167^{365/100} - 1 = 1.5294\%$, the equivalent semiannual yield is $1.015294^{1/2} - 1 = 0.7618\%$, and the annual yield on a semiannual bond basis is $2 \times 0.7618\% = 1.5236\%$.

Because the periodicity of the money market security, $365 / 100$, is greater than the periodicity of 2 for a semiannual-pay bond, the simple annual rate for the money market security, 1.5%, is less than the yield on a semiannual bond basis, which has a periodicity of 2.

MODULE QUIZ 44.3



1. A market rate of discount for a single payment to be made in the future is:
 - A. a spot rate.
 - B. a simple yield.
 - C. a forward rate.
2. Based on semiannual compounding, what would the YTM be on a 15-year, zero-coupon, \$1,000 par value bond that's currently trading at \$331.40?
 - A. 3.750%.
 - B. 5.151%.
 - C. 7.500%.
3. An analyst observes a Widget & Co. 7.125%, 4-year, semiannual-pay bond trading at 102.347% of par (where par is \$1,000). The bond is callable at 101 in two years. What is the bond's yield-to-call?
 - A. 3.167%.
 - B. 5.664%.
 - C. 6.334%.
4. A floating-rate note has a quoted margin of +50 basis points and a required margin of +75 basis points. On its next reset date, the price of the note will be:
 - A. equal to par value.
 - B. less than par value.
 - C. greater than par value.
5. Which of the following money market yields is a bond-equivalent yield?
 - A. Add-on yield based on a 365-day year.
 - B. Discount yield based on a 360-day year.
 - C. Discount yield based on a 365-day year.



MODULE 44.4: YIELD CURVES

Video covering
this content is
available online.

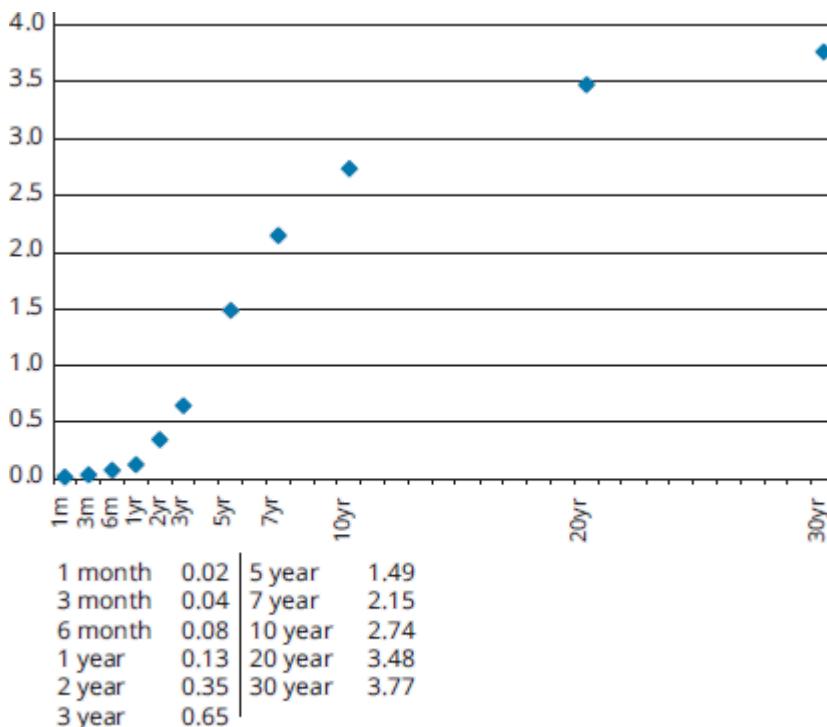
LOS 44.i: Define and compare the spot curve, yield curve on coupon bonds, par curve, and forward curve.

A **yield curve** shows yields by maturity. Yield curves are constructed for yields of various types and it's very important to understand exactly which yield is being shown. The **term structure** of interest rates refers to the yields at different maturities (terms) for like securities or interest rates. The yields on U.S. Treasury coupon bonds by maturity can be found at Treasury.gov, and several yield curves are available at Bloomberg.com.

The **spot rate yield curve** (spot curve) for U.S. Treasury bonds is also referred to as the *zero curve* (for zero-coupon) or *strip curve* (because zero-coupon U.S. Treasury bonds are also called *stripped Treasuries*). Recall that spot rates are the appropriate yields, and therefore appropriate discount rates, for single payments to be made in the future. Yields on zero-coupon government bonds are spot rates. Earlier in this reading, we calculated the value of a bond by discounting each separate payment by the spot rate corresponding to the time until the payment will be received. Spot rates are usually quoted on a semiannual bond basis, so they are directly comparable to YTMs quoted for coupon government bonds.

A **yield curve for coupon bonds** shows the YTMs for coupon bonds at various maturities. Yields are calculated for several maturities and yields for bonds with maturities between these are estimated by linear interpolation. Figure 44.4 shows a yield curve for coupon Treasury bonds constructed from yields on 1-month, 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, 20 year, and 30-year maturities. Yields are expressed on a semiannual bond basis.

Figure 44.4: U.S. Treasury Yield Curve as of August 1, 2013



Source: www.treasury.gov/resource-center

A **par bond yield curve**, or *par curve*, is not calculated from yields on actual bonds but is constructed from the spot curve. The yields reflect the coupon rate that a hypothetical bond at each maturity would need to have to be priced at par. Alternatively, they can be viewed as the YTM of a par bond at each maturity.

Consider a 3-year annual-pay bond and spot rates for one, two, and three years of S_1 , S_2 , and S_3 . The following equation can be used to calculate the coupon rate necessary for the bond to be trading at par.

$$\frac{\text{PMT}}{1 + S_1} + \frac{\text{PMT}}{(1 + S_2)^2} + \frac{\text{PMT} + 100}{(1 + S_3)^3} = 100$$

With spot rates of 1%, 2%, and 3%, a 3-year annual par bond will have a payment that will satisfy:

$$\frac{\text{PMT}}{1.01} + \frac{\text{PMT}}{(1.02)^2} + \frac{\text{PMT} + 100}{(1.03)^3} = 100, \text{ so the payment is } 2.96 \text{ and the par bond, so the payment is } 2.96 \text{ and the par bond}$$

coupon rate is 2.96%.

Forward rates are yields for future periods. The rate of interest on a 1-year loan that would be made two years from now is a forward rate. A **forward yield curve** shows the future rates for bonds or money market securities for the same maturities for annual periods in the future. Typically, the forward curve would show the yields of 1-year securities for each future year, quoted on a semiannual bond basis.

LOS 44.j: Define forward rates and calculate spot rates from forward rates, forward rates from spot rates, and the price of a bond using forward rates.

A forward rate is a borrowing/lending rate for a loan to be made at some future date. The notation used must identify both the length of the lending/borrowing period and when in the future the money will be loaned/borrowed. Thus, 1y1y is the rate for a 1-year loan one year from now; 2y1y is the rate for a 1-year loan to be made two years from now; 3y2y is the 2-year forward rate three years from now; and so on.

The Relationship Between Short-Term Forward Rates and Spot Rates

The idea here is that *borrowing for three years at the 3-year spot rate, or borrowing for one-year periods in three successive years, should have the same cost*. The S_i are the current spot rates for i periods.

This relation is illustrated as $(1 + S_3)^3 = (1 + S_1)(1 + 1y1y)(1 + 2y1y)$. Thus, $S_3 = [(1 + S_1)(1 + 1y1y)(1 + 2y1y)]^{1/3} - 1$, which is the geometric mean return we covered in Quantitative Methods.

EXAMPLE: Computing spot rates from forward rates

If the current 1-year spot rate is 2%, the 1-year forward rate one year from today (1y1y) is 3%, and the 1-year forward rate two years from today (2y1y) is 4%, what is the 3-year spot rate?

Answer:

$$S_3 = [(1.02)(1.03)(1.04)]^{1/3} - 1 = 2.997\%$$

This can be interpreted to mean that a dollar compounded at 2.997% for three years would produce the same ending value as a dollar that earns compound interest of 2% the first year, 3% the next year, and 4% for the third year.



PROFESSOR'S NOTE

You can get a very good approximation of the 3-year spot rate with the simple average of the forward rates. In the previous example, we calculated 2.997% and the simple average of the three annual rates is:

$$\frac{2 + 3 + 4}{3} = 3\%$$

Forward Rates Given Spot Rates

We can use the same relationships we use to calculate spot rates from forward rates to calculate forward rates from spot rates.

Our basic relation between forward rates and spot rates (for two periods) is:

$$(1 + S_2)^2 = (1 + S_1)(1 + 1y1y)$$

This again tells us that an investment has the same expected yield (borrowing has the same expected cost) whether we invest (borrow) for two periods at the 2-period spot rate, S_2 , or for one period at the current 1-year rate, S_1 , and for the next period at the forward rate, $1y1y$. Given two of these rates, we can solve for the other.

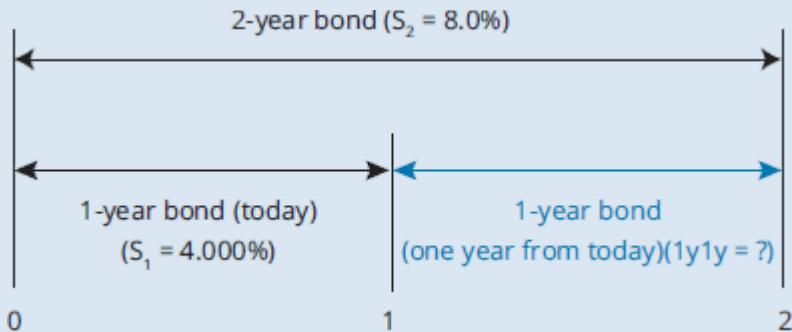
EXAMPLE: Computing a forward rate from spot rates

The 2-period spot rate, S_2 , is 8%, and the 1-period spot rate, S_1 , is 4%. Calculate the forward rate for one period, one period from now, $1y1y$.

Answer:

The following figure illustrates the problem.

Finding a Forward Rate



From our original equality, $(1 + S_2)^2 = (1 + S_1)(1 + 1y1y)$, we can get

$$\frac{(1 + S_2)^2}{(1 + S_1)} = (1 + 1y1y)$$

Or, because we know that both choices have the same payoff in two years:

$$(1.08)^2 = (1.04)(1 + 1y1y)$$

$$(1 + 1y1y) = \frac{(1.08)^2}{(1.04)}$$

$$1y1y = \frac{(1.08)^2}{(1.04)} - 1 = \frac{1.1664}{1.04} - 1 = 12.154\%$$

In other words, investors are willing to accept 4.0% on the 1-year bond today (when they could get 8.0% on the 2-year bond today) only because they can get 12.154% on a 1-year bond one year from today. This future rate that can be locked in today is a forward rate.

Similarly, we can back other forward rates out of the spot rates. We know that:

$$(1 + S_3)^3 = (1 + S_1)(1 + 1y1y)(1 + 2y1y)$$

And that:

$$(1 + S_2)^2 = (1 + S_1)(1 + 1y1y), \text{ so we can write } (1 + S_3)^3 = (1 + S_2)^2(1 + 2y1y)$$

This last equation says that investing for three years at the 3-year spot rate should produce the same ending value as investing for two years at the 2-year spot rate, and then for a third year at $2y1y$, the 1-year forward rate, two years from now.

Solving for the forward rate, $2y1y$, we get:

$$\frac{(1 + S_3)^3}{(1 + S_2)^2} - 1 = 2y1y$$

EXAMPLE: Forward rates from spot rates

Let's extend the previous example to three periods. The current 1-year spot rate is 4.0%, the current 2-year spot rate is 8.0%, and the current 3-year spot rate is 12.0%. Calculate the 1-year forward rates one and two years from now.

Answer:

We know the following relation must hold:

$$(1 + S_2)^2 = (1 + S_1)(1 + 1y1y)$$

We can use it to solve for the 1-year forward rate one year from now:

$$(1.08)^2 = (1.04)(1 + 1y1y), \text{ so } 1y1y = \frac{(1.08)^2}{(1.04)} - 1 = 12.154\%$$

We also know that the relations:

$$(1 + S_3)^3 = (1 + S_1)(1 + 1y1y)(1 + 2y1y)$$

and, equivalently $(1 + S_3)^3 = (1 + S_2)^2(1 + 2y1y)$ must hold.

Substituting values for S_3 and S_2 , we have:

$$(1.12)^3 = (1.08)^2 \times (1 + 2y1y)$$

so that the 1-year forward rate two years from now is:

$$2y1y = \frac{(1.12)^3}{(1.08)^2} - 1 = 20.45\%$$

We can check our results by calculating:

$$S_3 = [(1.04)(1.12154)(1.2045)]^{1/3} - 1 = 12.00\%$$

This may all seem a bit complicated, but the basic relation, that borrowing for successive periods at 1-period rates should have the same cost as borrowing at multiperiod spot rates, can be summed up as:

$$(1 + S_2)^2 = (1 + S_1)(1 + 1y1y) \text{ for two periods, and}$$

$$(1 + S_3)^3 = (1 + S_2)^2(1 + 2y1y) \text{ for three periods.}$$



PROFESSOR'S NOTE

Simple averages also give decent approximations for calculating forward rates from spot rates. In the preceding example, we had spot rates of 4% for one year and 8% for two years. Two years at 8% is 16%, so if the first-year rate is 4%, the second-year rate is close to $16 - 4 = 12\%$ (actual is 12.154). Given a 2-year spot rate of 8% and a 3-year spot rate of 12%, we could approximate the 1-year forward rate from time two to time three as $(3 \times 12) - (2 \times 8) = 20$. That may be close enough (actual is 20.45) to answer a multiple-choice question and, in any case, serves as a good check to make sure the exact rate you calculate is reasonable.

We can also calculate implied forward rates for loans for more than one period. Given spot rates of: 1-year = 5%, 2-year = 6%, 3-year = 7%, and 4-year = 8%, we can calculate $2y2y$.

The implied forward rate on a 2-year loan two years from now, $2y2y$, is:

$$\left[\frac{(1 + S_4)^4}{(1 + S_2)^2} \right]^{1/2} - 1 = \left(\frac{1.08^4}{1.06^2} \right)^{1/2} - 1 = 10.04\%.$$



PROFESSOR'S NOTE

The approximation works for multiperiod forward rates as well.

The difference between four years at 8% (= 32%) and two years at 6% (= 12%) is 20%. Because that difference is for two years, we divide by two to get an annual rate of 10%, $\frac{4 \times 8 - 6 \times 2}{2} = 10$, which is very close to the exact solution of 10.04%.

Valuing a Bond Using Forward Rates

EXAMPLE: Computing a bond value using forward rates

The current 1-year rate, S_1 , is 4%, the 1-year forward rate for lending from time = 1 to time = 2 is $1y1y = 5\%$, and the 1-year forward rate for lending from time = 2 to time = 3 is $2y1y = 6\%$. Value a 3-year annual-pay bond with a 5% coupon and a par value of \$1,000.

Answer:

$$\begin{aligned}
 \text{bond value} &= \frac{50}{1 + S_1} + \frac{50}{(1 + S_1)(1 + 1y1y)} + \frac{1,050}{(1 + S_1)(1 + 1y1y)(1 + 2y1y)} \\
 &= \frac{50}{1.04} + \frac{50}{(1.04)(1.05)} + \frac{1,050}{(1.04)(1.05)(1.06)} = \$1,000.98
 \end{aligned}$$



PROFESSOR'S NOTE

If you think this looks a little like valuing a bond using spot rates, as we did for arbitrage-free valuation, you are correct. The discount factors are equivalent to spot rate discount factors.

If we have a semiannual coupon bond, the calculation methods are the same, but we would use the semiannual discount rate rather than the annualized rate and the number of periods would be the number of semiannual periods.

MODULE 44.5: YIELD SPREADS



LOS 44.k: Compare, calculate, and interpret yield spread measures.

Video covering this content is available online.

A **yield spread** is the difference between the yields of two different bonds. Yield spreads are typically quoted in basis points.

A yield spread relative to a benchmark bond is known as a **benchmark spread**. For example, if a 5-year corporate bond has a yield of 6.25% and its benchmark, the 5-year Treasury note, has a yield of 3.50%, the corporate bond has a benchmark spread of $625 - 350 = 275$ basis points.

For fixed-coupon bonds, on-the-run government bond yields for the same or nearest maturity are frequently used as benchmarks. The benchmark may change during a bond's life. For a 5-year corporate bond, when issued, the benchmark spread is stated relative to a 5-year government bond yield, but two years later (when it has three years remaining to maturity) its benchmark spread will be stated relative to a 3-year government bond yield. A yield spread over a government bond is also known as a **G-spread**.

An alternative to using government bond yields as benchmarks is to use rates for interest rate swaps in the same currency and with the same tenor as a bond. Yield spreads relative to swap rates are known as **interpolated spreads** or **I-spreads**. I-spreads are frequently stated for bonds denominated in euros.



PROFESSOR'S NOTE

For bonds with tenors that do not match an on-the run government bond, yield spreads may be quoted relative to an "interpolated government bond yield." These are still G-spreads.

Yield spreads are useful for analyzing the factors that affect a bond's yield. If a corporate bond's yield increases from 6.25% to 6.50%, this may have been caused by factors that affect all bond yields (macroeconomic factors) or by firm-specific or industry-specific (microeconomic) factors. If a bond's yield increases but its yield spread remains the same, the yield on its

benchmark must have also increased, which suggests macroeconomic factors caused bond yields in general to increase. However, if the yield spread increases, this suggests the increase in the bond's yield was caused by microeconomic factors such as credit risk or the issue's liquidity.



PROFESSOR'S NOTE

Recall from our discussion of the Fisher effect in Economics that an interest rate is composed of the real risk-free rate, the expected inflation rate, and a risk premium. We can think of macroeconomic factors as those that affect the real risk-free rate and expected inflation, and microeconomic factors as those that affect the credit and liquidity risk premium.

Zero-Volatility and Option-Adjusted Spreads

A disadvantage of *G*-spreads and *I*-spreads is that they are theoretically correct only if the spot yield curve is flat so that yields are approximately the same across maturities. Normally, however, the spot yield curve is upward-sloping (i.e., longer-term yields are higher than shorter-term yields).

A method for deriving a bond's yield spread to a benchmark spot yield curve that accounts for the shape of the yield curve is to add an equal amount to each benchmark spot rate and value the bond with those rates. When we find an amount which, when added to the benchmark spot rates, produces a value equal to the market price of the bond, we have the appropriate yield curve spread. A yield spread calculated this way is known as a **zero-volatility spread** or **Z-spread**.

EXAMPLE: Zero-volatility spread

The 1-, 2-, and 3-year spot rates on Treasuries are 4%, 8.167%, and 12.377%, respectively. Consider a 3-year, 9% annual coupon corporate bond trading at 89.464. The YTM is 13.50%, and the YTM of a 3-year Treasury is 12%. Compute the *G*-spread and the *Z*-spread of the corporate bond.

Answer:

The *G*-spread is:

$$\text{G-spread} = \text{YTM}_{\text{Bond}} - \text{YTM}_{\text{Treasury}} = 13.50 - 12.00 = 1.50\%.$$

To compute the *Z*-spread, set the present value of the bond's cash flows equal to today's market price. Discount each cash flow at the appropriate zero-coupon bond spot rate *plus* a fixed spread *ZS*. Solve for *ZS* in the following equation and you have the *Z*-spread:

$$89.464 = \frac{9}{(1.04 + ZS)^1} + \frac{9}{(1.08167 + ZS)^2} + \frac{109}{(1.12377 + ZS)^3}$$

$$\Rightarrow ZS = 1.67\% \text{ or } 167 \text{ basis points}$$

Note that this spread is found by trial-and-error. In other words, pick a number "ZS," plug it into the right-hand side of the equation, and see if the result equals 89.464. If the right-hand side equals the left, then you have found the *Z*-spread. If not, adjust "ZS" in the appropriate direction and recalculate.

An **option-adjusted spread (OAS)** is used for bonds with embedded options. Loosely speaking, the option-adjusted spread takes the option yield component out of the *Z*-spread measure; the

OAS is the spread to the government spot rate curve that the bond would have if it were option-free.

If we calculate an OAS for a callable bond, it will be less than the bond's Z-spread. The difference is the extra yield required to compensate bondholders for the call option. That extra yield is the option value. Thus, we can write:

$$\text{option value} = \text{Z-spread} - \text{OAS}$$

$$\text{OAS} = \text{Z-spread} - \text{option value}$$

For example, if a callable bond has a Z-spread of 180 bp and the value of the call option is 60 bp, the bond's OAS is $180 - 60 = 120$ bp.

MODULE QUIZ 44.4, 44.5

1. Which of the following yield curves is *least likely* to consist of observed yields in the market?
 - A. Forward yield curve.
 - B. Par bond yield curve.
 - C. Coupon bond yield curve.
2. The 4-year spot rate is 9.45%, and the 3-year spot rate is 9.85%. What is the 1-year forward rate three years from today?
 - A. 8.258%.
 - B. 9.850%.
 - C. 11.059%.
3. Given the following spot and forward rates:
 - Current 1-year spot rate is 5.5%.
 - One-year forward rate one year from today is 7.63%.
 - One-year forward rate two years from today is 12.18%.
 - One-year forward rate three years from today is 15.5%.The value of a 4-year, 10% annual-pay, \$1,000 par value bond is *closest* to:
 - A. \$996.
 - B. \$1,009.
 - C. \$1,086.
4. A corporate bond is quoted at a spread of +235 basis points over an interpolated 12-year U.S. Treasury bond yield. This spread is:
 - A. a *G*-spread.
 - B. an *I*-spread.
 - C. a *Z*-spread.

KEY CONCEPTS

LOS 44.a

The price of a bond is the present value of its future cash flows, discounted at the bond's yield-to-maturity.

For an annual-coupon bond with N years to maturity:

$$\text{price} = \frac{\text{coupon}}{(1 + \text{YTM})} + \frac{\text{coupon}}{(1 + \text{YTM})^2} + \dots + \frac{\text{coupon} + \text{principal}}{(1 + \text{YTM})^N}$$

For a semiannual-coupon bond with N years to maturity:

$$\text{price} = \frac{\text{coupon}}{\left(1 + \frac{\text{YTM}}{2}\right)} + \frac{\text{coupon}}{\left(1 + \frac{\text{YTM}}{2}\right)^2} + \dots + \frac{\text{coupon} + \text{principal}}{\left(1 + \frac{\text{YTM}}{2}\right)^{N \times 2}}$$

LOS 44.b

A bond's price and YTM are inversely related. An increase in YTM decreases the price and a decrease in YTM increases the price.

A bond will be priced at a discount to par value if its coupon rate is less than its YTM, and at a premium to par value if its coupon rate is greater than its YTM.

Prices are more sensitive to changes in YTM for bonds with lower coupon rates and longer maturities, and less sensitive to changes in YTM for bonds with higher coupon rates and shorter maturities.

A bond's price moves toward par value as time passes and maturity approaches.

LOS 44.c

Spot rates are market discount rates for single payments to be made in the future.

The no-arbitrage price of a bond is calculated using (no-arbitrage) spot rates as follows:

$$\text{no-arbitrage price} = \frac{\text{coupon}}{(1 + s_1)} + \frac{\text{coupon}}{(1 + s_2)^2} + \dots + \frac{\text{coupon} + \text{principal}}{(1 + s_N)^N}$$

LOS 44.d

The full price of a bond includes interest accrued between coupon dates. The flat price of a bond is the full price minus accrued interest.

Accrued interest for a bond transaction is calculated as the coupon payment times the portion of the coupon period from the previous payment date to the settlement date.

Methods for determining the period of accrued interest include actual days (typically used for government bonds) or 30-day months and 360-day years (typically used for corporate bonds).

LOS 44.e

Matrix pricing is a method used to estimate the yield-to-maturity for bonds that are not traded or infrequently traded. The yield is estimated based on the yields of traded bonds with the same credit quality. If these traded bonds have different maturities than the bond being valued, linear interpolation is used to estimate the subject bond's yield.

LOS 44.f

The effective yield of a bond depends on its periodicity, or annual frequency of coupon payments. For an annual-pay bond the effective yield is equal to the yield-to-maturity. For bonds with greater periodicity, the effective yield is greater than the yield-to-maturity.

A YTM quoted on a semiannual bond basis is two times the semiannual discount rate.

LOS 44.g

Bond yields that follow street convention use the stated coupon payment dates. A true yield accounts for coupon payments that are delayed by weekends or holidays and may be slightly lower than a street convention yield.

Current yield is the ratio of a bond's annual coupon payments to its price. Simple yield adjusts current yield by using straight-line amortization of any discount or premium.

For a callable bond, a yield-to-call may be calculated using each of its call dates and prices. The lowest of these yields and YTM is a callable bond's yield-to-worst.

Floating rate notes have a *quoted margin* relative to a reference rate, typically LIBOR. The quoted margin is positive for issuers with more credit risk than the banks that quote LIBOR and may be negative for issuers that have less credit risk than loans to these banks. The *required margin* on a floating rate note may be greater than the quoted margin if credit quality has decreased, or less than the quoted margin if credit quality has increased.

LOS 44.h

For money market instruments, yields may be quoted on a discount basis or an add-on basis, and may use 360-day or 365-day years. A bond-equivalent yield is an add-on yield based on a 365-day year.

LOS 44.i

A yield curve shows the term structure of interest rates by displaying yields across different maturities.

The spot curve is a yield curve for single payments in the future, such as zero-coupon bonds or stripped Treasury bonds.

The par curve shows the coupon rates for bonds of various maturities that would result in bond prices equal to their par values.

A forward curve is a yield curve composed of forward rates, such as 1-year rates available at each year over a future period.

LOS 44.j

Forward rates are current lending/borrowing rates for short-term loans to be made in future periods.

A spot rate for a maturity of N periods is the geometric mean of forward rates over the N periods. The same relation can be used to solve for a forward rate given spot rates for two different periods.

To value a bond using forward rates, discount the cash flows at times 1 through N by the product of one plus each forward rate for periods 1 to N, and sum them.

For a 3-year annual-pay bond:

$$\text{price} = \frac{\text{coupon}}{(1 + S_1)} + \frac{\text{coupon}}{(1 + S_1)(1 + 1y1y)} + \frac{\text{coupon} + \text{principal}}{(1 + S_1)(1 + 1y1y)(1 + 2y1y)}$$

LOS 44.k

A yield spread is the difference between a bond's yield and a benchmark yield or yield curve. If the benchmark is a government bond yield, the spread is known as a government spread or *G*-spread. If the benchmark is a swap rate, the spread is known as an interpolated spread or *I*-spread.

A zero-volatility spread or Z-spread is the percent spread that must be added to each spot rate on the benchmark yield curve to make the present value of a bond equal to its price.

An option-adjusted spread or OAS is used for bonds with embedded options. For a callable bond, the OAS is equal to the Z-spread minus the call option value in basis points.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 44.1

1. **B** N = 20; I/Y = 15; FV = 1,000; PMT = 100; CPT → PV = -\$687.03. (LOS 44.a)
2. **A** N = 10; I/Y = 7.5; FV = 1,000; PMT = 50; CPT → PV = -\$828.40. (LOS 44.a)
3. **A** The price-yield relationship is inverse. If the required yield decreases, the bond's price will increase, and vice versa. (LOS 44.b)
4. **A** With 20 years to maturity, the value of the bond with an annual-pay yield of 6.5% is N = 20, PMT = 50, FV = 1,000, I/Y = 6.5, CPT → PV = -834.72. With N = 17, CPT → PV = -848.34, so the value will increase \$13.62. (LOS 44.a, 44.b)

Module Quiz 44.2

1. **A** bond value = $\frac{4,000}{1.032} + \frac{4,000}{(1.034)^2} + \frac{104,000}{(1.035)^3} = \$101,419.28$
(LOS 44.c)

2. **B** The full price includes accrued interest, while the flat price does not. Therefore, the flat (or clean) price is 1,059.04 – 23.54 = \$1,035.50. (LOS 44.d)
3. **C** Using linear interpolation, the yield on a bond with six years to maturity should be $6.40\% + (1 / 3)(7.20\% - 6.40\%) = 6.67\%$. A bond with a 7% coupon and a yield of 6.67% is at a premium to par value. (LOS 44.e)

Module Quiz 44.3

1. **A** A spot rate is a discount rate for a single future payment. Simple yield is a measure of a bond's yield that accounts for coupon interest and assumes straight-line amortization of a discount or premium. A forward rate is an interest rate for a future period, such as a 3-month rate six months from today. (LOS 44.g)

$$\begin{aligned} N &= 30; FV = 1,000; PMT = 0; PV = -331.40; CPT \rightarrow I/Y = 3.750 \times 2 \\ &= 7.500\%. \end{aligned}$$

2. **C**
Alternatively, $\left[\left(\frac{1,000}{331.4} \right)^{\frac{1}{30}} - 1 \right] \times 2 = 7.5\%$
(LOS 44.g)

3. **C** N = 4; FV = 1,010; PMT = 35.625; PV = -1,023.47; CPT → I/Y = $3.167 \times 2 = 6.334\%$. (LOS 44.g)

4. **B** If the required margin is greater than the quoted margin, the credit quality of the issue has decreased and the price on the reset date will be less than par value. (LOS 44.g)
5. **A** An add-on yield based on a 365-day year is a bond-equivalent yield. (LOS 44.h)

Module Quiz 44.4, 44.5

1. **B** Par bond yield curves are based on the theoretical yields that would cause bonds at each maturity to be priced at par. Coupon bond yields and forward interest rates can be observed directly from market transactions. (Module 44.4, LOS 44.i)

$$(1.0945)^4 = (1.0985)^3 \times (1 + 3y1y)$$

2. **A** $3y1y = \frac{(1.0945)^4}{(1.0985)^3} - 1 = 8.258\%$

Approximate forward rate = $4(9.45\%) - 3(9.85\%) = 8.25\%$. (Module 44.4, LOS 44.j)

$$\text{Bond value} = \frac{100}{1.055} + \frac{100}{(1.055)(1.0763)} + \frac{100}{(1.055)(1.0763)(1.1218)}$$

3. **B** $+ \frac{1,100}{(1.055)(1.0763)(1.1218)(1.155)} = 1,009.03$

(Module 44.4, LOS 44.j)

4. **A** G-spreads are quoted relative to an actual or interpolated government bond yield. I-spreads are quoted relative to swap rates. Z-spreads are calculated based on the shape of the benchmark yield curve. (Module 44.5, LOS 44.k)

READING 45

INTRODUCTION TO ASSET-BACKED SECURITIES

EXAM FOCUS

In this reading we introduce asset-backed securities, describing their benefits, legal structure, and characteristics. Our primary focus is residential mortgage-backed securities (RMBS). Candidates should understand the characteristics of mortgage pass-through securities and how and why collateralized mortgage obligations are created from them. Be prepared to compare and contrast agency RMBS, nonagency RMBS, and commercial MBS. Finally, candidates should know why collateralized debt obligations are created and how they differ from the other securitized debt securities covered.

MODULE 45.1: STRUCTURE OF MORTGAGE-BACKED SECURITIES



Video covering this content is available online.

LOS 45.a: Explain benefits of securitization for economies and financial markets.

Securitization refers to a process by which financial assets (e.g., mortgages, accounts receivable, or automobile loans) are purchased by an entity that then issues securities supported by the cash flows from those financial assets. The primary benefits of the securitization of financial assets are (1) a reduction in funding costs for firms selling the financial assets to the securitizing entity and (2) an increase in the liquidity of the underlying financial assets.

Consider a bank that makes mortgage loans to home buyers and retains and services these loans (i.e., collects the mortgage payments and performs the necessary recordkeeping functions). To gain exposure to a bank's mortgage loans, investors traditionally could only choose among investing in bank deposits, bank debt securities, or the common equity of banks.

Compared to this traditional structure, with the bank serving the function of financial intermediary between borrowers and lenders, securitization can provide the following benefits:

- Securitization reduces intermediation costs, which results in lower funding costs for borrowers and higher risk-adjusted returns for lenders (investors).
- With securitization, the investors' legal claim to the mortgages or other loans is stronger than it is with only a general claim against the bank's overall assets.

- When a bank securitizes its loans, the securities are actively traded, which increases the liquidity of the bank's assets compared to holding the loans.
- By securitizing loans, banks are able to lend more than if they could only fund loans with bank assets. When a loan portfolio is securitized, the bank receives the proceeds, which can then be used to make more loans.
- Securitization has led to financial innovation that allows investors to invest in securities that better match their preferred risk, maturity, and return characteristics. As an example, an investor with a long investment horizon can invest in a portfolio of long-term mortgage loans rather than in only bank bonds, deposits, or equities. The investor can gain exposure to long-term mortgages without having the specialized resources and expertise necessary to provide loan origination and loan servicing functions.
- Securitization provides diversification and risk reduction compared to purchasing individual loans (whole loans).

LOS 45.b: Describe securitization, including the parties involved in the process and the roles they play.

We can illustrate the basic structure of a *securitization transaction* with this simplified, fictitious example of Fred Motor Company.

Fred Motor Company sells most of its cars on retail sales installment contracts (i.e., auto loans). The customers buy the automobiles, and Fred loans the customers the money for the purchase (i.e., Fred *originates* the loans) with the autos as collateral and receives principal and interest payments on the loans until they mature. The loans have maturities of 48 to 60 months at various interest rates. Fred is also the *servicer* of the loans (i.e., it collects principal and interest payments, sends out delinquency notices, and repossesses and disposes of the autos if the customers do not make timely payments).

Fred has 50,000 auto loans totaling \$1 billion that it would like to remove from its balance sheet and use the proceeds to make more auto loans. It accomplishes this by selling the loan portfolio to a **special purpose entity (SPE)** called Auto Loan Trust for \$1 billion (Fred is called the *seller* or *depositor*). The SPE, which is set up for the specific purpose of buying these auto loans and selling asset-backed securities (ABS), is referred to as the *trust* or the *issuer*. The SPE then sells ABS to investors. The loan portfolio is the collateral supporting the ABS because the cash flows from the loans are the source of the funds to make the promised payments to investors. An SPE is sometimes also called a special purpose vehicle (SPV). The SPE is a separate legal entity from Fred.

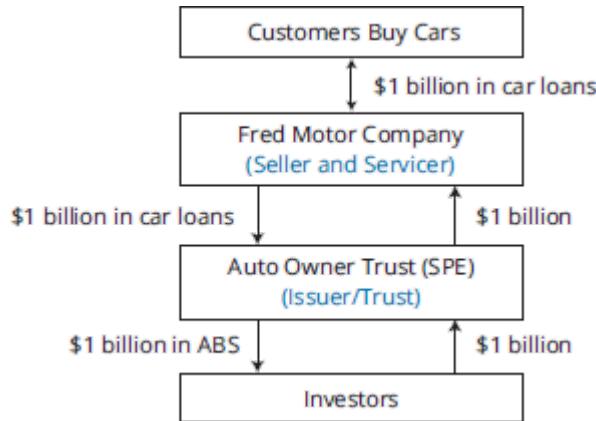
Let's review the parties to this transaction and their functions:

- The seller or depositor (Fred) originates the auto loans and sells the portfolio of loans to Auto Loan Trust, the SPE.
- The issuer/trust (Auto Loan Trust) is the SPE that buys the loans from the seller and issues ABS to investors.
- The servicer (Fred) services the loans.

- In this case, the seller and the servicer are the same entity (Fred Motor Company), but that is not always the case.

The structure of this securitization transaction is illustrated in Figure 45.1.

Figure 45.1: Structure of Fred Motor Company Asset Securitization



Subsequent to the initial transaction, the principal and interest payments on the original loans are allocated to pay servicing fees to the servicer and principal and interest payments to the owners of the ABS.

ABS are most commonly backed by automobile loans, credit card receivables, home equity loans, manufactured housing loans, student loans, Small Business Administration (SBA) loans, corporate loans, corporate bonds, emerging market bonds, and structured financial products. When the loans owned by the trust (SPE) are mortgages, we refer to the securities issued by the trust as **mortgage-backed securities (MBS)**.

Note that the SPE is a separate legal entity from Fred and the buyers of the ABS have no claim on other assets of Fred, only on the loans sold to the SPE.

If Fred had issued corporate bonds to raise the funds to make more auto loans, the bondholders would be subject to the financial risks of Fred. With the ABS structure, a decline in the financial position of Fred, its ability to make cash payments, or its bond rating do not affect the value of the claims of ABS owners to the cash flows from the trust collateral (loan portfolio) because it has been sold by Fred, which is now simply the servicer (not the owner) of the loans. The credit rating of the ABS securities may be higher than the credit rating of bonds issued by Fred, in which case the cost to fund the loans using the ABS structure is lower than if Fred funded additional loans by issuing corporate bonds.

LOS 45.c: Describe typical structures of securitizations, including credit tranching and time tranching.

Securitzations may involve a single class of ABS so the cash flows to the securities are the same for all security holders. They can also be structured with multiple classes of securities, each with a different claim to the cash flows of the underlying assets. The different classes are often referred to as **tranches**. With this structure, a particular risk of the ABS securities is

redistributed across the tranches. Some bear more of the risk and others bear less of the risk. The total risk is unchanged, simply reapportioned.

With **credit tranching**, the ABS tranches will have different exposures to the risk of default of the assets underlying the ABS. With this structure, also called a **senior/subordinated structure**, the subordinated tranches absorb credit losses as they occur (up to their principal values). The level of protection for the senior tranche increases with the proportion of subordinated bonds in the structure.

Let's look at an example to illustrate how a senior/subordinated structure redistributes the credit risk compared to a single-class structure. Consider an ABS with the following bond classes:

Senior Tranche	\$300,000,000
Subordinated Tranche A	\$80,000,000
Subordinated Tranche B	<u>\$30,000,000</u>
Total	\$410,000,000

Tranche B is first to absorb any losses (and is termed the *first-loss tranche*) until they exceed \$30 million in principal. Any losses from default of the underlying assets greater than \$30 million, and up to \$110 million, will be absorbed by Subordinated Tranche A. The Senior Tranche is protected from any credit losses of \$110 million or less and therefore will have the highest credit rating and offer the lowest yield of the three bond classes. This structure is also called a **waterfall** structure because in liquidation, each subordinated tranche would receive only the "overflow" from the more senior tranche(s) if they are repaid their principal value in full.

With **time tranching**, the first (sequential) tranche receives all principal repayments from the underlying assets up to the principal value of the tranche. The second tranche would then receive all principal repayments from the underlying assets until the principal value of this tranche is paid off. There may be other tranches with sequential claims to remaining principal repayments. Time tranching reapportions prepayment risk—the risk of receiving principal payments either sooner than expected (contraction risk) or later than expected (extension risk). The first tranche receives all principal payments until it is paid off, and so has the greatest contraction risk. The second tranche receives no prepayments until the first tranche is paid in full, then receives all principal payments until the entire issue is paid off, and so has more extension risk and less contraction risk compared with the first tranche.

Both credit tranching and time tranching are often included in the same structure. More detail about time tranching and the related planned amortization/support tranche structure is included later in this review when we discuss the structures of mortgage-backed securities.

LOS 45.d: Describe types and characteristics of residential mortgage loans that are typically securitized.

A **residential mortgage loan** is a loan for which the collateral that underlies the loan is residential real estate. If the borrower defaults on the loan, the lender has a legal claim to the collateral property. One key characteristic of a mortgage loan is its **loan-to-value ratio (LTV)**,

the percentage of the value of the collateral real estate that is loaned to the borrower. The lower the LTV, the higher the borrower's equity in the property.

For a lender, loans with lower LTVs are less risky because the borrower has more to lose in the event of default (so is less likely to default). Also, if the property value is high compared to the loan amount, the lender is more likely to recover the amount loaned if the borrower defaults and the lender repossesses and sells the property. In the United States, mortgages made to borrowers with good credit are termed *prime loans*. Mortgages to borrowers of lower credit quality, or that have a lower-priority claim to the collateral in event of default, are termed *subprime loans*.

Typical mortgage terms and structures differ across regions and countries. The key characteristics of mortgage loans include their maturity, the determination of interest charges, how the loan principal is amortized, the terms under which prepayments of loan principal are allowed, and the rights of the lender in the event of default by the borrower. We address each of the characteristics in more detail.

Maturity

The term of a mortgage loan is the time until the final loan payment is made. In the United States, mortgage loans typically have terms from 15 to 30 years. Terms are longer, 20 to 40 years, in many European countries and as long as 50 years in others. In Japan, mortgage loans may have terms of 100 years.

Interest Rate

A **fixed-rate mortgage** has an interest rate that is unchanged over the life of the mortgage.

An **adjustable-rate mortgage (ARM)**, also called a **variable-rate mortgage**, has an interest rate that can change over the life of the mortgage. An **index-referenced mortgage** has an interest rate that changes based on a market determined reference rate such as LIBOR or the one-year U.S. Treasury bill rate, although several other reference rates are used.

A mortgage loan may have an interest rate that is fixed for some initial period, but adjusted after that. If the loan becomes an adjustable-rate mortgage after the initial fixed-rate period it is called a *hybrid mortgage*. If the interest rate changes to a different fixed rate after the initial fixed-rate period it is called a *rollover* or *renegotiable mortgage*.

A **convertible mortgage** is one for which the initial interest rate terms, fixed or adjustable, can be changed at the option of the borrower, to adjustable or fixed, for the remaining loan period.

Amortization of Principal

With a **fully amortizing** loan, each payment includes both an interest payment and a repayment of some of the loan principal so there is no loan principal remaining after the last regular mortgage payment. When payments are fixed for the life of the loan, payments in the beginning of the loan term have a large interest component and a small principal repayment component, and payments at the end of the loan terms have a small interest component and large principal repayment component.

A loan is said to be **partially amortizing** when loan payments include some repayment of principal, but there is a lump sum of principal that remains to be paid at the end of the loan

period which is called a *balloon payment*. With an **interest-only mortgage**, there is no principal repayment for either an initial period or the life of the loan. If no principal is paid for the life of the loan it is an *interest-only lifetime* mortgage and the balloon payment is the original loan principal amount. Other interest-only mortgages specify that payments are interest-only over some initial period, with partial or full amortization of principal after that.

Prepayment Provisions

A partial or full repayment of principal in excess of the scheduled principal repayments required by the mortgage is referred to as a **prepayment**. If a homeowner sells her home during the mortgage term (a common occurrence), repaying the remaining principal is required and is one type of prepayment. A homeowner who *refinances* her mortgage prepays the remaining principal amount using the proceeds of a new, lower interest rate loan. Some homeowners prepay by paying more than their scheduled payments in order to reduce the principal outstanding, reduce their interest charges, and eventually pay off their loans prior to maturity.

Some loans have no penalty for prepayment of principal while others have a **prepayment penalty**. A prepayment penalty is an additional payment that must be made if principal is prepaid during an initial period after loan origination or, for some mortgages, prepaid anytime during the life of the mortgage. A prepayment penalty benefits the lender by providing compensation when the loan is paid off early because market interest rates have decreased since the mortgage loan was made (i.e., loans are refinanced at a lower interest rate).

Foreclosure

Some mortgage loans are **nonrecourse loans**, which means the lender has no claim against the assets of the borrower except for the collateral property itself. When this is the case, if home values fall so the outstanding loan principal is greater than the home value, borrowers sometimes voluntarily return the property to the lender in what is called a *strategic default*.

Other mortgage loans are **recourse loans** under which the lender has a claim against the borrower for the amount by which the sale of a repossessed collateral property falls short of the principal outstanding on the loan. Understandably, borrowers are more likely to default on nonrecourse loans than on recourse loans. In Europe, most residential mortgages are recourse loans. In the United States, they are recourse loans in some states and nonrecourse in others.

LOS 45.e: Describe types and characteristics of residential mortgage-backed securities, including mortgage pass-through securities and collateralized mortgage obligations, and explain the cash flows and risks for each type.

LOS 45.f: Define prepayment risk and describe the prepayment risk of mortgage-backed securities.

Residential mortgage-backed securities (RMBS) in the United States are termed **agency RMBS** or **nonagency RMBS**, depending on the issuer of the securities. Agency RMBS are issued by the Government National Mortgage Association (GNMA or Ginnie Mae), the Federal National Mortgage Association (Fannie Mae), and the Federal Home Loan Mortgage Corporation (Freddie Mac). Ginnie Mae securities are guaranteed by the GNMA and are considered to be backed by the full faith and credit of the U.S. government. Fannie Mae and Freddie Mac also guarantee the MBS

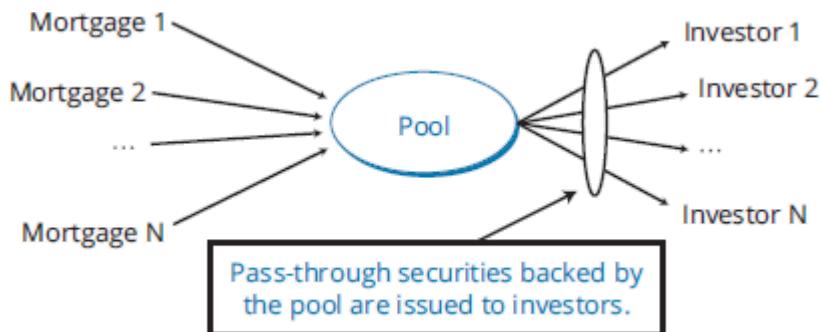
they issue but are *government-sponsored enterprises (GSE)*. While they are not considered to be backed by the full faith and credit of the U.S. government, these securities are considered to have very high credit quality.

Agency RMBS are **mortgage pass-through securities**. Each mortgage pass-through security represents a claim on the cash flows from a pool of mortgages. Any number of mortgages may be used to form the pool, and any mortgage included in the pool is referred to as a **securitized mortgage**. The mortgages in the pool typically have different maturities and different mortgage rates. The **weighted average maturity (WAM)** of the pool is equal to the weighted average of the final maturities of all the mortgages in the pool, weighted by each mortgage's outstanding principal balance as a proportion of the total outstanding principal value of all the mortgages in the pool. The **weighted average coupon (WAC)** of the pool is the weighted average of the interest rates of all the mortgages in the pool. The investment characteristics of mortgage pass-through securities are a function of their cash flow features and the strength of the guarantee provided.

In order to be included in agency MBS pools, loans must meet certain criteria, including a minimum percentage down payment, a maximum LTV ratio, maximum size, minimum documentation required, and insurance purchased by the borrower. Loans that meet the standards for inclusion in agency MBS are called *conforming loans*. Loans that do not meet the standards are called *nonconforming loans*. Nonconforming mortgages can be securitized by private companies for *nonagency RMBS*.

Investors in mortgage pass-through securities receive the monthly cash flows generated by the underlying pool of mortgages, less any servicing and guarantee/insurance fees. The fees account for the fact that **pass-through rates** (i.e., the coupon rate on the MBS, also called its *net interest* or *net coupon*) are less than the mortgage rate of the underlying mortgages in the pool.

Figure 45.2: Mortgage Pass-Through Cash Flow



The timing of the cash flows to pass-through security holders does not exactly coincide with the cash flows generated by the pool. This is due to the delay between the time the mortgage service provider receives the mortgage payments and the time the cash flows are *passed through* to the security holders.



MODULE QUIZ 45.1

1. Economic benefits of securitization *least likely* include:
 - A. reducing excessive lending by banks.
 - B. reducing funding costs for firms that securitize assets.
 - C. increasing the liquidity of the underlying financial assets.

2. In a securitization, the issuer of asset-backed securities is *best* described as:
 - A. the SPE.
 - B. the seller.
 - C. the servicer.
3. A mortgage-backed security with a senior/subordinated structure is said to feature:
 - A. time tranching.
 - B. credit tranching.
 - C. a pass-through structure.
4. A mortgage that has a balloon payment equal to the original loan principal is:
 - A. a convertible mortgage.
 - B. a fully amortizing mortgage.
 - C. an interest-only lifetime mortgage.
5. Residential mortgages that may be included in agency RMBS are *least likely* required to have:
 - A. a minimum loan-to-value ratio.
 - B. insurance on the mortgaged property.
 - C. a minimum percentage down payment.
6. The primary motivation for issuing collateralized mortgage obligations (CMOs) is to reduce:
 - A. extension risk.
 - B. funding costs.
 - C. contraction risk.

MODULE 45.2: PREPAYMENT RISK AND NON-MORTGAGE-BACKED ABS



Video covering this content is available online.

Prepayment Risk

An important characteristic of pass-through securities is their **prepayment risk**. Because the mortgage loans used as collateral for agency MBS have no prepayment penalty, the MBS themselves have significant prepayment risk. Recall that prepayments are principal repayments in excess of the scheduled principal repayments for amortizing loans. The risk that prepayments will be slower than expected is called *extension risk* and the risk that prepayments will be more rapid than expected is called *contraction risk*.

Prepayments cause the timing and amount of cash flows from mortgage loans and MBS to be uncertain; rapid prepayment reduces the amount of principal outstanding on the loans supporting the MBS so the total interest paid over the life of the MBS is reduced. Because of this, it is necessary to make specific assumptions about prepayment rates in order to value mortgage pass-through securities. The single monthly mortality rate (SMM) is the percentage by which prepayments reduce the month-end principal balance, compared to what it would have been with only scheduled principal payments (with no prepayments). The **conditional prepayment rate (CPR)** is an annualized measure of prepayments. Prepayment rates depend on the weighted average coupon rate of the loan pool, current interest rates, and prior prepayments of principal.

The Public Securities Association (PSA) *prepayment benchmark* assumes that the monthly prepayment rate for a mortgage pool increases as it ages (becomes *seasoned*). The PSA benchmark is expressed as a monthly series of CPRs. If the prepayment rate (CPR) of an MBS is expected to be the same as the PSA standard benchmark CPR, we say the PSA is 100 (100% of the benchmark CPR). A pool of mortgages may have prepayment rates that are faster or slower than PSA 100, depending on the current level of interest rates and the coupon rate of the issue. A PSA of 50 means that prepayments are 50% of the PSA benchmark CPR, and a PSA of 130 means that prepayments are 130% of the PSA benchmark CPR.

Based on an assumption about the prepayment rate for an MBS, we can calculate its **weighted average life**, or simply average life, which is the expected number of years until all the loan principal is repaid. Because of prepayments, the average life of an MBS will be less than its weighted average maturity. During periods of falling interest rates, the refinancing of mortgage loans will accelerate prepayments and reduce the average life of an MBS. A high PSA, such as 400, will reduce the average life of an MBS to only 4.5 years, compared to an average life of about 11 years for an MBS with a PSA of 100.

Collateralized Mortgage Obligations

Collateralized mortgage obligations (CMO) are securities that are collateralized by RMBS. Each CMO has multiple bond classes (CMO tranches) that have different exposures to prepayment risk. The total prepayment risk of the underlying RMBS is not changed; the prepayment risk is simply reapportioned among the various CMO tranches.

Institutional investors have different tolerances for prepayment risk. Some are primarily concerned with extension risk while others may want to minimize exposure to contraction risk. By partitioning and distributing the cash flows generated by RMBS into different risk packages to better match investor preferences, CMOs increase the potential market for securitized mortgages and perhaps reduce funding costs as a result.

CMOs are securities backed by mortgage pass-through securities (i.e., they are securities secured by other securities). Interest and principal payments from the mortgage pass-through securities are allocated in a specific way to different bond classes called tranches, so that each tranche has a different claim against the cash flows of the mortgage pass-throughs. Each CMO tranche has a different mixture of contraction and extension risk. Hence, CMO securities can be more closely matched to the unique asset/liability needs of institutional investors and investment managers.

The primary CMO structures include sequential-pay tranches, planned amortization class tranches (PACs), support tranches, and floating-rate tranches.

Sequential Pay CMO

One way to reapportion the prepayment risk inherent in the underlying pass-through MBS is to separate the cash flows into tranches that are retired sequentially (i.e., create a **sequential pay CMO**). As an example of this structure, we consider a simple CMO with two tranches. Both tranches receive interest payments at a specified coupon rate, but all principal payments (both scheduled payments and prepayments) are paid to Tranche 1 (the *short tranche*) until its principal is paid off. Principal payments then flow to Tranche 2 until its principal is paid off.

Contraction and extension risk still exist with this structure, but they have been redistributed to some extent between the two tranches. The short tranche, which matures first, offers investors relatively more protection against extension risk. The other tranche provides relatively more protection against contraction risk. Let's expand this example with some specific numbers to illustrate how sequential pay structures work.

Consider the simplified CMO structure presented in Figure 45.3. Payments to the two sequential-pay tranches are made first to Tranche A and then to Tranche B.

Figure 45.3: Sequential Pay CMO Structure

CMO Structure		
Tranche	Outstanding Par Value	Coupon Rate
A	\$200,000,000	8.50%
B	50,000,000	8.50%

Payments from the underlying collateral (which has a pass-through coupon rate of 8.5%) for the first five months, as well as months 183 through 187, are shown in Figure 45.4. These payments include scheduled payments plus estimated prepayments based on an assumed prepayment rate. (Note that some totals do not match due to rounding.)

Figure 45.4: CMO Projected Cash Flows

Month	Beginning Principal Balance	Principal Payment	Interest	Total Cash Flow = Principal Plus Interest
1	\$250,000,000	\$391,128	\$1,770,833	\$2,161,961
2	249,608,872	454,790	1,768,063	2,222,853
3	249,154,082	518,304	1,764,841	2,283,145
4	248,635,778	581,620	1,761,170	2,342,790
5	248,054,157	644,690	1,757,050	2,401,741
183	\$51,491,678	\$545,153	\$364,733	\$909,886
184	50,946,525	540,831	360,871	901,702
185	50,405,694	536,542	357,040	893,582
186	49,869,152	532,287	353,240	885,526
187	49,336,866	528,065	349,469	877,534



PROFESSOR'S NOTE

This example is provided as an illustration of how cash flows are allocated to sequential tranches. The LOS does not require you to do the calculations that underlie the numbers in Figure 45.4. The important point here is how the cash flows are allocated to each tranche.

Planned Amortization Class (PAC) CMO

Another CMO structure has one or more **planned amortization class (PAC)** tranches and **support tranches**. A PAC tranche is structured to make predictable payments, regardless of actual prepayments to the underlying MBS. The PAC tranches have both reduced contraction risk and reduced extension risk compared to the underlying MBS.

Reducing the prepayment risk of the PAC tranches is achieved by increasing the prepayment risk of the CMO's support tranches. If principal repayments are more rapid than expected, the support tranche receives the principal repayments in excess of those specifically allocated to the PAC tranches. Conversely, if the actual principal repayments are slower than expected, principal repayments to the support tranche are curtailed so the scheduled PAC payments can

be made. The larger the support tranche(s) relative to the PAC tranches, the smaller the probability that the cash flows to the PAC tranches will differ from their scheduled payments.

For a given CMO structure there are limits to how fast or slow actual prepayment experience can be before the support tranches can no longer either provide or absorb prepayments in the amounts required to keep the PAC payments to their scheduled amounts. The upper and lower bounds on the actual prepayment rates for which the support tranches are sufficient to either provide or absorb actual prepayments in order to keep the PAC principal repayments on schedule are called the **initial PAC collar**.

A PAC may have an initial collar given as 100 – 300 PSA. This means the PAC will make its scheduled payments to investors unless actual prepayment experience is outside these bounds (i.e., above 300 PSA or below 100 PSA). If the prepayment rate is outside of these bounds so payments to a PAC tranche are either sooner or later than promised, the PAC tranche is referred to as a **broken PAC**.

Support tranches have both more contraction risk and more extension risk than the underlying MBS and have a higher promised interest rate than the PAC tranche.

As an example, Figure 45.5 shows the average life for a hypothetical structure that includes a PAC I tranche and a support tranche at various PSA speeds, assuming the PSA speed stays at that level for the entire life of the PAC tranche.

Figure 45.5: Average Life Variability of PAC I Tranche vs. Support Tranche

PSA Speed	PAC I Tranche	Support Tranche
0	13.2	24.0
50	8.8	21.1
100	6.5	17.1
150	6.5	13.3
200	6.5	10.4
250	6.5	5.2
300	6.5	2.9
350	5.9	2.4
400	5.4	1.8
450	4.6	1.5
500	4.2	1.2

Figure 45.5 illustrates that the PAC I tranche has less prepayment risk than the support tranche because the variability of its average life is significantly lower.

- When prepayment speeds fall and prepayments decrease, the support tranche average life is significantly longer than the average life of the PAC I tranche. Thus, the support tranche has significantly more extension risk.
- When prepayment speeds rise and prepayments increase, the support tranche average life is much shorter than that of the PAC I tranche. Thus, the support tranche also has significantly more contraction risk.

- Within the initial PAC collar of 100 to 300 PSA, the average life of the PAC I tranche is constant at 6.5 years.

Nonagency RMBS

RMBS not issued by GNMA, Fannie Mae, or Freddie Mac are referred to as **nonagency RMBS**. They are not guaranteed by the government, so credit risk is an important consideration. The credit quality of a nonagency MBS depends on the credit quality of the borrowers as well as the characteristics of the loans, such as their LTV ratios. To be investment grade, most nonagency RMBS include some sort of **credit enhancement**. The level of credit enhancement is directly proportional to the credit rating desired by the issuer. Rating agencies determine the exact amount of credit enhancement necessary for an issue to hold a specific rating.

Credit tranching (subordination) is often used to enhance the credit quality of senior RMBS securities. A **shifting interest mechanism** is a method for addressing a decrease in the level of credit protection provided by junior tranches as prepayments or defaults occur in a senior/subordinated structure. If prepayments or credit losses decrease the credit enhancement of the senior securities, the shifting interest mechanism suspends payments to the subordinated securities for a period of time until the credit quality of the senior securities is restored.

LOS 45.g: Describe characteristics and risks of commercial mortgage-backed securities.

Commercial mortgage-backed securities (CMBS) are backed by income-producing real estate, typically in the form of:

- Apartments (multifamily).
- Warehouses (industrial use property).
- Shopping centers.
- Office buildings.
- Health care facilities.
- Senior housing.
- Hotel/resort properties.

An important difference between residential and commercial MBS is the obligations of the borrowers of the underlying loans. Residential MBS loans are repaid by homeowners; commercial MBS loans are repaid by real estate investors who, in turn, rely on tenants and customers to provide the cash flow to repay the mortgage loan. CMBS mortgages are structured as **nonrecourse loans**, meaning the lender can *only* look to the collateral as a means to repay a delinquent loan if the cash flows from the property are insufficient. In contrast, a residential mortgage lender with recourse can go back to the borrower personally in an attempt to collect any excess of the loan amount above the net proceeds from foreclosing on and selling the property.

For these reasons, the analysis of CMBS securities focuses on the credit risk of the property and not the credit risk of the borrower. The analysis of CMBS structures focuses on two key ratios to assess credit risk.

- 1. Debt service coverage ratio** is a basic cash flow coverage ratio of the amount of cash flow from a commercial property available to make debt service payments compared to the required debt service cost.

$$\text{debt-service coverage ratio} = \frac{\text{net operating income}}{\text{debt service}}$$

Net operating income (NOI) is calculated after the deduction for real estate taxes but before any relevant income taxes. This ratio, which is typically between one and two, indicates greater protection to the lender when it is higher. Debt service coverage ratios below one indicate that the borrower is not generating sufficient cash flow to make the debt payments and is likely to default. Remember: *the higher the better* for this ratio from the perspective of the lender and the MBS investor.

- 2. Loan-to-value ratio** compares the loan amount on the property to its current fair market or appraisal value.

$$\text{loan-to-value ratio} = \frac{\text{current mortgage amount}}{\text{current appraised value}}$$

The lower this ratio, the more protection the mortgage lender has in making the loan. Loan-to-value ratios determine the amount of collateral available, above the loan amount, to provide a cushion to the lender should the property be foreclosed on and sold. Remember: *the lower the better* for this ratio from the perspective of the lender and the MBS investor.

The basic **CMBS structure** is created to meet the risk and return needs of the CMBS investor. As with residential MBS securities, rating organizations such as S&P and Moody's assess the credit risk of each CMBS issue and determine the appropriate credit rating. Each CMBS is segregated into tranches. Losses due to default are first absorbed by the tranche with the lowest priority. Sometimes this most-junior tranche is not rated and is then referred to as the equity tranche, residual tranche, or first-loss tranche.

As with any fixed-rate security, call protection is valuable to the bondholder. In the case of MBS, call protection is equivalent to prepayment protection (i.e., restrictions on the early return of principal through prepayments). CMBS provide call protection in two ways: loan-level call protection provided by the terms of the individual mortgages and call protection provided by the CMBS structure.

There are several means of creating **loan-level call protection**:

- *Prepayment lockout.* For a specific period of time (typically two to five years), the borrower is prohibited from prepaying the mortgage loan.
- *Defeasance.* Should the borrower insist on making principal payments on the mortgage loan, the mortgage loan can be defeased. This is accomplished by using the prepaid principal to purchase a portfolio of government securities that is sufficient to make the remaining required payments on the CMBS. Given the high credit quality of government securities, defeased loans increase the credit quality of a CMBS loan pool.
- *Prepayment penalty points.* A penalty fee expressed in points may be charged to borrowers who prepay mortgage principal. Each point is 1% of the principal amount prepaid.

- *Yield maintenance charges.* The borrower is charged the amount of interest lost by the lender should the loan be prepaid. This *make whole* charge is designed to make lenders indifferent to prepayment, as cash flows are equivalent (at current market rates) whether the loan is prepaid or not.

With all loan call protection programs, any prepayment penalties received are distributed to the CMBS investors in a manner determined by the structure of the CMBS issue.

To create **CMBS-level call protection**, CMBS loan pools are segregated into tranches with a specific sequence of repayment. Those tranches with a higher priority will have a higher credit rating than lower priority tranches because loan defaults will first affect the lower tranches. A wide variety of features can be used to provide call protection to the more senior tranches of the CMBS.

Commercial mortgages are typically amortized over a period longer than the loan term; for example, payments for a 20-year commercial mortgage may be determined based on a 30-year amortization schedule. At the end of the loan term, the loan will still have principal outstanding that needs to be paid; this amount is called a **balloon payment**. If the borrower is unable to arrange refinancing to make this payment, the borrower is in default. This possibility is called balloon risk. The lender will be forced to extend the term of the loan during a workout period, during which time the borrower will be charged a higher interest rate. Because balloon risk entails extending the term of the loan, it is also referred to as extension risk for CMBS.

LOS 45.h: Describe types and characteristics of non-mortgage asset-backed securities, including the cash flows and risks of each type.

In addition to those backed by mortgages, there are ABS that are backed by various types of financial assets including small business loans, accounts receivable, credit card receivables, automobile loans, home equity loans, and manufactured housing loans. Each of these types of ABS has different risk characteristics and their structures vary to some extent as well. Here we explain the characteristics of two types, ABS backed by automobile loans and ABS backed by credit card receivables. These two have an important difference in that automobile loans are fully amortizing while credit card receivables are nonamortizing.

Auto Loan ABS

Auto loan-backed securities are backed by loans for automobiles. Auto loans have maturities from 36 to 72 months. Issuers include the financial subsidiaries of auto manufacturers, commercial banks, credit unions, finance companies, and other small financial institutions.

The cash flow components of auto loan-backed securities include interest payments, scheduled principal payments, and prepayments. Auto loans prepay if the cars are sold, traded in, or repossessed. Prepayments also occur if the car is stolen or wrecked and the loan is paid off from insurance proceeds. Finally, the borrower may simply use excess cash to reduce or pay off the loan balance.

Automobile loan ABS all have some sort of credit enhancement to make them attractive to institutional investors. Many have a senior-subordinated structure, with a junior tranche that absorbs credit risk. One or more internal credit enhancement methods, a reserve account, an

excess interest spread, or overcollateralization, is also often present in these structures. Just as with mortgages, prime loans refer to those made to borrowers with higher credit ratings and sub-prime loans refers to those made to borrowers with low credit ratings.

Credit Card ABS

Credit card receivable-backed securities are ABS backed by pools of credit card debt owed to banks, retailers, travel and entertainment companies, and other credit card issuers.

The cash flow to a pool of credit card receivables includes finance charges, annual fees, and principal repayments. Credit cards have periodic payment schedules, but because their balances are revolving (i.e., nonamortizing), the principal amount is maintained for a period of time. Interest on credit card ABS is paid periodically, but no principal is paid to the ABS holders during the **lockout period**, which may last from 18 months to 10 years after the ABS are created.

If the underlying credit card holders make principal payments during the lockout period, these payments are used to purchase additional credit card receivables, keeping the overall value of the receivables pool relatively constant. Once the lockout period ends, principal payments are passed through to security holders. Credit card ABS typically have an early (rapid) amortization provision that provides for earlier amortization of principal when it is necessary to preserve the credit quality of the securities.

Interest rates on credit card ABS are sometimes fixed but often they are floating. Interest payments may be monthly, quarterly, or for longer periods.

LOS 45.i: Describe collateralized debt obligations, including their cash flows and risks.

A **collateralized debt obligation (CDO)** is a structured security issued by an SPE for which the collateral is a pool of debt obligations. When the collateral securities are corporate and emerging market debt, they are called *collateralized bond obligations* (CBO). *Collateralized loan obligations* (CLO) are supported by a portfolio of leveraged bank loans. Unlike the ABS we have discussed, CDOs do not rely on interest payments from the collateral pool. CDOs have a **collateral manager** who buys and sells securities in the collateral pool in order to generate the cash to make the promised payments to investors.

Structured finance CDOs are those where the collateral is ABS, RMBS, other CDOs, and CMBS.

Synthetic CDOs are those where the collateral is a portfolio of credit default swaps on structured securities.



PROFESSOR'S NOTE

Credit default swaps are derivative securities that decrease (increase) in value as the credit quality of their reference securities increases (decreases).

CDOs issue three classes of bonds (tranches): senior bonds, mezzanine bonds, and subordinated bonds (sometimes called the equity or residual tranche). The subordinated tranche has characteristics more similar to those of equity investments than bond investments. In creating a

CDO, the structure must be able to offer an attractive return on the subordinated tranche, after accounting for the required yields on the senior and mezzanine bond classes.

An investment in the equity or residual tranche can be viewed as a leveraged investment where borrowed funds (raised from selling the senior and mezzanine tranches) are used to purchase the debt securities in the CDO's collateral pool. To the extent the collateral manager meets his goal of earning returns in excess of borrowing costs (the promised return to CDO investors), these excess returns are paid to the CDO manager and the equity tranche.

The CDO structure typically is to issue a floating-rate senior tranche that is 70%–80% of the total and a smaller mezzanine tranche that pays a fixed rate of interest. If the securities in the collateral pool pay a fixed rate of interest, the collateral manager may enter into an interest rate swap that pays a floating rate of interest in exchange for a fixed rate of interest in order to make the collateral yield more closely match the funding costs in an environment of changing interest rates. The term *arbitrage CDO* is used for CDOs structured to earn returns from the spread between funding costs and portfolio returns.

The collateral manager may use interest earned on portfolio securities, cash from maturing portfolio securities, and cash from the sale of portfolio securities to cover the promised payments to holders of the CDOs senior and mezzanine bonds.

LOS 45.j: Describe characteristics and risks of covered bonds and how they differ from other asset-backed securities.

Covered bonds are similar to asset-backed securities, but the underlying assets (the cover pool), although segregated, remain on the balance sheet of the issuing corporation (i.e., no SPE is created). Covered bonds are issued primarily by European and Asian financial institutions.

Special legislation protects the assets in the cover pool in the event of firm insolvency (they are bankruptcy remote). In contrast to an SPE structure, however, covered bonds also provide bondholders with recourse to the issuing firm. This increases their credit quality, which results in covered bonds generally having lower yields than comparable ABS.

Unlike an ABS, in which the pool of assets is fixed at issuance, a covered bond requires the issuer to replace or augment nonperforming or prepaid assets in the cover pool so that it always provides for the covered bond's promised interest and principal payments. Covered bonds typically are not structured with credit tranching.

Covered bonds may have different provisions in case their issuer defaults. A **hard-bullet covered bond** is in default if the issuer fails to make a scheduled payment. A **soft-bullet covered bond** may postpone the originally scheduled maturity date by as much as a year. A **conditional pass-through covered bond** converts to a pass-through bond on the maturity date if any payments remain due.



MODULE QUIZ 45.2

1. The risk that mortgage prepayments will occur more slowly than expected is *best* characterized as:
 - A. default risk.
 - B. extension risk.
 - C. contraction risk.

2. For investors in commercial mortgage-backed securities, balloon risk in commercial mortgages results in:
 - A. call risk.
 - B. extension risk.
 - C. contraction risk.
3. During the lockout period of a credit card ABS:
 - A. no new receivables are added to the pool.
 - B. investors do not receive interest payments.
 - C. investors do not receive principal payments.
4. A debt security that is collateralized by a pool of the sovereign debt of several developing countries is *most likely*:
 - A. a CMBS.
 - B. a CDO.
 - C. a CMO.
5. A covered bond is *most likely* to feature:
 - A. a fixed cover pool.
 - B. recourse to the issuer.
 - C. a special purpose entity.

KEY CONCEPTS

LOS 45.a

The primary benefits of the securitization of financial assets are:

- Reduce the funding costs for firms selling the financial assets to the securitizing entity.
- Increase the liquidity of the underlying financial assets.

LOS 45.b

Parties to a securitization are a seller of financial assets, a special purpose entity (SPE), and a servicer.

- The seller is the firm that is raising funds through the securitization.
- An SPE is an entity independent of the seller. The SPE buys financial assets from the seller and issues asset-backed securities (ABS) supported by these financial assets.
- The servicer carries out collections and other responsibilities related to the financial assets. The servicer may be the same entity as the seller but does not have to be.

The SPE may issue a single class of ABS or multiple classes with different priorities of claims to cash flows from the pool of financial assets.

LOS 45.c

Asset-backed securities (ABS) can be a single class of securities or multiple classes with differing claims to the cash flows from the underlying assets. Time tranching refers to classes that receive the principal payments from underlying securities sequentially as each prior tranche is repaid in full. With credit tranching, any credit losses are first absorbed by the tranche with the lowest priority, and after that by any other subordinated tranches, in order. Some structures have both time tranching and credit tranching.

LOS 45.d

Characteristics of residential mortgage loans include:

- Maturity.
- Interest rate: fixed-rate, adjustable-rate, or convertible.
- Amortization: full, partial, or interest-only.
- Prepayment penalties.
- Foreclosure provisions: recourse or nonrecourse.

The loan-to-value (LTV) ratio indicates the percentage of the value of the real estate collateral that is loaned. Lower LTVs indicate less credit risk.

LOS 45.e

Agency residential mortgage-backed securities (RMBS) are guaranteed and issued by GNMA, Fannie Mae, or Freddie Mac. Mortgages that back agency RMBS must be conforming loans that meet certain minimum credit quality standards. Nonagency RMBS are issued by private companies and may be backed by nonconforming mortgages.

Key characteristics of RMBS include:

- Pass-through rate, the coupon rate on the RMBS.
- Weighted average maturity (WAM) and weighted average coupon (WAC) of the underlying pool of mortgages.
- Conditional prepayment rate (CPR), which may be compared to the Public Securities Administration (PSA) benchmark for expected prepayment rates.

Nonagency RMBS typically include credit enhancement. External credit enhancement is a third-party guarantee. Internal credit enhancement includes reserve funds (cash or excess spread), overcollateralization, and senior/subordinated structures.

Collateralized mortgage obligations (CMOs) are collateralized by pools of residential MBS. CMOs are structured with tranches that have different exposures to prepayment risks.

In a sequential-pay CMO, all scheduled principal payments and prepayments are paid to each tranche in sequence until that tranche is paid off. The first tranche to be paid principal has the most contraction risk and the last tranche to be paid principal has the most extension risk.

A planned amortization class (PAC) CMO has PAC tranches that receive predictable cash flows as long as the prepayment rate remains within a predetermined range, and support tranches that have more contraction risk and more extension risk than the PAC tranches.

LOS 45.f

Prepayment risk refers to uncertainty about the timing of the principal cash flows from an ABS. Contraction risk is the risk that loan principal will be repaid more rapidly than expected, typically when interest rates have decreased. Extension risk is the risk that loan principal will be repaid more slowly than expected, typically when interest rates have increased.

LOS 45.g

Commercial mortgage-backed securities (CMBS) are backed by mortgages on income-producing real estate properties. Because commercial mortgages are nonrecourse loans, analysis of CMBS focuses on credit risk of the properties. CMBS are structured in tranches with credit losses absorbed by the lowest priority tranches in sequence.

Call (prepayment) protection in CMBS includes loan-level call protection such as prepayment lockout periods, defeasance, prepayment penalty points, and yield maintenance charges, and CMBS-level call protection provided by the lower-priority tranches.

LOS 45.h

Asset-backed securities may be backed by financial assets other than mortgages. Two examples are auto loan ABS and credit card ABS.

Auto loan ABS are backed by automobile loans, which are typically fully amortizing but with shorter maturities than residential mortgages. Prepayments result when autos are sold or traded in, stolen or wrecked and paid off from insurance proceeds, refinanced, or paid off from the borrower's excess cash.

Credit card ABS are backed by credit card receivables, which are revolving debt (nonamortizing). Credit card ABS typically have a lockout period during which only interest is paid to investors and principal payments on the receivables are used to purchase additional receivables.

LOS 45.i

Collateralized debt obligations (CDOs) are structured securities backed by a pool of debt obligations that is managed by a collateral manager. CDOs include:

- Collateralized bond obligations (CBOs) backed by corporate and emerging market debt.
- Collateralized loan obligations (CLOs) backed by leveraged bank loans.
- Structured finance CDOs backed by residential or commercial MBS, ABS, or other CLOs.
- Synthetic CDOs backed by credit default swaps on structured securities.

LOS 45.j

Covered bonds are similar to asset-backed securities, but instead of creating an SPE, the underlying assets remain on the balance sheet of the issuer. Covered bonds give bondholders recourse to the issuer as well as the asset pool, which increases the bonds' credit quality.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 45.1

1. **A** Banks that securitize loans they hold as assets receive cash with which they can make additional loans. The primary benefits of securitization to the economy include reducing firms' funding costs and increasing the liquidity of the financial assets that are securitized. (LOS 45.a)
2. **A** ABS are issued by a special purpose entity (SPE), which is an entity created for that specific purpose. In a securitization, the firm that is securitizing financial assets is described as the seller because it sells the assets to the SPE. The servicer is the entity that deals with collections on the securitized assets. (LOS 45.b)
3. **B** Senior and subordinated tranches are characteristics of a mortgage-backed security with credit tranching. (LOS 45.c)

4. **C** An interest-only lifetime mortgage includes no repayment of principal in its monthly payments so the balloon payment at maturity is equal to the original loan principal. A fully amortizing mortgage has no balloon payment at maturity. A convertible mortgage gives the borrower an option to change the loan from fixed-rate to adjustable-rate or from adjustable-rate to fixed-rate. (LOS 45.d)
5. **A** Conforming loans that may be securitized in agency RMBS have a *maximum* loan-to-value ratio, along with other requirements such as minimum percentage down payments and insurance on the mortgaged property. (LOS 45.e)
6. **B** Issuing CMOs may allow the issuer to raise funds at a lower cost by creating tranches that appeal to investors with different preferences for extension risk and contraction risk. CMOs do not reduce these risks compared to their pool of collateral; they only distribute the risks among the various CMO tranches. (LOS 45.e)

Module Quiz 45.2

1. **B** Extension risk is the risk that prepayments will be slower than expected. Contraction risk is the risk that prepayments will be faster than expected. (LOS 45.e, 45.f)
2. **B** Balloon risk is the possibility that a commercial mortgage borrower will not be able to refinance the principal that is due at the maturity date of the mortgage. This results in a default that is typically resolved by extending the term of the loan during a workout period. Thus, balloon risk is a source of extension risk for CMBS investors. (LOS 45.g)
3. **C** During the lockout period on a credit card receivables-backed ABS, no principal payments are made to investors. (LOS 45.h)
4. **B** A collateralized debt obligation (CDO) is backed by an underlying pool of debt securities, which may include emerging markets debt. Both collateralized mortgage obligations and commercial mortgage-backed securities are backed by mortgages only. (LOS 45.i)
5. **B** Covered bonds differ from ABS in that bondholders have recourse to the issuer as well as the cover pool. Covered bonds are not issued through special purpose entities. A covered bond issuer must maintain a dynamic cover pool, replacing any nonperforming or prepaid assets. (LOS 45.j)

READING 46

UNDERSTANDING FIXED-INCOME RISK AND RETURN

EXAM FOCUS

“Risk” in the title of this reading refers primarily to risk arising from uncertainty about future interest rates. Measurement of credit risk is addressed in the next reading. That said, there is a significant amount of testable material covered in this review. Calculations required by the learning outcomes include the sources of bond returns, three duration measures, money duration, the price value of a basis point, and approximate convexity. You must also be able to estimate a bond’s price change for a given change in yield based on its duration and convexity. Important concepts include how bond characteristics affect interest rate risk, factors that affect a bond’s reinvestment risk, and the interaction among price risk, reinvestment risk, and the investment horizon.

MODULE 46.1: SOURCES OF RETURNS, DURATION



Video covering this content is available online.

LOS 46.a: Calculate and interpret the sources of return from investing in a fixed-rate bond.

There are **three sources of returns** from investing in a fixed-rate bond:

1. Coupon and principal payments.
2. Interest earned on coupon payments that are reinvested over the investor’s holding period for the bond.
3. Any capital gain or loss if the bond is sold prior to maturity.

We will assume that a bond makes all of its promised coupon and principal payments on time (i.e., we are not addressing credit risk). Additionally, we assume that the *interest rate earned on reinvested coupon payments is the same as the YTM on the bond*. There are five results to gain from the analysis presented here.

Given the assumptions just listed:

1. An investor who holds a fixed-rate bond to maturity will earn an annualized rate of return equal to the YTM of the bond when purchased.
2. An investor who sells a bond prior to maturity will earn a rate of return equal to the YTM at purchase if the YTM at sale has not changed since purchase.

3. If the market YTM for the bond, our assumed reinvestment rate, increases (decreases) after the bond is purchased but before the first coupon date, a buy-and-hold investor's realized return will be higher (lower) than the YTM of the bond when purchased.
4. If the market YTM for the bond, our assumed reinvestment rate, *increases* after the bond is purchased but before the first coupon date, a bond investor will earn a rate of return that is lower than the YTM at bond purchase if the bond is held for a *short* period.
5. If the market YTM for the bond, our assumed reinvestment rate, *decreases* after the bond is purchased but before the first coupon date, a bond investor will earn a rate of return that is lower than the YTM at bond purchase if the bond is held for a *long* period.

We will present mathematical examples to demonstrate each of these results as well as some intuition as to why these results must hold.

A bond investor's **annualized holding period rate of return** is calculated as the compound annual return earned from the bond over the investor's holding period. This is the compound rate of return that, based on the purchase price of the bond, would provide an amount at the time of the sale or maturity of the bond equal to the sum of coupon payments, sale or maturity value, and interest earned on reinvested coupons.

We will illustrate this calculation (and the first result listed earlier) with a 6% annual-pay three-year bond purchased at a YTM of 7% and held to maturity.

With an annual YTM of 7%, the bond's purchase price is \$973.76.

$$N = 3; I/Y = 7; PMT = 60; FV = 1,000; CPT \rightarrow PV = -973.76$$

At maturity, the investor will have received coupon income and reinvestment income equal to the future value of an annuity of three \$60 coupon payments calculated with an interest rate equal to the bond's YTM. This amount is:

$$60(1.07)^2 + 60(1.07) + 60 = \$192.89$$

$$N = 3; I/Y = 7; PV = 0; PMT = 60; CPT \rightarrow FV = -192.89$$

We can easily calculate the amount earned from reinvestment of the coupons as:

$$192.89 - 3(60) = \$12.89$$

Adding the maturity value of \$1,000 to \$192.89, we can calculate the investor's rate of return over the three-year holding period as $\left(\frac{1,192.89}{973.76}\right)^{\frac{1}{3}} - 1 = 7\%$ and demonstrate that \$973.76 invested at a compound annual rate of 7% would return \$1,192.89 after three years.

We can calculate an investor's rate of return on the same bond purchased at a YTM of 5%.

Price at purchase:

$$N = 3; I/Y = 5; FV = 1,000; PMT = 60; CPT \rightarrow PV = -1,027.23$$

Coupons and reinvestment income:

$$60(1.05)^2 + 60(1.05) + 60 = \$189.15 \text{ or}$$

$$N = 3; I/Y = 5; PV = 0; PMT = 60; CPT \rightarrow FV = -189.15$$

Holding period return:

$$\left(\frac{1,189.15}{1,027.23}\right)^{\frac{1}{3}} - 1 = 5\%$$

With these examples, we have demonstrated our first result: that for a fixed-rate bond that does not default and has a reinvestment rate equal to the YTM, an investor who holds the bond until maturity will earn a rate of return equal to the YTM at purchase, regardless of whether the bond is purchased at a discount or a premium.

The intuition is straightforward. If the bond is selling at a discount, the YTM is greater than the coupon rate because together, the amortization of the discount and the higher assumed reinvestment rate on coupon income increase the bond's return. For a bond purchased at a premium, the YTM is less than the coupon rate because both the amortization of the premium and the reduction in interest earned on reinvestment of its cash flows decrease the bond's return.

Now let's examine the second result—that an investor who sells a bond prior to maturity will earn a rate of return equal to the YTM as long as the YTM has not changed since purchase. For such an investor, we call the time the bond will be held the investor's **investment horizon**. The value of a bond that is sold at a discount or premium to par will move to the par value of the bond by the maturity date. At dates between the purchase and the sale, the value of a bond at the same YTM as when it was purchased is its **carrying value** and reflects the amortization of the discount or premium since the bond was purchased.



PROFESSOR'S NOTE

Carrying value is a price along a bond's constant-yield price trajectory. We applied this concept in Financial Statement Analysis when we used the effective interest method to calculate the carrying value of a bond liability.

Capital gains or losses at the time a bond is sold are measured relative to this carrying value, as illustrated in the following example.

EXAMPLE: Capital gain or loss on a bond

An investor purchases a 20-year bond with a 5% semiannual coupon and a yield to maturity of 6%. Five years later the investor sells the bond for a price of 91.40. Determine whether the investor realizes a capital gain or loss, and calculate its amount.

Answer:

Any capital gain or loss is based on the bond's carrying value at the time of sale, when it has 15 years (30 semiannual periods) to maturity. The carrying value is calculated using the bond's YTM at the time the investor purchased it.

$$N = 30; I/Y = 3; PMT = 2.5; FV = 100; CPT \rightarrow PV = -90.20$$

Because the selling price of 91.40 is greater than the carrying value of 90.20, the investor realizes a capital gain of $91.40 - 90.20 = 1.20$ per 100 of face value.

Bonds held to maturity have no capital gain or loss. Bonds sold prior to maturity at the same YTM as at purchase will also have no capital gain or loss. Using the 6% three-year bond from

our earlier examples, we can demonstrate this for an investor with a two-year holding period (investment horizon).

When the bond is purchased at a YTM of 7% (for \$973.76), we have:

Price at sale (at end of year 2, YTM = 7%):

$$1,060 / 1.07 = 990.65 \text{ or}$$

$$N = 1; I/Y = 7; FV = 1,000; PMT = 60; CPT \rightarrow PV = -990.65$$

which is the carrying value of the bond.

Coupon interest and reinvestment income for two years:

$$60(1.07) + 60 = \$124.20 \text{ or}$$

$$N = 2; I/Y = 7; PV = 0; PMT = 60; CPT \rightarrow FV = -124.20$$

Investor's annual compound rate of return over the two-year holding period is:

$$\left(\frac{124.20 + 990.65}{973.76} \right)^{\frac{1}{2}} - 1 = 7\%$$

This result can be demonstrated for the case where the bond is purchased at a YTM of 5% (\$1,027.23) as well:

Price at sale (at end of year 2, YTM = 5%):

$$1,060 / 1.05 = 1,009.52 \text{ or}$$

$$N = 1; I/Y = 5; FV = 1,000; PMT = 60; CPT \rightarrow PV = -1,009.52$$

which is the carrying value of the bond.

Coupon interest and reinvestment income for two years:

$$60(1.05) + 60 = 123.00 \text{ or}$$

$$N = 2; I/Y = 5; PV = 0; PMT = 60; CPT \rightarrow FV = -123.00$$

Investor's annual compound rate of return over the two-year holding period is:

$$\left(\frac{123.00 + 1,009.52}{1,027.23} \right)^{\frac{1}{2}} - 1 = 5\%$$

For a bond investor with an investment horizon less than the bond's term to maturity, the annual holding period return will be equal to the YTM at purchase (under our assumptions), if the bond is sold at that YTM. The intuition here is that if a bond will have a rate of return equal to its YTM at maturity, which we showed, if we sell some of the remaining value of the bond discounted at that YTM, we will have earned that YTM up to the date of sale.

Now let's examine our third result: that if rates rise (fall) before the first coupon date, an investor who holds a bond to maturity will earn a rate of return greater (less) than the YTM at purchase.

Based on our previous result that an investor who holds a bond to maturity will earn a rate of return equal to the YTM at purchase if the reinvestment rate is also equal to the YTM at

purchase, the intuition of the third result is straightforward. If the YTM, which is also the reinvestment rate for the bond, increases (decreases) after purchase, the return from coupon payments and reinvestment income will increase (decrease) as a result and increase (decrease) the investor's rate of return on the bond above (below) its YTM at purchase. The following calculations demonstrate these results for the three-year 6% bond in our previous examples.

For a three-year 6% bond purchased at par (YTM of 6%), first assume that the YTM and reinvestment rate increases to 7% after purchase but before the first coupon payment date. The bond's annualized holding period return is calculated as:

Coupons and reinvestment interest:

$$60(1.07)^2 + 60(1.07) + 60 = \$192.89$$

$$N = 3; I/Y = 7; PV = 0; PMT = 60; CPT \rightarrow FV = -192.89$$

Investor's annual compound holding period return:

$$\left(\frac{1,192.89}{1,000}\right)^{\frac{1}{3}} - 1 = 6.06\%$$

which is greater than the 6% YTM at purchase.

If the YTM decreases to 5% after purchase but before the first coupon date, we have the following.

Coupons and reinvestment interest:

$$60(1.05)^2 + 60(1.05) + 60 = \$189.15$$

$$N = 3; I/Y = 5; PV = 0; PMT = 60; CPT \rightarrow FV = -189.15$$

Investor's annual compound holding period return:

$$\left(\frac{1,189.15}{1,000}\right)^{\frac{1}{3}} - 1 = 5.94\%$$

which is less than the 6% YTM at purchase.

Note that in both cases, the investor's rate of return is between the YTM at purchase and the assumed reinvestment rate (the new YTM).

We now turn our attention to the fourth and fifth results concerning the effects of the length of an investor's holding period on the rate of return for a bond that experiences an increase or decrease in its YTM before the first coupon date.

We have already demonstrated that when the YTM increases (decreases) after purchase but before the first coupon date, an investor who holds the bond to maturity will earn a rate of return greater (less) than the YTM at purchase. Now, we examine the rate of return earned by an investor with an investment horizon (expected holding period) less than the term to maturity under the same circumstances.

Consider a three-year 6% bond purchased at par by an investor with a one-year investment horizon. If the YTM increases from 6% to 7% after purchase and the bond is sold after one year, the rate of return can be calculated as follows.

Bond price just after first coupon has been paid with YTM = 7%:

$$N = 2; I/Y = 7; FV = 1,000; PMT = 60; CPT \rightarrow PV = -981.92$$

There is no reinvestment income and only one coupon of \$60 received so the holding period rate of return is simply:

$$\left(\frac{981.92 + 60}{1,000} \right) - 1 = 4.19\%$$

which is less than the YTM at purchase.

If the YTM *decreases* to 5% after purchase and the bond is sold at the end of one year, the investor's rate of return can be calculated as follows.

Bond price just after first coupon has been paid with YTM = 5%:

$$N = 2; I/Y = 5; FV = 1,000; PMT = 60; CPT \rightarrow PV = -1,018.59$$

And the holding period rate of return is simply:

$$\left(\frac{1,018.59 + 60}{1,000} \right) - 1 = 7.86\%$$

which is greater than the YTM at purchase.

The intuition of this result is based on the idea of a tradeoff between **market price risk** (the uncertainty about price due to uncertainty about market YTM) and **reinvestment risk** (uncertainty about the total of coupon payments and reinvestment income on those payments due to the uncertainty about future reinvestment rates).

Previously, we showed that for a bond held to maturity, the investor's rate of return increased with an increase in the bond's YTM and decreased with a decrease in the bond's YTM. For an investor who intends to hold a bond to maturity, there is no interest rate risk as we have defined it. Assuming no default, the bond's value at maturity is its par value regardless of interest rate changes so that the investor has only reinvestment risk. Her realized return will increase when interest earned on reinvested cash flows increases, and decrease when the reinvestment rate decreases.

For an investor with a short investment horizon, interest rate risk increases and reinvestment risk decreases. For the investor with a one-year investment horizon, there was no reinvestment risk because the bond was sold before any interest on coupon payments was earned. The investor had only market price risk so an increase in yield decreased the rate of return over the one-year holding period because the sale price is lower. Conversely, a decrease in yield increased the one-year holding period return to more than the YTM at purchase because the sale price is higher.

To summarize:

short investment horizon: market price risk > reinvestment risk

long investment horizon: reinvestment risk > market price risk

LOS 46.b: Define, calculate, and interpret Macaulay, modified, and effective durations.

Macaulay Duration

Duration is used as a measure of a bond's interest rate risk or sensitivity of a bond's *full* price to a change in its yield. The measure was first introduced by Frederick Macaulay and his formulation is referred to as **Macaulay duration**.

A bond's (annual) Macaulay duration is calculated as the weighted average of the number of years until each of the bond's promised cash flows is to be paid, where the weights are the present values of each cash flow as a percentage of the bond's full value.

Consider a newly issued three-year 4% annual-pay bond with a yield to maturity of 5%. The present values of each of the bond's promised payments, discounted at 5%, and their weights in the calculation of Macaulay duration, are shown in the following table.

$C_1 = 40$	$PV_1 = 40 / 1.05$	= 38.10	$W_1 = 38.10 / 972.77$	= 0.0392
$C_2 = 40$	$PV_2 = 40 / 1.05^2$	= 36.28	$W_2 = 36.28 / 972.77$	= 0.0373
$C_3 = 1,040$	$PV_3 = 1,040 / 1.05^3$	= 898.39 972.77	$W_3 = 898.39 / 972.77$	= 0.9235 1.0000

Note that the present values of all the promised cash flows sum to 972.77 (the full value of the bond) and the weights sum to 1.

Now that we have the weights, and because we know the time until each promised payment is to be made, we can calculate the Macaulay duration for this bond:

$$0.0392(1) + 0.0373(2) + 0.9235(3) = 2.884 \text{ years}$$

The Macaulay duration of a semiannual-pay bond can be calculated in the same way: as a weighted average of the number of *semiannual periods* until the cash flows are to be received. In this case, the result is the number of semiannual periods rather than years.

Because of the improved measures of interest rate risk described next, we say that Macaulay duration is the weighted-average time to the receipt of principal and interest payments, rather than our best estimate of interest rate sensitivity. Between coupon dates, the Macaulay duration of a coupon bond decreases with the passage of time and then goes back up significantly at each coupon payment date.

Modified Duration

Modified duration (ModDur) is calculated as Macaulay duration (MacDur) divided by one plus the bond's yield to maturity. For the bond in our earlier example, we have:

$$\text{ModDur} = 2.884 / 1.05 = 2.747$$

Modified duration provides an approximate percentage change in a bond's price for a 1% change in yield to maturity. The price change for a given change in yield to maturity can be calculated as:

$$\text{approximate percentage change in bond price} = -\text{ModDur} \times \Delta \text{YTM}$$

Based on a ModDur of 2.747, the price of the bond should fall by approximately $2.747 \times 0.1\% = 0.2747\%$ in response to a 0.1% increase in YTM. The resulting price estimate of \$970.098 is very close to the value of the bond calculated directly using a YTM of 5.1%, which is \$970.100.

For an annual-pay bond, the general form of modified duration is:

$$\text{ModDur} = \text{MacDur} / (1 + \text{YTM})$$

For a semiannual-pay bond with a YTM quoted on a semiannual bond basis:

$$\text{ModDur}_{\text{SEMI}} = \text{MacDur}_{\text{SEMI}} / (1 + \text{YTM} / 2)$$

This modified duration can be annualized (from semiannual periods to annual periods) by dividing by two, and then used as the approximate change in price for a 1% change in a bond's YTM.

Approximate Modified Duration

We can approximate modified duration directly using bond values for an increase in YTM and for a decrease in YTM of the same size.

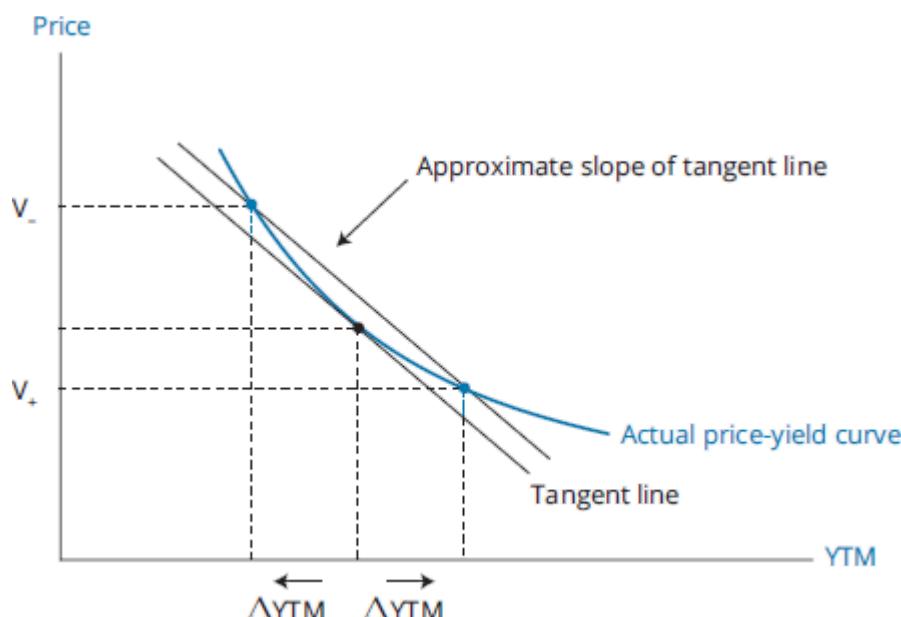
In Figure 46.1 we illustrate this method. The calculation of approximate modified duration is based on a given change in YTM. V_- is the price of the bond if YTM is *decreased* by ΔYTM and V_+ is the price of the bond if the YTM is *increased* by ΔYTM . Note that $V_- > V_+$. Because of the convexity of the price-yield relationship, the price increase (to V_-), for a given decrease in yield, is larger than the price decrease (to V_+).

$$\text{approximate modified duration} = \frac{V_- - V_+}{2 \times V_0 \times \Delta\text{YTM}}$$

The formula uses the average of the magnitudes of the price increase and the price decrease, which is why $V_- - V_+$ (in the numerator) is divided by 2 (in the denominator).

V_0 , the current price of the bond, is in the denominator to convert this average price change to a percentage, and the ΔYTM term is in the denominator to scale the duration measure to a 1% change in yield by convention. Note that the ΔYTM term in the denominator must be entered as a decimal (rather than in a whole percentage) to properly scale the duration estimate.

Figure 46.1: Approximate Modified Duration



EXAMPLE: Calculating approximate modified duration

A bond is trading at a full price of 980. If its yield to maturity increases by 50 basis points, its price will decrease to 960. If its yield to maturity decreases by 50 basis points, its price will increase to 1,002. Calculate the approximate modified duration.

Answer:

The approximate modified duration is $\frac{1,002 - 960}{2 \times 980 \times 0.005} = 4.29$, and the approximate change in price for a 1% change in YTM is 4.29%.

Note that modified duration is a *linear estimate* of the relation between a bond's price and YTM, whereas the actual relation is convex, not linear. This means that the modified duration measure provides good estimates of bond prices for small changes in yield, but increasingly poor estimates for larger changes in yield as the effect of the curvature of the price-yield curve is more pronounced.

Effective Duration

So far, all of our duration measures have been calculated using the YTM and prices of straight (option-free) bonds. This is straightforward because both the future cash flows and their timing are known with certainty. This is not the case with bonds that have embedded options, such as a callable bond or a mortgage-backed bond.

We say mortgage-backed bonds have a *prepayment option*, which is similar to a call option on a corporate bond. The borrowers (people who take out mortgages) typically have the option to pay off the principal value of their loans, in whole or in part, at any time. These prepayments accelerate when interest rates fall significantly because borrowers can refinance their home loans at a lower rate and pay off the remaining principal owed on an existing loan.

Thus, the pricing of bonds with embedded put, call, or prepayment options begins with the benchmark yield curve, not simply the current YTM of the bond. The appropriate measure of interest rate sensitivity for these bonds is **effective duration**.

The calculation of effective duration is the same as the calculation of approximate modified duration with the change in YTM, Δy , replaced by Δcurve , the change in the benchmark yield curve used with a bond pricing model to generate V_- and V_+ . The formula for calculating effective duration is:

$$\text{effective duration} = \frac{V_- - V_+}{2 \times V_0 \times \Delta\text{curve}}$$

Another difference between calculating effective duration and the methods we have discussed so far is that the effects of changes in benchmark yields and changes in the yield spread for credit and liquidity risk are separated. Modified duration makes no distinction between changes in the benchmark yield and changes in the spread. Effective duration reflects only the sensitivity of the bond's value to changes in the benchmark yield curve. Changes in the credit spread are sometimes addressed with a separate "credit duration" measure.

Finally, note that unlike modified duration, effective duration does not necessarily provide better estimates of bond prices for smaller changes in yield. It may be the case that larger changes in yield produce more predictable prepayments or calls than small changes.

LOS 46.c: Explain why effective duration is the most appropriate measure of interest rate risk for bonds with embedded options.

For bonds with embedded options, the future cash flows depend not only on future interest rates but also on the path that interest rates take over time (did they fall to a new level or rise to that level?). We must use effective duration to estimate the interest rate risk of these bonds. The effective duration measure must also be based on bond prices from a pricing model. The fact that bonds with embedded options have uncertain future cash flows means that our present value calculations for bond value based on YTM cannot be used.



MODULE QUIZ 46.1

1. The largest component of returns for a 7-year zero-coupon bond yielding 8% and held to maturity is:
 - A. capital gains.
 - B. interest income.
 - C. reinvestment income.
2. An investor buys a 10-year bond with a 6.5% annual coupon and a YTM of 6%. Before the first coupon payment is made, the YTM for the bond decreases to 5.5%. Assuming coupon payments are reinvested at the YTM, the investor's return when the bond is held to maturity is:
 - A. less than 6.0%.
 - B. equal to 6.0%.
 - C. greater than 6.0%.
3. Assuming coupon interest is reinvested at a bond's YTM, what is the interest portion of an 18-year, \$1,000 par, 5% annual coupon bond's return if it is purchased at par and held to maturity?
 - A. \$576.95.
 - B. \$1,406.62.
 - C. \$1,476.95.
4. An investor buys a 15-year, £800,000, zero-coupon bond with an annual YTM of 7.3%. If she sells the bond after three years for £346,333 she will have:
 - A. a capital gain.
 - B. a capital loss.
 - C. neither a capital gain nor a capital loss.
5. A 14% annual-pay coupon bond has six years to maturity. The bond is currently trading at par. Using a 25 basis point change in yield, the approximate modified duration of the bond is closest to:
 - A. 0.392.
 - B. 3.888.
 - C. 3.970.
6. Which of the following measures is *lowest* for a callable bond?
 - A. Macaulay duration.
 - B. Effective duration.
 - C. Modified duration.
7. Effective duration is more appropriate than modified duration for estimating interest rate risk for bonds with embedded options because these bonds:
 - A. tend to have greater credit risk than option-free bonds.
 - B. exhibit high convexity that makes modified duration less accurate.
 - C. have uncertain cash flows that depend on the path of interest rate changes.

MODULE 46.2: INTEREST RATE RISK AND MONEY DURATION



LOS 46.d: Define key rate duration and describe the use of key rate durations in measuring the sensitivity of bonds to changes in the shape of the benchmark yield curve.

Video covering this content is available online.

Recall that duration is an adequate measure of bond price risk only for parallel shifts in the yield curve. The impact of nonparallel shifts can be measured using a concept known as **key rate duration**. A key rate duration, also known as a **partial duration**, is defined as the sensitivity of the value of a bond or portfolio to changes in the spot rate for a specific maturity, holding other spot rates constant. A bond or portfolio will have a key rate duration for each maturity range on the spot rate curve. The sum of a bond's key rate durations equals its effective duration.

Key rate duration is particularly useful for measuring the effect of a nonparallel shift in the yield curve on a bond portfolio. We can use the key rate duration for each maturity to compute the effect on the portfolio of the interest rate change at that maturity. The effect on the overall portfolio is the sum of these individual effects.

LOS 46.e: Explain how a bond's maturity, coupon, and yield level affect its interest rate risk.

Other things equal, an *increase in a bond's maturity* will (usually) increase its interest rate risk. The present values of payments made further in the future are more sensitive to changes in the discount rate used to calculate present value than are the present values of payments made sooner.

We must say "usually" because there are instances where an increase in a discount coupon bond's maturity will decrease its Macaulay duration. For a discount bond, duration first increases with longer maturity and then decreases over a range of relatively long maturities until it approaches the duration of a perpetuity, which is $(1 + YTM) / YTM$.

Other things equal, an *increase in the coupon rate* of a bond will decrease its interest rate risk. For a given maturity and YTM, the duration of a zero-coupon bond will be greater than that of a coupon bond. Increasing the coupon rate means more of a bond's value will be from payments received sooner so that the value of the bond will be less sensitive to changes in yield.

Other things equal, an *increase (decrease) in a bond's YTM* will decrease (increase) its interest rate risk. To understand this, we can look to the convexity of the price-yield curve and use its slope as our proxy for interest rate risk. At lower yields, the price-yield curve has a steeper slope indicating that price is more sensitive to a given change in yield.

Adding either a put or a call provision will decrease a straight bond's interest rate risk as measured by effective duration. With a call provision, the value of the call increases as yields fall, so a decrease in yield will have less effect on the price of the bond, which is the price of a straight bond minus the value of the call option held by the issuer. With a put option, the bondholder's option to sell the bond back to the issuer at a set price reduces the negative impact of yield increases on price.

LOS 46.f: Calculate the duration of a portfolio and explain the limitations of portfolio duration.

There are two approaches to estimating the duration of a portfolio. The first is to calculate the weighted average number of periods until the portfolio's cash flows will be received. The second approach is to take a weighted average of the durations of the individual bonds in the portfolio.

The first approach is theoretically correct but not often used in practice. The yield measure for calculating portfolio duration with this approach is the **cash flow yield**, the IRR of the bond portfolio. This is inconsistent with duration capturing the relationship between YTM and price. This approach will not work for a portfolio that contains bonds with embedded options because the future cash flows are not known with certainty and depend on interest rate movements.

The second approach is typically used in practice. Using the durations of individual portfolio bonds makes it possible to calculate the duration for a portfolio that contains bonds with embedded options by using their effective durations. The weights for the calculation of portfolio duration under this approach are simply the full price of each bond as a proportion of the total portfolio value (using full prices). These proportions of total portfolio value are multiplied by the corresponding bond durations to get portfolio duration.

$$\text{portfolio duration} = W_1 D_1 + W_2 D_2 + \dots + W_N D_N$$

where:

W_i = full price of bond i divided by the total value of the portfolio

D_i = the duration of bond i

N = the number of bonds in the portfolio

One limitation of this approach is that for portfolio duration to "make sense" the YTM of every bond in the portfolio must change by the same amount. Only with this assumption of a **parallel shift** in the yield curve is portfolio duration calculated with this approach consistent with the idea of the percentage change in portfolio value per 1% change in YTM.

We can think of the second approach as a practical approximation of the theoretically correct duration that the first approach describes. This approximation is less accurate when there is greater variation in yields among portfolio bonds, but is the same as the portfolio duration under the first approach when the yield curve is flat.

LOS 46.g: Calculate and interpret the money duration of a bond and price value of a basis point (PVBP).

The **money duration** of a bond position (also called *dollar duration*) is expressed in currency units.

$$\text{money duration} = \text{annual modified duration} \times \text{full price of bond position}$$

Money duration is sometimes expressed as money duration per 100 of bond par value.

$$\text{money duration per 100 units of par value} = \text{annual modified duration} \times \text{full bond price per 100 of par value}$$

Multiplying the money duration of a bond times a given change in YTM (as a decimal) will provide the change in bond value for that change in YTM.

EXAMPLE: Money duration

1. Calculate the money duration on a coupon date of a \$2 million par value bond that has a modified duration of 7.42 and a full price of 101.32, expressed for the whole bond and per \$100 of face value.
2. What will be the impact on the value of the bond of a 25 basis points increase in its YTM?

Answer:

1. The money duration for the bond is modified duration times the full value of the bond:

$$7.42 \times \$2,000,000 \times 101.32\% = \$15,035,888$$

The money duration per \$100 of par value is:

$$7.42 \times 101.32 = \$751.79$$

$$\text{Or, } \$15,035,888 / (\$2,000,000 / \$100) = \$751.79$$

2. $\$15,035,888 \times 0.0025 = \$37,589.72$

The bond value decreases by \$37,589.72.

The **price value of a basis point** (PVBP) is the money change in the full price of a bond when its YTM changes by one basis point, or 0.01%. We can calculate the PVBP directly for a bond by calculating the average of the decrease in the full value of a bond when its YTM increases by one basis point and the increase in the full value of the bond when its YTM decreases by one basis point.

EXAMPLE: Calculating the price value of a basis point

A newly issued, 20-year, 6% annual-pay straight bond is priced at 101.39. Calculate the price value of a basis point for this bond assuming it has a par value of \$1 million.

Answer:

First we need to find the YTM of the bond:

$$N = 20; PV = -101.39; PMT = 6; FV = 100; CPT \rightarrow I/Y = 5.88$$

Now we need the values for the bond with YTMs of 5.89 and 5.87.

$$I/Y = 5.89; CPT \rightarrow PV = -101.273 (V_+)$$

$$I/Y = 5.87; CPT \rightarrow PV = -101.507 (V_-)$$

$$\text{PVBP (per \$100 of par value)} = (101.507 - 101.273) / 2 = 0.117$$

For the \$1 million par value bond, each 1 basis point change in the yield to maturity will change the bond's price by $0.117 \times \$1 \text{ million} \times 0.01 = \$1,170$.

MODULE QUIZ 46.2



1. A bond portfolio manager who wants to estimate the sensitivity of the portfolio's value to changes in the 5-year spot rate should use:
 - a key rate duration.
 - a Macaulay duration.
 - an effective duration.
2. Which of the following three bonds (similar except for yield and maturity) has the *least* Macaulay duration? A bond with:

- A. 5% yield and 10-year maturity.
 - B. 5% yield and 20-year maturity.
 - C. 6% yield and 10-year maturity.
3. Portfolio duration has limited usefulness as a measure of interest rate risk for a portfolio because it:
- A. assumes yield changes uniformly across all maturities.
 - B. cannot be applied if the portfolio includes bonds with embedded options.
 - C. is accurate only if the portfolio's internal rate of return is equal to its cash flow yield.
4. The current price of a \$1,000, 7-year, 5.5% semiannual coupon bond is \$1,029.23. The bond's price value of a basis point is *closest* to:
- A. \$0.05.
 - B. \$0.60.
 - C. \$5.74.

MODULE 46.3: CONVEXITY AND YIELD VOLATILITY

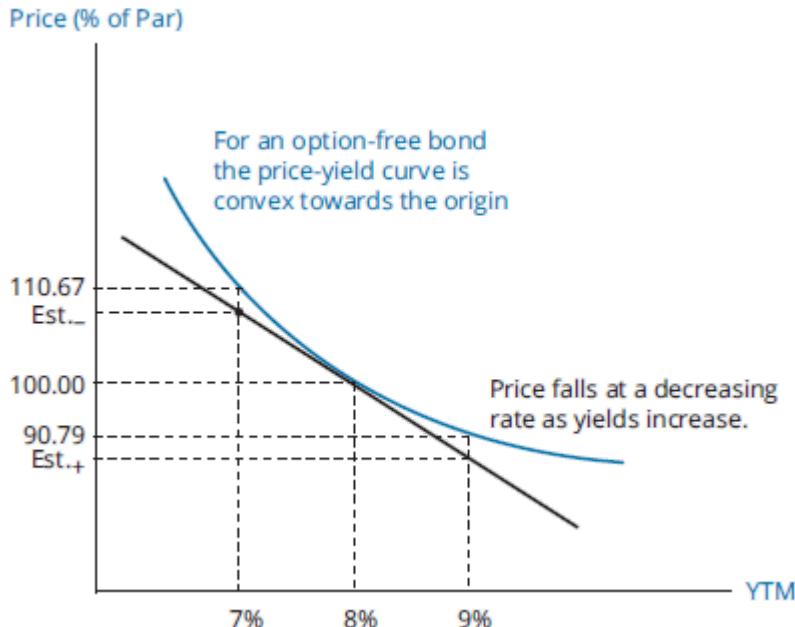


Video covering
this content is
available online.

LOS 46.h: Calculate and interpret approximate convexity and compare approximate and effective convexity.

Earlier we explained that modified duration is a linear approximation of the relationship between yield and price and that, because of the convexity of the true price-yield relation, duration-based estimates of a bond's full price for a given change in YTM will be increasingly different from actual prices. This is illustrated in Figure 46.2. Duration-based price estimates for a decrease and for an increase in YTM are shown as Est.₋ and Est.₊.

Figure 46.2: Price-Yield Curve for an Option-Free, 8%, 20-Year Bond



Estimates of the price impact of a change in yield based only on modified duration can be improved by introducing a second term based on the bond's convexity. **Convexity** is a measure

of the curvature of the price-yield relation. The more curved it is, the greater the convexity adjustment to a duration-based estimate of the change in price for a given change in YTM.

A bond's convexity can be estimated as:

$$\text{approximate convexity} = \frac{V_- + V_+ - 2V_0}{(\Delta \text{YTM})^2 V_0}$$

where:

the variables are the same as those we used in calculating approximate modified duration

Effective convexity, like effective duration, must be used for bonds with embedded options.

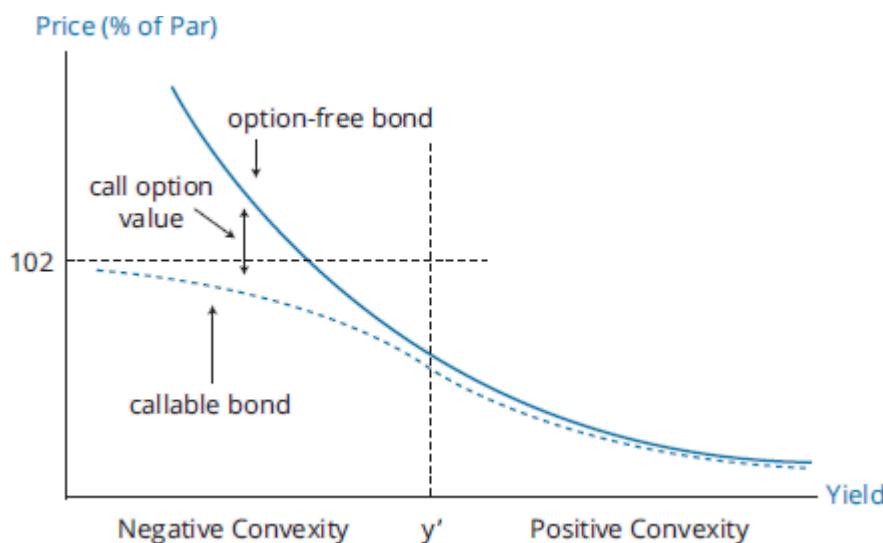
The calculation of effective convexity is the same as the calculation of approximate convexity, except that the change in the yield *curve*, rather than a change in the bond's YTM, is used.

$$\text{approximate effective convexity} = \frac{V_- + V_+ - 2V_0}{(\Delta \text{curve})^2 V_0}$$

A bond's convexity is increased or decreased by the same bond characteristics that affect duration. A longer maturity, a lower coupon rate, or a lower yield to maturity will all increase convexity, and vice versa. For two bonds with equal duration, the one with cash flows that are more dispersed over time will have the greater convexity.

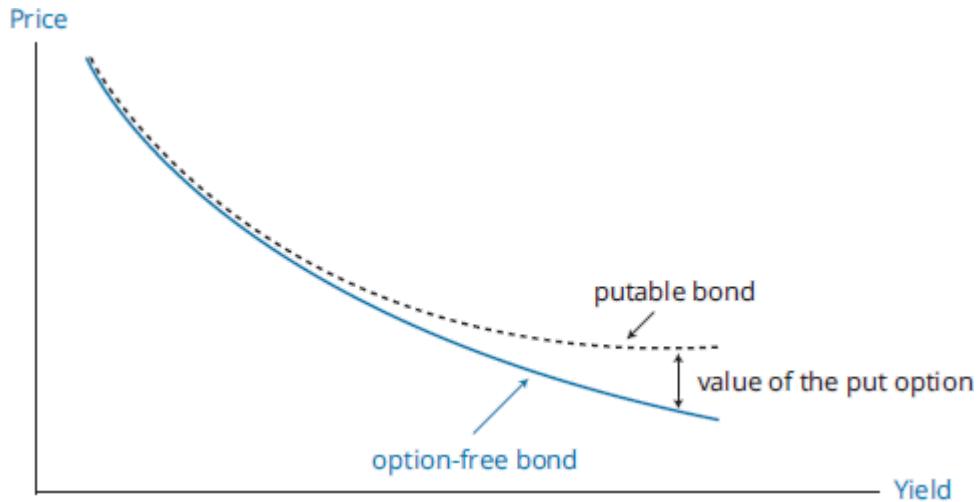
While the convexity of any option-free bond is positive, the convexity of a callable bond can be negative at low yields. This is because at low yields the call option becomes more valuable and the call price puts an effective limit on increases in bond value as shown in Figure 46.3. For a bond with negative convexity, the price increase from a decrease in YTM is *smaller* than the price decrease from an increase in YTM.

Figure 46.3: Price-Yield Function of a Callable vs. an Option-Free Bond



A putable bond has greater convexity than an otherwise identical option-free bond. In Figure 46.4 we illustrate the price-yield relation for a putable bond. At higher yields, the put becomes more valuable so that the value of the putable bond falls less than that of an option-free bond as yield increases.

Figure 46.4: Comparing the Price-Yield Curves for Option-Free and Putable Bonds



LOS 46.i: Calculate the percentage price change of a bond for a specified change in yield, given the bond's approximate duration and convexity.

By taking account of both a bond's duration (first-order effects) and convexity (second-order effects), we can improve an estimate of the effects of a change in yield on a bond's value, especially for larger changes in yield.

$$\begin{aligned} \text{change in full bond price} &= -\text{annual modified duration}(\Delta YTM) \\ &+ \frac{1}{2} \text{annual convexity}(\Delta YTM)^2 \end{aligned}$$

EXAMPLE: Estimating price changes with duration and convexity

Consider an 8% bond with a full price of \$908 and a YTM of 9%. Estimate the percentage change in the full price of the bond for a 30 basis point increase in YTM assuming the bond's duration is 9.42 and its convexity is 68.33.

Answer:

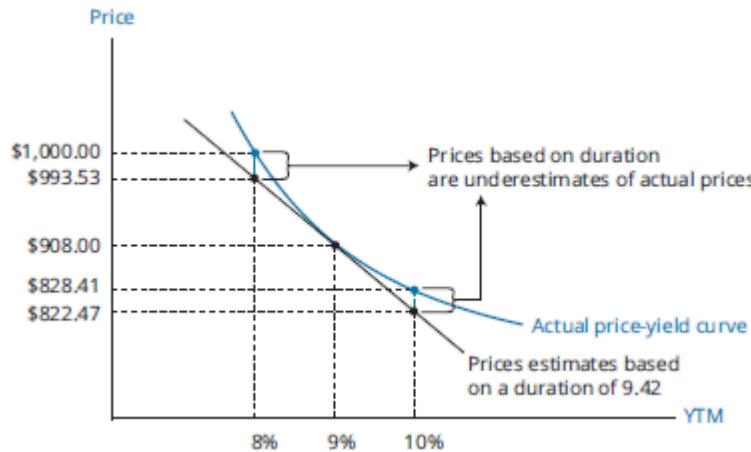
The duration effect is $-9.42 \times 0.003 = 0.02826 = -2.826\%$.

The convexity effect is $0.5 \times 68.33 \times (0.003)^2 = 0.000307 = 0.0307\%$.

The expected change in bond price is $(-0.02826 + 0.000307) = -2.7953\%$.

Note that the convexity adjustment to the price change is the same for both an increase and a decrease in yield. As illustrated in Figure 46.5, the duration-only based estimate of the increase in price resulting from a decrease in yield is too low for a bond with positive convexity, and is improved by a positive adjustment for convexity. The duration-only based estimate of the decrease in price resulting from an increase in yield is larger than the actual decrease, so it's also improved by a positive adjustment for convexity.

Figure 46.5: Duration-Based Price Estimates vs. Actual Bond Prices



LOS 46.j: Describe how the term structure of yield volatility affects the interest rate risk of a bond.

The **term structure of yield volatility** refers to the relation between the volatility of bond yields and their times to maturity. We have seen that the sensitivity of a bond's price with respect to a *given* change in yield depends on its duration and convexity. From an investor's point of view, it's the volatility of a bond's price that is of concern. The volatility of a bond's price has two components: the sensitivity of the bond's price to a given change in yield and the volatility of the bond's yield.

In calculating duration and convexity, we implicitly assumed that the yield curve shifted in a parallel manner. In practice, this is often not the case. For example, changes in monetary policy may have more of an effect on short-term interest rates than on longer-term rates.

It could be the case that a shorter-term bond has more price volatility than a longer-term bond with a greater duration because of the greater volatility of the shorter-term yield.

LOS 46.k: Describe the relationships among a bond's holding period return, its duration, and the investment horizon.

Macaulay duration has an interesting application in matching a bond to an investor's investment horizon. When the investment horizon and the bond's Macaulay duration are matched, a parallel shift in the yield curve prior to the first coupon payment will not (or will minimally) affect the investor's horizon return.

Earlier, we illustrated the effect of a change in yield that occurs prior to the first coupon payment. Our results showed that for an investor with a short investment horizon (anticipated holding period), the market price risk of the bond outweighs its reinvestment risk. Because of this, an increase in yield prior to the first coupon date was shown to reduce the horizon yield for a short investment horizon and increase the horizon yield for a longer-term investment horizon. For a longer investment horizon, the increase in reinvestment income from the yield increase was greater than the decrease in the sale price of the bond.

For a decrease in yield, an investor with a short investment horizon will have a capital gain and only a small decrease in reinvestment income. An investor with a long horizon will be more

affected by the decrease in reinvestment income and will have a horizon return that is less than the bond's original yield.

When the investment horizon just matches the Macaulay duration, the effect of a change in YTM on the sale price of a bond and on reinvestment income just offset each other. We can say that for such an investment, market price risk and reinvestment risk offset each other. The following example illustrates this result.

EXAMPLE: Investment horizon yields

Consider an eight-year, 8.5% bond priced at 89.52 to yield 10.5% to maturity. The Macaulay duration of the bond is 6. We can calculate the horizon yield for horizons of 3 years, 6 years, and 8 years, assuming the YTM falls to 9.5% prior to the first coupon date.

Answer:

Sale after 3 years

Bond price:

$$N = 5; PMT = 8.5; FV = 100; I/Y = 9.5; CPT \rightarrow PV = 96.16$$

Coupons and interest on reinvested coupons:

$$N = 3; PMT = 8.5; PV = 0; I/Y = 9.5; CPT \rightarrow FV = 28.00$$

Horizon return:

$$[(96.16 + 28.00) / 89.52]^{1/3} - 1 = 11.520\%$$

Sale after 6 years

Bond price:

$$N = 2; PMT = 8.5; FV = 100; I/Y = 9.5; CPT \rightarrow PV = 98.25$$

Coupons and interest on reinvested coupons:

$$N = 6; PMT = 8.5; PV = 0; I/Y = 9.5; CPT \rightarrow FV = 64.76$$

Horizon return:

$$[(98.25 + 64.76) / 89.52]^{1/6} - 1 = 10.505\%$$

Held to maturity, 8 years

Maturity value = 100

Coupons and interest on reinvested coupons:

$$N = 8; PMT = 8.5; PV = 0; I/Y = 9.5; CPT \rightarrow FV = 95.46$$

Horizon return:

$$[(100 + 95.46) / 89.52]^{1/8} - 1 = 10.253\%$$

For an investment horizon equal to the bond's Macaulay duration of 6, the horizon return is equal to the original YTM of 10.5%. For a shorter three-year investment horizon, the price increase from a reduction in the YTM to 9.5% dominates the decrease in reinvestment income so the horizon return, 11.520%, is greater than the original YTM. For an investor who holds the bond to maturity, there is no price effect and the decrease in reinvestment income reduces the horizon return to 10.253%, less than the original YTM.

The difference between a bond's Macaulay duration and the bondholder's investment horizon is referred to as a **duration gap**. A positive duration gap (Macaulay duration greater than the

investment horizon) exposes the investor to market price risk from increasing interest rates. A negative duration gap (Macaulay duration less than the investment horizon) exposes the investor to reinvestment risk from decreasing interest rates.

LOS 46.l: Explain how changes in credit spread and liquidity affect yield-to-maturity of a bond and how duration and convexity can be used to estimate the price effect of the changes.

The benchmark yield curve's interest rates have two components; the real rate of return and expected inflation. A bond's spread to the benchmark curve also has two components, a premium for credit risk and a premium for lack of liquidity relative to the benchmark securities.

Because we are treating the yields associated with each component as additive, a given increase or decrease in any of these components of yield will increase or decrease the bond's YTM by the same amount.

With a direct relationship between a bond's yield spread to the benchmark yield curve and its YTM, we can estimate the impact on a bond's value of a change in spread using the formula we introduced earlier for the price effects of a given change in YTM.

$$\% \Delta \text{ bond value} = -\text{duration}(\Delta \text{spread}) + \frac{1}{2} \text{convexity}(\Delta \text{spread})^2$$

EXAMPLE: Price effect of spread changes

Consider a bond that is valued at \$180,000 that has a duration of 8 and a convexity of 22. The bond's spread to the benchmark curve increases by 25 basis points due to a credit downgrade. What is the approximate change in the bond's market value?

Answer:

With $\Delta \text{spread} = 0.0025$ we have:

$$(-8 \times 0.0025) + (0.5 \times 22 \times 0.0025^2) = -1.99\% \text{ and the bond's value will fall by approximately } 1.99\% \times 180,000 = \$3,588.$$

LOS 46.m: Describe the difference between empirical duration and analytical duration.

The duration measures we have introduced in this reading, based on mathematical analysis, are often referred to as **analytical durations**. A different approach is to estimate **empirical durations** using the historical relationship between benchmark yield changes and bond price changes.

When we estimate corporate bond durations based on a shift in the benchmark (government) yield curve, we implicitly assume that the credit spread for the corporate bond remains unchanged (i.e., changes in the benchmark yield curve and a bond's yield spread are uncorrelated). When this assumption is not justified, estimates of empirical duration, based on the actual relationship between changes in the benchmark yield curve and bond values, may be more appropriate.

An example of such a situation is an increase in market uncertainty during which investor demand shifts sharply toward bonds with low credit risk (a “flight to quality”). Yields on government bonds decrease, but credit spreads increase at the same time. As a result, government bond prices increase but corporate bond prices increase by less or possibly not at all. For a corporate bond portfolio, an estimate of empirical duration that accounts for this effect would be lower (i.e., less price response to a decrease in benchmark yields) than an estimate of analytical duration would indicate. An analytical estimate of the duration of a portfolio consisting primarily of government debt securities, in this case, would still be appropriate, while an empirically derived estimate of duration would be more appropriate for a portfolio comprising corporate bonds (risky credits).



MODULE QUIZ 46.3

1. A bond has a convexity of 114.6. The convexity effect, if the yield decreases by 110 basis points, is *closest* to:
 - A. -1.673%.
 - B. +0.693%.
 - C. +1.673%.
2. The modified duration of a bond is 7.87. The approximate percentage change in price using duration only for a yield decrease of 110 basis points is *closest* to:
 - A. -8.657%.
 - B. +7.155%.
 - C. +8.657%.
3. Assume a bond has an effective duration of 10.5 and a convexity of 97.3. Using both of these measures, the estimated percentage change in price for this bond, in response to a decline in yield of 200 basis points, is *closest* to:
 - A. 19.05%.
 - B. 22.95%.
 - C. 24.89%.
4. Two bonds are similar in all respects except maturity. Can the shorter-maturity bond have greater interest rate risk than the longer-term bond?
 - A. No, because the shorter-maturity bond will have a lower duration.
 - B. Yes, because the shorter-maturity bond may have a higher duration.
 - C. Yes, because short-term yields can be more volatile than long-term yields.
5. An investor with an investment horizon of six years buys a bond with a modified duration of 6.0. This investment has:
 - A. no duration gap.
 - B. a positive duration gap.
 - C. a negative duration gap.
6. Which of the following *most accurately* describes the relationship between liquidity and yield spreads relative to benchmark government bond rates? All else being equal, bonds with:
 - A. less liquidity have lower yield spreads.
 - B. greater liquidity have higher yield spreads.
 - C. less liquidity have higher yield spreads.

KEY CONCEPTS

LOS 46.a

Sources of return from a bond investment include:

- Coupon and principal payments.
- Reinvestment of coupon payments.

- Capital gain or loss if bond is sold before maturity.

Changes in yield to maturity produce market price risk (uncertainty about a bond's price) and reinvestment risk (uncertainty about income from reinvesting coupon payments). An increase (a decrease) in YTM decreases (increases) a bond's price but increases (decreases) its reinvestment income.

LOS 46.b

Macaulay duration is the weighted average number of coupon periods until a bond's scheduled cash flows.

Modified duration is a linear estimate of the percentage change in a bond's price that would result from a 1% change in its YTM.

$$\text{approximate modified duration} = \frac{V_- - V_+}{2 \times V_0 \times \Delta \text{YTM}}$$

Effective duration is a linear estimate of the percentage change in a bond's price that would result from a 1% change in the benchmark yield curve.

$$\text{effective duration} = \frac{V_- - V_+}{2 \times V_0 \times \Delta \text{curve}}$$

LOS 46.c

Effective duration is the appropriate measure of interest rate risk for bonds with embedded options because changes in interest rates may change their future cash flows. Pricing models are used to determine the prices that would result from a given size change in the benchmark yield curve.

LOS 46.d

Key rate duration is a measure of the price sensitivity of a bond or a bond portfolio to a change in the spot rate for a specific maturity. We can use the key rate durations of a bond or portfolio to estimate its price sensitivity to changes in the shape of the yield curve.

LOS 46.e

Holding other factors constant:

- Duration increases when maturity increases.
- Duration decreases when the coupon rate increases.
- Duration decreases when YTM increases.

LOS 46.f

There are two methods for calculating portfolio duration:

- Calculate the weighted average number of periods until cash flows will be received using the portfolio's IRR (its cash flow yield). This method is better theoretically but cannot be used for bonds with options.
- Calculate the weighted average of durations of bonds in the portfolio (the method most often used). Portfolio duration is the percentage change in portfolio value for a 1% change in yield, only for parallel shifts of the yield curve.

LOS 46.g

Money duration is stated in currency units and is sometimes expressed per 100 of bond value.

money duration = annual modified duration × full price of bond position

money duration per 100 units of par value =
annual modified duration × full bond price per 100 of par value

The price value of a basis point is the change in the value of a bond, expressed in currency units, for a change in YTM of one basis point, or 0.01%.

PVBP = $[(V_- - V_+) / 2] \times \text{par value} \times 0.01$

LOS 46.h

Convexity refers to the curvature of a bond's price-yield relationship.

$$\text{approximate convexity} = \frac{V_- + V_+ - 2V_0}{(\Delta \text{YTM})^2 V_0}$$

Effective convexity is appropriate for bonds with embedded options:

$$\text{approximate effective convexity} = \frac{V_- + V_+ - 2V_0}{(\Delta \text{curve})^2 V_0}$$

LOS 46.i

Given values for approximate annual modified duration and approximate annual convexity, the percentage change in the full price of a bond can be estimated as:

$$\begin{aligned}\% \Delta \text{ full bond price} &= -\text{annual modified duration}(\Delta \text{YTM}) \\ &+ \frac{1}{2} \text{annual convexity}(\Delta \text{YTM})^2\end{aligned}$$

LOS 46.j

The term structure of yield volatility refers to the relationship between maturity and yield volatility. Short-term yields may be more volatile than long-term yields. As a result, a short-term bond may have more price volatility than a longer-term bond with a higher duration.

LOS 46.k

Over a short investment horizon, a change in YTM affects market price more than it affects reinvestment income.

Over a long investment horizon, a change in YTM affects reinvestment income more than it affects market price.

Macaulay duration may be interpreted as the investment horizon for which a bond's market price risk and reinvestment risk just offset each other.

duration gap = Macaulay duration – investment horizon

LOS 46.l

A bond's yield spread to the benchmark curve includes a premium for credit risk and a premium for illiquidity.

Given values for duration and convexity, the effect on the value of a bond from a given change in its yield spread (Δspread) can be estimated as:

$$\% \Delta \text{ bond value} = -\text{duration}(\Delta \text{spread}) + 1 - 2 \text{ convexity}(\Delta \text{spread})^2$$

LOS 46.m

Macaulay, modified, and effective duration are examples of analytical duration. Empirical duration is estimated from historical data using models. Empirical duration may differ from analytical duration in interest rate environments where the assumptions underlying analytical duration may not hold, such as for credit-risky bonds in a flight-to-quality scenario.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 46.1

1. **B** The increase in value of a zero-coupon bond over its life is interest income. A zero-coupon bond has no reinvestment risk over its life. A bond held to maturity has no capital gain or loss. (LOS 46.a)
2. **A** The decrease in the YTM to 5.5% will decrease the reinvestment income over the life of the bond so that the investor will earn less than 6%, the YTM at purchase. (LOS 46.a)
3. **B** The interest portion of a bond's return is the sum of the coupon payments and interest earned from reinvesting coupon payments over the holding period.

$$N = 18; PMT = 50; PV = 0; I/Y = 5%; CPT \rightarrow FV = -1,406.62$$

(LOS 46.a)

4. **A** The price of the bond after three years that will generate neither a capital gain nor a capital loss is the price if the YTM remains at 7.3%. After three years, the present value of the bond is $800,000 / 1.073^{12} = 343,473.57$, so she will have a capital gain relative to the bond's carrying value. (LOS 46.a)

5. **B** $V_- = 100.979$

$$N = 6; PMT = 14.00; FV = 100; I/Y = 13.75; CPT \rightarrow PV = -100.979$$

$$V_+ = 99.035$$

$$I/Y = 14.25; CPT \rightarrow PV = -99.035V_0 = 100.000$$

$$\Delta y = 0.0025$$

$$\text{Approximate modified duration} = \frac{V_- - V_+}{2V_0 \Delta \text{YTM}} = \frac{100.979 - 99.035}{2(100)(0.0025)} = 3.888$$

(LOS 46.b)

6. **B** The interest rate sensitivity of a bond with an embedded call option will be less than that of an option-free bond. Effective duration takes the effect of the call option into account and will, therefore, be less than Macaulay or modified duration. (LOS 46.b)
7. **C** Because bonds with embedded options have cash flows that are uncertain and depend on future interest rates, effective duration must be used. (LOS 46.c)

Module Quiz 46.2

1. **A** Key rate duration refers to the sensitivity of a bond or portfolio value to a change in one specific spot rate. (LOS 46.d)
2. **C** Other things equal, Macaulay duration is less when yield is higher and when maturity is shorter. The bond with the highest yield and shortest maturity must have the lowest Macaulay duration. (LOS 46.e)
3. **A** Portfolio duration is limited as a measure of interest rate risk because it assumes parallel shifts in the yield curve; that is, the discount rate at each maturity changes by the same amount. Portfolio duration can be calculated using effective durations of bonds with embedded options. By definition, a portfolio's internal rate of return is equal to its cash flow yield. (LOS 46.f)

4. **B** PVBP = initial price – price if yield is changed by 1 basis point.

First, we need to calculate the yield so we can calculate the price of the bond with a 1 basis point change in yield. Using a financial calculator: PV = -1,029.23; FV = 1,000; PMT = 27.5 = $(0.055 \times 1,000) / 2$; N = 14 = 2×7 years; CPT → I/Y = 2.49998, multiplied by 2 = 4.99995, or 5.00%.

Next, compute the price of the bond at a yield of 5.00% + 0.01%, or 5.01%. Using the calculator: FV = 1,000; PMT = 27.5; N = 14; I/Y = 2.505 (5.01 / 2); CPT → PV = \$1,028.63.

Finally, PVBP = \$1,029.23 – \$1,028.63 = \$0.60. (LOS 46.g)

Module Quiz 46.3

1. **B** Convexity effect = $1 / 2 \times \text{convexity} \times (\Delta \text{YTM})^2 = (0.5)(114.6)(0.011)^2 = 0.00693 = 0.693\%$
(LOS 46.h)
2. **C** $-7.87 \times (-1.10\%) = 8.657\%$
(LOS 46.i)
3. **B** Total estimated price change = (duration effect + convexity effect) $\{[-10.5 \times (-0.02)] + [1 / 2 \times 97.3 \times (-0.02)^2]\} \times 100 = 21.0\% + 1.95\% = 22.95\%$
(LOS 46.i)
4. **C** In addition to its sensitivity to changes in yield (i.e., duration), a bond's interest rate risk includes the volatility of yields. A shorter-maturity bond may have more interest rate risk than an otherwise similar longer-maturity bond if short-term yields are more volatile than long-term yields. (LOS 46.j)
5. **B** Duration gap is Macaulay duration minus the investment horizon. Because modified duration equals Macaulay duration / (1 + YTM), Macaulay duration is greater than modified duration for any YTM greater than zero. Therefore, this bond has a Macaulay duration greater than six years and the investment has a positive duration gap. (LOS 46.k)
6. **C** The less liquidity a bond has, the higher its yield spread relative to its benchmark. This is because investors require a higher yield to compensate them for giving up liquidity. (LOS 46.l)

READING 47

FUNDAMENTALS OF CREDIT ANALYSIS

EXAM FOCUS

This reading introduces credit analysis, primarily for corporate bonds, but considerations for credit analysis of high yield, sovereign, and non-sovereign government bonds are also covered. Focus on credit ratings, credit spreads, and the impact on return when ratings and spreads change.

MODULE 47.1: CREDIT RISK AND BOND RATINGS



LOS 47.a: Describe credit risk and credit-related risks affecting corporate bonds.

Video covering this content is available online.

LOS 47.b: Describe default probability and loss severity as components of credit risk.

Credit risk is the risk associated with losses stemming from the failure of a borrower to make timely and full payments of interest or principal. Credit risk has two components: *default risk* and *loss severity*.

- **Default risk** is the probability that a borrower (bond issuer) fails to pay interest or repay principal when due.
- **Loss severity**, or *loss given default*, refers to the value a bond investor will lose if the issuer defaults. Loss severity can be stated as a monetary amount or as a percentage of a bond's value (principal and unpaid interest).

The **expected loss** is equal to the default risk multiplied by the loss severity. Expected loss can likewise be stated as a monetary value or as a percentage of a bond's value.

The **recovery rate** is the percentage of a bond's value an investor will receive if the issuer defaults. Loss severity as a percentage is equal to one minus the recovery rate.

Bonds with credit risk trade at higher yields than bonds thought to be free of credit risk. The difference in yield between a credit-risky bond and a credit-risk-free bond of similar maturity is called its **yield spread**. For example, if a 5-year corporate bond is trading at a spread of +250 basis points to Treasuries and the yield on 5-year Treasury notes is 4.0%, the yield on the corporate bond is $4.0\% + 2.5\% = 6.5\%$.

Bond prices are inversely related to spreads; a wider spread implies a lower bond price and a narrower spread implies a higher price. The size of the spread reflects the creditworthiness of

the issuer and the liquidity of the market for its bonds. **Spread risk** is the possibility that a bond's spread will widen due to one or both of these factors.

- **Credit migration risk or downgrade risk** is the possibility that spreads will increase because the issuer has become less creditworthy. As we will see later in this reading, credit rating agencies assign ratings to bonds and issuers, and may upgrade or downgrade these ratings over time.
 - **Market liquidity risk** is the risk of receiving less than market value when selling a bond and is reflected in the size of the bid-ask spreads. Market liquidity risk is greater for the bonds of less creditworthy issuers and for the bonds of smaller issuers with relatively little publicly traded debt.
-

LOS 47.c: Describe seniority rankings of corporate debt and explain the potential violation of the priority of claims in a bankruptcy proceeding.

Each category of debt from the same issuer is ranked according to a **priority of claims** in the event of a default. A bond's priority of claims to the issuer's assets and cash flows is referred to as its **seniority ranking**.

Debt can be either **secured debt** or **unsecured debt**. Secured debt is backed by collateral, while unsecured debt or *debentures* represent a general claim to the issuer's assets and cash flows. Secured debt has higher priority of claims than unsecured debt.

Secured debt can be further distinguished as *first lien* or *first mortgage* (where a specific asset is pledged), *senior secured*, or *junior secured* debt. Unsecured debt is further divided into *senior*, *junior*, and *subordinated* gradations. The highest rank of unsecured debt is senior unsecured. Subordinated debt ranks below other unsecured debt.

The general seniority rankings for debt repayment priority are the following:

- First lien/senior secured.
- Second lien/secured.
- Senior unsecured.
- Senior subordinated.
- Subordinated.
- Junior subordinated.

All debt within the same category is said to rank **pari passu**, or have same priority of claims. All senior secured debt holders, for example, are treated alike in a corporate bankruptcy.

Recovery rates are highest for debt with the highest priority of claims and decrease with each lower rank of seniority. The lower the seniority ranking of a bond, the higher its credit risk. Investors require a higher yield to accept a lower seniority ranking.

In the event of a default or reorganization, senior lenders have claims on the assets before junior lenders and equity holders. A strict priority of claims, however, is not always applied in practice. Although in theory the priority of claims is absolute, in many cases lower-priority debt holders (and even equity investors) may get paid even if senior debt holders are not paid in full.

Bankruptcies can be costly and take a long time to settle. During bankruptcy proceedings, the value of a company's assets could deteriorate due to loss of customers and key employees, while legal expenses mount. A bankruptcy reorganization plan is confirmed by a vote among all classes of investors with less than 100% recovery rate. To avoid unnecessary delays, negotiation and compromise among various claimholders may result in a reorganization plan that does not strictly conform to the original priority of claims. By such a vote or by order of the bankruptcy court, the final plan may differ from absolute priority.

LOS 47.d: Compare and contrast corporate issuer credit ratings and issue credit ratings and describe the rating agency practice of “notching”.

Credit rating agencies assign ratings to categories of bonds with similar credit risk. Rating agencies rate both the issuer (i.e., the company issuing the bonds) and the debt issues, or the bonds themselves. Issuer credit ratings are called **corporate family ratings** (CFR), while issue-specific ratings are called **corporate credit ratings** (CCR). Issuer ratings are based on the overall creditworthiness of the company. The issuers are rated on their senior unsecured debt.

Figure 47.1 shows ratings scales used by Standard & Poor's, Moody's, and Fitch, three of the major credit rating agencies.

Figure 47.1: Credit Rating Categories

(a) Investment grade ratings		(b) Noninvestment grade ratings	
Moody's	Standard & Poor's, Fitch	Moody's	Standard & Poor's, Fitch
Aaa	AAA	Ba1	BB+
Aa1	AA+	Ba2	BB
Aa2	AA	Ba3	BB-
Aa3	AA-	B1	B+
A1	A+	B2	B
A2	A	B3	B-
A3	A-	Caa1	CCC+
Baa1	BBB+	Caa2	CCC
Baa2	BBB	Caa3	CCC-
Baa3	BBB-	Ca	CC
		C	C
		C	D

Triple A (AAA or Aaa) is the highest rating. Bonds with ratings of Baa3/BBB- or higher are considered **investment grade**. Bonds rated Ba1/BB+ or lower are considered **noninvestment grade** and are often called *high yield bonds* or *junk bonds*.

Bonds in default are rated D by Standard & Poor's and Fitch and are included in Moody's lowest rating category, C. When a company defaults on one of its several outstanding bonds, provisions in bond indentures may trigger default on the remaining issues as well. Such a provision is called a *cross default provision*.

The ratings of a firm's individual bonds can differ from its corporate (issuer) rating. The seniority and covenants (including collateral pledged) of an individual bond issue are the primary determinants of differences between an issuer's rating and the ratings of its individual bond issues. The assignment of individual issue ratings that are higher or lower than that of the issuer is referred to as **notching**.

Another example of a factor that rating agencies consider when notching an issue credit rating is **structural subordination**. In a holding company structure, both the parent company and the subsidiaries may have outstanding debt. A subsidiary's debt covenants may restrict the transfer of cash or assets "upstream" to the parent company before the subsidiary's debt is serviced. In such a case, even though the parent company's bonds are not junior to the subsidiary's bonds, the subsidiary's bonds have a priority claim to the subsidiary's cash flows. Thus the parent company's bonds are effectively subordinated to the subsidiary's bonds.

Notching is less common for highly-rated issuers than for lower-rated issuers. For firms with high overall credit ratings, differences in expected recovery rates among a firm's individual bonds are less important, so their bonds might not be notched at all. For firms with higher probabilities of default (lower ratings), differences in expected recovery rates among a firm's bonds are more significant. For this reason, notching is more likely for issues with lower creditworthiness in general. For a firm with speculative credit, its subordinated debt might be notched two ratings below its issuer rating.

LOS 47.e: Explain risks in relying on ratings from credit rating agencies.

Relying on ratings from credit rating agencies has some risks. Four specific risks are:

1. **Credit ratings are dynamic.** Credit ratings change over time. Rating agencies may update their default risk assessments during the life of a bond. Higher credit ratings tend to be more stable than lower credit ratings.
2. **Rating agencies are not perfect.** Ratings mistakes occur from time to time. During a past period, subprime mortgage securities were assigned much higher ratings than they deserved.
3. **Event risk is difficult to assess.** Risks that are specific to a company or industry are difficult to predict and incorporate into credit ratings. Litigation risk to tobacco companies is one example. Events that are difficult to anticipate, such as natural disasters, acquisitions, and equity buybacks using debt, are not easily captured in credit ratings.
4. **Credit ratings lag market pricing.** Market prices and credit spreads can change much faster than credit ratings. Additionally, two bonds with the same rating can trade at different yields. Market prices reflect expected losses, while credit ratings only assess default risk.

LOS 47.f: Explain the four Cs (Capacity, Collateral, Covenants, and Character) of traditional credit analysis.

A common way to categorize the key components of credit analysis is by the **four Cs of credit analysis**: capacity, collateral, covenants, and character.

Capacity

Capacity refers to a corporate borrower's ability repay its debt obligations on time. Analysis of capacity is similar to the process used in equity analysis. Capacity analysis entails three levels of assessment: (1) industry structure, (2) industry fundamentals, and (3) company fundamentals.

Industry structure

The first level of a credit analyst's assessment is industry structure. Industry structure can be described by Porter's five forces: threat of entry, power of suppliers, power of buyers, threat of substitution, and rivalry among existing competitors.



PROFESSOR'S NOTE

We describe industry analysis based on Porter's five forces in the readings on equity valuation.

Industry fundamentals

The next level of a credit analyst's assessment is industry fundamentals, including the influence of macroeconomic factors on an industry's growth prospects and profitability. Industry fundamentals evaluation focuses on:

- **Industry cyclicalities.** Cyclical industries are sensitive to economic performance. Cyclical industries tend to have more volatile earnings, revenues, and cash flows, which make them more risky than noncyclical industries.
- **Industry growth prospects.** Creditworthiness is most questionable for the weaker companies in a slow-growing or declining industry.
- **Industry published statistics.** Industry statistics provided by rating agencies, investment banks, industry periodicals, and government agencies can be a source for industry performance and fundamentals.

Company fundamentals

The last level of credit analysts' assessment is company fundamentals. A corporate borrower should be assessed on:

- **Competitive position.** Market share changes over time and cost structure relative to peers are some of the factors to analyze.
- **Operating history.** The performance of the company over different phases of business cycle, trends in margins and revenues, and current management's tenure.
- **Management's strategy and execution.** This includes the soundness of the strategy, the ability to execute the strategy, and the effects of management's decisions on bondholders.
- **Ratios and ratio analysis.** As we will discuss later in this reading, leverage and coverage ratios are important tools for credit analysis.

Collateral

Collateral analysis is more important for less creditworthy companies. The market value of a company's assets can be difficult to observe directly. Issues to consider when assessing

collateral values include:

- **Intangible assets.** Patents are considered high-quality intangible assets because they can be more easily sold to generate cash flows than other intangibles. Goodwill is not considered a high-quality intangible asset and is usually written down when company performance is poor.
- **Depreciation.** High depreciation expense relative to capital expenditures may signal that management is not investing sufficiently in the company. The quality of the company's assets may be poor, which may lead to reduced operating cash flow and potentially high loss severity.
- **Equity market capitalization.** A stock that trades below book value may indicate that company assets are of low quality.
- **Human and intellectual capital.** These are difficult to value, but a company may have intellectual property that can function as collateral.

Covenants

Covenants are the terms and conditions the borrowers and lenders agree to as part of a bond issue. Covenants protect lenders while leaving some operating flexibility to the borrowers to run the company. There are two types of covenants: (1) *affirmative covenants* and (2) *negative covenants*.

Affirmative covenants require the borrower to take certain actions, such as using the proceeds for the stated purpose; paying interest, principal, and taxes; carrying insurance on pledged assets; continuing in its current business activity; and following relevant laws and regulations. Affirmative covenants have, basically, administrative purposes.

Negative covenants restrict the borrower from taking certain actions that may reduce the value of the bondholders' claims. While affirmative covenants do not impose significant costs on the issuer (besides making the promised payments), negative covenants constrain the issuer's business activities and may thereby impose significant costs on the issuer. Examples of negative covenants include:

- Restrictions on the payment of dividends and share repurchases, for example, restricting distributions to shareholders to a certain percentage of net income.
- Restrictions on the amount of additional debt the borrower can issue, for example, setting a maximum debt-to-equity ratio or minimum interest coverage ratio.
- Restrictions on issuing any debt with a higher priority than the subject debt issue.
- Restrictions on pledging any collateral that is currently unencumbered as collateral for new borrowing.
- Restrictions on assets sales, for example, limiting asset sales to a certain percentage of total asset value.
- Restrictions on company investment, for example, requiring that a company not invest outside its current primary business activities.
- Restrictions on mergers and acquisitions.

Covenants that are overly restrictive of an issuer's operating activities may reduce the issuer's ability to repay; for example, prohibiting asset sales that could provide the cash to pay bond

interest and principal. On the other hand, covenants create a legally binding contractual framework for repayment of the debt obligation, which reduces uncertainty for the debt holders. A careful credit analysis should include an assessment of whether the covenants protect the interests of the bondholders without unduly constraining the borrower's operating activities.

Character

Character refers to management's integrity and its commitment to repay the loan. Factors such as management's business qualifications and operating record are important for evaluating character. Character analysis includes an assessment of:

- **Soundness of strategy.** Management's ability to develop a sound strategy.
- **Track record.** Management's past performance in executing its strategy and operating the company without bankruptcies, restructurings, or other distress situations that led to additional borrowing.
- **Accounting policies and tax strategies.** Use of accounting policies and tax strategies that may be hiding problems, such as revenue recognition issues, frequent restatements, and frequently changing auditors.
- **Fraud and malfeasance record.** Any record of fraud or other legal and regulatory problems.
- **Prior treatment of bondholders.** Benefits to equity holders at the expense of debt holders, through actions such as debt-financed acquisitions and special dividends, especially if they led to credit rating downgrades.



MODULE QUIZ 47.1

1. The two components of credit risk are:
 - A. default risk and yield spread.
 - B. default risk and loss severity.
 - C. loss severity and yield spread.
2. Expected loss can decrease with an increase in a bond's:
 - A. default risk.
 - B. loss severity.
 - C. recovery rate.
3. Absolute priority of claims in a bankruptcy might be violated because:
 - A. of the pari passu principle.
 - B. creditors negotiate a different outcome.
 - C. available funds must be distributed equally among creditors.
4. "Notching" is *best* described as a difference between:
 - A. an issuer credit rating and an issue credit rating.
 - B. a company credit rating and an industry average credit rating.
 - C. an investment grade credit rating and a noninvestment grade credit rating.
5. Which of the following statements is *least likely* a limitation of relying on ratings from credit rating agencies?
 - A. Credit ratings are dynamic.
 - B. Firm-specific risks are difficult to rate.
 - C. Credit ratings adjust quickly to changes in bond prices.
6. Ratio analysis is *most likely* used to assess a borrower's:
 - A. capacity.
 - B. character.

C. collateral.

MODULE 47.2: EVALUATING CREDIT QUALITY



Video covering
this content is
available online.

LOS 47.g: Calculate and interpret financial ratios used in credit analysis.

LOS 47.h: Evaluate the credit quality of a corporate bond issuer and a bond of that issuer, given key financial ratios of the issuer and the industry.

Ratio analysis is part of capacity analysis. Two primary categories of ratios for credit analysis are *leverage ratios* and *coverage ratios*. Credit analysts calculate company ratios to assess the viability of a company, to find trends over time, and to compare companies to industry averages and peers.

Profits and Cash Flows

Profits and cash flows are needed to service debt. Here we examine four profit and cash flow metrics commonly used in ratio analysis by credit analysts.

1. **Earnings before interest, taxes, depreciation, and amortization (EBITDA).** EBITDA is a commonly used measure that is calculated as operating income plus depreciation and amortization. A drawback to using this measure for credit analysis is that it does not adjust for capital expenditures and changes in working capital, which are necessary uses of funds for a going concern. Cash needed for these uses is not available to debt holders.
2. **Funds from operations (FFO).** Funds from operations are net income from continuing operations plus depreciation, amortization, deferred taxes, and noncash items. FFO is similar to cash flow from operations (CFO) except that FFO excludes changes in working capital.
3. **Free cash flow before dividends.** Free cash flow before dividends is net income plus depreciation and amortization minus capital expenditures minus increase in working capital. Free cash flow before dividends excludes nonrecurring items.
4. **Free cash flow after dividends.** This is free cash flow before dividends minus the dividends. If free cash flow after dividends is greater than zero, it represents cash that could pay down debt or accumulate on the balance sheet. Either outcome is a form of deleveraging, a positive indicator for creditworthiness.

Leverage Ratios

Analysts should adjust debt reported on the financial statements by including the firm's obligations such as underfunded pension plans (net pension liabilities) and off-balance-sheet liabilities such as operating leases.

The most common measures of leverage used by credit analysts are the debt-to-capital ratio, the debt-to-EBITDA ratio, the FFO-to-debt ratio, and the ratio of FCF after dividends to debt.

1. **Debt/capital.** Capital is the sum of total debt and shareholders' equity. The debt-to-capital ratio is the percentage of the capital structure financed by debt. A lower ratio indicates less credit risk. If the financial statements list high values for intangible assets

such as goodwill, an analyst should calculate a second debt-to-capital ratio adjusted for a write-down of these assets' after-tax value.

2. **Debt/EBITDA.** A higher ratio indicates higher leverage and higher credit risk. This ratio is more volatile for firms in cyclical industries or with high operating leverage because of their high variability of EBITDA.
3. **FFO/debt.** Because this ratio divides a cash flow measure by the value of debt, a higher ratio indicates lower credit risk.
4. **FCF after dividends/debt.** Greater values indicate a greater ability to service existing debt.

Coverage Ratios

Coverage ratios measure the borrower's ability to generate cash flows to meet interest payments. The two most commonly used are EBITDA-to-interest and EBIT-to-interest.

1. **EBITDA/interest expense.** A higher ratio indicates lower credit risk. This ratio is used more often than the EBIT-to-interest expense ratio. Because depreciation and amortization are still included as part of the cash flow measure, this ratio will be higher than the EBIT version.
2. **EBIT/interest expense.** A higher ratio indicates lower credit risk. This ratio is the more conservative measure because depreciation and amortization are subtracted from earnings.

Ratings agencies publish benchmark values for financial ratios that are associated with each ratings classification. Credit analysts can evaluate the potential for upgrades and downgrades based on subject company ratios relative to these benchmarks.

EXAMPLE: Credit analysis based on ratios

An analyst is assessing the credit quality of York, Inc. and Zale, Inc., relative to each other and their industry average. Selected financial information appears in the following table.

	York, Inc.	Zale, Inc.	Industry Average
Earnings before interest and taxes	\$550,000	\$2,250,000	\$1,400,000
Funds from operations	\$300,000	\$850,000	\$600,000
Interest expense	\$40,000	\$160,000	\$100,000
Total debt	\$1,900,000	\$2,700,000	\$2,600,000
Total capital	\$4,000,000	\$6,500,000	\$6,000,000

Explain how the analyst should evaluate the relative creditworthiness of York and Zale.

Answer:

Leverage and coverage ratios based on these data are as follows:

EBIT / interest:

$$\text{York: } \$550,000 / \$40,000 = 13.8\times$$

$$\text{Zale: } \$2,250,000 / \$160,000 = 14.1\times$$

$$\text{Industry average: } \$1,400,000 / \$100,000 = 14.0\times$$

Both York and Zale have interest coverage in line with their industry average.

FFO / total debt:

$$\text{York: } \$300,000 / \$1,900,000 = 15.8\%$$

$$\text{Zale: } \$850,000 / \$2,700,000 = 31.5\%$$

$$\text{Industry average: } \$600,000 / \$2,600,000 = 23.1\%$$

Zale's funds from operations relative to its debt level are greater than the industry average, while York is generating less FFO relative to its debt level.

Total debt / total capital:

$$\text{York: } \$1,900,000 / \$4,000,000 = 47.5\%$$

$$\text{Zale: } \$2,700,000 / \$6,500,000 = 41.5\%$$

$$\text{Industry average: } \$2,600,000 / \$6,000,000 = 43.3\%$$

York is more leveraged than Zale and the industry average. Based on these data, Zale appears to be more creditworthy than York.

LOS 47.i: Describe macroeconomic, market, and issuer-specific factors that influence the level and volatility of yield spreads.

We can think of the yield on an option-free corporate bond as the sum of the real risk-free interest rate, the expected inflation rate, a maturity premium, a liquidity premium, and a credit spread. All bond prices and yields are affected by changes in the first three of these components. The last two components are the yield spread:

$$\text{yield spread} = \text{liquidity premium} + \text{credit spread}$$

Yield spreads on corporate bonds are affected primarily by five interrelated factors:

1. **Credit cycle.** The market's perception of overall credit risk is cyclical. At the top of the credit cycle, the bond market perceives low credit risk and is generally bullish. Credit spreads narrow as the credit cycle improves. Credit spreads widen as the credit cycle deteriorates.
2. **Economic conditions.** Credit spreads narrow as the economy strengthens and investors expect firms' credit metrics to improve. Conversely, credit spreads widen as the economy weakens.
3. **Broker-dealer capital.** Because most bonds trade over the counter, investors need broker-dealers to provide market-making capital for bond markets to function. Yield spreads are narrower when broker-dealers provide sufficient capital but can widen when market-making capital becomes scarce.
4. **General market demand and supply.** Credit spreads narrow in times of high demand for bonds. Credit spreads widen in times of low demand for bonds. Excess supply conditions, such as large issuance in a short period of time, can lead to widening spreads.
5. **Issuer's financial performance.** Developments that are positive for the issuer's credit quality will narrow its yield spread, while developments that are negative for the issuer's credit quality will widen its credit spread.

Yield spreads on lower-quality issues tend to be more volatile than spreads on higher-quality issues.

LOS 47.j: Explain special considerations when evaluating the credit of high-yield, sovereign, and non-sovereign government debt issuers and issues.

High Yield Debt

High yield or *noninvestment grade* corporate bonds are rated below Baa3/BBB– by credit rating agencies. These bonds are also called *junk bonds* because of their higher perceived credit risk.

Reasons for noninvestment grade ratings may include:

- High leverage.
- Unproven operating history.
- Low or negative free cash flow.
- High sensitivity to business cycles.
- Low confidence in management.
- Unclear competitive advantages.
- Large off-balance-sheet liabilities.
- Industry in decline.

Because high yield bonds have higher default risk than investment grade bonds, credit analysts must pay more attention to loss severity. Special considerations for high yield bonds include their liquidity, financial projections, debt structure, corporate structure, and covenants.

Liquidity. Liquidity or availability of cash is critical for high yield issuers. High yield issuers have limited access to additional borrowings, and available funds tend to be more expensive for high yield issuers. Bad company-specific news and difficult financial market conditions can quickly dry up the liquidity of debt markets. Many high yield issuers are privately owned and cannot access public equity markets for needed funds.

Analysts focus on six sources of liquidity (in order of reliability):

1. Balance sheet cash.
2. Working capital.
3. Operating cash flow (CFO).
4. Bank credit.
5. Equity issued.
6. Sales of assets.

For a high yield issuer with few or unreliable sources of liquidity, significant amounts of debt coming due within a short time frame may indicate potential default. Running out of cash with no access to external financing to refinance or service existing debt is the primary reason why high yield issuers default. For high yield financial firms that are highly levered and depend on funding long-term assets with short-term liabilities, liquidity is critical.

Financial projections. Projecting future earnings and cash flows, including stress scenarios and accounting for changes in capital expenditures and working capital, are important for revealing potential vulnerabilities to the inability to meet debt payments.

Debt structure. High yield issuers' capital structures often include different types of debt with several levels of seniority and hence varying levels of potential loss severity. Capital structures typically include secured bank debt, second lien debt, senior unsecured debt, subordinated debt, and preferred stock. Some of these, especially subordinated debt, may be convertible to common shares.

A credit analyst will need to calculate leverage for each level of the debt structure when an issuer has multiple layers of debt with a variety of expected recovery rates.

High yield companies for which secured bank debt is a high proportion of the capital structure are said to be *top heavy* and have less capacity to borrow from banks in financially stressful periods. Companies that have top-heavy capital structures are more likely to default and have lower recovery rates for unsecured debt issues.

Corporate structure. Many high-yield companies use a holding company structure. A parent company receives dividends from the earnings of subsidiaries as its primary source of operating income. Because of structural subordination, subsidiaries' dividends paid upstream to a parent company are subordinate to interest payments. These dividends can be insufficient to pay the debt obligations of the parent, thus reducing the recovery rate for debt holders of the parent company.

Despite structural subordination, a parent company's credit rating may be superior to subsidiaries' ratings because the parent can benefit from having access to multiple cash flows from diverse subsidiaries.

Some complex corporate structures have intermediate holding companies that carry their own debt and do not own 100% of their subsidiaries' stock. These companies are typically a result of mergers, acquisitions, or leveraged buyouts.

Default of one subsidiary may not necessarily result in cross default. Analysts need to scrutinize bonds' indentures and other legal documents to fully understand the impact of complex corporate structures. To analyze these companies, analysts should calculate leverage ratios at each level of debt issuance and on a consolidated basis.

Covenants. Important covenants for high yield debt include:

- **Change of control put.** This covenant gives debt holders the right to require the issuer to buy back debt (typically for par value or a value slightly above par) in the event of an acquisition. For investment grade bonds, a change of control put typically applies only if an acquisition of the borrower results in a rating downgrade to below investment grade.
- **Restricted payments.** The covenant protects lenders by limiting the amount of cash that may be paid to equity holders.
- **Limitations on liens.** The covenant limits the amount of secured debt that a borrower can carry. Unsecured debt holders prefer the issuer to have less secured debt, which increases the recovery amount available to them in the event of default.

- **Restricted versus unrestricted subsidiaries.** Issuers can classify subsidiaries as restricted or unrestricted. Restricted subsidiaries' cash flows and assets can be used to service the debt of the parent holding company. This benefits creditors of holding companies because their debt is pari passu with the debt of restricted subsidiaries, rather than structurally subordinated. Restricted subsidiaries are typically the holding company's larger subsidiaries that have significant assets. Tax and regulatory issues can factor into the classification of subsidiary's restriction status. A subsidiary's restriction status is found in the bond indenture.

Bank covenants are often more restrictive than bond covenants, and when covenants are violated, banks can block additional loans until the violation is corrected. If a violation is not remedied, banks can trigger a default by accelerating the full repayment of a loan.

In terms of the factors that affect their return, high yield bonds may be viewed as a hybrid of investment grade bonds and equity. Compared to investment grade bonds, high yield bonds show greater price and spread volatility and are more highly correlated with the equity market.

High yield analysis can include some of the same techniques as equity market analysis, such as enterprise value. **Enterprise value (EV)** is equity market capitalization plus total debt minus excess cash. For high yield companies that are not publicly traded, comparable public company equity data can be used to estimate EV. Enterprise value analysis can indicate a firm's potential for additional leverage, or the potential credit damage that might result from a leveraged buyout. An analyst can compare firms based on the differences between their EV/EBITDA and debt/EBITDA ratios. Firms with a wider difference between these ratios have greater equity relative to their debt and therefore have less credit risk.

Sovereign Debt

Sovereign debt is issued by national governments. Sovereign credit analysis must assess both the government's ability to service debt and its willingness to do so. The assessment of willingness is important because bondholders usually have no legal recourse if a national government refuses to pay its debts.

A basic framework for evaluating and assigning a credit rating to sovereign debt includes five key areas:

1. **Institutional assessment** includes successful policymaking, minimal corruption, checks and balances among institutions, and a culture of honoring debts.
2. **Economic assessment** includes growth trends, income per capita, and diversity of sources for economic growth.
3. **External assessment** includes the country's foreign reserves, its external debt, and the status of its currency in international markets.
4. **Fiscal assessment** includes the government's willingness and ability to increase revenue or cut expenditures to ensure debt service, as well as trends in debt as a percentage of GDP.
5. **Monetary assessment** includes the ability to use monetary policy for domestic economic objectives (this might be lacking with exchange rate targeting or membership in a monetary union) and the credibility and effectiveness of monetary policy.

Credit rating agencies assign each national government two ratings: (1) a local currency debt rating and (2) a foreign currency debt rating. The ratings are assigned separately because

defaults on foreign currency denominated debt have historically exceeded those on local currency debt. Foreign currency debt typically has a higher default rate and a lower credit rating because the government must purchase foreign currency in the open market to make interest and principal payments, which exposes it to the risk of significant local currency depreciation. In contrast, local currency debt can be repaid by raising taxes, controlling domestic spending, or simply printing more money. Ratings can differ as much as two notches for local and foreign currency bonds.

Sovereign defaults can be caused by events such as war, political instability, severe devaluation of the currency, or large declines in the prices of the country's export commodities. Access to debt markets can be difficult for sovereigns in bad economic times.

Non-Sovereign Government Bonds

Non-sovereign government debt is issued by local governments (cities, states, and counties) and quasi-governmental entities. **Municipal bonds** are a significant part of the overall U.S. bond market. Interest payments from municipal bonds are most often exempt from national income taxes. Default rates for municipal bonds are very low relative to general corporate bonds.

Most municipal bonds can be classified as *general obligation bonds* or *revenue bonds*. **General obligation** (GO) bonds are unsecured bonds backed by the full faith credit of the issuing governmental entity, which is to say they are supported by its taxing power. **Revenue bonds** are issued to finance specific projects, such as airports, toll bridges, hospitals, and power generation facilities.

Unlike sovereigns, municipalities cannot use monetary policy to service their debt and usually must balance their operating budgets. Municipal governments' ability to service their general obligation debt depends ultimately on the local economy (i.e., the tax base). Economic factors to assess in evaluating the creditworthiness of GO bonds include employment, trends in per capita income and per capita debt, tax base dimensions (depth, breadth, and stability), demographics, and ability to attract new jobs (location, infrastructure). Credit analysts must also observe revenue variability through economic cycles. Relying on highly variable taxes that are subject to economic cycles, such as capital gains and sales taxes, can signal higher credit risk.

Municipalities may have long-term obligations such as underfunded pensions and post-retirement benefits. Inconsistent reporting requirements for municipalities are also an issue.

Revenue bonds often have higher credit risk than GO bonds because the project is the sole source of funds to service the debt. Analysis of revenue bonds combines analysis of the project, using techniques similar to those for analyzing corporate bonds, with analysis of the financing of the project.



MODULE QUIZ 47.2

1. Higher credit risk is indicated by a higher:
 - A. FFO/debt ratio.
 - B. debt/EBITDA ratio.
 - C. EBITDA/interest expense ratio.
2. Compared to other firms in the same industry, an issuer with a credit rating of AAA should have a lower:
 - A. FFO/debt ratio.
 - B. operating margin.
 - C. debt/capital ratio.

3. Credit spreads tend to widen as:
 - A. the credit cycle improves.
 - B. economic conditions worsen.
 - C. broker-dealers become more willing to provide capital.
4. Compared to shorter duration bonds, longer duration bonds:
 - A. have smaller bid-ask spreads.
 - B. are less sensitive to credit spreads.
 - C. have less certainty regarding future creditworthiness.
5. One key difference between sovereign bonds and municipal bonds is that sovereign issuers:
 - A. can print money.
 - B. have governmental taxing power.
 - C. are affected by economic conditions.

KEY CONCEPTS

LOS 47.a

Credit risk refers to the possibility that a borrower fails to make the scheduled interest payments or return of principal.

Spread risk is the possibility that a bond loses value because its credit spread widens relative to its benchmark. Spread risk includes credit migration or downgrade risk and market liquidity risk.

LOS 47.b

Credit risk is composed of default risk, which is the probability of default, and loss severity, which is the portion of the value of a bond or loan a lender or investor will lose if the borrower defaults. The expected loss is the probability of default multiplied by the loss severity.

LOS 47.c

Corporate debt is ranked by seniority or priority of claims. Secured debt is a direct claim on specific firm assets and has priority over unsecured debt. Secured or unsecured debt may be further ranked as senior or subordinated. Priority of claims may be summarized as follows:

- First lien/senior secured.
- Second lien/secured.
- Senior unsecured.
- Senior subordinated.
- Subordinated.
- Junior subordinated.

LOS 47.d

Issuer credit ratings, or corporate family ratings, reflect a debt issuer's overall creditworthiness and typically apply to a firm's senior unsecured debt.

Issue credit ratings, or corporate credit ratings, reflect the credit risk of a specific debt issue. Notching refers to the practice of adjusting an issue credit rating upward or downward from the issuer credit rating to reflect the seniority, covenants, and possibly the expected recovery in the event of a default of a debt issue.

LOS 47.e

Lenders and bond investors should not rely exclusively on credit ratings from rating agencies for the following reasons:

- Credit ratings can change during the life of a debt issue.
- Rating agencies cannot always judge credit risk accurately.
- Firms are subject to risk of unforeseen events that credit ratings do not reflect.
- Market prices of bonds often adjust more rapidly than credit ratings.

LOS 47.f

Components of traditional credit analysis are known as the four Cs:

- Capacity: The borrower's ability to make timely payments on its debt.
- Collateral: The value of assets pledged against a debt issue or available to creditors if the issuer defaults.
- Covenants: Provisions of a bond issue that protect creditors by requiring or prohibiting actions by an issuer's management.
- Character: Assessment of an issuer's management, strategy, quality of earnings, and past treatment of bondholders.

LOS 47.g

Credit analysts use profitability, cash flow, and leverage and coverage ratios to assess debt issuers' capacity.

- Profitability refers to operating income and operating profit margin, with operating income typically defined as earnings before interest and taxes (EBIT).
- Cash flow may be measured as earnings before interest, taxes, depreciation, and amortization (EBITDA); funds from operations (FFO); free cash flow before dividends; or free cash flow after dividends.
- Leverage ratios include debt-to-capital, debt-to-EBITDA, and FFO-to-debt.
- Coverage ratios include EBIT-to-interest expense and EBITDA-to-interest expense.

LOS 47.h

Lower leverage, higher interest coverage, and greater free cash flow imply lower credit risk and a higher credit rating for a firm. When calculating leverage ratios, analysts should include in a firm's total debt its obligations such as underfunded pensions and off-balance-sheet financing.

For a specific debt issue, secured collateral implies lower credit risk compared to unsecured debt, and higher seniority implies lower credit risk compared to lower seniority.

LOS 47.i

Corporate bond yields comprise the real risk-free rate, expected inflation rate, credit spread, maturity premium, and liquidity premium. An issue's yield spread to its benchmark includes its credit spread and liquidity premium.

The level and volatility of yield spreads are affected by the credit and business cycles, availability of capital from broker-dealers, the supply and demand for debt issues, and the

financial performance of the bond issuer. Yield spreads tend to narrow when the credit cycle is improving, the economy is expanding, and financial markets and investor demand for new debt issues are strong. Yield spreads tend to widen when the credit cycle, the economy, and financial markets are weakening, and in periods when the supply of new debt issues is heavy or broker-dealer capital is insufficient for market making.

LOS 47.j

High yield bonds are more likely to default than investment grade bonds, which increases the importance of estimating loss severity. Analysis of high yield debt should focus on liquidity, projected financial performance, the issuer's corporate and debt structures, and debt covenants.

Credit risk of sovereign debt includes the issuing country's ability and willingness to pay. Ability to pay is greater for debt issued in the country's own currency than for debt issued in a foreign currency. Willingness refers to the possibility that a country refuses to repay its debts.

Analysis of non-sovereign government debt is similar to analysis of sovereign debt, focusing on the strength of the local economy and its effect on tax revenues. Analysis of municipal revenue bonds is similar to analysis of corporate debt, focusing on the ability of a project to generate sufficient revenue to service the bonds.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 47.1

1. **B** Credit risk is composed of default risk and loss severity. Yield spreads reflect the credit risk of a borrower. (LOS 47.a)
2. **C** An increase in the recovery rate means that the loss severity has decreased, which decreases expected loss. (LOS 47.b)
3. **B** A negotiated bankruptcy settlement does not always follow the absolute priority of claims. (LOS 47.c)
4. **A** Notching refers to the credit rating agency practice of distinguishing between the credit rating of an issuer (generally for its senior unsecured debt) and the credit rating of particular debt issues from that issuer, which may differ from the issuer rating because of provisions such as seniority. (LOS 47.d)
5. **C** Bond prices and credit spreads change much faster than credit ratings. (LOS 47.e)
6. **A** Ratio analysis is used to assess a corporate borrower's capacity to repay its debt obligations on time. (LOS 47.f)

Module Quiz 47.2

1. **B** A higher debt/EBITDA ratio is sign of higher leverage and higher credit risk. Higher FFO/debt and EBITDA/interest expense ratios indicate lower credit risk. (LOS 47.g, 47.h)
2. **C** A low debt/capital ratio is an indicator of low leverage. An issuer rated AAA is likely to have a high operating margin and a high FFO/debt ratio compared to its industry group.

(LOS 47.g, 47.h)

3. **B** Credit spreads widen as economic conditions worsen. Spreads narrow as the credit cycle improves and as broker-dealers provide more capital to bond markets. (LOS 47.i)
4. **C** Longer duration bonds usually have longer maturities and carry more uncertainty of future creditworthiness. (LOS 47.i)
5. **A** Sovereign entities can print money to repay debt, while municipal borrowers cannot. Both sovereign and municipal entities have taxing powers, and both are affected by economic conditions. (LOS 47.j)

TOPIC QUIZ: FIXED INCOME

You have now finished the Fixed Income topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Fixed Income. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

READING 48

DERIVATIVE INSTRUMENT AND DERIVATIVE MARKET FEATURES

MODULE 48.1: DERIVATIVES MARKETS

LOS 48.a: Define a derivative and describe basic features of a derivative instrument.

A **derivative** is a security that *derives* its value from the value of another security or a variable (such as an interest rate or stock index value) at some specific future date. The security or variable that determines the value of a derivative security is referred to as the **underlying** for the derivative. The value of a derivative at a point in time is derived from the value of the underlying (asset or variable) on which the derivative contract is based.

A relatively simple example of a derivative is a **forward contract** that specifies the price at which one party agrees to buy or sell an underlying security at a specified future date. Consider a forward contract to buy 100 shares of Acme at \$30 per share three months from now:

- Acme shares are the **underlying asset** for the forward contract.
- \$30 is the **forward price** in the contract.
- The date of the future transaction, when the shares will be exchanged for cash, is referred to as the **settlement date** (maturity date) of the forward contract.
- 100 shares is the **contract size** of the forward contract.
- The forward price is set so that the forward contract has zero value to both parties at contract initiation; neither party pays at the initiation of the contract.

We can examine three outcomes for the price of Acme shares at settlement:

1. The **spot price**, or the market price of the underlying, is \$30, equal to the forward price of \$30.

Neither party has profits or losses on the forward contract. Ignoring transactions costs, the party selling Acme shares could buy them back at \$30 per share and the party buying the shares could sell them at \$30 per share.

2. The spot price of Acme shares is \$40, greater than the forward price of \$30.

The party buying 100 Acme shares for \$3,000 at settlement can sell those shares at the spot price for \$4,000, realizing a profit of \$1,000 on the forward contract. The party that must deliver the Acme shares delivers shares with a market value of \$4,000 and receives \$3,000, realizing a \$1,000 loss on the forward contract.

3. The spot price of Acme shares is \$25 at settlement, less than the forward price of \$30.

The party buying 100 Acme shares for \$3,000 at settlement can sell those shares at the market price of \$25 to get \$2,500 and realize a loss of \$500 on the forward contract. The party that must deliver the

Acme shares delivers shares with a market value of \$2,500 and receives \$3,000, realizing a \$500 gain on the forward contract.

To summarize, the buyer of the shares in a forward contract will have gains when the market price of the shares at settlement is greater than the forward price, and losses when the market price of the shares at settlement is less than the forward price. The party that must deliver the shares in a forward contract will have gains when the market price of the shares at settlement is less than the forward price, and losses when the market price of the shares at settlement is greater than the forward price. The gains of one party equal the losses of the other party at settlement.

We refer to the party that agrees to buy the underlying asset in a forward contract as the buyer of the forward. The buyer of the forward gains when the price of the underlying increases (and loses when it falls), similarly to a long position in the underlying. In this case we say the forward buyer has *long exposure* to the underlying, while the seller of the underlying has *short exposure* to the underlying, gaining when the price of the underlying decreases and losing when the value of the underlying increases.

In practice, a forward contract may be a **deliverable contract**. In our forward contract example, this means that the payment and the shares must be exchanged at the settlement date. A **cash-settled contract** specifies that only the gains and losses from the forward contract are exchanged at settlement. In our example above, with a share price of \$25, cash settlement would require the buyer of Acme shares to pay \$500 to the seller of Acme shares at settlement. Ignoring the transaction cost of buying and selling shares in the market, the gains and losses to the parties in our forward contract example are economically equivalent under the two alternative settlement methods.

We can view a derivatives contract as a way to transfer risk from one party to another. Consider a situation where the share seller in the forward contract owns 100 shares of Acme. She has existing risk because the future price of Acme shares is uncertain; the share price three months from now is a random variable. If she enters the forward contract from our example, she will receive \$3,000 for her shares at settlement, regardless of their market price. This effectively transfers her existing Acme price risk to the buyer of the shares in the forward contract.

When a party to a derivative contract has an existing risk that is transferred to another party, we say that party has **hedged** (offset, reduced) their existing risk. If the risk of the forward contract exactly matches an existing risk, then the forward contract can be used to fully hedge the existing risk. If a derivative is used to reduce, but not entirely offset, an existing risk, we say the existing risk is **partially hedged**.

If the Acme share buyer in our example has no existing Acme price risk, she clearly increases her risk by entering into the forward contract. In this situation, the Acme share buyer is said to be **speculating** on the future price of Acme shares.

You may have realized our share seller could have achieved her goal of eliminating her Acme price risk by simply selling her shares (a **cash market transaction**). Derivatives have potential advantages over cash market transactions:

- Investors can gain exposure to a risk at low cost, effectively creating a highly leveraged investment in the underlying.

- Transaction costs for a derivatives position may be significantly lower than for the equivalent cash market trade.
- Initiating a derivatives position may have less impact on market prices of the underlying relative to initiating an equivalent position in the underlying through a cash market transaction.

Underlying Assets and Variables

The underlying for a derivative is most often a stock or bond price, the level of a stock or bond index, or an interest rate. Here, we give examples of different underlying assets and variables for derivative contracts and the nature of the risks they transfer or modify. We present more details about derivatives based on these underlying assets in subsequent readings.

- A bond, for example a forward contract on a 30-year U.S. Treasury bond or other specific bond. The risk involved is the uncertainty about future bond prices.
- An index, for example the S&P 500 Index or the Citi Goldman Sachs Investment Grade Corporate Bond Index. The risk involved is the uncertainty about the future value of the index at a specific date. A portfolio manager can reduce the risk of a portfolio of large U.S. stocks for a period of time by selling a forward on the S&P 500 index. An investor can gain long exposure to a portfolio of high-grade corporate bonds, quickly and at low cost, by buying a forward on the index.
- A currency, for example British pounds (GBP). A U.S. manufacturer that expects a large payment in GBP in six months can offset the uncertainty about the USD value of this payment by selling a forward contract on the expected amount of GBP. A UK manufacturer that must make a large payment in USD in one year can offset the uncertainty about the GBP cost of the payment by buying a forward contract on the USD that is priced in GBP.
- An interest rate, for example the 1-year Treasury bill rate. Similar to derivatives based on bonds, except that a higher interest rate means gains for the buyer of an interest rate forward, whereas higher interest rates mean lower bond prices and losses for the buyer of a bond forward.
- Commodities, which are physical assets including **hard commodities** (typically mined or extracted, such as gold and oil) and **soft commodities** (typically grown, such as cotton, coffee, pork, and cattle). A farmer expecting a cotton crop in four months can reduce her cotton price risk by selling cotton forward. A utility that will require thousands of gallons of oil over the next year can reduce its oil price risk by buying oil forwards that settle at various times over the coming year.



PROFESSOR'S NOTE

It may help you to remember this common rule for hedging risk with futures, which are similar to forwards: "Do in the futures market what you must do in the future." A baking company that must buy wheat in the future should buy wheat futures (or forwards) to reduce the effects of the uncertainty about future wheat prices on their profits. A farmer who must sell wheat at harvest time should sell wheat futures (or forwards) to reduce price risk.

- Credit derivatives include credit default swaps (CDS), in which one party makes fixed periodic payments to another party, which will make a payment only if the underlying

credit instrument (or portfolio of such credit securities) suffers a loss in value due to a default by the issuer (borrower of funds) of the credit instrument.

- Derivative contracts are also created with the weather (for farmers, energy producers, travel and tourism companies), cryptocurrencies, or longevity (for life insurers or annuity providers) as the underlying asset.

Along with forwards and futures, derivatives types that we will cover in the remainder of our derivatives readings are:

- Options:* Put options give the buyer the right (but not the obligation) to sell the underlying for a specific price in the future. Call options give the buyer the right (but not the obligation) to buy the underlying for a specific price in the future.
- Swaps:* In a simple interest rate swap, one party agrees to make periodic payments on a given amount at a fixed interest rate, and the other agrees to make periodic interest payments on the same given amount, but at an interest rate based on a future market reference rate (MRR). The resulting cash flows are equivalent to one party (the fixed-rate payer) borrowing at a fixed rate and using the proceeds to buy a floating-rate bond.

LOS 48.b: Describe the basic features of derivative markets, and contrast over-the-counter and exchange-traded derivative markets.

Exchange-Traded Derivatives

Centralized physical exchanges provide markets for futures contracts, some options contracts, and some other derivative contracts. The largest by volume of trades are the National Securities Exchange (India), the B3 market (Brazil), and the CME Group (the United States).

Exchange-traded derivatives are standardized and backed by a central clearinghouse. The exchange specifies the terms of each of the derivative contracts that will be traded and rules for trading on the exchange.

A **central clearinghouse (CCH)** essentially takes the opposite position to each side of a trade (called **novation**), guaranteeing the payments promised under the contract. The CCH requires deposits from both participants when a trade is initiated, and additional deposits for accounts that decline in value, to support its guarantee and minimize counterparty credit risk.

Exchange members (dealers or market makers) buy and sell derivatives at slightly different prices and primarily earn trading profits from the bid/ask spreads between buy and sell prices, rather than from holding (speculating on) specific derivatives positions, although they may hold such positions from time to time to meet customer needs.

The standardization of contracts allows exchange-traded derivatives to be more liquid and more transparent to market participants, compared to customized derivatives. A market participant who has taken a position in exchange-traded derivatives can easily exit that position by entering into a contract with a position opposite to their existing derivatives. Standardization of contracts reduces trading costs compared to customized derivative contracts.

Standardization also facilitates the clearing and settlement of trades. Clearing refers to executing the trade, recording the participants, and handling the exchange of any required

payments. Settlement refers to the exchange of underlying assets or payments of the final amounts due at contract settlement (maturity).

Dealer (OTC) Markets for Derivatives

Forwards, most swaps, and some options are custom instruments created and traded by dealers in a market with no central location. Some dealer markets are quite structured (e.g., the Nasdaq market), while others are not. A dealer market with no central location is referred to as an over-the-counter (OTC) market. OTC markets are largely unregulated and less transparent than exchange markets. In OTC markets with no central clearinghouse, each side of a trade faces counterparty credit risk. Dealers (market makers) make derivatives trades with end users of derivatives and may also trade with each other to reduce their exposures to changes in prices of underlying assets.

OTC derivatives contracts can be customized to fit the needs of an end user regarding contract size, definition of the underlying, settlement date, whether the contract is deliverable or cash settled, and other relevant details. Users trying to gain or hedge a specific risk use OTC derivatives when a standardized derivative contract will not meet their needs (including a desire for privacy).

After the financial crisis of 2008, regulators worldwide instituted a **central clearing mandate** requiring that, for many swap trades, a **central counterparty (CCP)** takes on the counterparty credit risk of both sides of a trade, similar to the role of a central clearinghouse. As an example, multiple dealers record their swap trades on a **swap execution facility (SEF)**. When a dealer makes a swap trade, that information is sent to the SEF and the CCP replaces the trade with two trades, with the CCP as the counterparty to both of them, reducing counterparty risk. The downside of this structure is that counterparty risks are concentrated rather than distributed among financial intermediaries.

The following offers a summary of the primary differences between exchange-traded and OTC derivatives.

Exchange-traded derivatives are:

- Traded at a centralized location, an exchange.
- Traded by exchange members (market makers).
- Based on standardized contracts and have lower trading costs.
- Subject to the trading rules of the exchange (i.e., are more regulated).
- Backed by the central clearinghouse to minimize counterparty credit risk. They also require deposits by both parties at initiation, and additional deposits when a position decreases in value.
- More liquid.
- More transparent, as all transactions are known to the exchange and to regulators.

OTC derivatives (not subject to the central clearing mandate) are:

- Custom instruments.
- Less liquid and have higher transaction costs.
- Less transparent.

- Subject to counterparty risk.
- More difficult to clear and settle.
- Subject to higher trading costs.
- Not subject to requirements for the deposit of collateral.

Derivatives in dealer markets that are subject to the central clearing mandate have reduced counterparty risk, are subject to more disclosure of trades, and are easier to clear and settle, but are still customizable and are contracts with dealers or financial intermediaries.

MODULE QUIZ 48.1

1. Which of the following statements *most accurately* describes a derivative security? A derivative:
 - A. always increases risk.
 - B. has no expiration date.
 - C. has a payoff based on an asset value or interest rate.
2. Which of the following statements about exchange-traded derivatives is *least accurate*? Exchange-traded derivatives:
 - A. are liquid.
 - B. are standardized contracts.
 - C. carry significant default risk.

KEY CONCEPTS

LOS 48.a

A derivative is a security that derives its value from value of another security or variable at a specific future date. The security or variable that determines the value of a derivative security is referred to as the underlying.

Basic features of a derivative include the underlying, the price specified in the contract, the contract size, and the settlement date. The price is typically set so the contract has zero value at initiation to both parties. Contracts may be deliverable or cash-settled.

LOS 48.b

Exchange-traded derivatives are standardized and backed by a central clearinghouse that takes the opposite position to each side of a trade, guaranteeing the payments promised under the contract.

Over-the-counter (OTC) derivatives can be customized to fit the needs of the counterparties. OTC markets are largely unregulated and less transparent than exchange markets. Some OTC markets are subject to a central clearing mandate that reduces counterparty credit risk.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 48.1

1. C A derivative's value is derived from another asset or an interest rate.
(LOS 48.a)

2. C Exchange-traded derivatives have relatively low default risk (counterparty credit risk) because the clearinghouse stands between the counterparties involved, in most contracts. (LOS 48.b)

READING 49

FORWARD COMMITMENT AND CONTINGENT CLAIM FEATURES AND INSTRUMENTS

MODULE 49.1: FORWARDS AND FUTURES



LOS 49.a: Define forward contracts, futures contracts, swaps, options (calls and puts), and credit derivatives and compare their basic characteristics.

Video covering this content is available online.

Forward Contracts

In a **forward contract** between two parties, one party (the buyer) commits to buy and the other party (the seller) commits to sell a physical or financial asset at a specific price on a specific date (the settlement date) in the future.

The buyer has long exposure to the underlying asset in that he will make a profit on the forward if the price of the underlying at the settlement date exceeds the forward price, and have a loss if the price of the underlying at the settlement date is less than the forward price. The results are opposite for the seller of the forward, who has short exposure to the underlying asset.

Futures Contracts

A **futures contract** is quite similar to a forward contract but is standardized and exchange-traded. The primary ways in which forwards and futures differ are that futures trade in a liquid secondary market, are subject to greater regulation, and trade in markets with more disclosure (transparency). Futures are backed by a central clearinghouse and require daily cash settlement of gains and losses, so that counterparty credit risk is minimized.

On a futures exchange, **margin** is cash or other acceptable collaterals that both the buyer and seller must deposit. Unlike margin in bond or stock accounts, there is no loan involved and, consequently, there are no interest charges. This collateral provides protection for the clearinghouse. At the end of each trading day, the margin balance in a futures account is adjusted for any gains and losses in the value of the futures position based on the new settlement price, a process called the **mark-to-market** or marking-to-market. The settlement price is calculated as the average price of trades over a period at the end of the trading session.

Initial margin is the amount of cash or collateral that must be deposited in a futures account before a trade may be made. Initial margin per contract is relatively low and is approximately one day's maximum expected price fluctuation on the total value of the assets covered by the contract.

Maintenance margin is the minimum amount of margin that must be maintained in a futures account. If the margin balance in the account falls below the maintenance margin through daily mark to market from changes in the futures price, the account holder must deposit additional funds to bring the margin balance back up to the *initial* margin amount, or the exchange will close out the futures position. This is different from a margin call in an equity account, which requires investors only to bring the margin back up to the maintenance margin amount. Futures margin requirements are set by the exchange.

To illustrate the daily mark-to-market for futures, consider a contract for 100 ounces of gold that settles on May 15. The initial margin amount is \$5,000 and the maintenance margin is \$4,700.

On Day 0

- A buyer and seller make a trade at the end of the day at a price of \$1,950 per ounce and both parties deposit the initial margin of \$5,000 into their accounts.

On Day 1, the settlement price falls to \$1,947.50. The seller has gains and the buyer has losses.

- The exchange will credit the seller's account for $(1,950 - 1,947.50) \times 100 = \250 , increasing the margin balance to \$5,250.
- The exchange will deduct $(1,950 - 1,947.50) \times 100 = \250 from the buyer's account, decreasing the margin balance to \$4,750. Because \$4,750 is more than the maintenance (minimum) margin amount of \$4,700, no additional deposit is required.

On Day 2, the settlement price falls to \$1,945. Again, the seller has gains and the buyer has losses.

- The exchange will credit the seller's account for $(1,947.50 - 1,945) \times 100 = \250 , increasing their margin balance to \$5,500.
- The exchange will deduct $(1,947.50 - 1,945) \times 100 = \250 from the buyer's account, decreasing the margin balance to \$4,500. Because \$4,500 is less than the maintenance (minimum) margin amount of \$4,700, the buyer must deposit $5,000 - 4,500 = \$500$ into their margin account to return it to the initial margin amount of \$5,000.
- At the end of Day 2, both parties have futures positions at the new settlement price of \$1,945 per ounce.

Many futures contracts have **price limits**, which are exchange-imposed limits on how much each day's settlement price can change from the previous day's settlement price. Exchange members are prohibited from executing trades at prices outside these limits. If the equilibrium price at which traders would willingly trade is above the upper limit or below the lower limit, trades cannot take place. Some exchanges have **circuit breakers**; in this case, when a futures price reaches a limit price, trading is suspended for a short period.



MODULE QUIZ 49.1

1. Which type of contract always requires daily marking to market of gains and losses?
 - A. Futures contracts only.
 - B. Forward contracts only.
 - C. Both futures and forward contracts.
2. Compared to a futures contract, an otherwise identical forward contract *most likely* has greater:

- A. liquidity.
- B. transparency.
- C. counterparty risk.

MODULE 49.2: SWAPS AND OPTIONS



Video covering
this content is
available online.

Swaps

Swaps are agreements to exchange a series of payments on multiple settlement dates over a specified time period (e.g., quarterly payments for two years). At each settlement date, the two payments are netted so that only one net payment is made. The party with the greater liability at each settlement date pays the net difference to the other party.

Swaps trade in a dealer market and the parties are exposed to counterparty credit risk, unless the market has a central counterparty structure to reduce counterparty risk. In this case, margin deposits and mark-to-market payments may also be required to further reduce counterparty risk.

We can illustrate the basics of a swap with a simple fixed-for-floating interest rate swap for two years with quarterly interest payments based on a **notional principal** amount of \$10 million. In such a swap, one party makes quarterly payments at a fixed rate of interest (the **swap rate**) and the other makes quarterly payments based on a floating **market reference rate**.

The swap rate is set so that the swap has zero value to each party at its inception. As expectations of future values of the market reference rate change over time, the value of the swap can become positive for one party and negative for the other party.

Consider an interest rate swap with a notional principal amount of \$10 million, a fixed rate of 2%, and a floating rate of the 90-day secured overnight financing rate (SOFR). At each settlement date, the fixed-rate payment will be $\$10\text{ million} \times 0.02/4 = \$50,000$. The floating-rate payment at the end of the first quarter will be based on 90-day SOFR at the initiation of the swap, so that both payments are known at the inception of the swap.

If, at the end of the first quarter, 90-day SOFR is 1.6%, the floating-rate payment at the second quarterly settlement date will be $\$10\text{ million} \times 0.016 / 4 = \$40,000$. The fixed-rate payment is again \$50,000, so at the end of the second quarter the fixed-rate payer will pay the net amount of \$10,000 to the other party.

A company with 2-year floating-rate quarterly-pay note outstanding could enter such a swap as the fixed-rate payer, converting its floating-rate liability into a fixed-rate liability. It now makes fixed interest rate payments and can use the floating-rate payments from the counterparty to make the payments on its floating-rate debt. By entering into the swap, the company can hedge the interest rate risk (uncertainty about future quarterly rates) of their existing floating-rate liability.

As we will see in our reading on swap valuation, a swap can be constructed from a series of forward contracts in which the underlying is a floating rate and the forward price is a fixed rate. Each forward settles on one of the settlement dates of the swap. At each settlement date, the difference between the fixed and the floating rates would result in a net payment, just as with a

swap. Often, interest rate forwards settle at the beginning of the quarter rather than the end; the cash flows are the present value equivalents of the end-of-quarter swap payments.

Credit Swaps

One type of swap that is structured a bit differently is a **credit default swap (CDS)**. With a CDS, the protection buyer makes fixed payments on the settlement dates and the protection seller pays only if the underlying (a reference security) has a **credit event**. This could be a bond default, a corporate bankruptcy, or an involuntary restructuring.

When a credit event occurs, the protection seller must pay an amount that offsets the loss in value of the reference security. The fixed payments represent the yield premium on the reference bonds that compensates bondholders for the expected loss from default, the probability of default times the expected loss in the event of default (or other credit event). The protection buyer is essentially paying the yield premium on the reference security for insurance against default.

The holder of a risky bond can hedge its default risk by entering a CDS as the protection buyer. The protection seller receives the default risk premium (credit spread) and takes on the risk of default, resulting in risk exposure similar to that of holding the reference bond.

Options

The two types of options of interest to us here are **put options** and **call options** on an underlying asset. We introduce them using option contracts for 100 shares of a stock as the underlying asset.

A put option gives the buyer the right (but not the obligation) to sell 100 shares at a specified price (the **exercise price**, also referred to as the **strike price**) for specified period of time, the **time to expiration**. The put seller (also called the *writer* of the option) takes on the obligation to purchase the 100 shares at the price specified in the option, if the put buyer exercises the option.

Note the “one-way” nature of options. If the exercise price of the puts is \$25 at the expiration of the option, and the shares are trading at or above \$25, the put holder will not exercise the option. There is no reason to exercise the put and sell shares at \$25 when they can be sold for more than \$25 in the market. This is the outcome for any stock price greater than or equal to \$25. Regardless of whether the stock price at option expiration is \$25 or \$1,000, the put buyer lets the option expire, and the put seller keeps the proceeds from the sale.

If the stock price is below \$25, the put buyer will exercise the option and the put seller must purchase 100 shares for \$25 from the put buyer. On net, the put buyer essentially receives the difference between the stock price at expiration and \$25 (times 100 shares).

A call option gives the buyer the right (but not the obligation) to buy 100 shares at a specified price (the exercise price) for a specified period of time. The call seller (writer) takes on the obligation to sell the 100 shares at the exercise price, if the call buyer exercises the option.

LOS 49.b: Determine the value at expiration and profit from a long or a short position in a call or put option.

Unlike forwards, futures, and swaps, options are sold at a price (they do not have zero value at initiation). The price of an option is also referred to as the **option premium**.

At expiration the payoff (value) of a call option to the owner is $\text{Max}(0, S - X)$, where S is the price of the underlying at expiration and X is the exercise price of the call option. The $\text{Max}()$ function tells us that if $S < X$ at expiration, the option value is zero, that is, it expires worthless and will not be exercised.

At expiration the payoff (value) of a put option to the owner is $\text{Max}(0, X - S)$, where S is the price of the underlying at expiration and X is the exercise price of the put option. A put has a zero value at expiration unless $X - S$ is positive.

For the buyer of a put or call option, the profit at expiration is simply the difference between the value (payoff) of the option at expiration and the premium the investor paid for the option.

Because the seller (writer) of an option receives the option premium, the profit to the option seller at expiration is the amount of the premium received minus the option payoff at expiration. The writer loses the payoff at expiration and will have a loss on the option if the payoff is greater than the premium received.

Note the risk exposures of call and put buyers and writers. The buyer of a put or call has no further obligation, so the maximum loss to the buyer is simply the amount they paid for the option. The writer of a call option has exposure to an unlimited loss because the maximum price of the underlying, S , is (theoretically) unlimited, so that the payoff $S - X$ is unlimited. The payoff on a put option is $X - S$, so if the lower limit on S is zero, the maximum payoff on a put option is the exercise price, X .

Call Option Profits and Losses

Consider a call option with a premium of \$5 and an exercise price of \$50. This means the buyer pays \$5 to the writer. At expiration, if the price of the stock is less than or equal to the \$50 exercise price, the option has zero value, the buyer of the option is out \$5, and the writer of the option is ahead \$5. When the stock's price exceeds \$50, the option starts to gain (breakeven will come at \$55, when the value of the stock equals the exercise price plus the option premium).

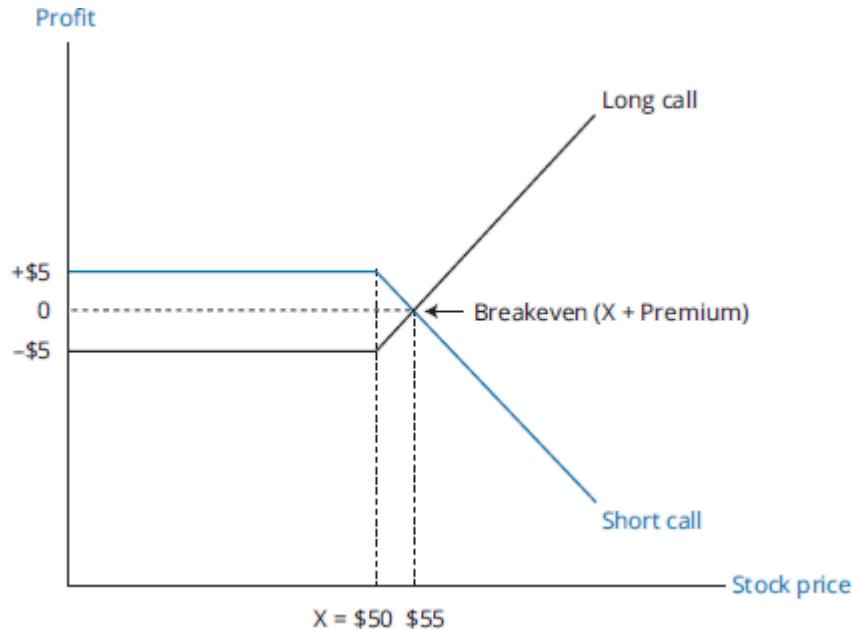
Conversely, as the price of the stock moves upward, the seller of the option starts to lose (negative figures will start at \$55, when the value of the stock equals the exercise price plus the option premium).

An illustration of the profit or loss at expiration for the buyer (long) and writer (short) of this call option, as a function of the stock price, is presented in Figure 49.1. This profit/loss diagram indicates the following:

- The maximum loss for the buyer of a call is the \$5 premium (at any $S \leq \$50$).
- The breakeven point for the buyer and seller is the exercise price plus the premium (at $S = \$55$).
- The profit potential to the buyer of the option is unlimited, and, conversely, the potential loss to the writer of the call option is unlimited.
- The call holder will exercise the option whenever the stock's price exceeds the exercise price at the expiration date.
- The greatest profit the writer can make is the \$5 premium (at any $S \leq \$50$).

- The sum of the profits between the buyer and seller of the call option is always zero; thus, trading options is a *zero-sum game*. One party's profits equal the other party's losses.

Figure 49.1: Profit/Loss Diagram for a Call Option



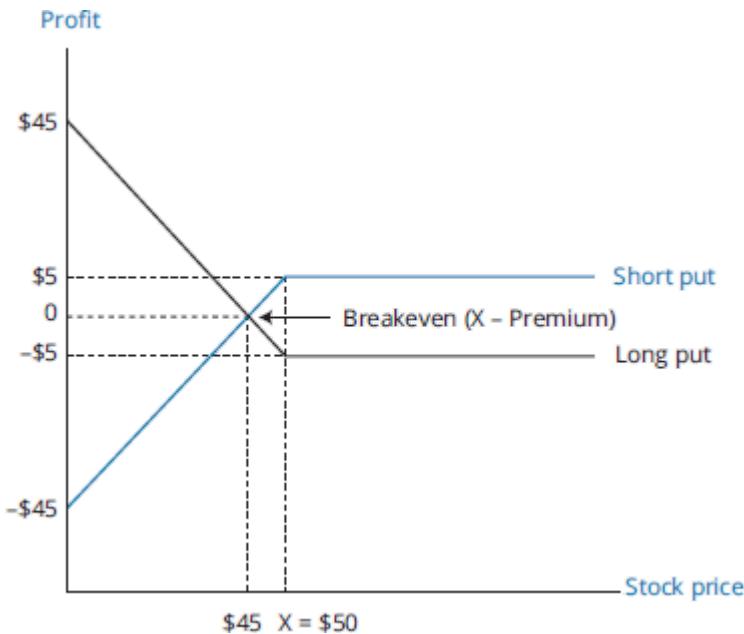
Put Option Profits and Losses

To examine the profits and losses associated with trading put options, consider a put option with a \$5 premium and a \$50 exercise price. The buyer pays \$5 to the writer. When the price of the stock at expiration is greater than or equal to the \$50 exercise price, the put has zero value. The buyer of the option has a loss of \$5 and the writer of the option has a gain of \$5. When the stock price is less than \$50, the put option has a positive payoff. Breakeven will come at \$45, when the value of the stock equals the exercise price less the option premium. At a stock price below \$45, the put seller will have a loss.

Figure 49.2 shows the profit/loss diagram for the buyer (long) and seller (short) of the put option that we have been discussing. This profit/loss diagram illustrates that:

- The maximum loss for the buyer of the put is the \$5 premium (at any $S \geq \$50$).
- The maximum gain to the buyer of the put is limited to the exercise price less the premium ($\$50 - \$5 = \$45$). The potential loss to the writer of the put is the same amount.
- The breakeven price for the put buyer (seller) is the exercise price minus the option premium ($\$50 - \$5 = \$45$).
- The maximum profit for the writer is the \$5 premium ($S \geq \50).
- The profit (loss) of the put buyer will always equal the loss (profit) of the put writer.

Figure 49.2: Profit/Loss Diagram for a Put Option



EXAMPLE: Option profit calculations

Suppose that both a call option and a put option have been written on a stock with an exercise price of \$40. The current stock price is \$42, and the call and put premiums are \$3 and \$0.75, respectively.

Calculate the profit to the long and short positions for both the put and the call with an expiration day stock price of \$35 and with a price at expiration of \$43.

Answer:

Profit will be computed as ending option value – initial option cost.

Stock at \$35:

- Long call: $\$0 - \$3 = -\$3$. The option has no value, so the buyer loses the premium paid.
- Short call: $\$3 - \$0 = \$3$. Because the option has no value, the call writer's gain equals the premium received.
- Long put: $\$5 - \$0.75 = \$4.25$. The buyer paid \$0.75 for an option that is now worth \$5.
- Short put: $\$0.75 - \$5 = -\$4.25$. The seller received \$0.75 for writing the option, but the option will be exercised so the seller will lose \$5 at expiration.

Stock at \$43:

- Long call: $-\$3 + \$3 = \$0$. The buyer paid \$3 for the option, and it is now in the money by \$3. Hence, the net profit is zero.
- Short call: $\$3 - \$3 = \$0$. The seller received \$3 for writing the option and now faces a $-\$3$ valuation because the buyer will exercise the option, for a net profit of zero.
- Long put: $-\$0.75 - \$0 = -\$0.75$. The buyer paid \$0.75 for the put option and the option now has no value.
- Short put: $\$0.75 - \$0 = \$0.75$. The seller received \$0.75 for writing the option and it has zero value at expiration.

A buyer of puts or a seller of calls has short exposure to the underlying (will profit when the price of the underlying asset decreases). A buyer of calls or a seller of puts has long exposure to the underlying (will profit when the price of the underlying asset increases).

LOS 49.c: Contrast forward commitments with contingent claims.

A **forward commitment** is a legally binding promise to perform some action in the future. Forward commitments include forward contracts, futures contracts, and most swaps.

A **contingent claim** is a claim (to a payoff) that depends on a particular event. Options are contingent claims; the event is the price of the underlying being above or below the exercise price. Credit default swaps are also considered contingent claims because the payment by the protection seller depends on a credit event occurring.

MODULE QUIZ 49.2

1. Interest rate swaps are:
 - A. highly regulated.
 - B. equivalent to a series of forward contracts.
 - C. contracts to exchange one asset for another.
2. A call option is:
 - A. the right to sell at a specific price.
 - B. the right to buy at a specific price.
 - C. an obligation to buy at a certain price.
3. At expiration, the exercise value of a put option is:
 - A. positive if the underlying asset price is less than the exercise price.
 - B. zero only if the underlying asset price is equal to the exercise price.
 - C. negative if the underlying asset price is greater than the exercise price.
4. At expiration, the exercise value of a call option is:
 - A. the underlying asset price minus the exercise price.
 - B. the greater of zero or the exercise price minus the underlying asset price.
 - C. the greater of zero or the underlying asset price minus the exercise price.
5. An investor writes a put option with an exercise price of \$40 when the stock price is \$42. The option premium is \$1. At expiration the stock price is \$37. The investor will realize:
 - A. a loss of \$2.
 - B. a loss of \$3.
 - C. a profit of \$1.
6. Which of the following derivatives is a forward commitment?
 - A. Stock option.
 - B. Interest rate swap.
 - C. Credit default swap.

KEY CONCEPTS

LOS 49.a

Forward contracts obligate one party to buy, and another to sell, a specific asset at a specific price at a specific time in the future.

Futures contracts are much like forward contracts, but are exchange-traded, liquid, and require daily settlement of any gains or losses.

A call option gives the holder the right, but not the obligation, to buy an asset at a specific price at some time in the future.

A put option gives the holder the right, but not the obligation, to sell an asset at a specific price at some time in the future.

In an interest rate swap, one party pays a fixed rate and the other party pays a floating rate, on a given amount of notional principal. Swaps are equivalent to a series of forward contracts based

on a floating rate of interest.

A credit default swap is a contract in which the protection seller provides a payment if a specified credit event occurs.

LOS 49.b

Call option value at expiration is $\text{Max}(0, \text{underlying price} - \text{exercise price})$ and profit or loss is $\text{Max}(0, \text{underlying price} - \text{exercise price}) - \text{option cost}$ (premium paid).

Put value at expiration is $\text{Max}(0, \text{exercise price} - \text{underlying price})$ and profit or loss is $\text{Max}(0, \text{exercise price} - \text{underlying price}) - \text{option cost}$.

A call buyer (call seller) benefits from an increase (decrease) in the value of the underlying asset.

A put buyer (put seller) benefits from a decrease (increase) in the value of the underlying asset.

LOS 49.c

A forward commitment is an obligation to buy or sell an asset or make a payment in the future. Forward contracts, futures contracts, and most swaps are forward commitments.

A contingent claim is a derivative that has a future payoff only if some future event takes place (e.g., asset price is greater than a specified price). Options and credit derivatives are contingent claims.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 49.1

1. **A** Futures contracts are marked to market daily. Forward contracts typically are not, but could be if there is central clearing party.
(LOS 49.a)
2. **C** Forward contracts involve counterparty risk; futures contracts trade through a clearinghouse. Because futures contracts trade on organized exchanges, they have greater liquidity and transparency than forward contracts. (LOS 49.a)

Module Quiz 49.2

1. **B** A swap is an agreement to buy or sell an underlying asset periodically over the life of the swap contract. It is equivalent to a series of forward contracts. (LOS 49.a)
2. **B** A call gives the owner the right to call an asset away (buy it) from the seller. (LOS 49.a)
3. **A** The exercise value of a put option is positive at expiration if the underlying asset price is less than the exercise price. Its exercise value is zero if the underlying asset price is greater than or equal to the exercise price. The exercise value of an option cannot be negative because the holder can allow it to expire unexercised. (LOS 49.b)
4. **C** If the underlying asset price is greater than the exercise price of a call option, the value of the option is equal to the difference. If the underlying asset price is less than the

exercise price, a call option expires with a value of zero. (LOS 49.b)

5. **A** Because the stock price at expiration is less than the exercise price, the buyer of the put option will exercise it against the writer. The writer will have to pay \$40 for the stock and can only sell it for \$37 in the market. However, the put writer collected the \$1 premium for writing the option, which reduces the net loss to \$2. (LOS 49.b)

6. **C** This type of custom contract is a forward commitment. (LOS 49.c)

READING 50

DERIVATIVE BENEFITS, RISKS, AND ISSUER AND INVESTOR USES

MODULE 50.1: USES, BENEFITS, AND RISKS OF DERIVATIVES



LOS 50.a: Describe benefits and risks of derivative instruments.

Video covering this content is available online.

Advantages of Derivatives

Derivative instruments offer several potential advantages over cash market transactions, including the following:

Ability to change risk allocation, transfer risk, and manage risk

We have discussed these benefits in our introduction to derivative contracts. Some examples of ways that risk exposures can be altered using derivatives, without any cash market securities transactions, are:

- A portfolio manager can increase or decrease exposure to the risk and return of a market index.
- A manufacturer can hedge the exchange rate risk of anticipated receipts or payments.
- The issuer of a floating-rate note can change that exposure to a fixed-rate obligation.

Derivative instruments can be used to create risk exposures that are not available in cash markets. Consider the following examples of changing an existing risk profile:

- The owner of common stock can buy puts that act as a floor on the sale price of their shares, reducing the downside risk of the stock by paying the cost of the puts.
- An investor can acquire the upside potential of an asset without taking on its downside risk by buying call options.

Information discovery

Derivatives prices and trading provide information that cash market transactions do not.

- Options prices depend on many things we can observe (interest rates, price of the underlying, time to expiration, and exercise price) and one we cannot, the expected future price volatility of the underlying. We can use values of the observable variables, together with current market prices of derivatives, to estimate the future price volatility of the underlying that market participants expect.

- Futures and forwards can be used to estimate expected prices of their underlying assets.
- Interest rate futures across maturities can be used to infer expected future interest rates and even the number of central bank interest rate changes over a future period.

Operational advantages

Compared to cash markets, derivatives markets have several operational advantages. Operational advantages of derivatives include greater ease of short selling, lower transaction costs, greater potential leverage, and greater liquidity.

- *Ease of short sales.* Taking a short position in an asset by selling a forward or a futures contract may be easy to do. Difficulty in borrowing an asset and restrictions on short sales may make short positions in underlying assets problematic or more expensive.
- *Lower transaction costs.* Transaction costs can be significantly lower with commodities derivatives, where transportation, storage, and insurance add costs to transactions in physical commodities. Entering a fixed-for-floating swap to change a floating-rate exposure to fixed rate is clearly less costly than retiring a floating-rate note and issuing a fixed-rate note.
- *Greater leverage.* The cash required to take a position in derivatives is typically much less than for an equivalent exposure in the cash markets.
- *Greater liquidity.* The low cash requirement for derivatives transactions makes very large transactions easier to handle.

Improved market efficiency

Low transaction costs, greater liquidity and leverage, and ease of short sales all make it less costly to exploit securities mispricing through derivatives transactions and improve the efficiency of market prices.

Risks of Derivatives

Implicit leverage

The implicit leverage in derivatives contracts gives them much more risk than their cash market equivalents. Just as we have shown regarding the leverage of an equity investment on margin, a lower cash requirement to enter a trade increases leverage. Futures margins, according to the CME Group, are typically in the 3% to 12% range, indicating leverage of 8:1 to 33:1. With required cash margin of 4%, a 1% decrease in the futures price decreases the cash margin by 25%.

A lack of transparency in derivatives contracts and securities that combine derivative and cash market exposures (structured securities) may lead to situations in which the purchasers do not well understand the risks of derivatives or securities with embedded derivatives.

Basis risk

Basis risk arises when the underlying of a derivative differs from a position being hedged with the derivative. For a manager with a portfolio of 50 large-cap U.S. stocks, selling a forward with the S&P 500 Index as the underlying (in an amount equal to the portfolio value) would hedge

portfolio risk, but would not eliminate it because of the possibility that returns on the portfolio and returns on the index may differ over the life of the forward. Basis risk also arises in a situation where an investor's horizon and the settlement date of the hedging derivative differ, such as hedging the value of a corn harvest that will occur on September 15 by selling corn futures that settle on October 1. Again the hedge may be effective but will not be perfect, and the corn producer is said to have basis risk.

Liquidity risk

Derivative instruments have a special type of liquidity risk when the cash flows from a derivatives hedge do not match the cash flows of the investor positions. As an example, consider a farmer who sells wheat futures to hedge the value of her wheat harvest. If the future price of wheat increases, losses on the short position essentially offset the extra income from the higher price that will come at harvest (as intended with a hedge), but these losses may also cause the farmer to get margin calls during the life of the contract. If the farmer does not have the cash (liquidity) to meet the margin calls, the position will be closed out and the value of the hedge will be lost.

Counterparty credit risk

We have discussed counterparty credit risk previously. Here we note additionally that different derivatives and positions have important differences in the existence or amount of counterparty risk. The seller of an option faces no counterparty credit risk; once the seller receives the option premium there is no circumstance in which the seller will be owed more at settlement. On the other hand, the buyer of an option will be owed money at settlement if the option is in the money; thus, the buyer faces counterparty credit risk. In contrast, both the buyer and seller of a forward on an underlying asset may face counterparty credit risk.

In futures markets the deposit of initial margin, the daily mark-to-market, and the guarantee of the central clearinghouse all reduce counterparty risk. With forwards there may be no guarantees, or the terms of the forward contract may specify margin deposits, a periodic mark-to-market, and a central clearing party to mitigate credit risk.

Systemic risk

Widespread impact on financial markets and institutions may arise from excessive speculation using derivative instruments. Market regulators attempt to reduce systemic risk through regulation, for example the central clearing requirement for swap markets to reduce counterparty credit risk.

LOS 50.b: Compare the use of derivatives among issuers and investors.

Derivatives Use by Issuers

Corporate users of derivative instruments are considered issuers of derivatives. A non-financial corporation may have risks associated with changes in asset and liability values as well as earnings volatility from changes in various underlying securities or interest rates. Some examples are:

- A corporation may have income in a foreign currency and hedge the exchange rate risk with forwards to smooth earnings reported in their domestic currency.
- A corporation may use fair value reporting for its fixed-rate debt, and that value changes as interest rates change. By entering an interest rate swap as the floating-rate payer, the corporation has essentially converted the fixed-rate liability to a floating-rate liability that has much lower duration so that its balance sheet value is less sensitive to changes in interest rates.
- A corporation with a commodity-like product may carry its inventory at fair market value, leading to fluctuations in the value reported on the balance sheet over time as the market price of their product changes. By selling forward contracts on an underlying that matches well with their product, the firm will have gains or losses on the forwards that offset decreases or increases in reported inventory value. With the market value of the forward position also reported on the balance sheet, total assets will have less variation from changes in the market price of their product.

Accounting rules may permit **hedge accounting**. Hedge accounting allows firms to recognize the gains and losses of qualifying derivative hedges at the same time they recognize the corresponding changes in the values of assets or liabilities being hedged. Issuer hedges against the effects of a changing price or value of a derivative's underlying are classified by their purpose.

- A hedge of the domestic currency value of future receipts in a foreign currency using forwards is termed a **cash flow hedge**. A swap that converts a floating-rate liability to a fixed-rate liability is also considered a cash flow hedge (cash flows for interest payments are more certain).
- A **fair value hedge** is one that reduces (offsets) changes in the values of the firm's assets or liabilities. Our examples of a firm that uses derivatives to hedge against changes in the balance sheet value of its inventory, and a firm that uses an interest rate swap to decrease the volatility of debt values on its balance sheet, are considered fair value hedges.
- A **net investment hedge** is one that reduces the volatility of the value of the equity of a company's foreign subsidiary reported on its balance sheet. Foreign currency forwards or futures can be used to hedge changes in the reported value of the subsidiary's equity due to changes in exchange rates.

Derivatives Use by Investors

As we have seen, investors can hedge, modify, or increase their exposure to the risk of an underlying asset or interest rate with derivatives positions, either forward commitments or contingent claims. Some examples are:

- An investor can buy silver forwards to gain exposure to the price of silver, with no or low funds initially required.
- An investor can increase the duration of their bond portfolio by entering an interest rate swap as the floating-rate payer/fixed-rate receiver, which is similar to issuing floating-rate debt and buying a fixed-rate bond with the proceeds.
- An equity portfolio manager can modify their market risk exposure temporarily at low cost, increasing it by buying equity index futures or decreasing it by selling equity index

futures. Alternatively, the portfolio manager could decrease downside risk and preserve upside potential by buying puts on an equity index.



MODULE QUIZ 50.1

1. Which of the following *most* accurately describes a risk of derivative instruments?
 - A. Derivatives make it easier for market participants to take short positions.
 - B. The underlying of a derivative might not fully match a position being hedged.
 - C. Volatility in underlying asset prices is implied by the prices of options on those assets.
2. Uses of derivatives by investors *most likely* include:
 - A. hedging against price risk for inventory held.
 - B. modifying the risk exposure of a securities portfolio.
 - C. stabilizing the balance sheet value of a foreign subsidiary.

KEY CONCEPTS

LOS 50.a

Advantages of derivatives include the ability to change or transfer risk; information discovery about the expected prices or volatility of underlying assets or interest rates; operational advantages such as ease of short sales, low transaction costs, and greater leverage and liquidity; and improved market efficiency.

Risks of derivatives include implicit leverage, basis risk from inexact hedges, liquidity risk from required cash flows, counterparty credit risk, and systemic risk for financial markets.

LOS 50.b

Derivatives uses by issuers include managing risks associated with changes in asset and liability values as well as earnings volatility from changes in various underlying securities or interest rates.

Derivatives uses by investors include hedging, modifying, or increasing their exposure to the risk of an underlying asset or interest rate.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 50.1

1. **B** Basis risk arises when the underlying of a derivative differs from a position being hedged. Ease of taking short positions with derivatives compared to their underlying assets, and the information about implied volatility that is revealed by option prices, are two of the advantages of derivative instruments. (LOS 50.a)
2. **B** Modifying the risk exposure of a securities portfolio is an example of derivatives use by investors. Hedging against price risk for inventory and stabilizing the balance sheet value of a foreign subsidiary are examples of derivatives use by issuers. (LOS 50.b)

READING 51

ARBITRAGE, REPLICATION, AND THE COST OF CARRY IN PRICING DERIVATIVES

MODULE 51.1: ARBITRAGE, REPLICATION, AND CARRYING COSTS



Video covering this content is available online.

LOS 51.a: Explain how the concepts of arbitrage and replication are used in pricing derivatives.

In contrast to valuing risky assets as the (risk-adjusted) present value of expected future cash flows, the valuation of derivative securities is based on a **no-arbitrage condition**. *Arbitrage* refers to a transaction in which an investor purchases one asset or portfolio of assets at one price and simultaneously sells an asset or portfolio of assets that has the same future payoffs, regardless of future events, at a higher price, realizing a risk-free gain on the transaction.

While arbitrage opportunities may be rare, the reasoning is that when they do exist, they will be exploited rapidly. Therefore, we can use a no-arbitrage condition to determine the current value of a derivative, based on the known value of a portfolio of assets that has the same future payoffs as the derivative, regardless of future events. Because there are transaction costs of exploiting an arbitrage opportunity, small differences in price may persist when the arbitrage gain is less than the transaction costs of exploiting it.

We can illustrate no-arbitrage pricing with a 1-year forward contract, with a forward price of $F_0(1)$, on an Acme share that pays no dividends and is trading at a current price, S_0 , of \$30.

Consider two strategies to own an Acme share at $t = 1$:

- *Portfolio 1:* Buy a pure discount bond with a yield of 5% that pays $F_0(1)$ at $t = 1$.
The current cost of the bond is $F_0(1)/1.05$. Additionally, enter a forward contract on one Acme share at $F_0(1)$ as the buyer. The forward has a zero cost, so the cost of Portfolio 1 is $F_0(1)/1.05$.
At $t = 1$ the bond pays $F_0(1)$, which will buy an Acme share at the forward price, so that the payoff on Portfolio 1 is the value of one share at $t = 1$, S_1 .
- *Portfolio 2:* Buy a share of Acme at $S_0 = 30$ and hold it for one year. Cost at $t = 0$ is \$30.
At $t = 1$ the value of the Acme share is S_1 and this is the payoff for Portfolio 2.

The no-arbitrage condition (law of one price) requires that two portfolios with the same payoff in the future for any future value of Acme have the same cost today. Because our two portfolios have a payoff of S_1 , they must have the same cost at $t = 0$ to prevent arbitrage. That is, $F_0(1)/1.05 = \$30$, so we can solve for the no-arbitrage forward price as $F_0(1) = 30(1.05) = 31.50$.

To better understand the no-arbitrage condition, we will consider two situations in which the forward price is not at its no-arbitrage value: $F_0(1) > 31.50$ and $F_0(1) < 31.50$.

- If the forward contract price is 32 ($F_0(1) > 31.50$), the profitable arbitrage is to sell the forward (because the forward price is “too high”) and buy a share of stock. At $t = 1$, deliver the share under the forward contract and receive 32, for a return of $32/30 - 1 = 6.67\%$, which is higher than the risk-free rate.

We can also view this transaction as borrowing 30 at the risk-free rate (5%) to buy the Acme share at $t = 0$, and at $t = 1$ paying 31.50 to settle the loan. The share delivered under the forward has a contract price of 32, so the arbitrageur has an arbitrage profit of $32 - 31.50 = 0.50$ with no risk and no initial cost.

- If the forward contract price is 31 ($F_0(1) < 31.50$), the profitable arbitrage is to buy the forward and sell short an Acme share at $t = 0$. The proceeds of the short sale, 30, can be invested at the risk-free rate to produce $30(1.05) = 31.50$ at $t = 1$. The forward contract requires the purchase of a share of Acme for 31, which the investor can return to close out the short position. The profit to an arbitrageur is $31.50 - 31 = 0.50$. With no cash investment at $t = 0$, the investor receives an arbitrage profit of 0.50 at $t = 1$.

When the forward price is “too high,” the arbitrage is to sell the forward and buy the underlying asset. When the forward price is “too low,” the arbitrage is to buy the forward and sell (short) the underlying asset. In either case, the actions of arbitrageurs will move the forward price toward its no-arbitrage level until arbitrage profits are no longer possible.

Replication refers to creating a portfolio with cash market transactions that has the same payoffs as a derivative for all possible future values of the underlying.

Our arbitrage example for Acme forwards will serve to illustrate replication.

A long forward on an Acme share can be replicated by borrowing 30 at 5% to purchase an Acme share, and repaying the loan on the settlement date of the forward. At settlement ($t = 1$), the payoff on the replication is $S_1 - 30(1.05) =$

$S_1 - 31.50$ (value of one share minus the repayment of the loan), the same as the payoff on a long forward at 31.50, for any value of Acme shares at settlement.

A short forward on an Acme share can be replicated by shorting an Acme share and investing the proceeds of 30 at 5%. At settlement the investor receives 31.50 from the investment of short sale proceeds, and must buy a share of Acme for S_1 . The payoff on the replicating portfolio is $31.50 - S_1$, the same as the payoff on a short forward at 31.50, for any value of Acme shares at settlement.

These replications allow us to calculate the no-arbitrage forward price of an asset, just as we did in our example using Acme shares. Because our replicating portfolio for a long forward has the same payoff as a long forward at time = T, the payoff at settlement on a portfolio that is long the replicating portfolio and short the forward must be zero to prevent arbitrage. For this strategy, when the forward is priced at its no-arbitrage value the payoff at time = T is:

$$S_T - S_0(1 + R_f)^T - [S_T - F_0(T)] = 0$$

$$\text{so that } S_0(1 + R_f)^T + F_0(T) = 0 \text{ and } F_0(T) = S_0(1 + R_f)^T.$$

For a portfolio that is short the replicating portfolio and long the forward, the payoff at time T is:

$$S_0(1 + R_f)^T - S_T + [S_T - F_0(T)] = 0$$

$$\text{so that } S_0(1 + R_f)^T - F_0(T) = 0 \text{ and } F_0(T) = S_0(1 + R_f)^T.$$

The forward price that will prevent arbitrage is $S_0(1 + R_f)^T$, just as we found in our example of a forward contract on an Acme share.

LOS 51.b: Explain the difference between the spot and expected future price of an underlying and the cost of carry associated with holding the underlying asset.

When we derived the no-arbitrage forward price for an asset as $F_0(T) = S_0(1 + R_f)^T$, we assumed there were no benefits of holding the asset and no costs of holding the asset, other than the opportunity cost of the funds to purchase the asset (the risk-free rate of interest).

Any additional costs or benefits of holding the underlying asset must be accounted for in calculating the no-arbitrage forward price. There may be additional costs of owning an asset, especially with commodities, such as storage and insurance costs. For financial assets, these costs are very low and not significant.

There may also be monetary benefits to holding an asset, such as dividend payments for equities and interest payments for debt instruments. Holding commodities may have non-monetary benefits, referred to as **convenience yield**. If an asset is difficult to sell short in the market, owning it may convey benefits in circumstances where selling the asset is advantageous. For example, a shortage of the asset may drive prices up temporarily, making sale of the asset in the short term profitable.

We denote the present value of any costs of holding the asset from time 0 to settlement at time T (e.g., storage, insurance, spoilage) as $PV_0(\text{cost})$, and the present value of any cash flows from the asset or convenience yield over the holding period as $PV_0(\text{benefit})$.

Consider first a case where there are storage costs of holding the asset, but no benefits. For an asset with no costs or benefits of holding the asset, we established the no-arbitrage forward price as $S_0(1 + R_f)^T$, the cost of buying and holding the underlying asset until time T. When there are storage costs to hold the asset until time T, an arbitrageur must both buy the asset and pay the present value of storage costs at $t = 0$. This increases the no-arbitrage price of a 1-year forward to $[S_0 + PV_0(\text{cost})](1 + R_f)^T$. Here we see that *costs of holding an asset increase its no-arbitrage forward price*.

Next consider a case where holding the asset has benefits, but no costs. Returning to our example of a 1-year forward on a share of Acme stock trading at 30, now consider the costs of buying and holding an Acme share that pays a dividend of \$1 during the life of the forward contract. In this case, an arbitrageur can now borrow the present value of the dividend

(discounted at R_f), and repay that loan when the dividend is received. The cost to buy and hold Acme stock with an annual dividend of \$1 is $[30 - PV_0(1)](1.05) = 30(1.05) - 1$. This illustrates that *benefits of holding an asset decrease its no-arbitrage forward price*.

The no-arbitrage price of a forward on an asset that has both costs and benefits of holding the asset is simply $[S_0 + PV_0(\text{costs}) - PV_0(\text{benefit})](1 + R_f)^T$.

We can also describe these relationships when costs and benefits are expressed as continuously compounded rates of return. Recall from Quantitative Methods that given a stated annual rate of r with continuous compounding, the effective annual return is $e^r - 1$, and the relationships between present and future values of S for a 1-year period are $FV = Se^r$ and $PV = Se^{-r}$. For a period of T years, $FV = Se^{rT}$ and $PV = Se^{-rT}$. With continuous compounding the following relationships hold:

- With no costs or benefits of holding the underlying asset, the no-arbitrage price of a forward that settles at time T is $S_0 e^{rT}$, where r is the stated annual risk-free rate with continuous compounding.
- With storage costs at a continuously compounded annual rate of c , the no-arbitrage forward price until time T is $S_0 e^{(r+c)T}$.
- With benefits, such as a dividend yield, expressed at a continuously compounded annual rate of b , the no-arbitrage forward price is until time T is $S_0 e^{(r+c-b)T}$.

EXAMPLE: No-arbitrage price with continuous compounding

Consider a stock index trading at 1,550 with a dividend yield of 1.3% (continuously compounded rate) when the risk-free rate is 3% (continuously compounded rate). Calculate the no-arbitrage 6-month forward price of the stock index.

Answer:

The no-arbitrage price of a long 6-month forward is $1,550 \times e^{(0.03 - 0.013)(0.5)} = 1,563.23$.

The **net cost of carry** (or simply **cost of carry** or **carry**) is the benefits of holding the asset minus the costs of holding the asset (including the opportunity cost of funds, R_f). When the benefits (cash flow yield or convenience yield) exceed the costs (including the opportunity cost of funds) of holding the asset, the forward price will be less than the spot price.

Forward Contracts on Currencies

Recall from Economics that we defined the no-arbitrage price of a forward on a currency as the forward price that satisfies the equality:

$$\text{forward exchange rate (p/b)} = \frac{1 + \text{interest rate}_{\text{price currency}}}{1 + \text{interest rate}_{\text{base currency}}} \times \text{spot exchange rate}$$

We can use this no-arbitrage forward rate to examine how an arbitrage profit can be made when the exchange rate in a forward contract is greater or less than the no-arbitrage forward exchange rate. The forward exchange rate depends on the spot exchange rate and the *difference* between the interest rates on the base and price currencies.

Consider a situation at $t = 0$ where the risk-free rate in euros is 3%, the risk-free rate in U.S. dollars is 2%, and the current USD/EUR exchange rate is 1.10. We will examine the arbitrage transactions that establish this relationship by looking at the trades for an investor based in the United States that seeks to profit from the higher interest rate on euros. The investor borrows 100 USD for one year at 2%, exchanges the USD for euros, invests the euros for one year at 3%, and then exchanges the resulting euros for USD. At the end of one year the arbitrageur will have $100/1.10 \times 1.03 = 93.64$ euros and owe $100(1.02) = 102$ USD.

As these transactions have no net cost, there should be no gain from this transaction relative to simply investing the USD for one year at 2%. If this is the case, the 93.64 euros should equal 102 USD. This is the case if the exchange rate at the end of the year is $102/93.64$, which equals a USD/EUR exchange rate of 1.0893. This is the no-arbitrage forward rate. From the formula we saw in Economics we can arrive at the same solution by $1.10 \times (1.02/1.03) = 1.0893$.

If the arbitrageur has a forward contract to buy USD with a price of $1/1.0893 = 0.9180$ euros, he can exchange the 93.64 euros for $93.64/0.9180 = 102$ USD, which is the amount owed on the original loan of 100 USD. The depreciation of the euro in the forward price just offsets the higher euro interest, and the arbitrage transaction returns zero. With a forward exchange rate greater than 1.0893, the arbitrage would have a profit, and with a forward exchange rate less than 1.0893, an arbitrageur could profit from the opposite transactions.

If we convert the effective annual rates to equivalent stated annual rates with continuous compounding, we get $R_{USD} = \ln 1.02 = 1.98\%$ and $R_{EUR} = \ln 1.03 = 2.96\%$. In this case we can say: Forward exchange rate = $1.10 \times e^{(0.0198-0.0296)} = 1.0893$.



MODULE QUIZ 51.1

1. Derivatives pricing models use the risk-free rate to discount future cash flows because these models:
 - A. are based on portfolios with certain payoffs.
 - B. assume that derivatives investors are risk-neutral.
 - C. assume that risk can be eliminated by diversification.
2. Arbitrage prevents:
 - A. market efficiency.
 - B. earning returns higher than the risk-free rate of return.
 - C. two assets with identical payoffs from selling at different prices.
3. The underlying asset of a derivative is *most likely* to have a convenience yield when the asset:
 - A. is difficult to sell short.
 - B. pays interest or dividends.
 - C. must be stored and insured.
4. An investor can replicate a forward on a stock that pays no dividends by:
 - A. selling the underlying short and investing the proceeds at the risk-free rate.
 - B. buying the underlying in the spot market and holding it.
 - C. borrowing at the risk-free rate to buy the underlying.
5. The forward price of a commodity will *most likely* be equal to the current spot price if the:
 - A. convenience yield equals the storage costs as a percentage.
 - B. convenience yield is equal to the risk-free rate plus storage costs as a percentage.
 - C. risk-free rate equals the storage costs as a percentage minus the convenience yield.

KEY CONCEPTS

LOS 51.a

Valuation of derivative securities is based on a no-arbitrage condition. When the forward price is too high, the arbitrage is to sell the forward and buy the underlying asset. When the forward price is too low, the arbitrage is to buy the forward and sell short the underlying asset. Arbitrage will move the forward price toward its no-arbitrage level.

Replication refers to creating a portfolio with cash market transactions that has the same payoffs as a derivative for all possible future values of the underlying. Replication allows us to calculate the no-arbitrage forward price of an asset.

LOS 51.b

Assuming no costs or benefits of holding the underlying asset, the forward price that will prevent arbitrage is the spot price compounded at the risk-free rate over the time until expiration.

The cost of carry is the benefits of holding the asset minus the costs of holding the asset.

Greater costs of holding an asset increase its no-arbitrage forward price.

Greater benefits of holding an asset decrease its no-arbitrage forward price.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 51.1

1. **A** Derivatives pricing models use the risk-free rate to discount future cash flows because they are based on arbitrage relationships that are theoretically riskless. (LOS 51.a)
2. **C** Arbitrage forces two assets with the same expected future value to sell for the same current price. (LOS 51.a)
3. **A** Convenience yield refers to nonmonetary benefits from holding an asset. One example of convenience yield is the advantage of owning an asset that is difficult to sell short when it is perceived to be overvalued. Interest and dividends are monetary benefits. Storage and insurance are carrying costs. (LOS 51.b)
4. **C** Borrowing S_0 at R_f to buy the underlying asset at S_0 has a zero cost and pays the spot price of the underlying asset minus the loan repayment of at time $= T$ of $S_0(1 + R_f)^T$, which is the same payoff as a long forward at $F_0 = S_0(1 + R_f)^T$, the no-arbitrage forward price. (LOS 51.a)
5. **B** When the opportunity cost of funds (R_f) and storage costs just offset the benefits of holding the commodity, the no-arbitrage forward price is equal to the current spot price of the underlying commodity. (LOS 51.b)

READING 52

PRICING AND VALUATION OF FORWARD CONTRACTS AND FOR AN UNDERLYING WITH VARYING MATURITIES

MODULE 52.1: FORWARD CONTRACT VALUATION



LOS 52.a: Explain how the value and price of a forward contract are determined at initiation, during the life of the contract, and at expiration.

Video covering this content is available online.

Consider a forward contract that is initially priced at its no-arbitrage value of $F_0(T) = S_0(1 + R_f)^T$. At initiation, the value of such a forward is:

$$V_0(T) = S_0 - F_0(T)(1 + R_f)^{-T} = 0.$$

At any time during its life, the value of the forward contract to the buyer will be $V_t(T) = S_t - F_0(T)(1 + R_f)^{-(T-t)}$. This is simply the current spot price of the asset minus the present value of the forward contract price.

This value can be realized by selling the asset short at S_t and investing $F_0(T)(1 + R_f)^{-(T-t)}$ in a pure discount bond at R_f . These transactions end any exposure to the forward; at settlement, the proceeds of the bond will cover the cost of the asset at the forward price, and the asset can be delivered to cover the short position.

At expiration, time T , the value of a forward to the buyer is $= S_T - F_0(T)(1 + R_f)^{-(T-T)} = S_T - F_0(T)$. The long buys an asset valued at S_T for the forward contract price of $F_0(T)$, gaining if $S_T > F_0(T)$, losing if $S_T < F_0(T)$. If the forward buyer has a gain, the forward seller has an equal loss, and vice versa.

In the more general case, when there are costs and benefits of holding the underlying asset, the value of a forward to the buyer at time $t < T$ is:

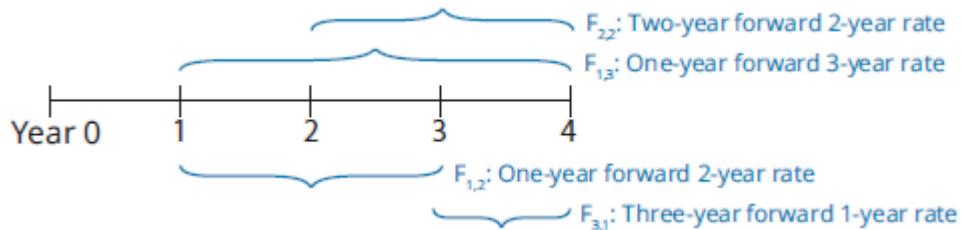
$$V_t(T) = [S_t + PV_t(\text{costs}) - PV_t(\text{benefit})] - F_0(T)(1 + R_f)^{-(T-t)}$$

LOS 52.b: Explain how forward rates are determined for an underlying with a term structure and describe their uses.

Forward rates are yields for future periods. The rate of interest on a 1-year loan to be made two years from today is a forward rate.

The notation for forward rates must identify both the length of the loan period and how far in the future the money will be loaned (or borrowed). 1y1y or $F_{1,1}$ is the rate for a 1-year loan one year from now; 2y1y or $F_{2,1}$ is the rate for a 1-year loan to be made two years from now; the 2-year forward rate three years from today is 3y2y or $F_{3,2}$; and so on.

Exhibit 52.1: Forward Rates



For money market rates the notation is similar, with 3m6m denoting a 6-month rate three months in the future.

Recall that spot rates are zero-coupon rates. We will denote the YTM (with annual compounding) on a zero-coupon bond maturing in n years as Z_n .

An **implied forward rate** is the forward rate for which the following two strategies have the same yield over the total period:

- Investing from $t = 0$ to the forward date, and rolling over the proceeds for the period of the forward.
- Investing from $t = 0$ until the end of the forward period.

As an example, lending for two years at Z_2 would produce the same ending value as lending for one year at Z_1 and, at $t = 1$, lending the proceeds of that loan for one year at $F_{1,1}$. That is, $(1 + Z_2)^2 = (1 + Z_1)(1 + F_{1,1})$. When this condition holds, $F_{1,1}$ is the implied (no-arbitrage) forward rate.

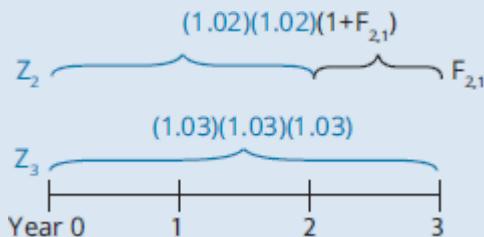
EXAMPLE: Implied forward rate

Consider two zero-coupon bonds, one that matures in two years and one that matures in three years, when $Z_2 = 2\%$ and $Z_3 = 3\%$. Calculate the implied 1-year forward rate two years from now, $F_{2,1}$.

Answer:

As illustrated in Exhibit 52.2, lending for three years at Z_3 should be equivalent to lending for two years at Z_2 and then for the third year at $F_{2,1}$.

Exhibit 52.2: Implied Forward Rate



Lending \$100 for two years at Z_2 (2%) results in a payment of $\$100(1.02)^2 = \104.04 at $t = 2$, while lending \$100 for three years at Z_3 (3%) results in a payment of $\$100(1.03)^3 = \109.27 . The forward interest rate $F_{2,1}$ must be $109.27/104.04 - 1 = 5.03\%$, the implied forward rate from $t = 2$ to $t = 3$.

An example of an interest rate derivative is a **forward rate agreement (FRA)**, in which the fixed-rate payer (long) will pay the forward rate on a notional amount of principal at a future date, and the floating-rate payer will pay a future reference rate times that same amount of principal. In practice, only the net amount is exchanged.

Consider a 3-month forward on a 6-month MRR (F_{3m6m}) with a notional principal of \$1 million. At settlement in three months, the buyer receives (or pays) the present value of (realized 6-month MRR – 1%)/2 × \$1 million.

We divide by 2 because MRRs are typically annualized rates. We take the present value of the difference in interest because the settlement payment is at the beginning of the 6-month period, whereas the interest savings would be at the end of the period.

Assume that the current 3-month MRR is 1.0% and 9-month MRR is 1.2%. Adjusting for periodicity, the no-arbitrage condition for the value of F_{3m6m} is:

$$1 + 0.012\left(\frac{9}{12}\right) = \left[1 + 0.01\left(\frac{3}{12}\right)\right] \left[1 + F_{3m,6m}\left(\frac{6}{12}\right)\right]$$

The implied forward rate, F_{3m6m} , as an annualized rate, is:

$$F_{3m,6m} = \left[\frac{1 + 0.012\left(\frac{9}{12}\right)}{1 + 0.01\left(\frac{3}{12}\right)} - 1 \right] \times \frac{12}{6} = 0.013$$

Now let's examine the payoff to the fixed-rate payer in an F_{3m6m} FRA with a notional principal of \$1 million when the 6-month MRR three months from now is 1.5%. Because the realized 6-month MRR is greater than the forward rate, the fixed-rate payer (floating-rate receiver) will have a gain.

The payment to the fixed-rate payer is the present value (discounted at 6-month MRR) of the interest differential between two 6-month loans, one at 1.3% and one at 1.5% (both annualized rates). The fixed-rate payer in the FRA receives:

$$\$1 \text{ million} \times \left(\frac{0.015 - 0.013}{2}\right) \left(\frac{1}{\frac{0.015}{2}}\right) = \$992.56.$$

FRAs are used primarily by financial institutions to manage the volatility of their interest-sensitive assets and liabilities. FRAs are also the building blocks of interest rate swaps over multiple periods. An FRA is equivalent to a single-period swap. Multiple-period swaps are used primarily by investors and issuers to manage interest rate risk.



MODULE QUIZ 52.1

- Two parties agree to a forward contract to exchange 100 shares of a stock one year from now for \$72 per share. Immediately after they initiate the contract, the price of the underlying stock increases to \$74 per share. This share price increase represents a gain for:
 - the buyer.
 - the seller.

- C. neither the buyer nor the seller.
2. The forward rate $F_{2,3}$ represents the interest rate on a loan for the period from:
 - A. year 2 to year 3.
 - B. year 2 to year 5.
 - C. year 3 to year 5.
 3. Given zero-coupon bond yields for 1, 2, and 3 years, an analyst can *least likely* derive an implied:
 - A. 1-year forward 1-year rate.
 - B. 2-year forward 1-year rate.
 - C. 2-year forward 2-year rate.

KEY CONCEPTS

LOS 52.a

The value of a forward contract at initiation is zero.

During its life, the value of a forward contract to the buyer is the spot price of the asset minus the present value of the forward contract price, and the value to the seller is the present value of the forward contract price minus the spot price of the asset.

At expiration, the value of a forward contract to the buyer is the spot price of the asset minus the forward contract price, and the value to the seller is the forward contract price minus the spot price of the asset.

LOS 52.b

An implied forward rate is the forward rate for which the following two strategies have the same yield over the total period:

- Investing from $t = 0$ to the forward date, and rolling over the proceeds for the period of the forward.
- Investing from $t = 0$ until the end of the forward period.

In a forward rate agreement (FRA), the fixed-rate payer (long) will pay the forward rate on a notional amount of principal at a future date, and the floating-rate payer will pay a future reference rate times that same amount of principal. FRAs are used primarily by financial institutions to manage the volatility of their interest-sensitive assets and liabilities.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 52.1

1. A If the value of the underlying is greater than the forward price, this increases the value of the forward contract, which represents a gain for the buyer and a loss for the seller. (LOS 52.a)
2. B $F_{2,3}$ is the 2-year forward 3-year rate, covering a period that begins two years from now and extends for three years after that. (LOS 52.b)

3. C The forward rate $F_{2,2}$ extends four years into the future and cannot be derived using zero-coupon yields that only extend three years. From zero-coupon bond yields for 1, 2, and 3 years, we can derive implied forward rates $F_{1,1}$, $F_{1,2}$, and $F_{2,1}$. (LOS 52.b)

READING 53

PRICING AND VALUATION OF FUTURES CONTRACTS

MODULE 53.1: FUTURES VALUATION



LOS 53.a: Compare the value and price of forward and futures contracts.

Video covering this content is available online.

While the *price* of a forward contract is constant over its life when no mark-to-market gains or losses are paid, its *value* will fluctuate with changes in the value of the underlying. The payment at settlement of the forward reflects the difference between the (unchanged) forward price and the spot price of the underlying.

The price and value of a futures contract *both* change when daily mark-to-market gains and losses are settled. Consider a futures contract on 100 ounces of gold at \$1,870 purchased on Day 0. The following illustrates the changes in contract price and value with daily mark-to-market payments.

Day 0	Price = Settlement Price of 1,870	MTM Value = 0
Day 1	Settlement Price = 1,875 \$500 addition to margin account	MTM value = \$500
	New futures price = 1,875	MTM value = 0
Day 2	Settlement price = 1,855 \$2,000 deduction from margin account	MTM value = -\$2,000
	New futures price = 1,855	MTM value = 0

The change in the futures price to the settlement price each day returns its value to zero. Prices of forward contracts for which mark-to-market gains and losses are settled daily will also be adjusted to the settlement price.

Interest rate futures contracts are available on many market reference rates. We may view these as exchange-traded equivalents to forward rate agreements. One key difference is that interest rate futures are quoted on a price basis. For a market reference rate from time A to time B, an interest rate futures price is stated as follows:

$$\text{futures price} = 100 - (100 \times \text{MRR}_{A, B-A})$$

For example, if the futures price for a 6-month rate six months from now is 97, then $\text{MRR}_{6m, 6m} = 3\%$.

Like other futures contracts, interest rate futures are subject to daily mark-to-market. The **basis point value (BPV)** of an interest rate futures contract is defined as:

$$\text{BPV} = \text{notional principal} \times \text{period} \times 0.01\%$$

If the contract in our example is based on notional principal of €1,000,000, its BPV is $\text{€}1,000,000 \times (0.0001 / 2) = \text{€}50$. This means a one basis point change in the MRR will change the futures contract value by €50.

LOS 53.b: Explain why forward and futures prices differ.

For pricing, the most important distinction between futures and forwards is that with futures, mark-to-market gains and losses are paid each day. Gains above initial margin can be withdrawn from a futures account and losses that reduce margin deposits below their maintenance level require payments into the account. Forwards most often have no mark-to-market cash flows, with gains or losses settled at contract expiration. Forwards typically do not require or provide funds in response to fluctuations in value during their lives.

If interest rates are constant or uncorrelated with futures prices over time, the prices of futures and forwards are the same. A positive correlation between interest rates and the futures price means that (for a long position) daily settlement provides funds (excess margin) when rates are high and they can earn more interest, and requires funds (margin deposits) when rates are low and opportunity cost of deposited funds is less. Because of this, futures are theoretically more attractive than forwards when interest rates and futures prices are positively correlated, and less attractive than forwards when interest rates and futures prices are negatively correlated.

Because of the short maturity of most forwards and the availability of funds at near risk-free rates, differences between equivalent forwards and futures are not observed in practice.

Additionally, derivative dealers in some markets with central clearing are required to post margin and may require derivative investors to post mark-to-market margin payments as well.

A separate issue arises for interest rate forwards and futures settlement payments. Recall that the payoff on an interest rate forward is the present value (at the beginning of the forward period) of any interest savings (at the end of the forward period) from the difference between the realized MRR and the forward MRR. Because the realized MRR is the discount rate for calculating the payment for a given amount of future interest savings, the payment for an increase in the MRR will be less than the payment for an equal decrease in the MRR, as the following example will illustrate.

Consider a \$1 million interest rate future on a 6-month MRR priced at 97.50 (an MRR of 2.5%) that settles six months from now. Each basis point change in the (annualized) MRR will change the value of the contract by $0.0001 \times 6/12 \times \$1 \text{ million} = \$50$. If the MRR at settlement is either 2.51% or 2.49%, the payoff on the future at the end of one year is either \$50 higher or \$50 lower than when the MRR at settlement is 2.5%.

Compare this result with the payoffs for an otherwise equivalent forward, F_{6m6m} , priced at 2.5%.

If the MRR at settlement is 2.51%, the long receives $50/(1 + 0.0251/2) = \$49.3803$.

If the MRR at settlement is 2.49%, the long must pay $50/(1 + 0.0249/2) = \$49.3852$.

The value of the forwards exhibit convexity. An increase in rates decreases the forward's value by less than a decrease in the interest rate increases the forward's value, just as we saw with bonds. Also just as with bonds, the convexity effect for the value of forwards increases for longer periods. The convexity of forwards is termed **convexity bias** and forwards and futures prices can be significantly different for longer-term interest rates.



MODULE QUIZ 53.1

1. For a forward contract on an asset that has no costs or benefits from holding it to have zero value at initiation, the arbitrage-free forward price must equal:
 - A. the expected future spot price.
 - B. the future value of the current spot price.
 - C. the present value of the expected future spot price.
2. For a futures contract to be more attractive than an otherwise equivalent forward contract, interest rates must be:
 - A. uncorrelated with futures prices.
 - B. positively correlated with futures prices.
 - C. negatively correlated with futures prices.

KEY CONCEPTS

LOS 53.a

For a forward contract on which no mark-to-market gains or losses are paid, the forward price is constant over its life, but the contract's value will fluctuate with changes in the value of the underlying.

For a futures contract, the price and value both change when daily mark-to-market gains and losses are settled. The change in the futures price to the settlement price each day returns its value to zero.

Unlike forward rate agreements, interest rate futures are quoted on a price basis:

$$\text{futures price} = 100 - (100 \times \text{MRR}_{A, B-A})$$

LOS 53.b

Because gains and losses on futures contracts are settled daily, prices of forwards and futures that have the same terms may be different if interest rates are correlated with futures prices. Futures are more valuable than forwards when interest rates and futures prices are positively correlated and less valuable when they are negatively correlated. If interest rates are constant or uncorrelated with futures prices, the prices of futures and forwards are the same.

Convexity bias can result in price differences between interest rate futures contracts and otherwise equivalent forward rate agreements.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 53.1

- 1. B** For an asset with no holding costs or benefits, the forward price must equal the future value of the current spot price, compounded at the risk-free rate over the term of the forward contract, for the contract to have a value of zero at initiation. Otherwise an arbitrage opportunity would exist. (LOS 53.a)
- 2. B** If interest rates are positively correlated with futures prices, interest earned on cash from daily settlement gains on futures contracts will be greater than the opportunity cost of interest on daily settlement losses, and a futures contract is more attractive than an otherwise equivalent forward contract that does not feature daily settlement. (LOS 53.b)

READING 54

PRICING AND VALUATION OF INTEREST RATES AND OTHER SWAPS

MODULE 54.1: SWAP VALUATION



LOS 54.a: Describe how swap contracts are similar to but different from a series of forward contracts.

Video covering this content is available online.

In a simple interest-rate swap, one party pays a floating rate and the other pays a fixed rate on a notional principal amount. Consider a 1-year swap with quarterly payments, one party paying a fixed rate and the other a floating rate equal to a 90-day market reference rate (MRR). At each payment date the difference between the swap fixed rate and the MRR is paid to the party that owes the least, that is, a net payment is made from one party to the other.

We can separate these payments into a known payment and three unknown payments that are equivalent to the payments on three forward rate agreements (FRAs). Let MRR_n represent the floating rate payment (based on the 90-day MRR) owed at the end of quarter n and F be the fixed payment owed at the end of each quarter. We can represent the swap payment to be received by the fixed-rate payer at the end of period n as $MRR_n - F$. We can replicate each of these payments to (or from) the fixed-rate payer in the swap with a forward contract, specifically a long position in an FRA with a contract rate equal to the swap fixed rate and a settlement value based on the 90-day MRR.

We illustrate this separation below for a 1-year fixed-for-floating swap with a fixed rate of F and floating-rate payments for period n of MRR_n . Note that if the fixed rate and MRR are quoted as annual rates, the payments will be $(MRR_n - F)$ times one-fourth of the notional principal.

First payment (90 days from now) = $MRR_1 - F$ which is known at time zero because the payment 90 days from now is based on the 90-day MRR at time zero and the swap fixed rate, F , both of which are known at the initiation of the swap.

Second payment (180 days from now) is equivalent to a long position in an FRA with contract rate F that settles in 180 days and pays $MRR_2 - F$.

Third payment (270 days from now) is equivalent to a long position in an FRA with contract rate F that settles in 270 days and pays $MRR_3 - F$.

Fourth payment (360 days from now) is equivalent to a long position in an FRA with contract rate F that settles in 360 days and pays $MRR_4 - F$.

Note that a forward on a 90-day MRR that settles 90 days from now, based on the 90-day MRR at that time, actually pays the present value of the difference between the fixed rate F and the

90-day MRR 90 days from now (times the notional principal amount). Thus, the forwards in our example actually pay on days 90, 180, and 270. However, the amounts paid are equivalent to the differences between the fixed-rate payment and floating-rate payment that are due when interest is actually due on days 180, 270, and 360, which are the amounts we used in the example.

Therefore, we can describe an interest-rate swap as equivalent to a series of forward contracts, specifically FRAs, each with a forward contract rate equal to the swap fixed rate. However, there is one important difference. Because the forward contract rates are all equal in the FRAs that are equivalent to the swap, these would not be zero-value forward contracts at the initiation of the swap. Recall that forward contracts are based on a contract rate for which the value of the forward contract at initiation is zero. There is no reason to suspect that the swap fixed rate results in a zero value forward contract for each of the future dates. Instead, a swap is most likely to consist of some forwards with positive values and some forwards with negative values. The sum of their values will equal zero at initiation.

Finding the swap fixed rate that gives the swap a zero value at initiation, which is also known as the **par swap rate**, is not difficult if we follow our principle of no-arbitrage pricing. The fixed rate payer in a swap can replicate that derivative position by borrowing at a fixed rate and lending the proceeds at a variable (floating) rate. For the swap in our example, borrowing at the fixed rate F and lending the proceeds at the 90-day MRR will produce the same cash flows as the swap. At each date, the payment due on the fixed-rate loan is F_n and the interest received on lending at the floating rate is MRR_n .

LOS 54.b: Contrast the value and price of swaps.

As with FRAs, the *price* of a swap is the fixed rate of interest specified in the swap contract (the par swap rate) and the *value* depends on how expected future floating rates change over time. At initiation, a swap has zero value because the present value of the fixed-rate payments equals the present value of the expected floating-rate payments.

We can solve for the no-arbitrage fixed rate, termed the **par swap rate**, from the following equality:

$$\frac{MRR_1}{1 + S_1} + \frac{MRR_2}{(1 + S_2)^2} + \frac{MRR_3}{(1 + S_3)^3} + \frac{MRR_4}{(1 + S_4)^4} = \frac{F}{1 + S_1} + \frac{F}{(1 + S_2)^2} + \frac{F}{(1 + S_3)^3} + \frac{F}{(1 + S_4)^4}$$

©2022 Kaplan, Inc.

where S_1 through S_4 are the current effective spot rates for 90, 180, 270, and 360 days, MRR_1 through MRR_4 are the forward 90-day rates implied by the spot rates, and F is the fixed rate payment.

Given the current spot rates (S_1 to S_4), we can calculate the implied (expected) forward rates (MRRs), and then solve for F , the fixed rate that will give the swap a value of zero.

An increase in expected future 90-day rates will produce an increase the value of the fixed-rate payer position in a swap, while a decrease in expected rates will decrease the value of that position. At any point in time, the value of the fixed-rate payer side of a swap will be the present value of the expected future floating-rate payments, minus the present value of the future fixed-rate payments. This calculation is based on the spot rates and implied future 90-day rates at that point in time and can be used for any required mark-to-market payments.



MODULE QUIZ 54.1

1. Which of the following is *most* similar to the floating-rate receiver position in a fixed-for-floating interest-rate swap?
 - A. Buying a fixed-rate bond and a floating-rate note.
 - B. Buying a floating-rate note and issuing a fixed-rate bond.
 - C. Issuing a floating-rate note and buying a fixed-rate bond.
2. The price of a fixed-for-floating interest-rate swap:
 - A. is specified in the swap contract.
 - B. is paid at initiation by the floating-rate receiver.
 - C. may increase or decrease during the life of the swap contract.

KEY CONCEPTS

LOS 54.a

In a simple interest-rate swap, one party pays a floating rate and the other pays a fixed rate on a notional principal amount. The first payment is known at initiation and the rest of the payments are unknown. The unknown payments are equivalent to the payments on FRAs. The par swap rate is the fixed rate at which the sum of the present values of these FRAs equals zero.

LOS 54.b

The price of a swap is the fixed rate of interest specified in the swap contract. The value depends on how expected future floating rates change over time. An increase in expected future short-term future rates will increase the value of the fixed-rate payer position in a swap, and a decrease in expected future rates will decrease the value of the fixed-rate payer position.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 54.1

1. **B** The floating-rate receiver (fixed-rate payer) in a fixed-for-floating interest-rate swap has a position similar to issuing a fixed-coupon bond and buying a floating-rate note. (LOS 54.a)
2. **A** The price of a fixed-for-floating interest-rate swap is defined as the fixed rate specified in the swap contract. Typically a swap will be priced such that it has a value of zero at initiation and neither party pays the other to enter the swap. (LOS 54.b)

READING 55

PRICING AND VALUATION OF OPTIONS

MODULE 55.1: OPTION VALUATION



LOS 55.a: Explain the exercise value, moneyness, and time value of an option.

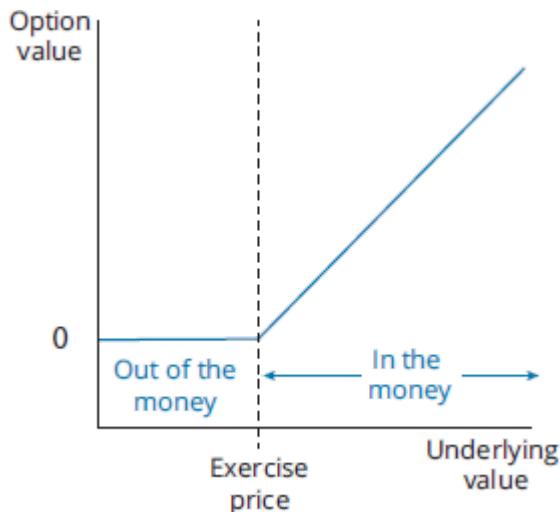
Video covering this content is available online.

Moneyness refers to whether an option is *in the money* or *out of the money*. If immediate exercise of the option would generate a positive payoff, it is in the money. If immediate exercise would result in a loss (negative payoff), it is out of the money. When the current asset price equals the exercise price, exercise will generate neither a gain nor loss, and the option is *at the money*.

The following describes the conditions for a **call option** to be in, out of, or at the money. S is the price of the underlying asset and X is the exercise price of the option.

- *In-the-money call options.* If $S - X > 0$, a call option is in the money. $S - X$ is the amount of the payoff a call holder would receive from immediate exercise, buying a share for X and selling it in the market for a greater price S .
- *Out-of-the-money call options.* If $S - X < 0$, a call option is out of the money.
- *At-the-money call options.* If $S = X$, a call option is said to be at the money.

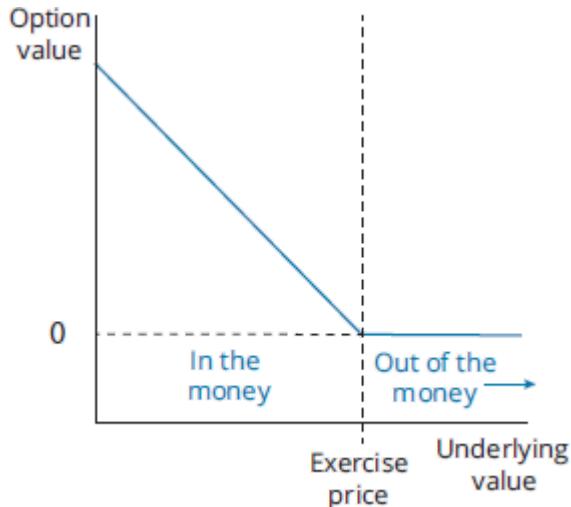
Figure 55.1: Call Option Moneyness



The following describes the conditions for a **put option** to be in, out of, or at the money.

- *In-the-money put options.* If $X - S > 0$, a put option is in the money. $X - S$ is the amount of the payoff from immediate exercise, buying a share for S and exercising the put to receive X for the share.
- *Out-of-the-money put options.* When the stock's price is greater than the exercise price, a put option is said to be out of the money. If $X - S < 0$, a put option is out of the money.
- *At-the-money put options.* If $S = X$, a put option is said to be at the money.

Figure 55.2: Put Option Moneyness



EXAMPLE: Moneyness

Consider a July 40 call and a July 40 put, both on a stock that is currently selling for \$37/share. Calculate how much these options are in or out of the money.



PROFESSOR'S NOTE

A July 40 call is a call option with an exercise price of \$40 and an expiration date in July.

Answer:

The call is \$3 out of the money because $S - X = -\$3.00$. The put is \$3 in the money because $X - S = \$3.00$.

We define the **exercise value** (or **intrinsic value**) of an option as the maximum of zero and the amount that the option is in the money. That is, the exercise value is the amount an option is in the money, if it is in the money, or zero if the option is at or out of the money. The exercise value is the value of the option if exercised immediately.

Prior to expiration, an option has **time value** in addition to its exercise value. The time value of an option is the amount by which the **option premium** (price) exceeds the exercise value and is sometimes called the *speculative value* of the option. This relationship can be written as:

$$\text{option premium} = \text{exercise value} + \text{time value}$$

At any point during the life of an option, its value will typically be greater than its exercise value. This is because there is some probability that the underlying asset price will change in an amount that gives the option a positive payoff at expiration greater than the current exercise

value. Recall that an option's exercise value (to a buyer) is the amount of the payoff at expiration and has a lower bound of zero.

When an option reaches expiration, there is no time remaining and the time value is zero. This means the value at expiration is either zero, if the option is at or out of the money, or its exercise value, if it is in the money.

LOS 55.b: Contrast the use of arbitrage and replication concepts in pricing forward commitments and contingent claims.

To model forward commitments, we used no-arbitrage pricing based on an initial value of zero to both parties. With options, however, the initial values of options are positive; the buyer pays a premium (the option price) to the writer (seller). Another difference is that while forward commitments have essentially unlimited gains or losses for both parties (except to the extent that prices are constrained by zero), options are one-sided: Potential losses for the buyer, and potential gains for the writer, are limited to the premium paid. For these reasons, the no-arbitrage approach we use for pricing contingent claims is different from the model we use for forward commitments.

The following is some terminology that we will use to determine the minimum and maximum values for European options:

S_t = the price of the underlying stock at time t

X = the exercise price of the option

$T-t$ = the time to expiration

c_t = the price of a European call at any time t prior to expiration at time $= T$

p_t = the price of a European put at any time t prior to expiration at time $= T$

R_f = the risk-free rate

Upper Bound for Call Options

The maximum value of a European call option at any time t is the time- t share price of the underlying stock. This makes sense because no one would pay more for the right to buy an asset than the asset's market value. It would be cheaper simply to buy the underlying asset. At time $t = 0$, the upper boundary condition for European call options is $c_0 \leq S_0$, and at any time t during a European call option's life, the upper boundary condition is $c_t \leq S_t$.

Upper Bound for Put Options

Logically the value of a put option cannot be more than its exercise price. This would be its exercise value if the underlying stock price goes to zero. However, because European puts cannot be exercised prior to expiration, their maximum value is the *present value* of the exercise price discounted at the risk-free rate. Even if the stock price goes to zero and is expected to stay at zero, the put buyer will not receive the intrinsic value, X , until the expiration date.

At time $t = 0$, the upper boundary condition can be expressed for European put options as:

$$p_0 \leq X (1 + R_f)^{-T}$$

At any time t during a European put option's life, the upper boundary condition is:

$$p_t \leq X (1 + R_f)^{-(T-t)}$$

Lower Bounds for Options

Theoretically, no option will sell for less than its intrinsic value and no option can take on a negative value. For European options, however, the lower bound is not so obvious because these options are not exercisable immediately.

To determine the lower bounds for European options, we can examine the value of a portfolio in which the option is combined with a long or short position in the stock and a pure discount bond.

For a *European call option*, construct the following portfolio:

- A long at-the-money European call option with exercise price X , expiring at time T .
- A long discount bond priced to yield the risk-free rate that pays X at option expiration.
- A short position in one share of the underlying stock priced at $S_0 = X$.

The current value of this portfolio is $c_0 - S_0 + X(1 + R_f)^{-T}$.

At expiration time T , this portfolio will pay $c_T - S_T + X$. That is, we will collect $c_T = \max[0, S_T - X]$ on the call option, pay S_T to cover our short stock position, and collect X from the maturing bond.

- If $S_T \geq X$, the call is in-the-money, and the portfolio will have a zero payoff because the call pays $S_T - X$, the bond pays $+X$, and we pay $-S_T$ to cover our short position. That is, the time $t = T$ payoff is: $S_T - X + X - S_T = 0$.
- If $S_T < X$, the call is out-of-the-money, and the portfolio has a positive payoff equal to $X - S_T$ because the call value, c_T , is zero, we collect X on the bond, and pay $-S_T$ to cover the short position. So, the time $t = T$ payoff is: $0 + X - S_T = X - S_T$.

No matter whether the option expires in-the-money, at-the-money, or out-of-the-money, the portfolio value will be equal to or greater than zero. We will never have to make a payment.

To prevent arbitrage, any portfolio that has no possibility of a negative payoff cannot have a negative value. Thus, we can state the value of the portfolio *at time t = 0* as:

$$c_0 - S_0 + X(1 + R_f)^{-T} \geq 0$$

which allows us to conclude that:

$$c_0 \geq S_0 - X(1 + R_f)^{-T}$$

Combining this result with the earlier minimum on the call value of zero, we can write:

$$c_0 \geq \max[0, S_0 - X(1 + R_f)^{-T}]$$

Note that $X(1 + R_f)^{-T}$ is the present value of a pure discount bond with a face value of X .

For a *European put option* we can derive the minimum value by forming the following portfolio at time $t = 0$:

- A long at-the-money European put option with exercise price X , expiring at T .
- A short position on a risk-free bond priced at $X(1 + R_f)^{-T}$, equivalent to borrowing $X(1 + R_f)^{-T}$.
- A long position in a share of the underlying stock priced at S_0 .

At expiration time T , this portfolio will pay $p_T + S_T - X$. That is, we will collect $p_T = \text{Max}[0, X - S_T]$ on the put option, receive S_T from the stock, and pay X on the bond (loan).

- If $S_T > X$, the payoff will equal: $p_T + S_T - X = S_T - X$.
- If $S_T \leq X$, the payoff will be zero.

Again, a no-arbitrage argument can be made that the portfolio value must be zero or greater, because there are no negative payoffs to the portfolio.

At time $t = 0$, this condition can be written as:

$$p_0 + S_0 - X(1 + R_f)^{-T} \geq 0$$

and rearranged to state the minimum value for a European put option at time $t = 0$ as:

$$p_0 \geq X(1 + R_f)^{-T} - S_0$$

We have now established the minimum bound on the price of a European put option as:

$$p_0 \geq \text{Max}[0, X(1 + R_f)^{-T} - S_0]$$

Figure 55.1 summarizes what we now know regarding the boundary prices for European options at any time t prior to expiration at time $t = T$.

Figure 55.1: Lower and Upper Bounds for Options

Option	Minimum Value	Maximum Value
European call	$c_t \geq \text{Max}[0, S_t - X(1 + R_f)^{-(T-t)}]$	S_t
European put	$p_t \geq \text{Max}[0, X / X(1 + R_f)^{-(T-t)} - S_t]^{(T-t)}$	$X(1 + R_f)^{-(T-t)}$



PROFESSOR'S NOTE:

For the exam, know the price limits in Figure 55.1. You will not be asked to derive them, but you may be expected to use them.

LOS 55.c: Identify the factors that determine the value of an option and describe how each factor affects the value of an option.

There are six factors that determine option prices.

1. Price of the underlying asset. For call options, the higher the price of the underlying, the greater its exercise value and the higher the value of the option. Conversely, the lower the price of the underlying, the less its exercise value and the lower the value of the call option. In general, call option values increase when the value of the underlying asset increases.

For put options this relationship is reversed. An increase in the price of the underlying reduces the value of a put option.

2. The exercise price. A higher exercise price decreases the values of call options and a lower exercise price increases the values of call options.

A higher exercise price increases the values of put options and a lower exercise price decreases the values of put options.

3. The risk-free rate of interest. An increase in the risk-free rate will increase call option values, and a decrease in the risk-free rate will decrease call option values.

An increase in the risk-free rate will decrease put option values, and a decrease in the risk-free rate will increase put option values.



PROFESSOR'S NOTE

One way to remember the effects of changes in the risk-free rate is to think about present values of the payments for calls and puts. These statements are strictly true only for in-the-money options, but it's a way to remember the relationships. The holder of a call option will pay in the future to exercise a call option and the present value of that payment is lower when the risk-free rate is higher, so a higher risk-free rate increases a call option's value. The holder of a put option will receive a payment in the future when the put is exercised and an increase in the risk-free rate decreases the present value of this payment, so a higher risk-free rate decreases a put option's value.

4. Volatility of the underlying. Volatility is what makes options valuable. If there were no volatility in the price of the underlying asset (its price remained constant), options would always be equal to their exercise values and time or speculative value would be zero. An increase in the volatility of the price of the underlying asset increases the values of both put and call options and a decrease in volatility of the price of the underlying decreases both put values and call values.

5. Time to expiration. Because volatility is expressed per unit of time, longer time to expiration effectively increases expected volatility and increases the value of a call option. Less time to expiration decreases the time value of a call option so that at expiration its value is simply its exercise value.

For most put options, longer time to expiration will increase option values for the same reasons. For some European put options, however, extending the time to expiration can decrease the value of the put. In general, the deeper a put option is in the money, the higher the risk-free rate, and the longer the current time to expiration, the more likely that extending the option's time to expiration will decrease its value.

To understand this possibility consider a put option at \$20 on a stock with a value that has decreased to \$1. The exercise value of the put is \$19 so the upside is very limited, the downside (if the price of the underlying subsequently increases) is significant, and because no payment will be received until the expiration date, the current option value reflects the present value of any expected payment. Extending the time to expiration would decrease that present value. While overall we expect a longer time to expiration to increase the

value of a European put option, in the case of a deep in-the-money put, a longer time to expiration could decrease its value.

6. Costs and benefits of holding the asset. If there are benefits of holding the underlying asset (dividend or interest payments on securities or a convenience yield on commodities), call values are decreased and put values are increased. The reason for this is most easily understood by considering cash benefits. When a stock pays a dividend, or a bond pays interest, this reduces the value of the asset. Decreases in the value of the underlying asset decrease call values and increase put values.

Positive storage costs make it more costly to hold an asset. We can think of this as making a call option more valuable because call holders can have long exposure to the asset without paying the costs of actually owning the asset. Puts, on the other hand, are less valuable when storage costs are higher.



MODULE QUIZ 55.1

1. The price of an out-of-the-money option is:
 - A. less than its time value.
 - B. equal to its time value.
 - C. greater than its time value.
2. The lower bound for the value of a European put option is:
 - A. $\text{Max}(0, S - X)$
 - B. $\text{Max}[0, X(1 + R_f)^{-(T-t)} - S]$
 - C. $\text{Max}[0, S - X(1 + R_f)^{-(T-t)}]$
3. A decrease in the risk-free rate of interest will:
 - A. increase put and call option prices.
 - B. decrease put option prices and increase call option prices.
 - C. increase put option prices and decrease call option prices.

KEY CONCEPTS

LOS 55.a

If immediate exercise of an option would generate a positive payoff, the option is in the money. If immediate exercise would result in a negative payoff, the option is out of the money. An option's exercise value is the greater of zero or the amount it is in the money. Time value is the amount by which an option's price is greater than its exercise value. Time value is zero at expiration.

LOS 55.b

The approach for pricing contingent claims is different from the model for forward commitments because contingent claims have one-sided payoffs and values at initiation that are not equal to zero. A replication model for European options is based on the value of a portfolio in which the option is combined with a pure discount bond and a long or short position in the underlying.

LOS 55.c

Factors that determine the value of an option:

<u>Increase in:</u>	Effect on Call Option Values	Effect on Put Option Values
Price of underlying asset	Increase	Decrease
Exercise price	Decrease	Increase
Risk-free rate	Increase	Decrease
Volatility of underlying asset	Increase	Increase
Time to expiration	Increase	Increase, except some European puts
Costs of holding underlying asset	Increase	Decrease
Benefits of holding underlying asset	Decrease	Increase

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 55.1

1. **B** Because an out-of-the-money option has an exercise value of zero, its price is its time value. (LOS 55.a)
2. **B** The lower bound for a European put ranges from zero to the present value of the exercise price less the current stock price, where the exercise price is discounted at the risk-free rate. (LOS 55.b)
3. **C** A decrease in the risk-free rate will decrease call option values and increase put option values. (LOS 55.c)

READING 56

OPTION REPLICATION USING PUT–CALL PARITY

MODULE 56.1: PUT–CALL PARITY



LOS 56.a: Explain put–call parity for European options.

Video covering this content is available online.

Our derivation of **put–call parity** for European options is based on the payoffs of two portfolio combinations: a fiduciary call and a protective put.

A *fiduciary call* is a combination of a call with exercise price X and a pure-discount, riskless bond that pays X at maturity (option expiration). The payoff for a fiduciary call at expiration is X when the call is out of the money, and $X + (S - X) = S$ when the call is in the money.

A *protective put* is a share of stock together with a put option on the stock. The expiration date payoff for a protective put is $(X - S) + S = X$ when the put is in the money, and S when the put is out of the money.



PROFESSOR'S NOTE

When working with put–call parity, it is important to note that the exercise prices on the put and the call and the face value of the riskless bond are all equal to X .

If at expiration S is greater than or equal to X :

- The protective put pays S on the stock while the put expires worthless, so the payoff is S .
- The fiduciary call pays X on the bond portion while the call pays $(S - X)$, so the payoff is $X + (S - X) = S$.

If at expiration X is greater than S :

- The protective put pays S on the stock while the put pays $(X - S)$, so the payoff is $S + (X - S) = X$.
- The fiduciary call pays X on the bond portion while the call expires worthless, so the payoff is X .

In either case, the payoff on a protective put is the same as the payoff on a fiduciary call. Our no-arbitrage condition holds that portfolios with identical payoffs regardless of future conditions must sell for the same price to prevent arbitrage. We can express the put–call parity relationship as:

$$c + X(1 + R_f)^{-T} = S + p$$

Equivalencies for each of the individual securities in the put–call parity relationship can be expressed as:

$$S = c - p + X(1 + R_f)^{-T}$$

$$p = c - S + X(1 + R_f)^{-T}$$

$$c = S + p - X(1 + R_f)^{-T}$$

$$X(1 + R_f)^{-T} = S + p - c$$

Note that the options must be European style and the puts and calls must have the same exercise price and time to expiration for these relations to hold.

The single securities on the left-hand side of the equations all have exactly the same payoffs as the portfolios on the right-hand side. The portfolios on the right-hand side are the **synthetic** equivalents of the securities on the left. For example, to synthetically produce the payoff for a long position in a share of stock, use the following relationship:

$$S = c - p + X(1 + R_f)^{-T}$$

This means that the payoff on a long stock can be synthetically created with a long call, a short put, and a long position in a risk-free discount bond.

The other securities in the put–call parity relationship can be constructed in a similar manner.



PROFESSOR'S NOTE

After expressing the put–call parity relationship in terms of the security you want to synthetically create, the sign on the individual securities will indicate whether you need a long position (+ sign) or a short position (- sign) in the respective securities.

EXAMPLE: Call option valuation using put–call parity

Suppose that the current stock price is \$52 and the risk-free rate is 5%. You have found a quote for a 3-month put option with an exercise price of \$50. The put price is \$1.50, but due to light trading in the call options, there was not a listed quote for the 3-month, \$50 call. Estimate the price of the 3-month call option.

Answer:

Rearranging put–call parity, we find that the call price is:

$$\text{call} = \text{put} + \text{stock} - \text{present value (X)}$$

$$\text{call} = \$1.50 + \$52 - \frac{\$50}{1.05^{0.25}} = \$4.11$$

This means that if a 3-month, \$50 call is available, it should be priced at (within transaction costs of) \$4.11 per share.

LOS 56.b: Explain put–call **forward** parity for European options.

Put–call–forward parity is derived with a forward contract rather than the underlying asset itself. Consider a forward contract on an asset at time T with a contract price of $F_0(T)$. At contract initiation the forward contract has zero value. At time T , when the forward contract

settles, the long must purchase the asset for $F_0(T)$. The purchase (at time = 0) of a pure discount bond that will pay $F_0(T)$ at maturity (time = T) will cost $F_0(T)(1 + R_f)^{-T}$.

By purchasing such a pure discount bond and simultaneously taking a long position in the forward contract, an investor has created a synthetic asset. At time = T the proceeds of the bond are just sufficient to purchase the asset as required by the long forward position. Because there is no cost to enter into the forward contract, the total cost of the synthetic asset is the present value of the forward price, $F_0(T)(1 + R_f)^{-T}$.

The put–call-forward parity relationship is derived by substituting the synthetic asset for the underlying asset in the put–call parity relationship. Substituting $F_0(T)(1 + R_f)^{-T}$ for the asset price S_0 in $S + p = c + X(1 + R_f)^{-T}$ gives us:

$$F_0(T)(1 + R_f)^{-T} + p_0 = c_0 + X(1 + R_f)^{-T}$$

which is put–call-forward parity at time 0, the initiation of the forward contract, based on the principle of no arbitrage. By rearranging the terms, put–call forward parity can also be expressed as:

$$p_0 - c_0 = [X - F_0(T)](1 + R_f)^{-T}$$

Application of Options Theory to Corporate Finance

We can view the claims of a firm's equity holders and debt holders as a call option and a put option, respectively. Consider a firm that has a value of V_t at time = t and has issued debt in the form of a zero-coupon bond that will pay D at time = T . At time = T , if $V_T > D$ the equity holders receive $V_T - D$ and if $V_T < D$, the firm is insolvent and equity holders receive nothing. The payoff to the equity holders at time = T can be written as $\text{Max}(0, V_T - D)$ which is equivalent to a call option with the firm value as the underlying and an exercise price of D .

At time = T , if $V_T > D$ the debt holders receive D and if $V_T < D$, the firm is insolvent and debt holders receive V_T . The payoff to the debt holders at time = T can be written as $\text{Max}(V_T, D)$. This is equivalent to a portfolio that is long a risk-free bond that pays D at $t = T$, and short (has sold) a put option on the value of the firm, V_T , with an exercise price of D . If $V_T > D$ the portfolio pays D and the put expires worthless, and if $V_T < D$ the portfolio pays $D - (D - V_T) = V_T$ and the debtholders effectively pay $D - V_T$ on the short put position.



MODULE QUIZ 56.1

1. The put–call parity relationship for European options must hold because a protective put will have the same payoff as:
 - A. a covered call.
 - B. a fiduciary call.
 - C. an uncovered call.
2. The put–call-forward parity relationship *least likely* includes:
 - A. a risk-free bond.
 - B. call and put options.
 - C. the underlying asset.

KEY CONCEPTS

LOS 56.a

A fiduciary call (a call option and a risk-free zero-coupon bond that pays the strike price X at expiration) and a protective put (a share of stock and a put at X) have the same payoffs at expiration, so arbitrage will force these positions to have equal prices: $c + X(1 + R_f)^{-T} = S + p$. This establishes put-call parity for European options.

Based on the put-call parity relation, a synthetic security (stock, bond, call, or put) can be created by combining long and short positions in the other three securities.

$$c = S + p - X(1 + R_f)^{-T}$$

$$p = c - S + X(1 + R_f)^{-T}$$

$$S = c - p + X(1 + R_f)^{-T}$$

$$X(1 + R_f)^{-T} = S + p - c$$

LOS 56.b

Because we can replicate the payoff on an asset by lending the present value of the forward price at the risk-free rate and taking a long position in a forward, we can write put-call-forward parity as:

$$c_0 + X(1 + R_f)^{-T} = F_0(T)(1 + R_f)^{-T} + p_0$$

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 56.1

1. **B** Given call and put options on the same underlying asset with the same exercise price and expiration date, a protective put (underlying asset plus a put option) will have the same payoff as a fiduciary call (call option plus a risk-free bond that will pay the exercise price on the expiration date) regardless of the underlying asset price on the expiration date. (LOS 56.a)
2. **C** The put-call-forward parity relationship is $F_0(T)(1 + R_f)^{-T} + p_0 = c_0 + X(1 + R_f)^{-T}$, where $X(1 + R_f)^{-T}$ is a risk-free bond that pays the exercise price on the expiration date, and $F_0(T)$ is the forward price of the underlying asset. (LOS 56.b)

READING 57

VALUING A DERIVATIVE USING A ONE-PERIOD BINOMIAL MODEL

MODULE 57.1: BINOMIAL MODEL FOR OPTION VALUES



Video covering this content is available online.

LOS 57.a: Explain how to value a derivative using a one-period binomial model.

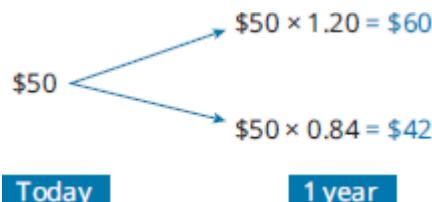
Recall from Quantitative Methods that a **binomial model** is based on the idea that, over the next period, some value will change to one of two possible values (binomial). To construct a one-period binomial model for pricing an option, we need:

- A value for the underlying at the beginning of the period.
- An exercise price for the option. The exercise price can be different from the value of the underlying. We assume the option expires one period from now.
- Returns that will result from an up-move and a down-move in the value of the underlying over one period.
- The risk-free rate over the period.

For now we do not need to consider the probabilities of an up-move or a down-move. Later in this reading we will examine one-period binomial models with risk-neutral probabilities.

As an example, we can model a call option with an exercise price of \$55 on a stock that is currently valued (S_0) at \$50. Let us assume that in one period the stock's value will either increase (S_1^u) to \$60 or decrease (S_1^d) to \$42. We state the return from an up-move (R^u) as $\$60/\$50 = 1.20$, and the return from a down-move (R^d) as $\$42/\$50 = 0.84$.

Figure 57.1: One-Period Binomial Tree



The call option will be in the money after an up-move or out of the money after a down-move. Its value at expiration after an up-move, c_1^u , is $\text{Max}(0, \$60 - \$55) = \$5$. Its value after a down-move, c_1^d , is $\text{Max}(0, \$42 - \$55) = 0$.

Now we can use no-arbitrage pricing to determine the initial value of the call option (c_0). We do this by creating a portfolio of the option and the underlying stock, such that the portfolio will have the same value following either an up-move (V_1^u) or a down-move (V_1^d) in the stock. For our example, we would write the call option and buy a number of shares of the stock that we will denote as h . We must solve for the h that results in $V_1^u = V_1^d$.

- The initial value of our portfolio, V_0 , is $hS_0 - c_0$ (remember we are short the call option).
- The portfolio value after an up-move, V_1^u , is $hS_1^u - c_1^u$.
- The portfolio value after a down-move, V_1^d , is $hS_1^d - c_1^d$.

In our example, $V_1^u = h(\$60) - \5 , and $V_1^d = h(\$42) - 0$. Setting $V_1^u = V_1^d$ and solving for h , we get:

$$h(\$60) - \$5 = h(\$42)$$

$$h(\$60) - h(\$42) = \$5$$

$$h = \$5 / (\$60 - \$42) = 0.278$$

This result, the number of shares of the underlying we would buy for each call option we would write, is known as the **hedge ratio** for this option.

With $V_1^u = V_1^d$, the value of the portfolio after one period is known with certainty. This means we can say that either V_1^u or V_1^d must equal V_0 compounded at the risk-free rate for one period. In this example, $V_1^d = 0.278(\$42) = \11.68 , or $V_1^u = 0.278(\$60) - \$5 = \$11.68$. Let us assume the risk-free rate over one period is 3%. Then $V_0 = \$11.68 / 1.03 = \11.34 .

Now we can solve for the value of the call option, c_0 . Recall that $V_0 = hS_0 - c_0$, so $c_0 = hS_0 - V_0$. Here, $c_0 = 0.278(\$50) - \$11.34 = \$2.56$.

LOS 57.b: Describe the concept of risk neutrality in derivatives pricing.

Another approach to constructing a one-period binomial model involves risk-neutral probabilities of an up-move or a down-move. Consider a share of stock currently priced at \$30. The size of the possible price changes, and the probabilities of these changes occurring, are as follows:

$$R^u = \text{up-move factor} = 1.15$$

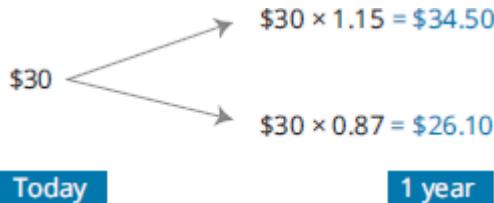
$$R^d = \text{down-move factor} = \frac{1}{R^u} = \frac{1}{1.15} = 0.87$$

$$\pi_U = \text{risk-neutral probability of an up-move} = 0.715$$

$$\pi_D = \text{risk-neutral probability of a down-move} = 1 - \pi_U = 1 - 0.715 = 0.285$$

Note that the down-move factor is the reciprocal of the up-move factor, and the probability of an up-move is one minus the probability of a down-move. The one-period binomial tree for the stock is shown in Figure 57.2. The beginning stock value of \$30 is to the left, and to the right are the two possible end-of-period stock values, $30 \times 1.15 = \$34.50$ and $30 \times 0.87 = \$26.10$.

Figure 57.2: One-Period Binomial Tree



The risk-neutral probabilities of an up-move and a down-move are calculated from the sizes of the moves and the risk-free rate:

$$\pi_U = \text{risk-neutral probability of an up-move} = \frac{1 + R_f - R^d}{R^u - R^d}$$

$$\pi_D = \text{risk-neutral probability of a down-move} = 1 - \pi_U$$

where:

R_f = risk-free rate

R^u = size of an up-move

R^d = size of a down-move



PROFESSOR'S NOTE

These two probabilities are not the actual probabilities of the up- and down-moves.

They are risk-neutral pseudo probabilities. The calculation of risk-neutral probabilities does not appear to be required for the Level I exam.

We can calculate the value of an option on the stock by:

- Calculating the payoffs of the option at expiration for the up-move and down-move prices.
- Calculating the expected payoff of the option in one year as the (risk-neutral) probability-weighted average of the up-move and down-move payoffs.
- Calculating the PV of the expected payoff by discounting at the risk-free rate.

EXAMPLE: Calculating call option value with risk-neutral probabilities

Use the binomial tree in Figure 57.2 to calculate the value today of a 1-year call option on a stock with an exercise price of \$30. Assume the risk-free rate is 7%, the current value of the stock is \$30, and the up-move factor is 1.15.

Answer:

First, we need to calculate the down-move factor and risk-neutral the probabilities of the up- and down-moves:

$$R^d = \text{size of down-move} = \frac{1}{R^u} = \frac{1}{1.15} = 0.87$$

$$\pi_U = \text{risk-neutral probability of an up-move} = \frac{1 + 0.07 - 0.87}{1.15 - 0.87} = 0.715$$

$$\pi_D = \text{risk-neutral probability of a down-move} = 1 - 0.715 = 0.285$$

Next, determine the payoffs on the option in each state. If the stock moves up to \$34.50, a call option with an exercise price of \$30 will pay \$4.50. If the stock moves down to \$26.10, the call option will expire worthless. The option payoffs are illustrated in the following figure.

Let the stock values for the up-move and down-move be S_1^u and S_1^d and for the call values, c_1^u and c_1^d .

One-Period Call Option With X = \$30

The expected value of the option in one period is:

$$\begin{array}{l}
 S_0 = \$30 \\
 \begin{array}{ll}
 \pi_u = 0.715 & S_1^u = \$30 \times 1.15 = \$34.50 \\
 & C_1^u = \max(0, \$34.50 - \$30) = \$4.50 \\
 \pi_d = 0.285 & S_1^d = \$30 \times 0.87 = \$26.10 \\
 & C_1^d = \max(0, \$26.10 - \$30) = \$0
 \end{array}
 \end{array}$$

Today

1 year

$$E(\text{call option value in 1 year}) = (\$4.50 \times 0.715) + (\$0 \times 0.285) = \$3.22$$

The value of the option today, discounted at the risk-free rate of 7%, is:

$$C_0 = \frac{\$3.22}{1.07} = \$3.01$$

We can use the same basic framework to value a one-period put option. The only difference is that the payoff to the put option will be different from the call payoffs.

EXAMPLE: Valuing a one-period put option on a stock

Use the information in the previous example to calculate the value of a put option on the stock with an exercise price of \$30.

Answer:

If the stock moves up to \$34.50, a put option with an exercise price of \$30 will expire worthless. If the stock moves down to \$26.10, the put option will be worth \$3.90.

The risk-neutral probabilities are 0.715 and 0.285 for an up- and down-move, respectively. The expected value of the put option in one period is:

$$E(\text{put option value in 1 year}) = (\$0 \times 0.715) + (\$3.90 \times 0.285) = \$1.11$$

The value of the option today, discounted at the risk-free rate of 7%, is:

$$P_0 = \frac{\$1.11}{1.07} = \$1.04$$

In practice, we would construct a binomial model with many short periods and have many possible outcomes at expiration. However, the one-period model is sufficient to illustrate the concept and method.

Note that the actual probabilities of an up-move and a down-move do not enter directly into our calculation of option value. The size of the up-move and down-move, along with the risk-free rate, determines the risk-neutral probabilities we use to calculate the expected payoff at option expiration. Remember, the risk-neutral probabilities come from constructing a hedge that creates a certain payoff. Because their calculation is based on an arbitrage relationship, we can discount the expected payoff based on risk-neutral probabilities, using the risk-free rate.



MODULE QUIZ 57.1

- To construct a one-period binomial model for valuing an option, are probabilities of an up-move or a down-move in the underlying price required?
 - No.
 - Yes, but they can be calculated from the returns on an up-move and a down-move.
 - Yes, the model requires estimates for the actual probabilities of an up-move and a down-move.

2. In a one-period binomial model based on risk neutrality, the value of an option is *best* described as the present value of:
 - A. a probability-weighted average of two possible outcomes.
 - B. a probability-weighted average of a chosen number of possible outcomes.
 - C. one of two possible outcomes based on a chosen size of increase or decrease.
3. A one-period binomial model for option pricing uses risk-neutral probabilities because:
 - A. the model is based on a no-arbitrage relationship.
 - B. they are unbiased estimators of the actual probabilities.
 - C. the buyer can let an out-of-the-money option expire unexercised.

KEY CONCEPTS

LOS 57.a

A one-period binomial model for pricing an option requires the underlying asset's value at the beginning of the period, an exercise price for the option, the asset prices that will result from an up-move and a down-move, and the risk-free rate.

A portfolio of the underlying asset hedged with a position in an option can be created such that the portfolio has the same value for both an up-move and a down-move. Because the portfolio's value at the end of the period is certain, that value must be the portfolio's initial value compounded at the risk-free rate. The number of units of the underlying required to construct such portfolios is the hedge ratio.

LOS 57.b

To determine the value of an option using the concept of risk neutrality, we calculate its payoffs for both an up-move and a down-move, calculate the expected payoff as a weighted average using the risk-neutral probabilities of an up-move and a down-move, and discount this expected payoff for one period at the risk-free rate.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 57.1

1. **A** A one-period binomial model can be constructed based on replication and no-arbitrage pricing, without regard to the probabilities of an up-move or a down-move. (LOS 57.a)
2. **A** In a one-period binomial model based on risk-neutral probabilities, the value of an option is the present value of a probability-weighted average of two possible option payoffs at the end of a single period, during which the price of the underlying asset is assumed to move either up or down to specific values. (LOS 57.b)
3. **A** Because a one-period binomial model is based on a no-arbitrage relationship, we can discount the expected payoff at the risk-free rate. (LOS 57.b)

TOPIC QUIZ: DERIVATIVES

You have now finished the Derivatives topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Derivatives. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

READING 58

CATEGORIES, CHARACTERISTICS, AND COMPENSATION STRUCTURES OF ALTERNATIVE INVESTMENTS

MODULE 58.1: ALTERNATIVE INVESTMENT STRUCTURES



Video covering this content is available online.

LOS 58.a: Describe types and categories of alternative investments.

Alternative investments comprise various types of investments that do not fall under the heading of **traditional investments**, which refers to long-only investments in cash or publicly traded stocks and bonds.

Managers of alternative investment portfolios may use derivatives and leverage and short securities. Structures that employ such strategies but utilize publicly traded stocks and bonds are also termed alternative investments. Many types of real estate investment are considered alternatives to traditional investment as well.

Types of alternative investment structures include hedge funds, private equity funds, and various types of real estate investments. Alternative investments typically are actively managed and may include investments in commodities, infrastructure, and illiquid securities.

Fee structures for alternative investments are different from those of traditional investments, with higher management fees on average and often with additional incentive fees based on performance.

Compared with traditional investments, alternative investments typically exhibit several of the following characteristics:

- Less liquidity of assets held.
- More specialization by investment managers.
- Less regulation and transparency.
- More problematic and less available historical return and volatility data.
- Different legal issues and tax treatments.
- Relatively low correlations with returns of traditional investments.
- Relatively higher management fees and incentive fees based on performance.
- Restrictions on redemptions.
- Relatively more concentrated portfolios.

The perceived benefits of including alternative investments in portfolios are risk reduction from diversification (due to low correlations of alternative investments with traditional investments) and possible higher returns from holding illiquid securities and from markets for some alternative investments possibly being less efficient than those for traditional investments. Although correlations of returns on alternative investments with returns on traditional investments may be low on average, these correlations may increase significantly during periods of economic stress.

We will examine several categories of alternative investments in detail in our reading on Private Capital, Real Estate, Infrastructure, Natural Resources, and Hedge Funds. Here we introduce those categories:

1. **Hedge funds.** These funds may use leverage, hold long and short positions, use derivatives, and invest in illiquid assets. Managers of hedge funds use a great many different strategies in attempting to generate investment gains. They do not necessarily hedge risk as the name might imply.
2. **Private equity.** As the name suggests, private equity funds invest in the equity of companies that are not publicly traded, or in the equity of publicly traded firms that the funds intend to take private. Leveraged buyout (LBO) funds use borrowed money to purchase equity in established companies and comprise the majority of private equity investment funds. Venture capital funds invest in young, unproven companies at various early stages in their lives.
3. **Private debt.** Funds may make loans directly to companies, lend to early stage (venture) firms, or invest in the debt of firms that are financially distressed (struggling to make their debt payments or have entered into bankruptcy).
4. **Real estate.** Real estate investments include residential or commercial properties, as well as real estate-backed debt. These investments are held in a variety of structures, including full or leveraged ownership of individual properties, individual real estate-backed loans, private and publicly traded securities backed by pools of properties or mortgages, and limited partnerships.
5. **Commodities.** To gain exposure to changes in commodity prices, investors can own physical commodities, commodity derivatives, or the equity of commodity-producing firms. Some funds seek exposure to the returns on various commodity indices, often by holding derivatives contracts (futures) that are expected to track a specific commodity index.
6. **Farmland.** Agricultural land can produce income from leasing it out for farming or from raising crops or livestock for harvest and sale.
7. **Timberland.** Forested land is purchased or trees are planted for harvesting, which provides cash flows.
8. **Infrastructure.** Infrastructure refers to long-lived assets that provide public services. These include economic infrastructure assets, such as roads, airports, and utility grids, and social infrastructure assets, such as schools and hospitals. While often financed and constructed by governmental entities, infrastructure investments have more recently been undertaken by public-private partnerships, with each holding a significant stake in the infrastructure assets constructed. Various deal structures are employed, and the asset may revert to public ownership at some future date.

LOS 58.b: Describe characteristics of direct investment, co-investment, and fund investment methods for alternative investments.

Direct investing refers to an investor that purchases assets itself, rather than pooling its funds with others or using a specialized outside manager. Larger, more knowledgeable investors may purchase private companies or real estate on their own. For example, a sovereign wealth fund may have its own specialized managers to invest in real estate, agricultural land, or companies in the venture stage.

Direct investing has advantages in that there are no fees to outside managers and the investor has more choice and control over the investments made. Disadvantages include the possibility of less diversification across investments, higher minimum investment amounts, and greater investor expertise required to evaluate deals and do their own due diligence.

Fund investing refers to investing in a pool of assets along with other investors, using a fund manager that selects and manages a pool of investments according to an agreed-upon strategy. In this case, the individual investors do not control the selection of assets for investment or their subsequent management and sale. The manager typically receives a percentage of the investable funds (management fee), as well as a percentage of the investment gains (incentive fee).

The advantages of investing through a fund, compared with direct investment, include having the expertise of the fund manager, less involvement and expertise required of the investor, diversification across investments made by the fund, and lower minimum investment requirements. Disadvantages are the cost of fund manager fees and the possibility that the fund manager may perform poorly even when selected through careful due diligence.

With **co-investing**, an investor contributes to a pool of investment funds (as with fund investing) but also has the right to invest directly alongside the fund manager in some of the assets in which the manager invests. Co-investing can reduce overall fees, while benefiting from the manager's expertise, and provide the investor with more control over asset selection and management. Compared with fund investing, it requires greater expertise, due diligence, and involvement with asset management. Co-investment opportunities may be subject to **adverse selection** if fund managers choose to make the full investment through the fund for assets they are quite confident about, and offer co-investment opportunities on assets they are less confident about.

LOS 58.c: Describe investment and compensation structures commonly used in alternative investments.

Alternative investments are often structured as **limited partnerships**. In a limited partnership, the **general partner (GP)** is the fund manager and makes investment decisions. The **limited partners (LPs)** are the investors, who own a partnership share proportional to their investment. The limited partners typically have no say in how the fund is managed and no liability beyond their investment in the partnership. The general partner takes on the liabilities of the partnership, including the repayment of any partnership debt.

Limited partners commit to an investment amount and in some cases only contribute a portion of that initially, providing the remaining funds over time as required by the general partner (as fund investments are made). General partnerships are less regulated than publicly traded companies, and limited partnership shares are typically only available to accredited investors (those with sufficient wealth to bear the risk of the investment and enough investment sophistication to understand the risks involved).

The rules and operational details that govern a partnership are contained in the **limited partnership agreement**. Special terms that apply to one limited partner but not to others are stated in **side letters**.

Fee Structures

The total fees paid by investors in alternative investment funds often consist of a **management fee** and an **incentive fee** (or **performance fee**). The fund manager earns the management fee regardless of investment performance. Management fees are typically between 1% and 2%. For hedge funds, the management fees are calculated as a percentage of assets under management (AUM), typically the net asset value of the fund's investments. For private equity funds, the management fee is calculated as a percentage of **committed capital**, not capital invested. Committed capital is typically not all invested immediately but is "drawn down" (invested) as securities are identified and added to the portfolio. Committed capital is usually drawn down over three to five years, but the drawdown period is at the discretion of the fund manager. Committed capital that has not yet been drawn down is referred to as **dry powder**.

Incentive fees are a portion of profits on fund investments. Most often, the partnership agreement will specify a **hurdle rate** of return that must be met or exceeded before any incentive fees are paid. Hurdle rates can be defined in two ways: either "hard" or "soft." If a **soft hurdle rate** is met, incentive fees are a percentage of the total increase in the value of each partner's investment. With a **hard hurdle rate**, incentive fees are based only on gains above the hurdle rate.

As an example, consider a fund with a hurdle rate of 8% that has produced returns of 12% for the year so that it has exceeded the hurdle rate. We will use an incentive fee structure of 20% of gains. If the 8% is a soft hurdle rate, the incentive fee will be 20% of the gains of 12%, or 2.4%. If the 8% is a hard hurdle rate, the incentive fee will be 20% of the gains above the hurdle rate ($12\% - 8\% = 4\%$), which would be 0.8%.

Typically, incentive fees are paid at the end of each year based on the increase in the value of fund investments, after management fees and other charges, which may include consulting and monitoring fees that are charged to individual portfolio companies.

A **catch-up clause** in a partnership agreement is based on a hurdle rate and is similar in its effect to a soft hurdle rate. Consider a fund with returns of 14%, a hurdle rate of 8%, and a 20% incentive fee. A catch-up clause would result in the first 8% of gains going to the limited partners and the next 2% going to the general partner, allowing the general partner to "catch up" to receiving 20% of the first 10% of gains. After the catch-up, further gains are split 80/20 between the limited partners and the general partner.

Another feature that is often included is a **high-water mark**. This means that no incentive fee is paid on gains that only offset prior losses. Thus, incentive fees are only paid to the extent that

the current value of an investor's account is above the highest net-of-fees value previously recorded (at the end of a payment period). This feature ensures that investors will not be charged incentive fees twice on the same gains in their portfolio values. Because investors invest in a fund at different times, they each may have a different high-water mark value. Investors should check carefully whether incentive fees are calculated correctly.

A partnership's **waterfall** refers to the way in which payments are allocated to the GP and the LPs as profits and losses are realized on deals. With a **deal-by-deal waterfall** (or **American waterfall**), profits are distributed as each fund investment is sold and shared according to the partnership agreement. This favors the GP because incentive fees are paid before 100% of the limited partners' original investment plus the hurdle rate is returned to them. With a **whole-of-fund waterfall** (or **European waterfall**), the LPs receive all distributions until they have received 100% of their initial investment plus the hurdle rate (typically after all fund investments have been sold).

A **clawback provision** stipulates that if the GP accrues or receives incentive payments on gains that are subsequently reversed as the partnership exits deals, the LPs can recover previous (excess) incentive payments. With a deal-by-deal waterfall, it may be the case that successful deals are exited initially and losses are realized later. A clawback provision would allow the LPs to recover these incentive fees to the extent that the subsequent losses negate prior gains on which incentive fees had been paid.



MODULE QUIZ 58.1

1. Compared with managers of traditional investments, managers of alternative investments are likely to have fewer restrictions on:
 - A. holding cash.
 - B. buying stocks.
 - C. using derivatives.
2. Compared with alternative investments, traditional investments tend to:
 - A. be less liquid.
 - B. be less regulated.
 - C. require lower fees.
3. An investor that wants to gain exposure to alternative investments but does not have the in-house expertise to perform due diligence on individual deals is *most likely* to engage in:
 - A. co-investing.
 - B. fund investing.
 - C. direct investing.
4. Management fees for a private capital fund are determined as a percentage of:
 - A. invested capital.
 - B. committed capital.
 - C. assets under management.
5. For an investor in a private equity fund, the *least* advantageous of the following limited partnership terms is:
 - A. a clawback provision.
 - B. a European-style waterfall provision.
 - C. an American-style waterfall provision.

KEY CONCEPTS

LOS 58.a

Alternative investments comprise various types of investments that do not fall under the heading of traditional investments. Categories of alternative investments include:

- Hedge funds.
- Private capital (private equity and private debt).
- Real estate.
- Natural resources (commodities, farmland, and timberland).
- Infrastructure.

LOS 58.b

Direct investing refers to purchasing assets independently, rather than pooling funds with others or using a specialized outside manager. Its advantages include not paying fees to outside managers and having control over which investments to make. Its disadvantages include high minimum investment amounts, possible lack of diversification, and the need to have expertise to evaluate deals.

Fund investing refers to pooling assets along with other investors, using a fund manager that selects and manages investments according to an agreed-upon strategy. Its advantages include having the expertise of the fund manager, less involvement and expertise required of the investor, diversification across investments, and lower minimum investment requirements. Its disadvantages include fund manager fees and the possibility that the fund manager may perform poorly.

Co-investing refers to fund investing that includes the right to invest additional capital directly alongside the fund manager. Its advantages include reduced overall fees, while benefiting from the manager's expertise, and more control over asset selection. Its disadvantages are that it requires greater expertise, due diligence, and involvement than fund investing.

LOS 58.c

Many alternative investments are structured as limited partnerships, in which the general partner is the fund manager and the limited partners are the investors. They are less regulated than publicly traded companies and typically only available to accredited investors. Limited partners may commit to an investment amount and, in some cases only, contribute a portion of that initially, providing the remaining funds over time as required by the general partner.

Fees in alternative investment funds often consist of a management fee and an incentive fee. For hedge funds, management fees are a percentage of assets under management. For private capital, management fees are a percentage of committed capital rather than capital invested.

Incentive fees are a portion of profits on fund investments. Typically, a hurdle rate must be exceeded before incentive fees are paid. With a soft hurdle rate, incentive fees are based on the total increase in the value of each partner's investment. With a hard hurdle rate, incentive fees are based only on gains above the hurdle rate. A high-water mark is a provision that no incentive fees are paid on gains that only offset prior losses.

A waterfall refers to the way payments are allocated to the general partner and the limited partners. With a deal-by-deal or American waterfall, profits are distributed as each fund investment is sold. With a whole-of-fund or European waterfall, the limited partners receive all distributions until they have received 100% of their initial investment plus the hurdle rate.

With a clawback provision, if the general partner receives incentive payments on gains that are subsequently reversed, the limited partners can recover excess incentive payments.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 58.1

1. **C** Traditional managers can hold cash and buy stocks but may be restricted from using derivatives. (LOS 58.a)
2. **C** Traditional investments typically require lower fees, are more regulated, and are more liquid than alternative investments. (LOS 58.a)
3. **B** With fund investing, due diligence on the fund's portfolio investments is a responsibility of the fund manager rather than the fund investors. Direct investing and co-investing require greater due diligence of individual deals on the part of the investor. (LOS 58.b)
4. **B** For a private capital fund, management fees are a percentage of committed capital rather than invested capital. For a hedge fund, management fees are a percentage of assets under management. (LOS 58.c)
5. **C** An American-style waterfall structure has a deal-by-deal calculation of incentive fees to the general partner. In this case, a successful deal where incentive fees are paid, followed by the sale of a holding that has losses in the same year, can result in incentive fees greater than those calculated with a European-style (whole-of-fund) waterfall. A clawback provision benefits the limited partner investors by allowing them to recover incentive fees paid earlier if the fund realizes losses later. A clawback provision coupled with an American-style waterfall will result in the same overall incentive fees as a European-style waterfall if the transactions occur in subsequent years. (LOS 58.c)

READING 59

PERFORMANCE CALCULATION AND APPRAISAL OF ALTERNATIVE INVESTMENTS

MODULE 59.1: PERFORMANCE APPRAISAL AND RETURN CALCULATIONS



Video covering this content is available online.

LOS 59.a: Describe issues in performance appraisal of alternative investments.

Alternative investments are typically exposed to some risks that unleveraged long-only traditional investments are not. These additional risks arise from the following:

- Lack of transparency.
- Illiquidity, including restrictions on and performance impact of redemptions.
- Complexity of positions and strategies employed.
- Use of derivatives.
- Use of securities that are marked to market.
- Use of leverage.
- Variety of manager strategies and areas of expertise.
- Cash drag from significant drawdown periods (e.g., with private equity).

Ideally, returns on alternative investments should be adjusted for these risks, although that may be difficult in practice. It is clear, however, that evaluating alternative investment returns (or expected returns) without considering these additional risks would be naïve and possibly quite misleading.

One commonly reported risk-adjusted return measure for alternative investments is the **Sharpe ratio**, which is calculated as excess return per unit of risk:

$$\text{Sharpe ratio} = \frac{\text{return} - \text{risk free rate}}{\text{standard deviation of returns}}$$

The biggest issue with using this statistic is that standard deviation may not be the appropriate measure of risk. Returns on alternative investments might not be normally distributed. When return skewness is significantly different from zero or return kurtosis is significantly different from three, standard deviation is not the most appropriate measure of risk, so Sharpe ratios are not the most appropriate measure of risk-adjusted returns. Additionally, if values used to calculate mean returns are based on model pricing (rather than market values), returns are

smoothed relative to market-based return measures, so estimated standard deviation will understate the actual volatility of returns (uncertainty about actual asset values over time).

Some analysts prefer the **Sortino ratio** to the Sharpe ratio as a risk-adjusted return measure. The Sortino ratio substitutes downside deviation, which is based on only deviations below the mean (expected) return, for standard deviation in the Sharpe ratio. One limitation of both the Sharpe and Sortino ratios is that they do not take into account the diversification benefits from low correlations with returns to traditional investments.

The **Treynor ratio** uses beta (a measure of systematic risk that is based on return correlations) as the denominator. A lower beta (resulting from low correlation of returns with traditional assets) results in a higher Treynor ratio. A higher Sharpe, Sortino, or Treynor ratio indicates greater excess returns per unit of risk taken.

The **Calmar ratio** is computationally less complex. It is calculated as the average annual compound return divided by the maximum drawdown (decrease in value from peak to trough). Larger values are preferred, as again we are measuring returns relative to risk.

Lower correlations of returns among alternative investments in a portfolio are preferred to higher correlations because of diversification (risk-reduction) benefits. One caveat is that during periods of financial stress or crisis, these correlations, along with the correlation to returns of traditional investments, tend to increase toward one (that is, at such times asset prices all tend to decrease together).

Performance Appraisal for Private Capital and Real Estate

Both private capital and real estate investments are often characterized by inflows as well as outflows over an investment's life, and by cash returned to investors at various times over an investment's life. One simple measure of investment success is the **multiple of invested capital (or money multiple)**—the ratio of total capital returned plus the value of any remaining assets, to the total capital paid in over the life of the investment. Because this measure does not consider the timing of cash inflows and outflows, which can affect annual returns on invested capital significantly, it can be considered somewhat naïve. For this reason, internal rates of return, which are dependent on the timing of cash flows, are often used as performance measures for private equity and real estate investments. Comparing investments based on IRRs requires specific assumptions about both the cost of investment capital and the reinvestment returns that can be earned on cash distribution.

For funds investing in private equity or real estate where initial committed capital is drawn down over time as individual investments are made, if returns are calculated on committed capital, there is a fee drag on calculated returns. Typically, investments are made over a significant period of time and returns are staggered over subsequent periods. This results in a pattern of quite low returns on investment early in the fund's life and relatively higher returns over the final years. In real estate funds, significant investments in property improvements after acquisition may amplify this pattern of returns on investment.

Given the variability of cash flows over a fund's life and the importance of management decisions in the timing and magnitude of after-tax cash flows, an IRR over the life of a fund is the most appropriate measure of after-tax investment performance.

LOS 59.b: Calculate and interpret returns of alternative investments both before and after fees.

Before-fee returns on alternative investments are calculated the same way we calculate returns on any investment. The calculation of holding period returns and periodic rates of return, both with and without interim cash distributions, is described in Quantitative Methods and expanded on by application in other topic areas.

Calculating after-fee returns simply requires adjustment of the cash flows or values for the various fees involved, typically management and incentive fees. Some examples will illustrate the application of various fees and the relevant terminology.

EXAMPLE: Hedge fund fees

BJI Funds is a hedge fund with a value of \$110 million at initiation. BJI Funds charges a 2% management fee based on assets under management at the beginning of the year and a 20% incentive fee with a 5% soft hurdle rate, and it uses a high-water mark. Incentive fees are calculated on gains net of management fees. The year-end values before fees are as follows:

- Year 1: \$100.2 million
- Year 2: \$119.0 million

Calculate the total fees and the investor's after-fee return for both years.

Answer:

Year 1:

Management fee: $110.0 \text{ million} \times 2\% = \2.2 million

Gross value end of year (given): \$100.2 million

Return net of management fees: $\frac{100.2 - 2.2}{110} - 1 = -10.9\%$

There is no incentive fee because the return after the management fee is less than the 5% hurdle rate.

Total fees: \$2.2 million

Ending value net of fees: $100.2 \text{ million} - \$2.2 \text{ million} = \$98.0 \text{ million}$

Year 1 after-fees return: $\frac{98.0}{110.0} - 1 = -10.9\%$

Year 2:

Management fees: $98.0 \text{ million} \times 2\% = \1.96 million

Year-end value net of management fees: $119.0 - \$1.96 = \117.04 million

The high-water mark is \$110 million.

Year 2 value net of management fees, above high-water mark: $117.04 \text{ million} - 110.0 \text{ million} = \7.04 million

Year 2 return net of management fees, above high-water mark: $\frac{7.04}{110} = 6.4\%$

Note that the incentive fee is calculated based on gains in value above \$110 million because that is the high-water mark.

The incentive fee is calculated on the entire gain above the high-water mark because 6.4% is greater than the soft hurdle rate. If the 5% was a hard hurdle rate, the incentive fee would be calculated only on the gains more than 5% above the high-water mark.

Incentive fee: $7.04 \times 0.20 = \$1.41 \text{ million}$

Total fees: $\$1.96 \text{ million} + \$1.41 \text{ million} = \$3.37 \text{ million}$

Year 2 year-end value after fees: $119.0 - 3.37 = \$115.63 \text{ million}$

Year 2 after-fee return: $\frac{115.63}{98.0} - 1 = 18.0\%$

EXAMPLE: Fund-of-funds

An investor makes a total investment of \$60 million in a fund-of-funds that has a “1 and 10” fee structure, with management and incentive fees calculated independently based on year-end values. \$40 million of the investment was allocated to the Alpha fund, and \$20 million was allocated to the Beta fund. One year later, the value of the Alpha fund investment is \$45 million and the value of the Beta fund investment is \$28 million, both net of fund fees. Calculate the investor’s return for the year net of fees.

Answer:

At year-end, the gross value of the investor’s investment is $\$45 + \$28 = \$73$ million.

The fund-of-funds management fee is 1% of \$73 million, which is \$0.73 million.

The investor’s gain for the year before fund-of-funds fees is $\$73 - \$60 = \$13$ million.

The fund-of-funds manager’s incentive fee is 10% of \$13 million, which is \$1.3 million.

The year-end value of the investor’s fund-of-funds investment is $\$73 - \$0.73 - \$1.3 = \70.97 million.

The investor’s one-year return after fees is $\frac{70.97}{60} - 1 = 18.3\%$

Note that the same investments made directly with the Alpha and Beta funds would have returned

$$\frac{73}{60} - 1 = 21.7\%$$

EXAMPLE: Waterfall structure and clawback provision

A private equity fund invests \$100 million in a venture company that is sold for \$130 million.

It also invests \$100 million in an LBO that goes poorly and is liquidated for \$80 million.

1. If the carried interest incentive fee for the GP is 20% and there is no clawback provision, what is the investor’s return after incentive fees, assuming the investment outcomes are realized in the same year?
 - a. under an American-style (deal-by-deal) waterfall structure?
 - b. under a European-style (whole-of-fund) waterfall structure?
2. How would the answers be affected if the venture investment was sold in year 1 and the LBO investment was sold in year 2?
3. How would including a clawback provision affect investor returns calculated in question 1?

Answer:

1. Under an American-style (deal-by-deal) waterfall structure, an incentive fee of $20\% \times (\$130 - \$100) = \$6$ million would be paid on the venture investment. Because there is a loss on the LBO investment, no incentive fee is paid.

Investor’s return on investment: $\frac{130 + 80 - 6}{200} - 1 = 2\%$

Under a European-style (whole-of-fund) waterfall structure, the gain for the period is $130 + 80 - 200 = \$10$ million and the incentive fee is $20\% \times 10 = \$2$ million.

Investor’s return on investment: $\frac{130 + 80 - 2}{200} - 1 = 4\%$

2. The European-style waterfall structure would have the same overall return as the American-style structure, as the incentive fee for the venture investment of \$6 million would be paid in year 1 and no incentive fee would be received on the LBO investment.
3. With a clawback provision, after the LBO investment is sold, the incentive fee of \$6 million paid on the venture investment is more than 20% of the return on the total investment. It is 60% of the total (net)

gain of \$10 million. The investor could “claw back” \$4 million of the \$6 million paid as an incentive fee on the venture investment so that the total incentive fee is reduced to 20% of the \$10 million gain.

Negotiated fee structures

Although “2 and 20” and “1 and 10” were at one time fairly standard fee structures for fund and fund-of-funds investments, these fee structures have been under competitive pressure. Investors making larger commitments can negotiate lower fees. There can also be a trade-off between liquidity provisions and fees. Investors can negotiate for lower fees or better liquidity (shorter lockups and notice periods). Hurdle rates, hard versus soft hurdles, and catch-up provisions may also be subject to negotiation.

Annual investor fees can also be **either-or fees**, the maximum of the management fee or the incentive fee. Under such a structure, with a 1% management fee and a 30% incentive fee, investor fees each year would be simply the management fee unless the calculated incentive fee is greater. Such a structure may also stipulate that the 1% management fees paid be subtracted from incentive fees when they are paid in a subsequent year.

Investors should be aware that other investors may receive terms that differ from those in the partnership agreement. Customized fee structures are contained in **side letters** to individual investors detailing how their terms differ from those in the standard offering documents.

Early investors in a fund may also receive lower fees or better liquidity terms as an incentive to invest at the inception of a fund. The investment interests of early investors that receive such relatively better terms are called **founder's shares**.



MODULE QUIZ 59.1

1. Standard deviation is *least likely* an appropriate measure of risk for:
 - A. hedge funds.
 - B. publicly traded REITs.
 - C. exchange-traded funds.
2. A hedge fund has a return of 30% before fees in its first year. The fund has a management fee of 1.5% on end-of-year fund value and a 15% incentive fee, with an 8% hard hurdle rate on gains net of the management fee. The return after fees for an investor in this fund is *closest to*:
 - A. 20.5%.
 - B. 21.5%.
 - C. 25.0%.
3. A private equity fund has a “2 and 20” fee structure with the incentive fee independent of management fees. The fund will sell a holding for a profit of 9%. The hurdle rate is specified as 8%. The provision that would result in an incentive fee of 1% is:
 - A. a hard hurdle rate.
 - B. a soft hurdle rate.
 - C. a catch-up provision.

KEY CONCEPTS

LOS 59.a

Alternative investments are typically exposed to risks that traditional investments are not, including lack of transparency, illiquidity, complexity of positions and strategies, use of derivatives, securities that are marked to market, use of leverage, variety of manager strategies,

and cash drag from significant drawdown periods. Return measures should be adjusted for these risks.

The Sharpe ratio (excess return per unit of risk) may not be appropriate because it measures risk as standard deviation, but alternative investment returns are unlikely to be normally distributed. Other measures available include the Sortino ratio (which uses downside deviations), the Treynor ratio (which uses beta as a risk measure), and the Calmar ratio (average annual compound return divided by the maximum drawdown).

Return measures for real estate and private capital include internal rates of return and the multiple of invested capital (ratio of total capital returned plus the value of remaining assets, to total capital paid in).

LOS 59.b

Before-fee returns on alternative investments are calculated just as we calculate fees on any investment. Calculating after-fee returns simply requires adjustment of the cash flows or values for the various fees involved, typically management and incentive fees.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 59.1

1. **A** Hedge funds may hold illiquid assets that may use estimated values to calculate returns. Risk as measured by standard deviation could be understated. For publicly traded securities, such as REITs and ETFs, standard definitions of risk are more applicable. (LOS 59.a)

2. **C** This is easiest to see with an assumed initial value. Assume 100 was invested.

$$\text{Year-end gross value} = 130$$

$$\text{Management fee} = 1.5\% \times 130 = 1.95$$

$$\text{Gains net of management fee} = 30 - 1.95 = 28.05$$

$$\text{Hurdle gains} = 8\% \times 100 = 8$$

$$\text{Incentive fee} = 15\% \times (30 - 1.95 - 8) = 3.0075$$

$$\text{Total fees} = 1.95 + 3.0075 = 4.9575$$

$$\text{Ending value after fees} = 130 - 4.9575 = 125.0425$$

$$\text{Return after fees} = 125.0425/100 - 1 = 25.0425\%$$

(LOS 59.b)

3. **C** With a catch-up provision, the limited partners get the first 8% of gross return and the general partner gets all returns above that to a maximum of 2%, and gains above that are shared 80% to the limited partners and 20% to the general partner.

With a soft hurdle rate of 8%, the incentive fee would be 20% of 9%, or 1.8%.

With a hard hurdle rate of 8%, the incentive fee would be 20% of (9% - 8%), or 0.2%.

(LOS 59.b)

READING 60

PRIVATE CAPITAL, REAL ESTATE, INFRASTRUCTURE, NATURAL RESOURCES, AND HEDGE FUNDS

MODULE 60.1: PRIVATE EQUITY AND PRIVATE DEBT



Video covering this content is available online.

LOS 60.a: Explain investment characteristics of private equity.

Private capital refers to entities that provide financing to companies without issuing securities in the public markets. The two categories of private capital we will review here are private equity and private debt.

Private Equity

The majority of private equity funds invest either in private companies or public companies they intend to take private (leveraged buyout funds) or in early stage companies (venture capital funds). The companies in which a private equity fund invests are known as its **portfolio companies**. Two additional, but smaller, categories of private equity funds are distressed investment funds and developmental capital funds.

A private equity fund may also charge fees for arranging buyouts, for a deal that does not happen, or for handling asset divestitures after a buyout.

Private equity strategies

Leveraged buyouts (LBOs) are the most common type of private equity fund investment. “Leveraged” refers to the fact that the fund’s purchase of the portfolio company is funded primarily by debt. When a private equity fund acquires a publicly traded company in an LBO, the company is said to be “going private.”

Two types of LBOs are **management buyouts (MBOs)**, in which the existing management team is involved in the purchase, and **management buy-ins (MBIs)**, in which an external management team will replace the existing management team.

In an LBO, the private equity firm seeks to increase the value of the firm through some combination of new management, management incentives, restructuring, cost reduction, or revenue enhancement. Firms with high cash flow are attractive LBO candidates because their cash flow can be used to service and eventually pay down the debt taken on for acquisition.

Developmental capital or minority equity investing refers to the provision of capital for business growth or restructuring. The firms financed may be public or private. In the case of public companies, such financing is called **private investment in public equities (PIPEs)**.

Venture capital funds invest in companies in the early stages of their development. The investment often is in the form of equity but can be in convertible preferred shares or convertible debt. While the risk of start-up companies is often great, returns on successful companies can be very high. This is often the case when a company has grown to the point where it is sold (at least in part) to the public via an IPO.

Venture capital fund managers are actively involved in the development of their portfolio companies, often sitting on their boards or filling key management roles.

Categorization of venture capital investments is based on the company's stage of development. Terminology used to identify venture firm investment at different stages of the company's life includes the following:

1. The **formative stage** refers to investments made during a firm's earliest period and comprises three distinct phases.
 - **Angel investing or pre-seed capital** refers to investments made very early in a firm's life, often the idea stage, and the investment funds are used for business plans and assessing market potential. The funding source is usually individuals ("angels") rather than venture capital funds.
 - The **seed stage or seed capital** refers to investments made for product development, marketing, and market research. This is typically the stage during which venture capital funds make initial investments, through ordinary or convertible preferred shares.
 - The **early stage or start-up stage** refers to investments made to fund initial commercial production and sales.
2. **Later-stage investment or expansion venture capital** refers to the stage of development where a company already has production and sales and is operating as a commercial entity. Investment funds provided at this stage are typically used for expansion of production and/or increasing sales through an expanded marketing campaign.
3. **Mezzanine-stage financing** refers to capital provided to prepare the firm for an IPO. The term refers to the timing of the financing (between private company and public company) rather than the type of financing.

Private equity exit strategies

The average holding period for companies in private equity portfolios is five years. There are several primary methods of exiting an investment in a portfolio company:

1. **Trade sale.** Sell a portfolio company to a competitor or another strategic buyer.
2. **IPO.** Sell all or some shares of a portfolio company to the public.
3. **Recapitalization.** The company issues debt to fund a dividend distribution to equity holders (the fund). This is not an exit, in that the fund still controls the company, but is often a step toward an exit.

4. **Secondary sale.** Sell a portfolio company to another private equity firm or a group of investors.
5. **Write-off/liquidation.** Reassess and adjust to take losses from an unsuccessful outcome.

Private equity potential benefits and risks

There is evidence that over the last 20 years, returns on private equity funds have been higher on average than overall stock returns. Less-than-perfect correlation of private equity returns with traditional investment returns suggests that there may be portfolio diversification benefits from including private equity in portfolios.

The standard deviation of private equity returns has been higher than the standard deviation of equity index returns, suggesting greater risk. As with hedge fund returns data, private equity returns data may suffer from survivorship bias and backfill bias (both lead to overstated returns). Because portfolio companies are revalued infrequently, reported standard deviations of returns and correlations of returns with equity returns may both be biased downward.

LOS 60.b: Explain investment characteristics of private debt.

Private debt refers to lending to private entities. The terms used include the following:

- **Direct lending.** Loans made directly to a private entity without an intermediary. A **leveraged loan** refers to a loan made by a private debt fund using money borrowed from another source. That is, the fund's portfolio of loans is leveraged through the use of fund borrowing to magnify fund returns.
- **Venture debt.** Lending to venture firms (start-up or early stage firms that are not yet profitable). Venture debt is often convertible to the venture firm's common stock or combined with warrants granting the right to buy the venture firm's common stock at a given price for a specified period.
- **Mezzanine loans.** Private debt that is subordinated—that is, has a lower priority of claims than the borrower's existing (more senior) debt. May have special features such as conversion rights or warrants as compensation for the additional risk involved.
- **Distressed debt.** Purchasing the debt of firms in bankruptcy, in default on existing loans, or for which default seems imminent. Often, the fund becomes active in implementing a plan for the debtor firm to restructure its existing debt or make other changes that will result in an increase in the value of the acquired debt.

Private debt firms may also invest in a wide variety of other types of debt. Some private capital firms invest in both the debt and equity of their portfolio companies.

Risks and Returns of Private Capital

Given that private capital encompasses a variety of securities, from start-up equity to the senior debt of a mature company, the range of risks across the spectrum of private capital funds is very large. Additionally, investment in private equity or private debt funds may have quite limited liquidity.

Returns on private debt are necessarily higher than those on the debt of publicly traded companies because lack of liquidity increases the required returns on both private debt and private equity. Most often, both private equity and private debt funds are more active in the operations of the portfolio firms. Both venture capital and private equity firms often work as partners with a portfolio company's management, or as part of its management, to develop and implement a business strategy. Fund returns are quite dependent on the ability of the fund manager to both select and manage fund investments.

It is important for investors to understand the various risks that an investment in a given private capital fund entails and not to view returns above those of traditional debt and equity investments as a "free lunch." Diligence in manager selection is an important part of investment success.

MODULE QUIZ 60.1

1. In which stage of a firm's development is a venture capital fund *most likely* to make its initial investment?
 - A. Start-up.
 - B. Seed capital.
 - C. Angel investing.
2. In a secondary sale, a private capital firm sells one of its portfolio companies to:
 - A. the public.
 - B. a competitor in its industry.
 - C. another private capital fund.

MODULE 60.2: REAL ESTATE, INFRASTRUCTURE, AND NATURAL RESOURCES



Video covering this content is available online.

LOS 60.c: Explain investment characteristics of real estate.

Investment in **real estate** can provide income in the form of rents, as well as the potential for capital gains. Real estate as an asset class can provide diversification benefits to an investor's portfolio and a potential inflation hedge because rents and real estate values tend to increase with inflation. Real estate investments can be differentiated according to their underlying assets. Assets included under the heading of real estate investments include the following:

- Residential property—single-family homes.
- Commercial property—produces income (e.g., office buildings).
- Loans with residential or commercial property as collateral—mortgages ("whole loans"), construction loans.

Residential property is considered a direct investment in real estate. Some buyers pay cash, but most take on a mortgage (borrow) to purchase. The issuer (lender) of the mortgage has a direct investment in a whole loan and is said to "hold the mortgage." Issuers often sell the mortgages they originate, which are then pooled (securitized) as publicly traded mortgage-backed securities (MBS) that represent an indirect investment in the mortgage loan pool. Property purchased with a mortgage is called a leveraged investment, and the owner's equity is the property value minus the outstanding loan amount. Changes in property value over time, therefore, affect the property owner's equity in the property.

Commercial real estate properties generate income from rents. Homes purchased for rental income are considered investments in commercial property. Large properties (e.g., an office building) are a form of direct investment for institutions or wealthy individuals, either purchased for cash or leveraged (a mortgage loan is taken for a portion of the purchase price). Long-time horizons, illiquidity, the large size of investments needed, and the complexity of the investments make commercial real estate inappropriate for many investors. Commercial real estate properties can also be held by a limited partnership, in which the partners have limited liability and the general partner manages the investment and the properties, or by a real estate investment trust.

As with residential mortgages, whole loans (commercial property mortgages) are considered a direct investment, but loans can be pooled into commercial mortgage-backed securities (CMBS) that represent an indirect investment.

Real estate investment trusts (REITs) issue shares that trade publicly like shares of stock. REITs are often identified by the type of real estate assets they hold: mortgages, hotel properties, malls, office buildings, cell phone towers, or other commercial properties. Income is used to pay dividends. Typically, 90% of income must be distributed to shareholders to avoid taxes on this income that would have to be paid by the REIT before distribution to shareholders.

Potential Benefits and Risks of Real Estate

Real estate performance is measured by three different types of indices:

- An **appraisal index**, such as those prepared by the National Council of Real Estate Investment Fiduciaries (NCREIF), is based on periodic estimates of property values. Appraisal index returns are smoother than those based on actual sales and have the lowest standard deviation of returns of the various index methods.
- A **repeat sales index** is based on price changes for properties that have sold multiple times. The sample of properties sold and thus included in the index is not necessarily random and may not be representative of the broad spectrum of properties available (an example of sample selection bias).
- **REIT indices** are based on the actual trading prices of REIT shares, similar to equity indices.

Historically, REIT index returns and global equity returns have had a relatively strong correlation (on the order of 0.6) because business cycles affect REITs and global equities similarly. The correlation between global bond returns and REIT returns has been very low historically. In either case, diversification benefits can result from including real estate in an investor's portfolio. However, the method of index construction (e.g., appraisal or repeat sales indices) may be a factor in the low reported correlations, in which case actual diversification benefits may be less than expected.

Real Estate Investment Due Diligence

Property values fluctuate because of global and national economic factors, local market conditions, and interest rate levels. Other specific risks include variation in the abilities of managers to select and manage properties and changes in regulations. Decisions regarding selecting, financing, and managing real estate projects directly affect performance. The degree of

leverage used in a real estate investment is important because leverage amplifies losses as well as gains.

Distressed-property investing has additional risk factors compared with investing in properties with sound financials and stable operating histories.

Real estate development has additional risk factors, including regulatory issues such as zoning, permitting, and environmental considerations or remediation, and economic changes and financing decisions over the development period. The possible inability to get long-term financing at the appropriate time for properties initially developed with temporary (short-term) financing presents another risk.

LOS 60.d: Explain investment characteristics of infrastructure.

Infrastructure investments include transportation assets, such as roads, airports, ports, and railways, as well as utility assets, such as gas distribution facilities, electric generation and distribution facilities, and waste disposal and treatment facilities. Other categories of infrastructure investments are communications (e.g., broadcast assets and cable systems) and social (e.g., prisons, schools, and healthcare facilities).

Investments in infrastructure assets that are already constructed are referred to as **brownfield investments**, and investments in infrastructure assets that are to be constructed are referred to as **greenfield investments**. In general, investing in brownfield investments provides stable cash flows and relatively high yields but offers little potential for growth. Investing in greenfield investments is subject to more uncertainty and may provide relatively lower yields but offers greater growth potential.

In addition to categorizing infrastructure investments by type or by whether or not construction of the assets is complete, they may be categorized by their geographic location.

Investment in infrastructure can be made by constructing the assets and either selling or leasing them to the government or by directly operating the assets. Alternatively, investment in infrastructure can be made by purchasing existing assets from the government to lease back to the government or operate directly. Infrastructure investments can also be made by a public-private partnership.

Infrastructure assets typically have a long life and are quite large in cost and scale, so direct investment in them has low liquidity. However, more-liquid investments backed by infrastructure assets are available through exchange-traded funds (ETFs), mutual funds, private equity funds, or master limited partnerships (MLPs). Publicly traded vehicles for investing in infrastructure are a small part of the overall universe of infrastructure investments and are relatively concentrated in a few categories of assets.

Investing in infrastructure assets can provide diversification benefits, but investors should be aware that they are often subject to regulatory risk, risk from financial leverage, and the possibility that cash flows will be less than expected. Investors who construct infrastructure assets have construction risk. When the assets are owned and operated by a private owner, operational risk must also be considered.

LOS 60.e: Explain investment characteristics of natural resources.

Commodities

While it is possible to invest directly in commodities such as grain and gold, the most commonly used instruments to gain exposure to commodity prices are derivatives. Commodities themselves are physical goods and thus incur costs for storage and transportation. Returns are based on price changes and not on income streams.

Futures, forwards, options, and swaps are all available forms of commodity derivatives. Futures trade on exchanges; some options trade on exchanges while others trade over the counter; and forwards and swaps are over-the-counter instruments originated by dealers. Futures and forwards are contractual obligations to buy or sell a commodity at a specified price and time. Options convey the right, but not the obligation, to buy or sell a commodity at a specified price and time.

Other methods of exposure to commodities include the following:

- Exchange-traded funds (**commodity ETFs**) are suitable for investors who are limited to buying equity shares. ETFs can invest in commodities or commodity futures and can track prices or indexes.
- **Managed futures funds**, such as commodity trading advisers (CTAs), are actively managed. Some managers concentrate on specific sectors (e.g., agricultural commodities), while others are more diversified. Managed futures funds can be structured as limited partnerships with fees like those of hedge funds (e.g., 2 and 20) and restrictions on the number, net worth, and liquidity of the investors. They can also be structured like mutual funds with shares that are publicly traded so that retail investors can also benefit from professional management. Additionally, such a structure allows a lower minimum investment and greater liquidity compared with a limited partnership structure.
- Specialized funds in specific commodity sectors can be organized under any of the structures we have discussed and focus on certain commodities, such as oil and gas, grains, precious metals, or industrial metals.

Potential benefits and risks of commodities

Returns on commodities over time have been lower than returns on global stocks or bonds. As with other investments, speculators can earn high returns over short periods when their expectations about short-term commodity price movements are correct and they act on them.

Historically, correlations of commodity returns with those of global equities and global bonds have been low, typically less than 0.2, so adding commodities to a traditional portfolio can provide diversification benefits. Because commodity prices tend to move with inflation rates, holding commodities can act as a hedge of inflation risk. To the extent that commodity prices move with inflation, the real return over time would be zero, although futures contracts may offer positive real returns.

Commodity prices and investments

Spot prices for commodities are a function of supply and demand. Demand is affected by the value of the commodity to end-users and by global economic conditions and cycles. Supply is affected by production and storage costs and existing inventories. Both supply and demand are affected by the purchases and sales of nonhedging investors (speculators).

For many commodities, supply is inelastic in the short run because of long lead times to alter production levels (e.g., drill oil wells, plant crops or decide to plant less of them). As a result, commodity prices can be volatile when demand changes significantly over the economic cycle. Production of some commodities, especially agricultural commodities, can be significantly affected by the weather, leading to high prices when production is low and low prices when production is high. Costs of extracting oil and minerals increase as more expensive methods or more remote areas are used. To estimate future needs, commodity producers analyze economic events, government policy, and forecasts of future supply. Investors analyze inventory levels, forecasts of production, changes in government policy, and expectations of economic growth in order to forecast commodity prices.

Commodity valuation

Wheat today and wheat six months from today are different products. Purchasing the commodity today will give the buyer the use of it if needed, while contracting for wheat to be delivered six months from today avoids storage costs and having cash tied up. An equation that considers these aspects is:

$$\text{futures price} \approx \text{spot price} \times (1 + \text{risk-free rate}) + \text{storage costs} - \text{convenience yield}$$

Convenience yield is the value of having the physical commodity for use over the period of the futures contract. If there is little or no convenience yield, futures prices will be higher than spot prices, a situation termed *contango*. When the convenience yield is high, futures prices will be less than spot prices, a situation referred to as *backwardation*.

Farmland and Timberland

Two additional assets under the heading of natural resources are timberland and farmland, for which one component of returns comes from sales of timber or agricultural products. Timberland returns also include price changes on the land, which depend on expectations of lumber prices and how much timber has been harvested. Farmland returns are based on land price changes, changes in farm commodity prices, and the quality and quantity of the crops produced.

While most agricultural crops must be harvested within a short period, timber is different in that the choice of when to harvest is based on current prices and expected growth rates. Because agricultural crops (including trees) consume carbon, they are attractive to investors with an ESG focus on climate change.

Risks of investing in farmland and timberland include low liquidity, high fixed costs of production, variable cash flows that depend on weather, and potential losses from natural disasters such as wildfires.



MODULE QUIZ 60.2

1. Direct commercial real estate ownership *least likely* requires investing in:
 - A. large amounts.

- B. illiquid assets.
 - C. a short time horizon.
2. Compared with purchasing commodities, long positions in commodity derivatives offer the benefit of:
 - A. no storage costs.
 - B. convenience yield.
 - C. better correlation with spot prices.
 3. Greenfield investments in infrastructure are *most accurately* described as investments in assets:
 - A. that are operating profitably.
 - B. that have not yet been constructed.
 - C. related to environmental technology.

MODULE 60.3: HEDGE FUNDS



Video covering
this content is
available online.

LOS 60.f: Explain investment characteristics of hedge funds.

Hedge funds generally:

- Use leverage.
- Take both long and short positions.
- Use derivatives for speculation or hedging portfolio risk.

In addition to the structures for limited partnerships and types of fees paid to the general partner we have covered, hedge funds typically have restrictions on limited partner redemptions. A **lockup period** is the time after initial investment over which limited partners either cannot request redemptions or incur significant fees for redemptions (a **soft lockup**). A **notice period** (typically between 30 and 90 days) is the amount of time a fund has to fulfill a redemption request made after the lockup period has passed.

Hedge fund managers often incur significant transaction costs when they redeem shares. Redemption fees can offset these costs. Notice periods allow time for managers to reduce positions in an orderly manner. Redemptions often increase when hedge fund performance is poor over a period, and the costs of honoring redemptions may further decrease the value of the remaining partnership interests. This is an additional source of risk for hedge fund investors.

A **fund-of-funds** is an investment company that invests in hedge funds. Fund-of-funds investing can give investors diversification among hedge fund strategies, can provide expertise in selecting individual hedge funds, and can provide smaller investors with access to hedge funds in which they may not be able to invest directly.

Fund-of-funds managers charge an additional layer of fees beyond the fees charged by the individual hedge funds in the portfolio. Historically, these additional fees have been a 1% management fee and a 10% incentive fee. Because these fees to the fund-of-funds manager are on top of fees charged by the individual funds, they can significantly reduce investor net returns.

Recently, there has been market pressure to reduce hedge fund fees. Rather than the previous standard of 2 and 20, average hedge fund fees have fallen closer to 1.3% in management fees and 15% in incentive fees. Fund-of-funds fees have also fallen from 1 and 10 and some may charge only a management fee or a lower management fee combined with a reduced incentive fee.

Hedge Fund Strategies

Similar to categorizing alternative investments, classifying hedge funds can also be challenging. According to Hedge Fund Research, Inc., there are four main classifications of hedge fund strategies:

1. **Event-driven strategies** are typically based on a corporate restructuring or acquisition that creates profit opportunities for long or short positions in common equity, preferred equity, or debt of a specific corporation. Event-driven funds are typically long-biased.

Subcategories are as follows:

- **Merger arbitrage.** Buy the shares of a firm being acquired and sell short the firm making the acquisition. Although term “arbitrage” is used, such a strategy is not risk free because deal terms may change or an announced merger may not take place.
- **Distressed/restructuring.** Buy the (undervalued) securities of firms in financial distress when analysis indicates that value will be increased by a successful restructuring; possibly short overvalued securities at the same time.
- **Activist shareholder.** Buy sufficient equity shares to influence a company’s policies, with the goal of increasing company value (e.g., by restructuring, change in strategy/management, or return of capital to equity holders).
- **Special situations.** Invest in the securities of firms that are issuing or repurchasing securities, spinning off divisions, selling assets, or distributing capital.

2. **Relative value strategies** involve buying a security and selling short a related security, with the goal of profiting when a perceived pricing discrepancy between the two is resolved.

- **Convertible arbitrage fixed income.** Exploit pricing discrepancies between convertible bonds and the common stock of the issuing companies and options on the common shares.
- **Asset-backed fixed income.** Exploit pricing discrepancies among various MBS or asset-backed securities (ABS).
- **General fixed income.** Exploit pricing discrepancies between fixed-income securities of various issuers and types.
- **Volatility.** Exploit pricing discrepancies arising from differences between returns volatility implied by options prices and manager expectations of future volatility.
- **Multistrategy.** Exploit pricing discrepancies among securities in asset classes different from those previously listed and across asset classes and markets.

3. **Macro strategies** are based on global economic trends and events and may involve long or short positions in equities, fixed income, currencies, or commodities. Managed futures funds may focus on trading commodity futures (these funds are known as **commodity trading advisers**, or CTAs) or incorporate financial futures.

4. **Equity hedge fund strategies** seek to profit from long or short positions in publicly traded equities and derivatives with equities as their underlying assets.

- **Market neutral.** Use technical or fundamental analysis to select undervalued equities to be held long and to select overvalued equities to be sold short, in approximately

equal amounts to profit from their relative price movements without exposure to market risk. Leverage may be used.

- **Fundamental long/short growth.** Use fundamental analysis to find high-growth companies. Identify and buy equities of companies that are expected to sustain relatively high rates of capital appreciation, and short equities of companies expected to have low or no revenue growth.
- **Fundamental value.** Buy equity shares that are believed to be undervalued based on fundamental analysis and sometimes short an index or companies believed to be overvalued. Exposures to value stocks and small-cap stocks often result.
- **Sector specific.** Identify opportunities within a sector, such as health care, biotech, technology, and financial services. Manager expertise within a specific sector is believed to lead to superior returns.
- **Short bias.** Employ technical and fundamental analysis and take predominantly short positions in overvalued equities, possibly with smaller long positions but with negative market exposure overall.

Hedge Fund Potential Benefits and Risks

Hedge fund returns have tended to be better than those of global equities in down equity markets and to lag the returns of global equities in up markets. Different hedge fund strategies have the best returns during different time periods. Statements about the performance and diversification benefits of hedge funds are problematic because of the great variety of strategies used. Less-than-perfect correlation with global equity returns may offer some diversification benefits, but correlations tend to increase during periods of financial crisis.

Characteristics of hedge fund indexes may bias returns and correlations with traditional investment returns. Because hedge funds might not be included in an index until they have been in existence for a given time period or until they reach a given size, index returns may exhibit **survivorship bias**. Funds that have been successful, so that they have stayed in business for multiple years or reached a specific level of assets under management, tend to be overrepresented in a hedge fund index, which biases returns upward. **Backfill bias** refers to the effect on historical index returns of adding fund returns for prior years to index returns when a fund is added to an index.

Model values and appraisal values are typically less volatile than market values. To the extent that funds use models or appraisals for asset valuation and return calculations, both standard deviations of fund returns and correlations of fund returns with those of traditional investments will be biased downward. Investors must understand these potential biases when using index returns to evaluate the risk and return characteristics of hedge funds.



MODULE QUIZ 60.3

1. An investor who chooses a fund-of-funds as an alternative to a single hedge fund is *most likely* to benefit from:
 - A. lower fees.
 - B. higher returns.
 - C. more due diligence.
2. Diversification benefits from adding hedge funds to an equity portfolio may be limited because:
 - A. correlations tend to increase during periods of financial crisis.
 - B. hedge fund returns are less than perfectly correlated with global equities.

- C. hedge funds tend to perform better when global equity prices are declining.
- 3. A hedge fund that operates as an activist shareholder is *most likely* engaging in:
 - A. a macro strategy.
 - B. a relative value strategy.
 - C. an event-driven strategy.

KEY CONCEPTS

LOS 60.a

Private equity funds usually invest in the equity of private companies or companies wanting to become private, financing their assets with high levels of debt.

- Venture capital funds provide capital to companies early in their development. Stages of venture capital investing include the formative stage (composed of the angel investing, seed, and early stages), the later stage (expansion), and the mezzanine stage (prepare for IPO).
- Leveraged buyouts (LBOs) include management buyouts, in which the existing management team is involved in the purchase, and management buy-ins, in which an external management team replaces the existing management.
- Developmental capital or minority equity investing refers to providing capital for business growth or restructuring. The firms financed may be public or private. In the case of public companies, such financing is referred to as private investment in public equities.

Methods for exiting investments in portfolio companies include trade sale (sell to a competitor or another strategic buyer), IPO (sell some or all shares to investors), recapitalization (issue portfolio company debt), secondary sale (sell to another private equity firm or other investors), or write-off/liquidation.

LOS 60.b

Private debt refers to lending to private entities. Private debt investments include direct lending, venture debt, mezzanine loans, and distressed debt. Some private capital firms invest in both equity and debt.

LOS 60.c

Real estate as an asset class includes residential and commercial real estate, individual mortgages, and pools of mortgages or properties. It includes direct investment in single properties or loans, as well as indirect investment in limited partnerships, which are private securities, and mortgage-backed securities and real estate investment trusts, which are publicly traded.

Reasons to invest in real estate include potential long-term total returns, income from rent payments, diversification benefits, and hedging against inflation.

Real estate investment categories include residential properties, commercial real estate, real estate investment trusts, and mortgage-backed securities.

LOS 60.d

Infrastructure refers to long-lived assets that provide public services and are often built or operated by governments. Infrastructure investments may be classified as greenfield (assets to

be built) or brownfield (existing assets).

Liquidity is low for direct investments in infrastructure because the assets are long-lived and tend to be large scale. However, some liquid investment vehicles exist that are backed by infrastructure assets.

LOS 60.e

Natural resource investments include commodities, farmland, and timberland.

Exposure to commodity prices is most commonly achieved with derivatives. Other methods of exposure include commodity ETFs, managed futures funds, and specialized funds in specific commodity sectors.

Commodity valuation is described by the following equation:

$$\text{futures price} \approx \text{spot price} \times (1 + \text{risk-free rate}) + \text{storage costs} - \text{convenience yield}$$

Returns on farmland and timberland include sales of the products raised on the land.

Timberland returns also include price changes on the land, which depend on expected lumber prices and how much timber has been harvested. Farmland returns also depend on land price changes, changes in farm commodity prices, and the quality and quantity of the crops produced.

LOS 60.f

Hedge funds generally use leverage, take both long and short positions, and use derivatives. They typically have restrictions on limited partner redemptions, including a lockup period during which limited partners either cannot request redemptions or incur significant fees for redemptions and a notice period to fulfill a redemption request after the lockup period.

A fund-of-funds invests in other hedge funds. Its advantages are that it can give investors diversification among hedge fund strategies, provides expertise in selecting hedge funds, and provides access to hedge funds that are otherwise unavailable. Its disadvantage is an additional layer of fees beyond the fees charged by the hedge funds in the portfolio.

Event-driven strategies include merger arbitrage, distressed/restructuring, activist, and special situations.

Relative value strategies seek profits from unusual pricing issues.

Macro strategies are “top-down” strategies based on global economic trends.

Equity hedge strategies are “bottom-up” strategies that take long and short positions in equities and equity derivatives. Strategies include market neutral, fundamental growth, fundamental value, quantitative directional, short bias, and sector specific.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 60.1

1. B Venture capital funds typically make their initial investments during a firm's seed stage for product development, marketing, and market research. At the angel investing stage, the

funding source is usually individuals rather than venture capital funds. The start-up stage or early stage follows the seed stage and refers to investments made to fund initial commercial production and sales. (LOS 60.a)

2. **C** In a secondary sale, a private capital firm sells one of its portfolio companies to another private capital fund or group of private investors. Selling a portfolio company to a competitor in its industry is known as a trade sale. Selling a portfolio company to the public requires an initial public offering. (LOS 60.a)

Module Quiz 60.2

1. **C** Commercial real estate ownership requires long-time horizons and purchasing illiquid assets that require large investment amounts. (LOS 60.c)
2. **A** While commodity futures retain the risk and correlation characteristics of the underlying commodities, the investor does not incur storage costs. Derivatives cannot have higher correlation with spot prices than the commodity itself, as its price is the spot price. Convenience yield is a benefit of owning the actual commodities. (LOS 60.e)
3. **B** Greenfield investments refer to infrastructure assets that are yet to be constructed. (LOS 60.d)

Module Quiz 60.3

1. **C** A fund-of-funds manager is expected to provide more due diligence and better redemption terms. Funds-of-funds charge an additional layer of fees. Investing in funds-of-funds may provide more diversification, but may not necessarily provide higher returns. (LOS 60.f)
2. **A** Adding hedge funds to traditional portfolios may not provide the expected diversification to an equity portfolio because return correlations tend to increase during periods of financial crisis. (LOS 60.f)
3. **C** Activist shareholder strategies are a subcategory of event-driven strategies. (LOS 60.f)

TOPIC QUIZ: ALTERNATIVE INVESTMENTS

You have now finished the Alternative Investments topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Alternative Investments. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares with the scores of others who entered their answers.

APPENDIX

Rates, Returns, and Yields

A **holding period return (HPR)**, or holding period yield (HPY), can be for a period of any length and is simply the percentage increase in value over the period, which is calculated as:

$$\text{HPR} = \text{ending value} / \text{beginning value} - 1$$

1. If an investor puts \$2,000 into an account and 565 days later it has grown in value to \$2,700, the 565-day HPY is $2,700 / 2,000 - 1 = 35\%$.
2. If an investor buys a share of stock for \$20/share, receives a \$0.40 dividend, and sells the shares after nine months, the nine-month HPY is $(22 + 0.40) / 20 - 1 = 12\%$.

An HPR for a given period is also the **effective yield** for that period.

An **effective annual yield** is the HPR for a one-year investment or the HPY for a different period converted to its annual equivalent yield.

3. If the six-month HPR is 2%, the effective annual yield is $1.02^2 - 1 = 4.040\%$.
4. If the 125-day HPR is 1.5%, the effective annual yield is $1.015^{365/125} - 1 = 4.443\%$.
5. If the two-year HPR (two-year effective rate) is 9%, the effective annual yield is $1.09^{1/2} - 1 = 4.4031\%$.

Compounding Frequency

Sometimes the “rate” on an investment is expressed as a **simple annual rate** (or *stated rate*)—the annual rate with no compounding of returns. The number of compounding periods per year is called the **periodicity** of the rate. For a periodicity of one, the stated rate and the effective annual rate are the same. When the periodicity is greater than one (more than one compounding period per year), the effective annual rate is the effective rate for the sub-periods, compounded for the number of sub-periods.

6. A bank CD has a stated annual rate of 6% with annual compounding (periodicity of 1); the effective annual rate is 6% and a \$1,000 investment will return $\$1,000(1.06) = \$1,060$ at the end of one year.
7. A bank CD has a stated annual rate of 6% with semiannual compounding (periodicity of 2); the effective annual rate is $(1 + 0.06 / 2)^2 = 1.03^2 - 1 = 6.09\%$ and a \$1,000 investment will return $\$1,000(1.0609) = \$1,060.90$ at the end of one year.
8. A bank CD has a stated annual rate of 6% with quarterly compounding (periodicity of 4); the effective annual rate is $(1 + 0.06 / 4)^4 = 1.015^4 - 1 = 6.136\%$ and a \$1,000 investment will return $\$1,000(1.06136) = \$1,061.36$ at the end of one year.

Note that increasing compounding frequency increases the effective annual yield for any given stated rate. In the limit, as compounding periods get shorter (more frequent), compounding is *continuous*. A stated rate of $r\%$, with continuous compounding, results in an effective annual return of $e^r - 1$.

9. A bank CD has a stated annual rate of 6%, continuously compounded; its effective annual yield is $e^{0.06} - 1 = 6.184\%$ and a \$1,000 investment will return \$1,061.84 at the end of one year.

Bond Quotations and Terminology

The **stated (coupon) rate** on a bond is the total cash coupon payments made over one year as a percentage of face value.

10. A bond with a face value of \$1,000 that pays a coupon of \$50 once each year (an annual-pay bond) has a stated (coupon) rate of $50 / 1,000 = 5\%$ and we say it has a periodicity of 1.
11. A bond with a face value of \$1,000 that pays a coupon of \$25 twice each year (a semiannual-pay bond) has a stated (coupon) rate of $(25 + 25) / 1,000 = 5\%$ and we say it has a periodicity of 2.
12. A bond with a face value of \$1,000 that pays a coupon of \$12.50(1.25%) four times each year (a quarterly-pay bond) has a coupon rate of $(12.50 + 12.50 + 12.50 + 12.50) / 1,000 = 5\%$ and we say it has a periodicity of 4.

The **current yield** on a bond is the stated (coupon) rate divided by the bond price as a percentage of face value or, alternatively, the sum of the coupon payments for one year divided by the bond price.

13. A bond with a stated coupon rate of 5% that is selling at 98.54% of face value has a current yield of $5 / 98.54 = 5.074\%$.
14. A bond that is trading at \$1,058 and makes annual coupon payments that sum to \$50 has a current yield of $50 / 1,058 = 4.726\%$.

The **yield to maturity** (YTM) of a bond, on an *annual basis*, is the effective annual yield and is used for bonds that pay an annual coupon. For bonds that pay coupons semiannually, we often quote the YTM on a *semiannual basis*, that is, two times the effective semiannual yield. To compare the yields of two bonds, we must calculate their YTMs on the same basis.

15. A bond with a YTM of 5% on a semiannual basis has a YTM on an annual basis (effective annual yield) of $(1 + 0.05 / 2)^2 - 1 = 5.0625\%$.
16. A bond with a YTM of 5% on an annual basis has a YTM on a semiannual basis of $(1.05^{1/2} - 1) \times 2 = 4.939\%$.

Internal Rate of Return (IRR)

The internal rate of return is the discount rate that makes the PV of a series of cash flows equal to zero. This calculation must be done with a financial calculator. We use the IRR for calculating the return on a capital project, the YTM on a bond, and the money weighted rate of return for a portfolio.

17. For the YTM of an annual-pay bond (YTM on an annual basis) on a coupon date with N years remaining until maturity, we calculate the annual IRR that satisfies:

$$-\text{bond price} + \frac{\text{coupon 1}}{1 + \text{IRR}} + \frac{\text{coupon 2}}{(1 + \text{IRR})^2} + \dots + \frac{\text{coupon N} + \text{face value}}{(1 + \text{IRR})^N} = 0$$

18. For the YTM of a semiannual-pay bond on a coupon date with N years remaining until maturity, we calculate the IRR that satisfies:

$$-\text{bond price} + \frac{\text{coupon 1}}{1 + \frac{\text{IRR}}{2}} + \frac{\text{coupon 2}}{(1 + \frac{\text{IRR}}{2})^2} + \dots + \frac{\text{coupon } 2N + \text{face value}}{(1 + \frac{\text{IRR}}{2})^{2N}} = 0$$

After solving for $\text{IRR} / 2$, which is the IRR for semiannual periods, we must multiply it by 2 to get the bond's YTM on a semiannual basis.

19. For a capital project, the (annual) IRR satisfies:

$$-\text{initial outlay} + \frac{\text{CF}_1}{1 + \text{IRR}} + \frac{\text{CF}_2}{(1 + \text{IRR})^2} + \dots + \frac{\text{CF}_N}{(1 + \text{IRR})^N} = 0$$

where annual cash flows (CF) can be positive or negative (when a future expenditure is required). Note that if the sign of the cash flows changes more than once, there may be more than one IRR that satisfies the equation.

Money Market Securities

For some money market securities, such as U.S. T-bills, price quotations are given on a bond discount (or simply discount) basis. The bond discount yield (BDY) is the percentage discount from face value of a T-bill, annualized based on a 360-day year, and is therefore not an effective yield but simply an annualized discount from face value.

20. A T-bill that will pay \$1,000 at maturity in 180 days is selling for \$984, a discount of $1 - 984 / 1,000 = 1.6\%$. The annualized discount is $1.6\% \times 360 / 180 = 3.2\%$.

21. A 120-day T-bill is quoted at a BDY of 2.83%, its price is $[1 - (0.0283 \times 120 / 360)] \times 1,000 = \990.57 . Its 120-day *holding period return* is $1,000 / 990.57 - 1 = 0.952\%$. Its *effective annual yield* is $(1,000 / 990.57)^{365/120} - 1 = 2.924\%$.

LIBOR (London Interbank Offered Rate) is an add-on rate quoted for several currencies and for several periods of one year or less, as an annualized rate.

22. HPY on a 30-day loan at a quoted LIBOR rate of 1.8% is $0.018 \times 30 / 360 = 0.15\%$ so the interest on a \$10,000 loan is $10,000 \times 0.0015 = \$15$.

A related yield is the **money market yield (MMY)**, which is HPY annualized based on a 360-day year.

23. A 120-day discount security with a maturity value of \$1,000 that is priced at \$995 has a money market yield of $(1,000 / 995 - 1) \times 360 / 120 = 1.5075\%$.

Forward rates are rates for a loan to be made in a future period. They are quoted based on the period of the loan. For loans of one year, we write 1y1y for a 1-year loan to be made one year from today and 2y1y for a 1-year loan to be made two years from today.

Spot rates are discount rates for single payments to be made in the future (such as for zero-coupon bonds).

24. Given a 3-year spot rate expressed as a compound annual rate (S_3) of 2%, a 3-year bond that makes a single payment of \$1,000 in three years has a current value of $1,000 / (1 + 0.02)^3 = \$942.32$.

An N -year spot rate is the geometric mean of the individual annual forward rates:

$$S_N = [(1 + S_1)(1 + 1y1y)(1 + 2y1y)\dots(1 + Ny1y)]^{1/N} - 1$$

and the annualized forward rate for $M - N$ periods, N periods from now is:

$$Ny(M-N)y = \left[\frac{(1 + S_M)^M}{(1 + S_N)^N} \right]^{\frac{1}{M-N}} - 1$$

25. Given $S_5 = 2.4\%$ and $S_7 = 2.6\%$, $5y2y = [(1.026)^7 / (1.024)^5]^{1/2} - 1 = 3.1017\%$, which is approximately equal to $(7 \times 2.6\% - 5 \times 2.4\%) / 2 = 3.1\%$.

FORMULAS

Quoted add-on yield = HPY × 365/days to maturity

Quoted discount yield = discount on the security × 360/days to maturity

for an annual-coupon bond with N years to maturity:

$$\text{price} = \frac{\text{coupon}}{(1 + \text{YTM})} + \frac{\text{coupon}}{(1 + \text{YTM})^2} + \dots + \frac{\text{coupon} + \text{principal}}{(1 + \text{YTM})^N}$$

for a semiannual-coupon bond with N years to maturity:

$$\text{price} = \frac{\text{coupon}}{\left(1 + \frac{\text{YTM}}{2}\right)} + \frac{\text{coupon}}{\left(1 + \frac{\text{YTM}}{2}\right)^2} + \dots + \frac{\text{coupon} + \text{principal}}{\left(1 + \frac{\text{YTM}}{2}\right)^{N \times 2}}$$

bond value using spot rates:

$$\text{no-arbitrage price} = \frac{\text{coupon}}{(1 + S_1)} + \frac{\text{coupon}}{(1 + S_2)^2} + \dots + \frac{\text{coupon} + \text{principal}}{(1 + S_N)^N}$$

full price between coupon payment dates:

$$(\text{Bond value at last coupon date based on the current YTM}) \times (1 + \text{YTM}/\#)^{t/T}$$

where # is the number of coupon periods per year, t is the number of days from the last coupon payment date until the date the bond trade will settle, and T is the number of days in the coupon period.

flat price = full price - accrued interest

$$\text{current yield} = \frac{\text{annual cash coupon payment}}{\text{bond price}}$$

forward and spot rates: $(1 + S_2)^2 = (1 + S_1)(1 + 1y1y)$

option-adjusted spread: OAS = Z-spread - option value

$$\text{modified duration} = \frac{\text{Macaulay duration}}{1 + \text{YTM}}$$

$$\text{approximate modified duration} = \frac{V_- - V_+}{2 \times V_0 \times \Delta \text{YTM}}$$

$$\text{effective duration} = \frac{V_- - V_+}{2 \times V_0 \times \Delta \text{curve}}$$

money duration = annual modified duration × full price of bond position

money duration per 100 units of par value =

annual modified duration × full bond price per 100 of par value

price value of a basis point: PVBP = $[(V_- - V_+) / 2] \times \text{par value} \times 0.01$

$$\text{approximate convexity} = \frac{V_- + V_+ - 2V_0}{(\Delta YTM)^2 V_0}$$

$$\text{approximate effective convexity} = \frac{V_- + V_+ - 2V_0}{(\Delta \text{curve})^2 V_0}$$

$$\% \Delta \text{ full bond price} = -\text{annual modified duration}(\Delta YTM) + \frac{1}{2} \text{annual convexity}(\Delta YTM)^2$$

$$\% \Delta \text{ bond value} = -\text{duration}(\Delta \text{spread}) + \frac{1}{2} \text{convexity}(\Delta \text{spread})^2$$

duration gap = Macaulay duration – investment horizon

$$\text{return impact} \approx -\text{duration} \times \Delta \text{spread} + \frac{1}{2} \text{convexity} \times (\Delta \text{spread})^2$$

no-arbitrage forward price: $F_0(T) = S_0 (1 + R_f)^T$

payoff to long forward at expiration = $S_T - F_0(T)$

value of forward at time t : $V_t(T) = [S_t + PV_t(\text{costs}) - PV_t(\text{benefit})] - F_0(T) (1 + R_f)^{-(T-t)}$

exercise value of a call = $\text{Max}[0, S - X]$

exercise value of a put = $\text{Max}[0, X - S]$

option value = exercise value + time value

put-call parity: $c + X(1 + R_f)^{-T} = S + p$

put-call-forward parity: $F_0(T)(1 + R_f)^{-T} + p_0 = c_0 + X(1 + R_f)^{-T}$

no-arbitrage forward price: $F_0(T) = S_0 (1 + R_f)^T$

payoff to long forward at expiration = $S_T - F_0(T)$

value of forward at time t : $V_t(T) = [S_t + PV_t(\text{costs}) - PV_t(\text{benefit})] - F_0(T) (1 + R_f)^{-(T-t)}$

exercise value of a call = $\text{Max}[0, S - X]$

exercise value of a put = $\text{Max}[0, X - S]$

option value = exercise value + time value

put-call parity: $c + X(1 + R_f)^{-T} = S + p$

put-call-forward parity: $F_0(T)(1 + R_f)^{-T} + p_0 = c_0 + X(1 + R_f)^{-T}$

INDEX

A

activist shareholder, 251
add-on yield, 25
adjustable-rate mortgage (ARM), 71
adverse selection, 225
affirmative covenant, 4, 125
agency bonds, 23
agency RMBS, 73
alternative investments, 223
American waterfall, 227
amortizing loan, 8
angel investing, 242
annualized holding period rate of return, 92
appraisal index, 246
approximate modified duration, 98
asset-backed fixed income, 251
at par, 3
average life variability, 78

B

backfill bias, 252
backup lines of credit, 25
balloon payment, 8, 81
bankruptcy remote vehicles, 5
basis points, 10
basis point value (BPV), 186

bearer, 5
benchmark bonds, 23
benchmark spread, 58
best efforts offering, 21
bilateral loan, 24
binomial model, 213
bond indenture, 4
bridge financing, 25
broken PAC, 78
brownfield investments, 247
bullet, 8

C

call option, 12, 154, 197
Calmar ratio, 234
cap, 10
capital-indexed bonds, 11
capital market securities, 3, 20
carry, 174
carrying value, 93
cash flow hedge, 167
cash flow yield, 103
cash market transaction, 143
cash-settled contract, 142
catch-up clause, 227
central bank funds market, 28
central bank funds rates, 28
central clearinghouse (CCH), 145
central clearing mandate, 146
central counterparty (CCP), 146

certificates of deposit (CDs), 28
circuit breakers, 153
clawback provision, 227
clean price, 41
CMBS-level call protection, 81
CMBS structure, 80
co-investing, 225
collateral, 6
collateralized debt obligations (CDO), 83
collateralized mortgage obligations (CMO), 76
collateral manager, 83
collateral trust bonds, 6
commercial mortgage-backed securities (CMBS), 79
commercial paper, 24
commercial real estate, 245
committed capital, 226
commodities, 224
commodity ETFs, 248
commodity trading advisers, 251
conditional prepayment rate (CPR), 75
constant-yield price trajectory, 38
contingency provision, 12
contingent claim, 158
contract size, 141
convenience yield, 173
convenience yield, 249
conventional bonds, 3
conversion price, 14
conversion ratio, 14
conversion value, 14
convertible arbitrage fixed income, 251

convertible mortgage, 71
convexity, 106
convexity bias, 187
corporate bonds, 25
corporate credit ratings, 121
corporate family ratings, 121
corporations, 2
covered bonds, 6
covered bonds, 84
credit default swap (CDS, 154
credit enhancement, 6, 79
credit event, 154
credit-linked coupon bond, 10
credit-linked note (CLN), 26
credit migration risk, 120
credit rating, 121
credit risk, 119
credit spread, 129
credit tranching, 70
currency option bond, 3
current yield, 47, 259

D

deal-by-deal waterfall, 227
debentures, 6
debt service coverage ratio, 80
default risk, 119
deferred coupon bond, 11
deliverable contract, 142
derivative, 141

developed markets, 20
developmental capital, 242
direct investing, 225
direct lending, 243
dirty price, 41
discount, 3
discount margin, 49
distressed debt, 244
distressed/restructuring, 251
domestic bonds, 4
downgrade risk, 120
dry powder, 226
dual-currency bond, 3
duration, 97
duration gap, 111

E

early stage, 242
effective annual yield, 258
effective convexity, 107
effective duration, 100
effective yield, 258
either-or fees, 238
embedded options, 12
emerging markets, 20
enterprise value, 132
equipment trust certificates, 6
equity hedge fund strategies, 252
eurobonds, 4
European waterfall, 227

event-driven strategies, 251
exchange-traded derivatives, 145
exercise price, 154
exercise value, 199
expansion venture capital, 242
expected loss, 120

F

fair value hedge, 167
farmland, 224
fiduciary call, 207
fixed-rate mortgage, 71
flat price, 41
floaters, 10
floating-rate notes, 10
floating-rate note yields, 48
floor, 10
foreign bonds, 4
formative stage, 242
forward commitment, 158
forward contract, 141, 151
forward price, 141
forward rate agreement (FRA), 181
forward rates, 53
forward yield curve, 53
founder's shares, 238
four Cs of credit analysis, 123
fully amortizing, 8, 72
fundamental long/short growth, 252
fundamental value, 252

fund investing, 225

fund-of-funds, 250

futures contract, 151

G

general fixed income, 251

general obligation bonds, 133

general partner (GP), 226

global bonds, 5

greenfield investments, 247

grey market, 21

G-spread, 58

guarantee certificate, 27

H

haircut, 29

hard commodities, 144

hard hurdle rate, 226

hedge accounting, 167

hedged, 143

hedge funds, 224

hedge ratio, 214

high-water mark, 227

high yield, 130

holding period return, 258

hurdle rate, 226

I

implied forward rate, 180

incentive fee, 226
income yield, 47
indexed-annuity bonds, 11
indexed zero-coupon bonds, 11
index-linked bonds, 11, 20
index-referenced mortgage, 71
inflation-linked bonds, 11
infrastructure, 224
infrastructure investments, 246
initial margin, 152
initial PAC collar, 78
interbank funds, 28
interest-indexed bonds, 11
interest-only mortgage, 72
interpolated spreads (I-spreads), 58
intrinsic value, 199
inverse floater, 10
investment grade, 122
investment horizon, 93
IPO, 243

K

key rate duration, 102

L

later-stage investment, 242
leveraged buyouts (LBOs), 241
leveraged loan, 243
LIBOR, 261
limited partnership agreement, 226

limited partnerships, 226
limited partners (LPs), 226
linkers, 11
loan-level call protection, 81
loan-to-value ratio (LTV), 71, 80
lockout period, 82
lockup period, 250
loss severity, 119

M

Macaulay duration, 97
macro strategies, 251
maintenance margin, 152
make-whole provisions, 13
managed futures funds, 248
management buy-ins (MBIs), 242
management buyouts (MBOs), 242
management fee, 226
margin, 151
market liquidity risk, 120
market neutral, 252
market reference rate, 20, 153
mark-to-market, 152
material adverse change, 25
matrix pricing, 42
medium-term notes (MTNs), 26
merger arbitrage, 251
mezzanine loans, 243
mezzanine-stage financing, 242
minority equity investing, 242

modified duration, 98
money duration, 104
money market securities, 3, 20
money market yield, 261
money multiple, 234
moneyness, 197
mortgage-backed security (MBS), 6, 69
mortgage pass-through securities, 73
multiple of invested capital, 234
multistrategy, 251
municipal bonds, 20, 133

N

national bond market, 4
negative covenant, 4, 125
negotiable certificates of deposit, 28
net cost of carry , 174
net investment hedge, 167
no-arbitrage, 171
no-arbitrage condition, 171
no-arbitrage price, 40
nonagency RMBS, 73
noninvestment grade, 122
nonrecourse loans, 72, 73
non-sovereign government debt, 23, 133
non-sovereign governments, 2
notching, 122
notice period, 250
notional principal, 153
novation, 145

O

option-adjusted spread (OAS), 60

option-adjusted yield, 48

option-free bonds, 12

option premium, 155, 199

original issue discount, 7

overnight repo, 28

P

parallel shift, 103

par bond yield curve, 53

pari passu, 121

par swap rate, 192

partial duration, 102

partially amortizing, 9, 72

partially hedged, 143

par value, 3

pass-through rates, 73

payment-in-kind (PIK) bond, 11

performance fee, 226

periodicity, 45, 258

perpetual bonds, 2

plain vanilla, 3

planned amortization class (PAC), 77

portfolio companies, 241

premium, 3

prepayment, 72

prepayment option, 100

prepayment penalty, 72

prepayment risk, 75

pre-seed capital, 242
price limits, 152
price value of a basis point (PVBP), 104
primary dealers, 22
primary market, 21
principal amount, 3
principal protected bonds, 11
priority of claims, 120
private capital, 241
private debt, 224, 243
private equity, 224
private investment in public equities (PIPEs), 242
private placement, 21
public offering, 21
pure discount bonds, 3
put-call-forward parity, 209
put-call parity, 207
put option, 13, 198
put options, 154

Q

quasi-government bonds, 23
quasi-government entities, 2
quoted margin, 49
quoted price, 41

R

rating agencies, 121
real estate, 245
real estate, 224

real estate investment trusts (REITs), 245

recapitalization, 243

recourse loans, 72

recovery rate, 120

reference rate, 10

registered bonds, 5

REIT indices, 246

relative value strategies, 251

repeat sales index, 246

replication, 172

repo date, 29

repo margin, 29

repo rate, 28

repurchase (repo) agreement, 28

required margin, 49

residential mortgage loan, 71

residential property, 245

revenue bonds, 134

reverse repo agreement, 29

rollover risk, 25

running yield, 47

S

secondary market, 22

secondary sale, 243

sector specific, 252

secured bonds, 6

secured debt, 120

securitization, 67

securitized bonds, 5

securitized mortgage, 73
seed capital, 242
seed stage, 242
semiannual bond basis, 45
seniority ranking, 120
senior/subordinated structure, 70
sequential pay CMO, 76
serial bond issue, 25
settlement date, 141
Sharpe ratio, 233
shelf registration, 22
shifting interest mechanism, 79
short bias, 252
side letters, 226, 238
simple annual rate, 258
simple yield, 47
simply cost of carry , 174
sinking fund provisions, 9
soft commodities, 144
soft hurdle rate, 226
soft lockup, 250
Sortino ratio, 234
sovereign bonds, 22, 133
sovereign national governments, 2
special purpose entity (SPE), 5, 68
special purpose vehicle (SPV), 68
special situations, 251
speculating, 143
split coupon bond, 11
spot price, 142
spot rates, 40

spot rate yield curve, 52
spread risk, 120
start-up stage, 242
stated (coupon) rate, 259
step-up coupon bonds, 10
street convention, 46
strike price, 154
strip curve, 52
structural subordination, 122
structured finance CDOs, 83
support tranches, 77
supranational bonds, 23
supranational entities, 2
survivorship bias, 252
swap execution facility (SEF), 146
swap rate, 153
swaps, 153
syndicate, 21
syndicated loan, 24
synthetic, 208
synthetic CDOs, 83

T

tenor, 2
term maturity structure, 26
term repo, 28
term structure of interest rates, 52
term structure of yield volatility, 109
term to maturity, 2
timberland, 224

time to expiration, 154

time tranching, 70

time value, 199

trade sale, 243

traditional investments, 223

tranche, 70

Treynor ratio, 234

true yield, 46

trust deed, 3

U

underlying, 141

underlying asset, 141

underwritten offering, 21

unsecured bonds, 6

unsecured debt, 120

V

value a one-period put option, 216

variable-rate mortgage, 71

venture capital, 242

venture debt, 243

volatility, 251

W

warrants, 14

waterfall, 70, 227

weighted average coupon (WAC), 73

weighted average life, 75

weighted average maturity (WAM), 73

whole-of-fund waterfall, 227

write-off/liquidation, 243

Y

yield curve, 52

yields for money market instruments, 50

yield spread, 58, 120

yield-to-call, 47

yield-to-maturity (YTM), 35, 259

yield-to-worst, 47

Z

zero-coupon bonds, 3

zero-coupon rates, 40

zero curve, 52

zero-volatility spread (Z-spread), 59



2023 CFA®

Exam Prep

SchweserNotes™

Portfolio Management and Ethical
and Professional Standards



LEVEL I BOOK 5

KAPLAN SCHWEISER

Book 5: Portfolio Management and Ethical and Professional Standards

SchweserNotes™ 2023

Level I CFA®



SCHWESERNOTES™ 2023 LEVEL I CFA® BOOK 5: PORTFOLIO MANAGEMENT AND ETHICAL AND PROFESSIONAL STANDARDS

©2022 Kaplan, Inc. All rights reserved.

Published in 2022 by Kaplan, Inc.

10 9 8 7 6 5 4 3 2 1

ISBN: 978-1-0788-2592-4

These materials may not be copied without written permission from the author. The unauthorized duplication of these notes is a violation of global copyright laws and the CFA Institute Code of Ethics. Your assistance in pursuing potential violators of this law is greatly appreciated.

Required CFA Institute disclaimer: Kaplan Schweser is a CFA Institute Prep Provider. Only CFA Institute Prep Providers are permitted to make use of CFA Institute copyrighted materials which are the building blocks of the exam. We are also required to update our materials every year and this is validated by CFA Institute.

CFA Institute does not endorse, promote, review or warrant the accuracy or quality of the product and services offered by Kaplan Schweser. CFA Institute®, CFA® and "Chartered Financial Analyst®" are trademarks owned by CFA Institute.

Certain materials contained within this text are the copyrighted property of CFA Institute. The following is the copyright disclosure for these materials: "Copyright, 2022, CFA Institute. Reproduced and republished from 2023 Learning Outcome Statements, Level I, II, and III questions from CFA® Program Materials, CFA Institute Standards of Professional Conduct, and CFA Institute's Global Investment Performance Standards with permission from CFA Institute. All Rights Reserved."

Disclaimer: The SchweserNotes should be used in conjunction with the original readings as set forth by CFA Institute in their 2023 Level I CFA Study Guide. The information contained in these Notes covers topics contained in the readings referenced by CFA Institute and is believed to be accurate. However, their accuracy cannot be guaranteed nor is any warranty conveyed as to your ultimate exam success. The authors of the referenced readings have not endorsed or sponsored these Notes.

CONTENTS

Learning Outcome Statements (LOS)

PORFOLIO MANAGEMENT

READING 61

Portfolio Management: An Overview

Exam Focus

Module 61.1: Portfolio Management Process

Module 61.2: Asset Management and Pooled Investments

Key Concepts

Answer Key for Module Quizzes

READING 62

Portfolio Risk and Return: Part I

Exam Focus

Module 62.1: Returns Measures

Module 62.2: Risk Aversion

Module 62.3: Portfolio Standard Deviation

Module 62.4: The Efficient Frontier

Key Concepts

Answer Key for Module Quizzes

READING 63

Portfolio Risk and Return: Part II

Exam Focus

Module 63.1: Systematic Risk and Beta

Module 63.2: The CAPM and the SML

Key Concepts

Answer Key for Module Quizzes

READING 64

Basics of Portfolio Planning and Construction

Exam Focus

Module 64.1: Portfolio Planning and Construction

Key Concepts

Answer Key for Module Quizzes

READING 65

The Behavioral Biases of Individuals

Exam Focus

Module 65.1: Cognitive Errors vs. Emotional Biases

Module 65.2: Emotional Biases

Key Concepts

Answer Key for Module Quizzes

READING 66

Introduction to Risk Management

Exam Focus

Module 66.1: Introduction to Risk Management

Key Concepts

Answer Key for Module Quizzes

READING 67

Technical Analysis

Exam Focus

Module 67.1: Technical Analysis

Key Concepts

Answer Key for Module Quizzes

READING 68

Fintech in Investment Management

Exam Focus

Module 68.1: Fintech in Investment Management

Key Concepts

Answer Key for Module Quizzes

Topic Quiz: Portfolio Management

ETHICAL AND PROFESSIONAL STANDARDS

READING 69

Ethics and Trust in the Investment Profession

Exam Focus

Module 69.1: Ethics and Trust

Key Concepts

Answer Key for Module Quizzes

READING 70

Code of Ethics and Standards of Professional Conduct

Exam Focus

Module 70.1: Code and Standards

Answer Key for Module Quizzes

READING 71

Guidance for Standards I–VII

Exam Focus

Module 71.1: Guidance for Standards I(A) and I(B)

Module 71.2: Guidance for Standards I(C) and I(D)

Module 71.3: Guidance for Standard II

Module 71.4: Guidance for Standards III(A) and III(B)

Module 71.5: Guidance for Standards III(C), III(D), and III(E)

Module 71.6: Guidance for Standard IV

Module 71.7: Guidance for Standard V

Module 71.8: Guidance for Standard VI

Module 71.9: Guidance for Standard VII

Answer Key for Module Quizzes

READING 72

Introduction to the Global Investment Performance Standards (GIPS)

Exam Focus

Module 72.1: Introduction to GIPS

Key Concepts

Answer Key for Module Quizzes

READING 73

Ethics Application

Exam Focus

Module 73.1: Ethics Application

Answer Key for Module Quizzes

Topic Quiz: Ethical and Professional Standards

Formulas

Index

LEARNING OUTCOME STATEMENTS (LOS)

61. Portfolio Management: An Overview

The candidate should be able to:

- a. describe the portfolio approach to investing.
- b. describe the steps in the portfolio management process.
- c. describe types of investors and distinctive characteristics and needs of each.
- d. describe defined contribution and defined benefit pension plans.
- e. describe aspects of the asset management industry.
- f. describe mutual funds and compare them with other pooled investment products.

62. Portfolio Risk and Return: Part I

The candidate should be able to:

- a. calculate and interpret major return measures and describe their appropriate uses.
- b. compare the money-weighted and time-weighted rates of return and evaluate the performance of portfolios based on these measures.
- c. describe characteristics of the major asset classes that investors consider in forming portfolios.
- d. explain risk aversion and its implications for portfolio selection.
- e. explain the selection of an optimal portfolio, given an investor's utility (or risk aversion) and the capital allocation line.
- f. calculate and interpret the mean, variance, and covariance (or correlation) of asset returns based on historical data.
- g. calculate and interpret portfolio standard deviation.
- h. describe the effect on a portfolio's risk of investing in assets that are less than perfectly correlated.
- i. describe and interpret the minimum-variance and efficient frontiers of risky assets and the global minimum-variance portfolio.

63. Portfolio Risk and Return: Part II

The candidate should be able to:

- a. describe the implications of combining a risk-free asset with a portfolio of risky assets.
- b. explain the capital allocation line (CAL) and the capital market line (CML).
- c. explain systematic and nonsystematic risk, including why an investor should not expect to receive additional return for bearing nonsystematic risk.
- d. explain return generating models (including the market model) and their uses.
- e. calculate and interpret beta.
- f. explain the capital asset pricing model (CAPM), including its assumptions, and the security market line (SML).
- g. calculate and interpret the expected return of an asset using the CAPM.
- h. describe and demonstrate applications of the CAPM and the SML.
- i. calculate and interpret the Sharpe ratio, Treynor ratio, M^2 , and Jensen's alpha.

64. Basics of Portfolio Planning and Construction

The candidate should be able to:

- a. describe the reasons for a written investment policy statement (IPS).
- b. describe the major components of an IPS.
- c. describe risk and return objectives and how they may be developed for a client.
- d. explain the difference between the willingness and the ability (capacity) to take risk in analyzing an investor's financial risk tolerance.
- e. describe the investment constraints of liquidity, time horizon, tax concerns, legal and regulatory factors, and unique circumstances and their implications for the choice of portfolio assets.
- f. explain the specification of asset classes in relation to asset allocation.
- g. describe the principles of portfolio construction and the role of asset allocation in relation to the IPS.
- h. describe how environmental, social, and governance (ESG) considerations may be integrated into portfolio planning and construction.

65. The Behavioral Biases of Individuals

The candidate should be able to:

- a. compare and contrast cognitive errors and emotional biases.
- b. discuss commonly recognized behavioral biases and their implications for financial decision making.

- c. describe how behavioral biases of investors can lead to market characteristics that may not be explained by traditional finance.

66. Introduction to Risk Management

The candidate should be able to:

- a. define risk management.
- b. describe features of a risk management framework.
- c. define risk governance and describe elements of effective risk governance.
- d. explain how risk tolerance affects risk management.
- e. describe risk budgeting and its role in risk governance.
- f. identify financial and non-financial sources of risk and describe how they may interact.
- g. describe methods for measuring and modifying risk exposures and factors to consider in choosing among the methods.

67. Technical Analysis

The candidate should be able to:

- a. explain principles and assumptions of technical analysis.
- b. describe potential links between technical analysis and behavioral finance.
- c. compare principles of technical analysis and fundamental analysis.
- d. describe and interpret different types of technical analysis charts.
- e. explain uses of trend, support, and resistance lines.
- f. explain common chart patterns.
- g. explain common technical indicators.
- h. describe principles of intermarket analysis.
- i. explain technical analysis applications to portfolio management.

68. Fintech in Investment Management

The candidate should be able to:

- a. describe "fintech".
- b. describe Big Data, artificial intelligence, and machine learning.
- c. describe fintech applications to investment management.
- d. describe financial applications of distributed ledger technology.

69. Ethics and Trust in the Investment Profession

The candidate should be able to:

- a. explain ethics.
- b. describe the role of a code of ethics in defining a profession.
- c. describe professions and how they establish trust.
- d. describe the need for high ethical standards in investment management.
- e. explain professionalism in investment management.
- f. identify challenges to ethical behavior.
- g. compare and contrast ethical standards with legal standards.
- h. describe a framework for ethical decision making.

70. Code of Ethics and Standards of Professional Conduct

The candidate should be able to:

- a. describe the structure of the CFA Institute Professional Conduct Program and the process for the enforcement of the Code and Standards.
- b. identify the six components of the Code of Ethics and the seven Standards of Professional Conduct.
- c. explain the ethical responsibilities required by the Code and Standards, including the sub-sections of each Standard.

71. Guidance for Standards I–VII

The candidate should be able to:

- a. demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity.
- b. recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.
- c. identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards.

72. Introduction to the Global Investment Performance Standards (GIPS)

The candidate should be able to:

- a. explain why the GIPS standards were created, who can claim compliance, and who benefits from compliance.
- b. describe the key concepts of the GIPS Standards for Firms.
- c. explain the purpose of composites in performance reporting.
- d. describe the fundamentals of compliance, including the recommendations of the GIPS standards with respect to the definition of the firm and the firm's definition of discretion.
- e. describe the concept of independent verification.

73. Ethics Application

The candidate should be able to:

- a. evaluate practices, policies, and conduct relative to the CFA Institute Code of Ethics and Standards of Professional Conduct.
- b. explain how the practices, policies, and conduct do or do not violate the CFA Institute Code of Ethics and Standards of Professional Conduct.

READING 61

PORTFOLIO MANAGEMENT: AN OVERVIEW

EXAM FOCUS

Here, we introduce the portfolio management process and the investment policy statement. In this reading, you will learn the investment needs of different types of investors, as well as the different kinds of pooled investments. Later, our reading on Basics of Portfolio Planning and Construction will provide more detail on investment policy statements and investor objectives and constraints.

MODULE 61.1: PORTFOLIO MANAGEMENT PROCESS



Video covering this content is available online.

LOS 61.a: Describe the portfolio approach to investing.

The **portfolio perspective** refers to evaluating individual investments by their contribution to the risk and return of an investor's portfolio. The alternative to taking a portfolio perspective is to examine the risk and return of individual investments in isolation. An investor who holds all his wealth in a single stock because he believes it to be the best stock available is not taking the portfolio perspective—his portfolio is very risky compared to holding a diversified portfolio of stocks. Modern portfolio theory concludes that the extra risk from holding only a single security is not rewarded with higher expected investment returns. Conversely, diversification allows an investor to reduce portfolio risk without necessarily reducing the portfolio's expected return.

In the early 1950s, the research of Professor Harry Markowitz provided a framework for measuring the risk-reduction benefits of diversification. Using the standard deviation of returns as the measure of investment risk, he investigated how combining risky securities into a portfolio affected the portfolio's risk and expected return. One important conclusion of his model is that unless the returns of the risky assets are perfectly positively correlated, risk is reduced by diversifying across assets.

In the 1960s, professors Treynor, Sharpe, Mossin, and Lintner independently extended this work into what has become known as modern portfolio theory (MPT). MPT results in equilibrium expected returns for securities and portfolios that are a linear function of each security's or portfolio's market risk (the risk that cannot be reduced by diversification).

One measure of the benefits of diversification is the **diversification ratio**. It is calculated as the ratio of the risk of an equally weighted portfolio of n securities (measured by its standard

deviation of returns) to the risk of a single security selected at random from the n securities. If the average standard deviation of returns for the n stocks is 25%, and the standard deviation of returns for an equally weighted portfolio of the n stocks is 18%, the diversification ratio is $18 / 25 = 0.72$. If the standard deviation of returns for an equally weighted portfolio is 25%, there are no diversification benefits and the diversification ratio equals one. A *lower* diversification ratio indicates a *greater* risk-reduction benefit from diversification.

While the diversification ratio provides a quick measure of the potential benefits of diversification, an equal-weighted portfolio is not necessarily the portfolio that provides the greatest reduction in risk. Computer optimization can calculate the portfolio weights that will produce the lowest portfolio risk (standard deviation of returns) for a given group of securities.

Portfolio diversification works best when financial markets are operating normally; diversification provides less reduction of risk during market turmoil, such as the credit contagion of 2008. During periods of financial crisis, correlations tend to increase, which reduces the benefits of diversification.

LOS 61.b: Describe the steps in the portfolio management process.

There are three major steps in the portfolio management process:

Step 1: The **planning step** begins with an analysis of the investor's risk tolerance, return objectives, time horizon, tax exposure, liquidity needs, income needs, and any unique circumstances or investor preferences.

This analysis results in an **investment policy statement (IPS)** that details the investor's investment objectives and constraints. It should also specify an objective benchmark (such as an index return) against which the success of the portfolio management process will be measured. The IPS should be updated at least every few years and any time the investor's objectives or constraints change significantly.

Step 2: The **execution step** involves an analysis of the risk and return characteristics of various asset classes to determine how funds will be allocated to the various asset types. Often, in what is referred to as a *top-down* analysis, a portfolio manager will examine current economic conditions and forecasts of such macroeconomic variables as GDP growth, inflation, and interest rates, in order to identify the asset classes that are most attractive. The resulting portfolio is typically diversified across such asset classes as cash, fixed-income securities, publicly traded equities, hedge funds, private equity, and real estate, as well as commodities and other real assets.

Once the asset class allocations are determined, portfolio managers may attempt to identify the most attractive securities within the asset class. Security analysts use model valuations for securities to identify those that appear undervalued in what is termed *bottom-up* security analysis.

Step 3: The **feedback step** is the final step. Over time, investor circumstances will change, risk and return characteristics of asset classes will change, and the actual weights of the assets in the portfolio will change with asset prices. The portfolio manager must monitor these changes and **rebalance** the portfolio periodically in response, adjusting the

allocations to the various asset classes back to their desired percentages. The manager must also measure portfolio performance and evaluate it relative to the return on the benchmark portfolio identified in the IPS.

LOS 61.c: Describe types of investors and distinctive characteristics and needs of each.

Individual investors save and invest for a variety of reasons, including purchasing a house or educating their children. In many countries, special accounts allow citizens to invest for retirement and to defer any taxes on investment income and gains until the funds are withdrawn. Defined contribution pension plans are popular vehicles for these investments. Pension plans are described later in this reading.

Many types of **institutions** have large investment portfolios. An **endowment** is a fund that is dedicated to providing financial support on an ongoing basis for a specific purpose. For example, in the United States, many universities have large endowment funds to support their programs. A **foundation** is a fund established for charitable purposes to support specific types of activities or to fund research related to a particular disease. A typical foundation's investment objective is to fund the activity or research on a continuing basis without decreasing the real (inflation adjusted) value of the portfolio assets. Foundations and endowments typically have long investment horizons, high risk tolerance, and, aside from their planned spending needs, little need for additional liquidity.

The investment objective of a **bank**, simply put, is to earn more on the bank's loans and investments than the bank pays for deposits of various types. Banks seek to keep risk low and need adequate liquidity to meet investor withdrawals as they occur.

Insurance companies invest customer premiums with the objective of funding customer claims as they occur. Life insurance companies have a relatively long-term investment horizon, while property and casualty (P&C) insurers have a shorter investment horizon because claims are expected to arise sooner than for life insurers.

Investment companies manage the pooled funds of many investors. **Mutual funds** manage these pooled funds in particular styles (e.g., index investing, growth investing, bond investing) and restrict their investments to particular subcategories of investments (e.g., large-firm stocks, energy stocks, speculative bonds) or particular regions (emerging market stocks, international bonds, Asian-firm stocks).

Sovereign wealth funds refer to pools of assets owned by a government. For example, the Abu Dhabi Investment Authority, a sovereign wealth fund in the United Arab Emirates funded by Abu Dhabi government surpluses, has approximately USD 700 billion in assets.¹

Figure 61.1 provides a summary of the risk tolerance, investment horizon, liquidity needs, and income objectives for different types of investors.

Figure 61.1: Characteristics of Different Types of Investors

Investor	Risk Tolerance	Investment Horizon	Liquidity Needs	Income Needs
Individuals	Depends on individual	Depends on individual	Depends on individual	Depends on individual
Banks	Low	Short	High	Pay interest
Endowments	High	Long	Low	Spending level
Insurance	Low	Long—life Short—P&C	High	Low
Mutual funds	Depends on fund	Depends on fund	High	Depends on fund
Defined benefit pensions	High	Long	Low	Depends on age

LOS 61.d: Describe defined contribution and defined benefit pension plans.

A **defined contribution pension plan** is a retirement plan in which the firm contributes a sum each period to the employee's retirement account. The firm's contribution can be based on any number of factors, including years of service, the employee's age, compensation, profitability, or even a percentage of the employee's contribution. In any event, the firm makes no promise to the employee regarding the future value of the plan assets. The investment decisions are left to the employee, who assumes all of the investment risk.

In a **defined benefit pension plan**, the firm promises to make periodic payments to employees after retirement. The benefit is usually based on the employee's years of service and the employee's compensation at, or near, retirement. For example, an employee might earn a retirement benefit of 2% of her final salary for each year of service. Consequently, an employee with 20 years of service and a final salary of \$100,000, would receive \$40,000 ($\$100,000 \text{ final salary} \times 2\% \times 20 \text{ years of service}$) each year upon retirement until death. Because the employee's future benefit is defined, the employer assumes the investment risk. The employer makes contributions to a fund established to provide the promised future benefits. Poor investment performance will increase the amount of required employer contributions to the fund.



MODULE QUIZ 61.1

- Compared to investing in a single security, diversification provides investors a way to:
 - increase the expected rate of return.
 - decrease the volatility of returns.
 - increase the probability of high returns.
- Which of the following is *least likely* to be considered an appropriate schedule for reviewing and updating an investment policy statement?
 - At regular intervals (e.g., every year).
 - When there is a major change in the client's constraints.
 - Frequently, based on the recent performance of the portfolio.
- A top-down security analysis begins by:
 - analyzing a firm's business prospects and quality of management.
 - identifying the most attractive companies within each industry.

- C. examining economic conditions.
- 4. Portfolio diversification is *least likely* to protect against losses:
 - A. during severe market turmoil.
 - B. when markets are operating normally.
 - C. when the portfolio securities have low return correlation.
- 5. Low risk tolerance and high liquidity requirements *best* describe the typical investment needs of:
 - A. a defined-benefit pension plan.
 - B. a foundation.
 - C. an insurance company.
- 6. A long time horizon and low liquidity requirements *best* describe the investment needs of:
 - A. an endowment.
 - B. an insurance company.
 - C. a bank.
- 7. In a defined contribution pension plan:
 - A. the employee accepts the investment risk.
 - B. the plan sponsor promises a predetermined retirement income to participants.
 - C. the plan manager attempts to match the fund's assets to its liabilities.
- 8. In a defined benefit pension plan:
 - A. the employee assumes the investment risk.
 - B. the employer contributes to the employee's retirement account each period.
 - C. the plan sponsor promises a predetermined retirement income to participants.

MODULE 61.2: ASSET MANAGEMENT AND POOLED INVESTMENTS



Video covering
this content is
available online.

LOS 61.e: Describe aspects of the asset management industry.

The asset management industry comprises firms that manage investments for clients. Asset management firms include both independent managers and divisions of larger financial services companies. They are referred to as **buy-side firms**, in contrast with **sell-side firms** such as broker-dealers and investment banks.

Full-service asset managers are those that offer a variety of investment styles and asset classes. **Specialist asset managers** may focus on a particular investment style or a particular asset class. A **multi-boutique firm** is a holding company that includes a number of different specialist asset managers.

A key distinction is between firms that use active management and those that use passive management. **Active management** attempts to outperform a chosen benchmark through manager skill, for example by using fundamental or technical analysis. **Passive management** attempts to replicate the performance of a chosen benchmark index. This may include traditional broad market index tracking or a **smart beta** approach that focuses on exposure to a particular market risk factor.

Passive management represents about one-fifth of assets under management. Its share of industry revenue is even smaller because fees for passive management are lower than fees for active management.

Asset management firms may also be classified as traditional or alternative, based on the asset classes they manage. Traditional asset managers focus on equities and fixed-income securities. Alternative asset managers focus on asset classes such as private equity, hedge funds, real estate, or commodities. Profit margins tend to be higher for the alternative asset classes. As a result, many traditional asset managers have been moving into this area, somewhat blurring the distinction between these types of firms.

Some trends in the asset management industry are worth noting:

- The market share for passive management has been growing over time. This is due in part to the lower fees passive managers charge investors, and in part to questions about whether active managers are actually able to add value over time on a risk-adjusted basis, especially in developed markets that are believed to be relatively efficient.
- The amount of data available to asset managers has grown exponentially in recent years. This has encouraged them to invest in information technology and third-party services to process these data, attempting to capitalize on information quickly to make investment decisions.
- **Robo-advisors** are a technology that can offer investors advice and recommendations based on their investment requirements and constraints, using a computer algorithm. These advisors increasingly appeal to younger investors and those with smaller portfolios than have typically been served by asset management firms. They have also lowered the barriers to entry into the asset management industry for firms such as insurance companies.



PROFESSOR'S NOTE

Robo-advisors and issues related to Big Data are discussed further in our reading on Fintech in Investment Management.

LOS 61.f: Describe mutual funds and compare them with other pooled investment products.

Mutual funds are one form of **pooled investments** (i.e., a single portfolio that contains investment funds from multiple investors). Each investor owns shares representing ownership of a portion of the overall portfolio. The total net value of the assets in the fund (pool) divided by the number of such shares issued is referred to as the **net asset value (NAV)** of each share.

With an **open-end fund**, investors can buy newly issued shares at the NAV. Newly invested cash is invested by the mutual fund managers in additional portfolio securities. Investors can **redeem** their shares (sell them back to the fund) at NAV as well. All mutual funds charge a fee for the ongoing management of the portfolio assets, which is expressed as a percentage of the net asset value of the fund. **No-load funds** do not charge additional fees for purchasing shares (up-front fees) or for redeeming shares (redemption fees). **Load funds** charge either up-front fees, redemption fees, or both.

Closed-end funds are professionally managed pools of investor money that do not take new investments into the fund or redeem investor shares. The shares of a closed-end fund trade like equity shares (on exchanges or over-the-counter). As with open-end funds, the portfolio management firm charges ongoing management fees.

Types of Mutual Funds

Money market funds invest in short-term debt securities and provide interest income with very low risk of changes in share value. Fund NAVs are typically set to one currency unit, but there have been instances over recent years in which the NAV of some funds declined when the securities they held dropped dramatically in value. Funds are differentiated by the types of money market securities they purchase and their average maturities.

Bond mutual funds invest in fixed-income securities. They are differentiated by bond maturities, credit ratings, issuers, and types. Examples include government bond funds, tax-exempt bond funds, high-yield (lower rated corporate) bond funds, and global bond funds.

A great variety of **stock mutual funds** are available to investors. **Index funds are passively managed**; that is, the portfolio is constructed to match the performance of a particular index, such as the Standard & Poor's 500 Index. **Actively managed** funds refer to funds where the management selects individual securities with the goal of producing returns greater than those of their benchmark indexes. Annual management fees are higher for actively managed funds, and actively managed funds have higher turnover of portfolio securities (the percentage of investments that are changed during the year). This leads to greater tax liabilities compared to passively managed index funds.

Other Forms of Pooled Investments

Exchange-traded funds (ETFs) are similar to closed-end funds in that purchases and sales are made in the market rather than with the fund itself. There are important differences, however. While closed-end funds are often actively managed, ETFs are most often invested to match a particular index (passively managed). With closed-end funds, the market price of shares can differ significantly from their NAV due to imbalances between investor supply and demand for shares at any point in time. Special redemption provisions for ETFs are designed to keep their market prices very close to their NAVs.

ETFs can be sold short, purchased on margin, and traded at intraday prices, whereas open-end funds are typically sold and redeemed only daily, based on the share NAV calculated with closing asset prices. Investors in ETFs must pay brokerage commissions when they trade, and there is a spread between the bid price at which market makers will buy shares and the ask price at which market makers will sell shares. With most ETFs, investors receive any dividend income on portfolio stocks in cash, while open-end funds offer the alternative of reinvesting dividends in additional fund shares. One final difference is that ETFs may produce less capital gains liability compared to open-end index funds. This is because investor sales of ETF shares do not require the fund to sell any securities. If an open-end fund has significant redemptions that cause it to sell appreciated portfolio shares, shareholders incur a capital gains tax liability.

A **separately managed account** is a portfolio that is owned by a single investor and managed according to that investor's needs and preferences. No shares are issued, as the single investor owns the entire account.

Hedge funds are pools of investor funds that are not regulated to the extent that mutual funds are. Hedge funds are limited in the number of investors who can invest in the fund and are often sold only to qualified investors who have a minimum amount of overall portfolio wealth. Minimum investments can be quite high, often between \$250,000 and \$1 million.

Private equity and **venture capital** funds invest in portfolios of companies, often with the intention to sell them later in public offerings. Managers of funds may take active roles in managing the companies in which they invest.



PROFESSOR'S NOTE

Hedge funds, private equity, and venture capital are addressed in the Alternative Investments topic area.



MODULE QUIZ 61.2

1. Compared to exchange-traded funds (ETFs), open-end mutual funds are typically associated with lower:
 - A. brokerage costs.
 - B. minimum investment amounts.
 - C. management fees.
2. Private equity and venture capital funds:
 - A. expect that only a small percentage of investments will pay off.
 - B. play an active role in the management of companies.
 - C. restructure companies to increase cash flow.
3. Hedge funds *most likely*:
 - A. have stricter reporting requirements than a typical investment firm because of their use of leverage and derivatives.
 - B. hold equal values of long and short securities.
 - C. are not offered for sale to the general public.

KEY CONCEPTS

LOS 61.a

A diversified portfolio produces reduced risk for a given level of expected return, compared to investing in an individual security. Modern portfolio theory concludes that investors that do not take a portfolio perspective bear risk that is not rewarded with greater expected return.

LOS 61.b

The three steps in the portfolio management process are:

1. **Planning:** Determine client needs and circumstances, including the client's return objectives, risk tolerance, constraints, and preferences. Create, and then periodically review and update, an investment policy statement (IPS) that spells out these needs and circumstances.
2. **Execution:** Construct the client portfolio by determining suitable allocations to various asset classes based on the IPS and on expectations about macroeconomic variables such as inflation, interest rates, and GDP growth (top-down analysis). Identify attractively priced securities within an asset class for client portfolios based on valuation estimates from security analysts (bottom-up analysis).
3. **Feedback:** Monitor and rebalance the portfolio to adjust asset class allocations and securities holdings in response to market performance. Measure and report performance relative to the performance benchmark specified in the IPS.

LOS 61.c

Types of investment management clients and their characteristics:

Investor Type	Risk Tolerance	Investment Horizon	Liquidity Needs	Income Needs
Individuals	Depends on individual	Depends on individual	Depends on individual	Depends on individual
Banks	Low	Short	High	Pay interest
Endowments	High	Long	Low	Spending level
Insurance	Low	Long—life Short—P&C	High	Low
Mutual funds	Depends on fund	Depends on fund	High	Depends on fund
Defined benefit pension	High	Long	Low	Depends on age

LOS 61.d

In a defined contribution plan, the employer contributes a certain sum each period to the employee's retirement account. The employer makes no promise regarding the future value of the plan assets; thus, the employee assumes all of the investment risk.

In a defined benefit plan, the employer promises to make periodic payments to the employee after retirement. Because the employee's future benefit is defined, the employer assumes the investment risk.

LOS 61.e

The asset management industry comprises buy-side firms that manage investments for clients. Asset management firms include both independent managers and divisions of larger financial services companies and may be full-service or specialist firms offering investments in traditional or alternative asset classes.

Active management attempts to outperform a chosen benchmark through manager skill. Passive management attempts to replicate the performance of a chosen benchmark index. Most assets under management are actively managed, but the market share for passive management has been increasing.

LOS 61.f

Mutual funds combine funds from many investors into a single portfolio that is invested in a specified class of securities or to match a specific index. Many varieties exist, including money market funds, bond funds, stock funds, and balanced (hybrid) funds. Open-ended shares can be bought or sold at the net asset value. Closed-ended funds have a fixed number of shares that trade at a price determined by the market.

Exchange-traded funds are similar to mutual funds, but investors can buy and sell ETF shares in the same way as shares of stock. Management fees are generally low, though trading ETFs results in brokerage costs.

Separately managed accounts are portfolios managed for individual investors who have substantial assets. In return for an annual fee based on assets, the investor receives personalized investment advice.

Hedge funds are available only to accredited investors and are exempt from most reporting requirements. Many different hedge fund strategies exist. A typical annual fee structure is 20% of excess performance plus 2% of assets under management.

Buyout funds involve taking a company private by buying all available shares, usually funded by issuing debt. The company is then restructured to increase cash flow. Investors typically exit the investment within three to five years.

Venture capital funds are similar to buyout funds, except that the companies purchased are in the start-up phase. Venture capital funds, like buyout funds, also provide advice and expertise to the start-ups.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 61.1

1. **B** Diversification provides an investor reduced risk. However, the expected return is generally similar or less than that expected from investing in a single risky security. Very high or very low returns become less likely. (LOS 61.a)
2. **C** An IPS should be updated at regular intervals and whenever there is a major change in the client's objectives or constraints. Updating an IPS based on portfolio performance is not recommended. (LOS 61.b)
3. **C** A top-down analysis begins with an analysis of broad economic trends. After an industry that is expected to perform well is chosen, the most attractive companies within that industry are identified. A bottom-up analysis begins with criteria such as firms' business prospects and quality of management. (LOS 61.b)
4. **A** Portfolio diversification has been shown to be relatively ineffective during severe market turmoil. Portfolio diversification is most effective when the securities have low correlation and the markets are operating normally. (LOS 61.a)
5. **C** Insurance companies need to be able to pay claims as they arise, which leads to insurance firms having low risk tolerance and high liquidity needs. Defined benefit pension plans and foundations both typically have high risk tolerance and low liquidity needs. (LOS 61.c)
6. **A** An endowment has a long time horizon and low liquidity needs, as an endowment generally intends to fund its causes perpetually. Both insurance companies and banks require high liquidity. (LOS 61.c)
7. **A** In a defined contribution pension plan, the employee accepts the investment risk. The plan sponsor and manager neither promise a specific level of retirement income to participants nor make investment decisions. These are features of a defined benefit plan. (LOS 61.d)
8. **C** In a defined benefit plan, the employer promises a specific level of benefits to employees when they retire. Thus, the employer bears the investment risk. (LOS 61.d)

Module Quiz 61.2

1. **A** Open-end mutual funds do not have brokerage costs, as the shares are purchased from and redeemed with the fund company. Minimum investment amounts and management fees are typically higher for mutual funds. (LOS 61.f)
2. **B** Private equity and venture capital funds play an active role in the management of companies. Private equity funds other than venture capital expect that the majority of investments will pay off. Venture capital funds do not typically restructure companies. (LOS 61.f)

3. **C** Hedge funds may not be offered for sale to the general public; they can be sold only to qualified investors who meet certain criteria. Hedge funds that hold equal values of long and short securities today make up only a small percentage of funds; many other kinds of hedge funds exist that make no attempt to be market neutral. Hedge funds have reporting requirements that are less strict than those of a typical investment firm. (LOS 61.f)

¹ Source: SWF Institute (<https://www.swfinstitute.org/>).

READING 62

PORTFOLIO RISK AND RETURN: PART I

EXAM FOCUS

This reading makes use of many of the statistical and returns measures we covered in Quantitative Methods. You should understand the historical return and risk rankings of the major asset classes and how the correlation (covariance) of returns between assets and between various asset classes affects the risk of portfolios. An investor's degree of risk aversion describes an investor's preferences regarding the tradeoff between risk and return. These preferences, along with the risk and return characteristics of available portfolios, can be used to identify an optimal portfolio for an investor, that is, the portfolio that maximizes the investor's expected utility.

MODULE 62.1: RETURNS MEASURES



LOS 62.a: Calculate and interpret major return measures and describe their appropriate uses.

Video covering this content is available online.

Holding period return (HPR) is simply the percentage increase in the value of an investment over a given time period:

$$\begin{aligned}\text{holding period return} &= \frac{\text{end-of-period value}}{\text{beginning-of-period value}} - 1 = \frac{P_t + \text{Div}_t}{P_0} - 1 \\ &= \frac{P_t - P_0 + \text{Div}_t}{P_0}\end{aligned}$$

If a stock is valued at €20 at the beginning of the period, pays €1 in dividends over the period, and at the end of the period is valued at €22, the HPR is:

$$\text{HPR} = (22 + 1) / 20 - 1 = 0.15 = 15\%$$

Average Returns

The **arithmetic mean return** is the simple average of a series of periodic returns. It has the statistical property of being an unbiased estimator of the true mean of the underlying distribution of returns:

$$\text{arithmetic mean return} = \frac{(R_1 + R_2 + R_3 + \dots + R_n)}{n}$$

The **geometric mean return** is a compound annual rate. When periodic rates of return vary from period to period, the geometric mean return will have a value less than the arithmetic mean return:

$$\text{geometric mean return} = \sqrt[n]{(1 + R_1) \times (1 + R_2) \times (1 + R_3) \times \dots \times (1 + R_n)} - 1$$

For example, for returns R_t over three annual periods, the geometric mean return is calculated as the following example shows.

EXAMPLE: Return measures

An investor purchased \$1,000 of a mutual fund's shares. The fund had the following total returns over a 3-year period: +5%, -8%, +12%. Calculate the value at the end of the 3-year period, the holding period return, the mean annual return, and the geometric mean annual return.

Answer:

$$\text{ending value} = (1,000)(1.05)(0.92)(1.12) = \$1,081.92$$

$$\text{holding period return} = (1.05)(0.92)(1.12) - 1 = 0.08192 = 8.192\%, \text{ which can also be calculated as } 1,081.92 / 1,000 - 1 = 8.192\%$$

$$\text{arithmetic mean return} = (5\% - 8\% + 12\%) / 3 = 3\%$$

$$\text{geometric mean return} = \sqrt[3]{(1.05)(0.92)(1.12)} - 1 = 0.02659 = 2.66\%,$$

$$\text{which can also be calculated as geometric mean return} = \sqrt[3]{1 + \text{HPR}} - 1 = \sqrt[3]{1.08192} - 1 = 2.66\%.$$

Other Return Measures

Gross return refers to the total return on a security portfolio before deducting fees for the management and administration of the investment account. **Net return** refers to the return after these fees have been deducted. Note that commissions on trades and other costs that are necessary to generate the investment returns are deducted in both gross and net return measures.

Pretax nominal return refers to the return prior to paying taxes. Dividend income, interest income, short-term capital gains, and long-term capital gains may all be taxed at different rates.

After-tax nominal return refers to the return after the tax liability is deducted.

Real return is nominal return adjusted for inflation. Consider an investor who earns a nominal return of 7% over a year when inflation is 2%. The investor's approximate real return is simply $7 - 2 = 5\%$. The investor's exact real return is slightly lower, $1.07 / 1.02 - 1 = 0.049 = 4.9\%$.

Real return measures the increase in an investor's purchasing power: how much more goods she can purchase at the end of one year due to the increase in the value of her investments. If she invests \$1,000 and earns a nominal return of 7%, she will have \$1,070 at the end of the year. If the price of the goods she consumes has gone up 2%, from \$1.00 to \$1.02, she will be able to consume $1,070 / 1.02 = 1,049$ units. She has given up consuming 1,000 units today but instead is able to purchase 1,049 units at the end of one year. Her purchasing power has gone up 4.9%; this is her real return.

A leveraged return refers to a return to an investor that is a multiple of the return on the underlying asset. The leveraged return is calculated as the gain or loss on the investment as a percentage of an investor's cash investment. An investment in a derivative security, such as a futures contract, produces a leveraged return because the cash deposited is only a fraction of the

value of the assets underlying the futures contract. Leveraged investments in real estate are very common: investors pay only a portion of a property's cost in cash and borrow the rest.

LOS 62.b: Compare the money-weighted and time-weighted rates of return and evaluate the performance of portfolios based on these measures.

The **money-weighted return** applies the concept of IRR to investment portfolios. The money-weighted rate of return is defined as the internal rate of return on a portfolio, taking into account all cash inflows and outflows. The beginning value of the account is an inflow, as are all deposits into the account. All withdrawals from the account are outflows, as is the ending value.

EXAMPLE: Money-weighted rate of return

Assume an investor buys a share of stock for \$100 at $t = 0$ and at the end of the year ($t = 1$), she buys an additional share for \$120. At the end of Year 2, the investor sells both shares for \$130 each. At the end of each year in the holding period, the stock paid a \$2.00 per share dividend. What is the money-weighted rate of return?

Step 1: Determine the timing of each cash flow and whether the cash flow is an inflow (+), into the account, or an outflow (-), available from the account.

$t = 0:$ purchase of first share = +\$100.00 inflow to account

$t = 1:$ purchase of second share = +\$120.00

dividend from first share = -\$2.00

Subtotal, $t = 1$ +\$118.00 inflow to account

$t = 2:$ dividend from two shares = -\$4.00

proceeds from selling shares = -\$260.00

Subtotal, $t = 2$ -\$264.00 outflow from account

Step 2: Net the cash flows for each time period and set the PV of cash inflows equal to the present value of cash outflows.

$$PV_{\text{inflows}} = PV_{\text{outflows}}$$

$$\$100 + \frac{\$118}{(1+r)} = \frac{\$264}{(1+r)^2}$$

Step 3: Solve for r to find the money-weighted rate of return. This can be done using trial and error or by using the IRR function on a financial calculator or spreadsheet.

The intuition here is that we deposited \$100 into the account at $t = 0$, then added \$118 to the account at $t = 1$ (which, with the \$2 dividend, funded the purchase of one more share at \$120), and ended with a total value of \$264.

To compute this value with a financial calculator, use these net cash flows and follow the procedure(s) described to calculate the IRR.

$$\begin{aligned} \text{Net cash flows: } CF_0 &= +100; CF_1 = +120 - 2 = +118; \\ CF_2 &= -260 + -4 = -264 \end{aligned}$$

Calculating money-weighted return with the TI Business Analyst II Plus®

Note that the values for F01, F02, etc., are all equal to one.

Key Strokes	Explanation	Display
[CF] [2nd][CLR WORK]	Clear Cash Flow Registers	CF0 = 0.00000
100 [ENTER]	Initial Cash Outlay	CF0 = +100.00000
[↓] 118 [ENTER]	Period 1 Cash Flow	C01 = +118.00000
[↓] [↓] 264 [+/-] [ENTER]	Period 2 Cash Flow	C02 = -264.00000
[IRR] [CPT]	Calculate IRR	IRR = 13.86122

The money-weighted rate of return for this problem is 13.86%.



PROFESSOR'S NOTE

In the preceding example, we entered the flows into the account as positive and the ending value as a negative (the investor could withdraw this amount from the account). Note that there is no difference in the solution if we enter the cash flows into the account as negative values (out of the investor's pocket) and the ending value as a positive value (into the investor's pocket). As long as payments into the account and payments out of the account (including the ending value) are entered with opposite signs, the computed IRR will be correct.

Time-weighted rate of return measures compound growth. It is the rate at which \$1 compounds over a specified performance horizon. Time-weighting is the process of averaging a set of values over time. The *annual* time-weighted return for an investment may be computed by performing the following steps:

Step 1: Value the portfolio immediately preceding significant additions or withdrawals. Form subperiods over the evaluation period that correspond to the dates of deposits and withdrawals.

Step 2: Compute the holding period return (HPR) of the portfolio for each subperiod.

Step 3: Compute the product of $(1 + \text{HPR})$ for each subperiod to obtain a total return for the entire measurement period [i.e., $(1 + \text{HPR}_1) \times (1 + \text{HPR}_2) \dots (1 + \text{HPR}_n)$] - 1. If the total investment period is greater than one year, you must take the geometric mean of the measurement period return to find the annual time-weighted rate of return.

EXAMPLE: Time-weighted rate of return

An investor purchases a share of stock at $t = 0$ for \$100. At the end of the year, $t = 1$, the investor buys another share of the same stock for \$120. At the end of Year 2, the investor sells both shares for \$130 each. At the end of both years 1 and 2, the stock paid a \$2 per share dividend. What is the annual time-weighted rate of return for this investment? (This is the same investment as the preceding example.)

Answer:

Step 1: Break the evaluation period into two subperiods based on timing of cash flows.

Holding period 1:	Beginning value	= \$100
	Dividends paid	= \$2
	Ending value	= \$120
Holding period 2:	Beginning value	= \$240 (2 shares)
	Dividends paid	= \$4 (\$2 per share)
	Ending value	= \$260 (2 shares)

Step 2: Calculate the HPR for each holding period.

$$HPR_1 = [(\$120 + 2) / \$100] - 1 = 22\%$$

$$HPR_2 = [(\$260 + 4) / \$240] - 1 = 10\%$$

Step 3: Find the compound annual rate that would have produced a total return equal to the return on the account over the 2-year period.

$$(1 + \text{time-weighted rate of return})^2 = (1.22)(1.10)$$

$$\text{time-weighted rate of return} = [(1.22)(1.10)]^{0.5} - 1 = 15.84\%$$

In the investment management industry, *the time-weighted rate of return is the preferred method of performance measurement, because it is not affected by the timing of cash inflows and outflows.*

In the preceding examples, the time-weighted rate of return for the portfolio was 15.84%, while the money-weighted rate of return for the same portfolio was 13.86%. The results are different because the money-weighted rate of return gave a larger weight to the Year 2 HPR, which was 10%, versus the 22% HPR for Year 1. This is because there was more money in the account at the beginning of the second period.

If funds are contributed to an investment portfolio just before a period of relatively poor portfolio performance, the money-weighted rate of return will tend to be lower than the time-weighted rate of return. On the other hand, if funds are contributed to a portfolio at a favorable time (just prior to a period of relatively high returns), the money-weighted rate of return will be higher than the time-weighted rate of return. The use of the time-weighted return removes these distortions and thus provides a better measure of a manager's ability to select investments over the period. If the manager has complete control over money flows into and out of an account, the money-weighted rate of return would be the more appropriate performance measure.

LOS 62.c: Describe characteristics of the major asset classes that investors consider in forming portfolios.

An examination of the returns and standard deviation of returns for the major investable asset classes supports the idea of a tradeoff between risk and return. Using U.S. data over the period 1926–2017 as an example, shown in Figure 62.1, small-capitalization stocks have had the greatest average returns and greatest risk over the period. T-bills had the lowest average returns and the lowest standard deviation of returns.

Figure 62.1: Risk and Return of Major Asset Classes in the United States (1926–2017)¹

Assets Class	Average Annual Return (Geometric Mean)	Standard Deviation (Annualized Monthly)
Small-cap stocks	12.1%	31.7%
Large-cap stocks	10.2%	19.8%
Long-term corporate bonds	6.1%	8.3%
Long-term government bonds	5.5%	9.9%
Treasury bills	3.4%	3.1%
Inflation	2.9%	4.0%

Results for other markets around the world are similar: asset classes with the greatest average returns also have the highest standard deviations of returns.

The annual nominal return on U.S. equities has varied greatly from year to year, ranging from losses greater than 40% to gains of more than 50%. We can approximate the real returns over the period by subtracting inflation. The asset class with the least risk, T-bills, had a real return of only approximately 0.5% over the period, while the approximate real return on U.S. large-cap stocks was 7.3%. Because annual inflation fluctuated greatly over the period, real returns have been much more stable than nominal returns.

Evaluating investments using expected return and variance of returns is a simplification because returns do not follow a normal distribution; distributions are negatively skewed, with greater kurtosis (fatter tails) than a normal distribution. The negative skew reflects a tendency towards large downside deviations, while the positive excess kurtosis reflects frequent extreme deviations on both the upside and downside. These non-normal characteristics of skewness ($\neq 0$) and kurtosis ($\neq 3$) should be taken into account when analyzing investments.

Liquidity is an additional characteristic to consider when choosing investments because liquidity can affect the price and, therefore, the expected return of a security. Liquidity can be a major concern in emerging markets and for securities that trade infrequently, such as low-quality corporate bonds.



MODULE QUIZ 62.1

- An investor buys a share of stock for \$40 at time $t = 0$, buys another share of the same stock for \$50 at $t = 1$, and sells both shares for \$60 each at $t = 2$. The stock paid a dividend of \$1 per share at $t = 1$ and at $t = 2$. The periodic money-weighted rate of return on the investment is *closest* to:
 - 22.2%.
 - 23.0%.
 - 23.8%.
- Which of the following asset classes has historically had the highest returns and standard deviation of returns?
 - Small-cap stocks.
 - Large-cap stocks.
 - Long-term corporate bonds.

MODULE 62.2: RISK AVERSION



LOS 62.d: Explain risk aversion and its implications for portfolio selection.

Video covering this content is

A **risk-averse** investor is simply one that dislikes risk (i.e., prefers less risk to more risk). Given two investments that have equal expected returns, a risk-averse investor will choose the one with less risk (standard deviation, σ). Financial models assume all investors are risk averse.

A **risk-seeking** (risk-loving) investor would actually prefer more risk to less and, given equal expected returns, would prefer the more risky investment. A **risk-neutral** investor would have no preference regarding risk and would therefore be indifferent between any two investments with equal expected returns.

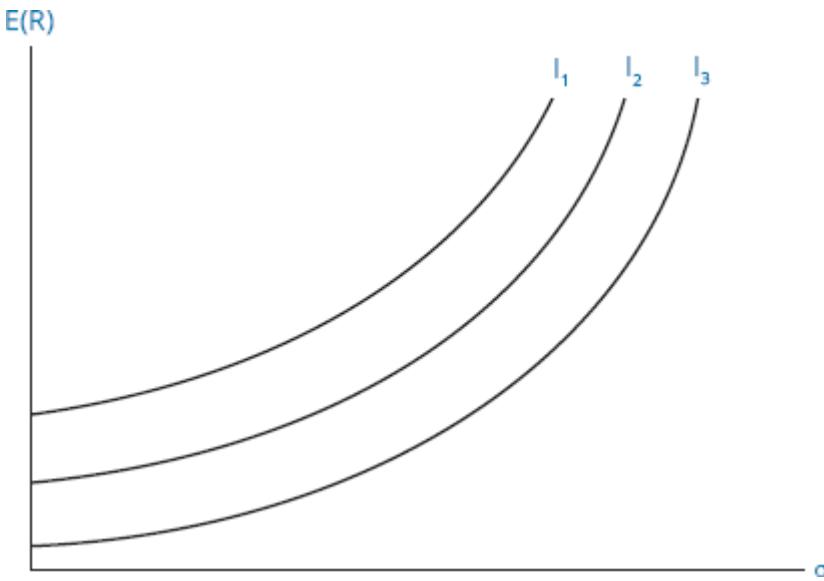
Consider this gamble: A coin will be flipped; if it comes up heads, you receive \$100; if it comes up tails, you receive nothing. The expected payoff is $0.5(\$100) + 0.5(\$0) = \$50$. A risk-averse investor would choose a payment of \$50 (a certain outcome) over the gamble. A risk-seeking investor would prefer the gamble to a certain payment of \$50. A risk-neutral investor would be indifferent between the gamble and a certain payment of \$50.

When the expected returns on two portfolios are equal, a risk-averse investor will always prefer the less risky portfolio. Those who choose high-risk portfolios feel that the increase in expected portfolio returns is adequate compensation for their portfolio's higher risk.

LOS 62.e: Explain the selection of an optimal portfolio, given an investor's utility (or risk aversion) and the capital allocation line.

Investors' **utility functions** represent their preferences regarding the tradeoff between risk and return (i.e., their degrees of risk aversion). An **indifference curve** is a tool from economics that, in this application, plots combinations of risk (standard deviation) and expected returns among which an investor is indifferent. In constructing indifference curves for portfolios based on only their expected return and standard deviation of returns, we are assuming that these are the only portfolio characteristics that investors care about. In Figure 62.2, we show three indifference curves for an investor. The investor's expected utility is the same for all points (portfolios) along any single indifference curve. Portfolios along indifference curve I_1 in Figure 62.2 are preferred to all portfolios along I_2 , which are preferred to all portfolios along I_3 .

Figure 62.2: Risk-Averse Investor's Indifference Curves



Indifference curves slope upward for risk-averse investors because they will only take on more risk (standard deviation of returns) if they are compensated with greater expected returns. An investor who is more risk averse requires a greater increase in expected return to compensate for a given increase in risk than a less risk-averse investor. In other words, the indifference curves of a more risk-averse investor will be steeper than those of a less risk-averse investor, reflecting a higher **risk aversion coefficient**.

In our previous illustration of efficient portfolios available in the market, we included only risky assets. Now we will introduce a risk-free asset into our universe of available assets, and we will examine the risk and return characteristics of a portfolio that combines a portfolio of risky assets and a risk-free asset. As we have seen, we can calculate the expected return and standard deviation of a portfolio with weight W_A allocated to risky Asset A and weight W_B allocated to risky Asset B using the following formulas:

$$E(R_{\text{portfolio}}) = W_A E(R_A) + W_B E(R_B)$$

$$\sigma_{\text{portfolio}} = \sqrt{W_A^2 \sigma_A^2 + W_B^2 \sigma_B^2 + 2 W_A W_B \rho_{AB} \sigma_A \sigma_B}$$

Allow Asset B to be the risk-free asset and Asset A to be the risky asset portfolio. Because a risk-free asset has zero standard deviation and zero correlation of returns with those of a risky portfolio, this results in the reduced equation:

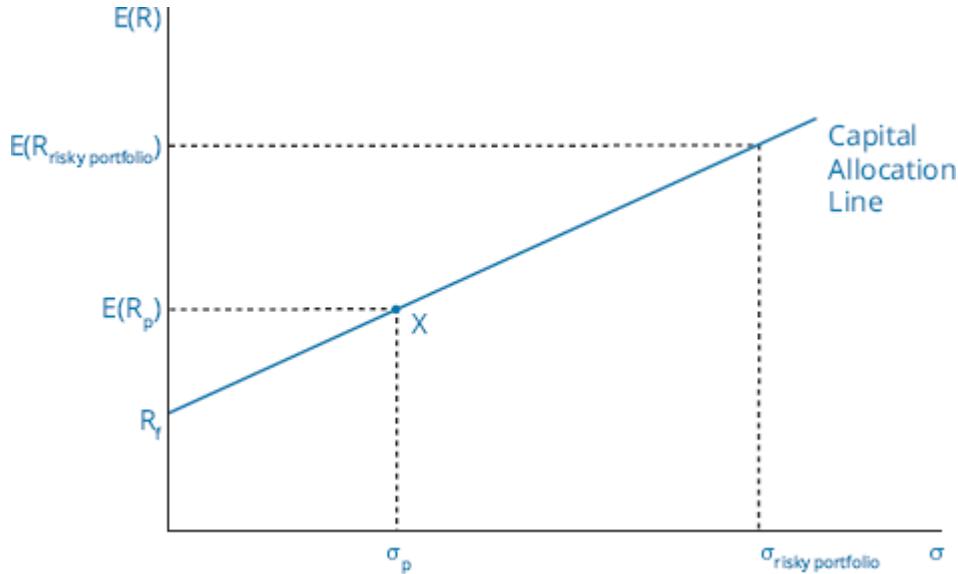
$$\sigma_{\text{portfolio}} = \sqrt{W_A^2 \sigma_A^2} = W_A \sigma_A$$

The intuition of this result is straightforward: If we put X% of our portfolio into the risky asset, and the rest into the risk-free asset, our portfolio will have X% of the risk of the risky asset. The relationship between portfolio risk and return for various portfolio allocations is linear, as illustrated in Figure 62.3.

Combining a risky portfolio with a risk-free asset is the process that supports the **two-fund separation theorem**, which states that all investors' optimal portfolios will be made up of some combination of the optimal portfolio of risky assets and the risk-free asset. The line representing these possible combinations of risk-free assets and the optimal risky asset portfolio is referred to as the **capital allocation line**.

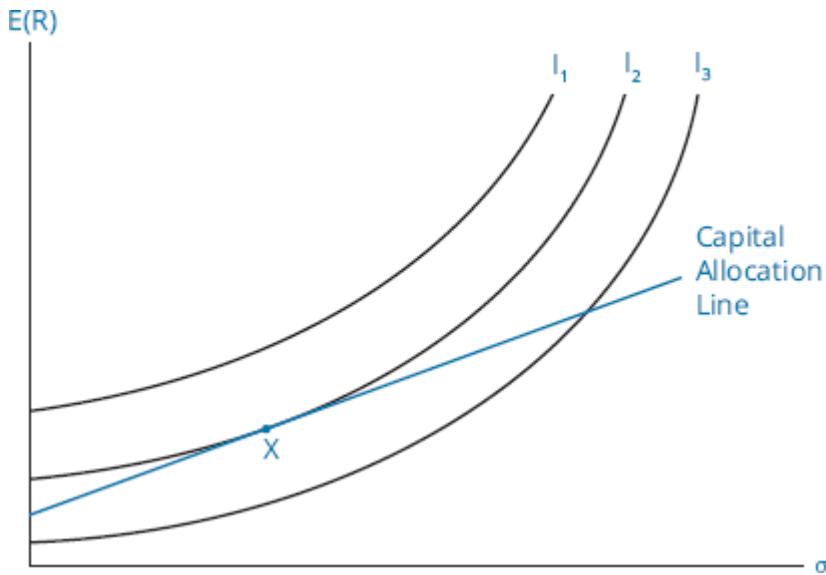
Point X on the capital allocation line in Figure 62.3 represents a portfolio that is 40% invested in the risky asset portfolio and 60% invested in the risk-free asset. Its expected return will be $0.40[E(R_{\text{risky asset portfolio}})] + 0.60(R_f)$, and its standard deviation will be $0.40(\sigma_{\text{risky asset portfolio}})$.

Figure 62.3: Capital Allocation Line and Risky Asset Weights



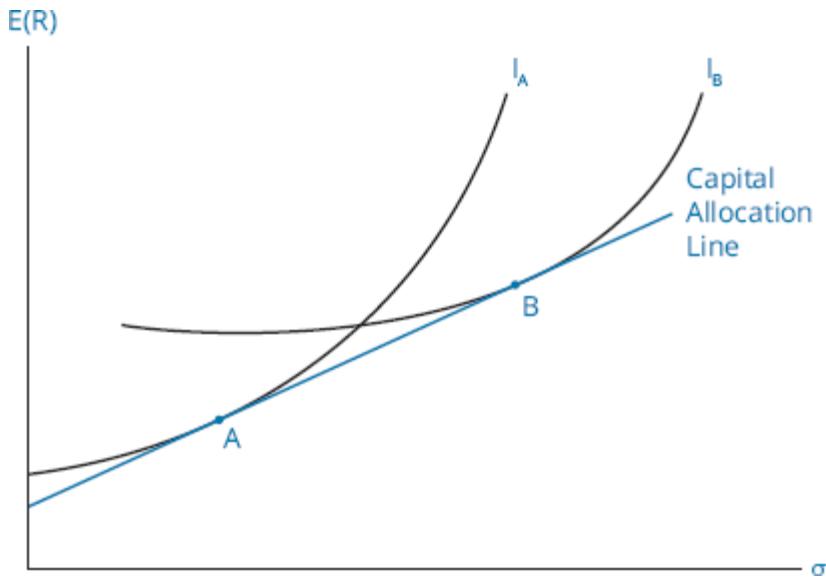
Now that we have constructed a set of the possible efficient portfolios (the capital allocation line), we can combine this with indifference curves representing an individual's preferences for risk and return to illustrate the logic of selecting an optimal portfolio (i.e., one that maximizes the investor's expected utility). In Figure 62.4, we can see that Investor A, with preferences represented by indifference curves I_1 , I_2 , and I_3 , can reach the level of expected utility on I_2 by selecting Portfolio X. This is the optimal portfolio for this investor, as any portfolio that lies on I_2 is preferred to all portfolios that lie on I_3 (and in fact to any portfolios that lie between I_2 and I_3). Portfolios on I_1 are preferred to those on I_2 , but none of the portfolios that lie on I_1 are available in the market.

Figure 62.4: Risk-Averse Investor's Indifference Curves



The final result of our analysis here is not surprising; investors who are less risk averse will select portfolios with more risk. Recall that the lower an investor's risk aversion, the flatter his indifference curves. As illustrated in Figure 62.5, the flatter indifference curve for Investor B (I_B) results in an optimal (tangency) portfolio that lies to the right of the one that results from a steeper indifference curve, such as that for Investor A (I_A). An investor who is less risk averse should optimally choose a portfolio with more invested in the risky asset portfolio and less invested in the risk-free asset.

Figure 62.5: Portfolio Choices Based on Two Investors' Indifference Curves



MODULE QUIZ 62.2

1. Which of the following statements about risk-averse investors is *most accurate*? A risk-averse investor:
- seeks out the investment with minimum risk, while return is not a major consideration.
 - will take additional investment risk if sufficiently compensated for this risk.
 - avoids participating in global equity markets.

2. The capital allocation line is a line from the risk-free return through:
 - A. the global maximum-return portfolio.
 - B. the optimal risky portfolio.
 - C. the global minimum-variance portfolio.

MODULE 62.3: PORTFOLIO STANDARD DEVIATION



Video covering this content is available online.

LOS 62.f: Calculate and interpret the mean, variance, and covariance (or correlation) of asset returns based on historical data.

Variance (Standard Deviation) of Returns for an Individual Security

In finance, the variance and standard deviation of returns are common measures of investment risk. Both of these are measures of the variability of a distribution of returns about its mean or expected value.

We can calculate the population variance, σ^2 , when we know the return R_t for each period, the total number periods (T), and the mean or expected value of the population's distribution (μ), as follows:

$$\sigma^2 = \frac{\sum_{t=1}^T (R_t - \mu)^2}{T}$$

In the world of finance, we are typically analyzing only a sample of returns data, rather than the entire population. To calculate sample variance, s^2 , using a sample of T historical returns and the mean, \bar{R} , of the observations, we use the following formula:

$$s^2 = \frac{\sum_{t=1}^T (R_t - \bar{R})^2}{T - 1}$$

Covariance and Correlation of Returns for Two Securities

Covariance measures the extent to which two variables move together over time. A positive covariance means that the variables (e.g., rates of return on two stocks) tend to move together. Negative covariance means that the two variables tend to move in opposite directions. A covariance of zero means there is no linear relationship between the two variables. To put it another way, if the covariance of returns between two assets is zero, knowing the return for the next period on one of the assets tells you nothing about the return of the other asset for the period.

Here we will focus on the calculation of the covariance between two assets' returns using **historical data**. The calculation of the sample covariance is based on the following formula:

$$\text{Cov}_{1,2} = \frac{\sum_{t=1}^n \{ [R_{t,1} - \bar{R}_1] [R_{t,2} - \bar{R}_2] \}}{n - 1}$$

where:

$R_{t,1}$ = return on Asset 1 in period t

$R_{t,2}$ = return on Asset 2 in period t

\bar{R}_1 = mean return on Asset 1

\bar{R}_2 = mean return on Asset 2

n = number of periods

The magnitude of the covariance depends on the magnitude of the individual stocks' standard deviations and the relationship between their co-movements. Covariance is an absolute measure and is measured in return units squared.

The covariance of the returns of two securities can be standardized by dividing by the product of the standard deviations of the two securities. This standardized measure of co-movement is called **correlation** and is computed as:

$$\rho_{1,2} = \frac{\text{Cov}_{1,2}}{\sigma_1 \sigma_2}$$

The relation can also be written as:

$$\text{Cov}_{1,2} = \rho_{1,2} \sigma_1 \sigma_2$$

The term $\rho_{1,2}$ is called the *correlation coefficient* between the returns of securities 1 and 2. The correlation coefficient has no units. It is a pure measure of the co-movement of the two stocks' returns and is bounded by -1 and +1.

How should you interpret the correlation coefficient?

- A correlation coefficient of +1 means that deviations from the mean or expected return are always proportional in the same direction. That is, they are perfectly positively correlated.
- A correlation coefficient of -1 means that deviations from the mean or expected return are always proportional in opposite directions. That is, they are perfectly negatively correlated.
- A correlation coefficient of zero means that there is no linear relationship between the two stocks' returns. They are uncorrelated. One way to interpret a correlation (or covariance) of zero is that, in any period, knowing the actual value of one variable tells you nothing about the value of the other.

EXAMPLE: Calculating mean return, returns variance, returns covariance, and correlation

Given three years of percentage returns for Assets A and B in the following table, calculate the mean return and sample standard deviation for each asset, the sample covariance, and the correlation of returns.

Year	Asset A	Asset B
1	5%	7%
2	-2%	-4%
3	12%	18%

Answer:

$$\text{mean return for Asset A} = (5\% - 2\% + 12\%) / 3 = 5\%$$

$$\text{mean return for Asset B} = (7\% - 4\% + 18\%) / 3 = 7\%$$

$$\text{sample variance of returns for Asset A} = \frac{(5-5)^2 + (-2-5)^2 + (12-5)^2}{3-1} = 49$$

$$\text{sample standard deviation for Asset A} = \sqrt{49} = 7\%$$

$$\text{sample variance of returns for Asset B} = \frac{(7-7)^2 + (-4-7)^2 + (18-7)^2}{3-1} = 121$$

$$\text{sample standard deviation for Asset B} = \sqrt{121} = 11\%$$

sample covariance of returns for Assets A and B

$$= \frac{(5-5)(7-7) + (-2-5)(-4-7) + (12-5)(18-7)}{3-1} = 77$$

$$\text{correlation of returns for Assets A and B} = \frac{77}{7 \times 11} = 1$$

In this example, the returns on Assets A and B are perfectly positively correlated.

LOS 62.g: Calculate and interpret portfolio standard deviation.



The variance of returns for a portfolio of two risky assets is calculated as follows:

Video covering this content is available online.

$$\text{Var}_{\text{portfolio}} = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \text{Cov}_{12}$$

where w_1 is the proportion of the portfolio invested in Asset 1, and w_2 is the proportion of the portfolio invested in Asset 2. w_2 must equal $(1 - w_1)$.

Previously, we established that the correlation of returns for two assets is calculated as:

$$\rho_{12} = \frac{\text{Cov}_{12}}{\sigma_1 \sigma_2}, \text{ so that we can also write } \text{Cov}_{12} = \rho_{12} \sigma_1 \sigma_2.$$

Substituting this term for Cov_{12} in the formula for the variance of returns for a portfolio of two risky assets, we have the following:

$$\text{Var}_{\text{portfolio}} = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \rho_{12} \sigma_1 \sigma_2$$

Writing the formula in this form allows us to easily see the effect of the correlation of returns between the two assets on portfolio risk.

EXAMPLE: Calculating portfolio standard deviation

A portfolio is 30% invested in stocks that have a standard deviation of returns of 20% and is 70% invested in bonds that have a standard deviation of returns of 12%. The correlation of bond returns with stock returns is 0.60. What is the standard deviation of portfolio returns? What would it be if stock and bond returns were perfectly positively correlated?

Answer:

portfolio standard deviation

$$\begin{aligned} &= \sqrt{(0.3^2)(0.2^2) + (0.7^2)(0.12^2) + 2(0.3)(0.7)(0.6)(0.2)(0.12)} \\ &= 12.9\% \end{aligned}$$

If stock and bond returns were perfectly positively correlated, portfolio standard deviation would simply be the weighted average of the assets' standard deviations: $0.3(20\%) + 0.7(12\%) = 14.4\%$.

MODULE QUIZ 62.3



1. In a 5-year period, the annual returns on an investment are 5%, -3%, -4%, 2%, and 6%. The standard deviation of annual returns on this investment is *closest* to:
 - A. 4.0%.
 - B. 4.5%.
 - C. 20.7%.
2. A measure of how the returns of two risky assets move in relation to each other is:
 - A. the range.
 - B. the covariance.
 - C. the standard deviation.
3. Which of the following statements about correlation is *least accurate*?
 - A. Diversification reduces risk when correlation is less than +1.
 - B. If the correlation coefficient is 0, a zero-variance portfolio can be constructed.
 - C. The lower the correlation coefficient, the greater the potential benefits from diversification.
4. The variance of returns is 0.09 for Stock A and 0.04 for Stock B. The covariance between the returns of A and B is 0.006. The correlation of returns between A and B is:
 - A. 0.10.
 - B. 0.20.
 - C. 0.30.

Use the following data to answer Questions 5 and 6.

A portfolio was created by investing 25% of the funds in Asset A (standard deviation = 15%) and the balance of the funds in Asset B (standard deviation = 10%).

5. If the correlation coefficient is 0.75, what is the portfolio's standard deviation?
 - A. 10.6%.
 - B. 12.4%.
 - C. 15.0%.
6. If the correlation coefficient is -0.75, what is the portfolio's standard deviation?
 - A. 2.8%.
 - B. 4.2%.
 - C. 5.3%.

MODULE 62.4: THE EFFICIENT FRONTIER



LOS 62.h: Describe the effect on a portfolio's risk of investing in assets that are less than perfectly correlated.

Video covering this content is available online.

If two risky asset returns are perfectly positively correlated, $\rho_{12} = +1$, then the square root of portfolio variance (the portfolio standard deviation of returns) is equal to:

$$\sigma_{\text{portfolio}} = \sqrt{\text{Var}_{\text{portfolio}}} = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2(1)} = w_1\sigma_1 + w_2\sigma_2$$



PROFESSOR'S NOTE

This might be easier to see by examining the algebra in reverse. If $w_1\sigma_1 + w_2\sigma_2$ equals the square root of the term under the radical in this special case, then $(w_1\sigma_1 + w_2\sigma_2)^2$ should equal the term under the radical. If we expand $(w_1\sigma_1 + w_2\sigma_2)^2$, we get:

$$\begin{aligned}(w_1\sigma_1 + w_2\sigma_2)^2 &= (w_1\sigma_1)^2 + (w_1\sigma_1)(w_2\sigma_2) + (w_2\sigma_2)(w_1\sigma_1) + (w_2\sigma_2)^2 \\ &= (w_1\sigma_1)^2 + (w_2\sigma_2)^2 + 2(w_1\sigma_1)(w_2\sigma_2) \\ &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1\sigma_1w_2\sigma_2\end{aligned}$$

In this unique case, with $\rho_{12} = 1$, the portfolio standard deviation is simply a weighted average of the standard deviations of the individual asset returns. A portfolio 25% invested in Asset 1 and 75% invested in Asset 2 will have a standard deviation of returns equal to 25% of the standard deviation (σ_1) of Asset 1's return, plus 75% of the standard deviation (σ_2) of Asset 2's return.

Focusing on returns correlation, we can see that the greatest portfolio risk results when the correlation between asset returns is +1. For any value of correlation less than +1, portfolio variance is reduced. Note that for a correlation of zero, the entire third term in the portfolio variance equation is zero. For negative values of correlation ρ_{12} , the third term becomes negative and further reduces portfolio variance and standard deviation.

We will illustrate this property with an example.

EXAMPLE: Portfolio risk as correlation varies

Consider two risky assets that have returns variances of 0.0625 and 0.0324, respectively. The assets' standard deviations of returns are then 25% and 18%, respectively. Calculate the variances and standard deviations of portfolio returns for an equal-weighted portfolio of the two assets when their correlation of returns is 1, 0.5, 0, and -0.5.

The calculations are as follows:

$$\text{variance}_{\text{portfolio}} = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \rho_{12} \sigma_1 \sigma_2$$

$$\sigma_{\text{portfolio}} = \sqrt{\text{variance}_{\text{portfolio}}}$$

$$\sigma_{\text{portfolio}} = \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \rho_{12} \sigma_1 \sigma_2}$$

$\rho = \text{correlation} = +1:$

$$\sigma = \text{portfolio standard deviation} = 0.5(25\%) + 0.5(18\%) = 21.5\%$$

$$\sigma^2 = \text{portfolio variance} = 0.215^2 = 0.046225$$

$\rho = \text{correlation} = 0.5:$

$$\sigma^2 = (0.5^2)0.0625 + (0.5^2)0.0324 + 2(0.5)(0.5)(0.5)(0.25)(0.18) = 0.034975$$

$$\sigma = 18.70\%$$

$\rho = \text{correlation} = 0:$

$$\sigma^2 = (0.5^2)0.0625 + (0.5^2)0.0324 = 0.023725$$

$$\sigma = 15.40\%$$

$\rho = \text{correlation} = -0.5:$

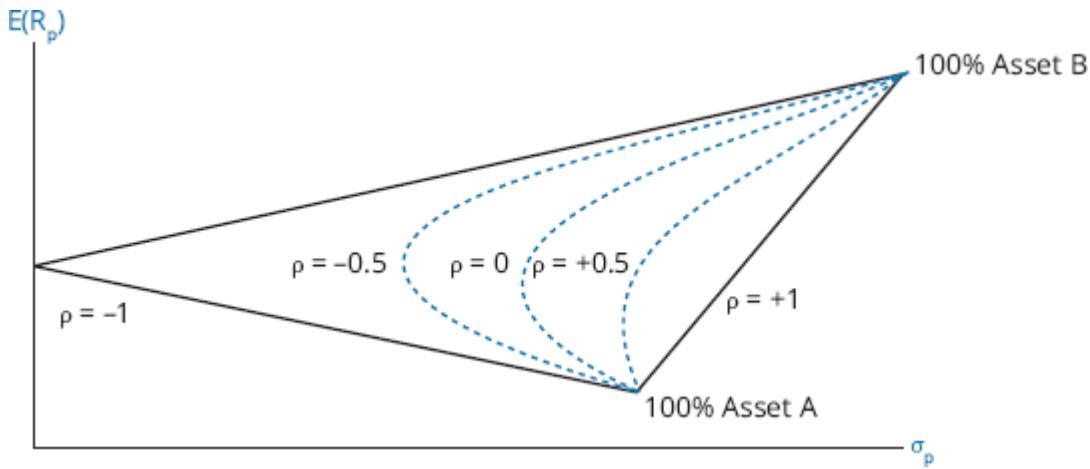
$$\sigma^2 = (0.5^2)0.0625 + (0.5^2)0.0324 + 2(0.5)(0.5)(-0.5)(0.25)(0.18) = 0.012475$$

$$\sigma = 11.17\%$$

Note that portfolio risk decreases as the correlation between the assets' returns decreases. This is an important result of the analysis of portfolio risk: The lower the correlation of asset returns, the greater the risk reduction (diversification) benefit of combining assets in a portfolio. If asset returns were perfectly negatively correlated, portfolio risk could be eliminated altogether for a specific set of asset weights.

We show these relations graphically in Figure 62.6 by plotting the portfolio risk and return for all portfolios of two risky assets, for specific values of the assets' returns correlation.

Figure 62.6: Risk and Return for Different Values of ρ



From these analyses, the risk reduction benefits of investing in assets with low return correlations should be clear. The desire to reduce risk is what drives investors to invest in not just domestic stocks, but also bonds, foreign stocks, real estate, and other asset classes.

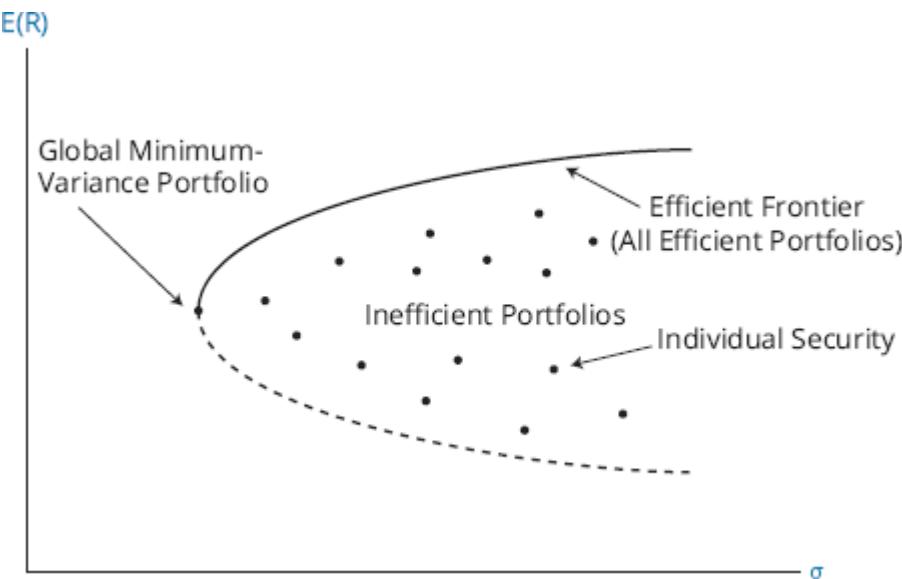
LOS 62.i: Describe and interpret the minimum-variance and efficient frontiers of risky assets and the global minimum-variance portfolio.

For each level of expected portfolio return, we can vary the portfolio weights on the individual assets to determine the portfolio that has the least risk. These portfolios that have the lowest standard deviation of all portfolios with a given expected return are known as **minimum-variance portfolios**. Together they make up the **minimum-variance frontier**.

Assuming that investors are risk averse, investors prefer the portfolio that has the greatest expected return when choosing among portfolios that have the same standard deviation of returns. Those portfolios that have the greatest expected return for each level of risk (standard deviation) make up the **efficient frontier**. The efficient frontier coincides with the top portion of the minimum-variance frontier. A risk-averse investor would only choose portfolios that are on the efficient frontier because all available portfolios that are not on the efficient frontier have lower expected returns than an efficient portfolio with the same risk. The portfolio on the efficient frontier that has the least risk is the **global minimum-variance portfolio**.

These concepts are illustrated in Figure 62.7.

Figure 62.7: Minimum-Variance and Efficient Frontiers



MODULE QUIZ 62.4



- Which of the following statements about covariance and correlation is *least accurate*?
 - A zero covariance implies there is no linear relationship between the returns on two assets.
 - If two assets have perfect negative correlation, the variance of returns for a portfolio that consists of these two assets will equal zero.
 - The covariance of a 2-stock portfolio is equal to the correlation coefficient times the standard deviation of one stock's returns times the standard deviation of the other stock's returns.
- Which of the following available portfolios *most likely* falls below the efficient frontier?

<u>Portfolio</u>	<u>Expected return</u>	<u>Expected standard deviation</u>
A. A	7%	14%
B. B	9%	26%
C. C	12%	22%

KEY CONCEPTS

LOS 62.a

Holding period return is used to measure an investment's return over a specific period. Arithmetic mean return is the simple average of a series of periodic returns. Geometric mean return is a compound annual rate.

Gross return is total return after deducting commissions on trades and other costs necessary to generate the returns, but before deducting fees for the management and administration of the investment account. Net return is the return after management and administration fees have been deducted.

Pretax nominal return is the numerical percentage return of an investment, without considering the effects of taxes and inflation. After-tax nominal return is the numerical return after the tax liability is deducted, without adjusting for inflation. Real return is the increase in an investor's purchasing power, roughly equal to nominal return minus inflation. Leveraged return is the gain or loss on an investment as a percentage of an investor's cash investment.

LOS 62.b

The money-weighted rate of return is the IRR calculated using periodic cash flows into and out of an account and is the discount rate that makes the PV of cash inflows equal to the PV of cash outflows.

The time-weighted rate of return measures compound growth. It is the rate at which \$1 compounds over a specified performance horizon.

If funds are added to a portfolio just before a period of poor performance, the money-weighted return will be lower than the time-weighted return. If funds are added just prior to a period of high returns, the money-weighted return will be higher than the time-weighted return.

The time-weighted return is the preferred measure of a manager's ability to select investments. If the manager controls the money flows into and out of an account, the money-weighted return is the more appropriate performance measure.

LOS 62.c

As predicted by theory, asset classes with the greatest average returns have also had the highest risk.

Some of the major asset classes that investors consider when building a diversified portfolio include small-capitalization stocks, large-capitalization stocks, long-term corporate bonds, long-term Treasury bonds, and Treasury bills.

In addition to risk and return, when analyzing investments, investors also take into consideration an investment's liquidity, as well as non-normal characteristics such as skewness and kurtosis.

LOS 62.d

A risk-averse investor is one that dislikes risk. Given two investments that have equal expected returns, a risk-averse investor will choose the one with less risk. However, a risk-averse investor will hold risky assets if he feels that the extra return he expects to earn is adequate compensation for the additional risk. Assets in the financial markets are priced according to the preferences of risk-averse investors.

A risk-seeking (risk-loving) investor prefers more risk to less and, given investments with equal expected returns, will choose the more risky investment.

A risk-neutral investor would be indifferent to risk and would be indifferent between two investments with the same expected return regardless of the investments' standard deviation of returns.

LOS 62.e

An indifference curve plots combinations of risk and expected return that provide the same expected utility. Indifference curves for risk and return slope upward because risk-averse investors will only take on more risk if they are compensated with greater expected returns. A more risk-averse investor will have steeper indifference curves.

Flatter indifference curves (less risk aversion) result in an optimal portfolio with higher risk and higher expected return. An investor who is less risk averse will optimally choose a portfolio with more invested in the risky asset portfolio and less invested in the risk-free asset, compared to a more risk-averse investor.

LOS 62.f

We can calculate the population variance, σ^2 , when we know the return R_t for period t , the total number T of periods, and the mean μ of the population's distribution:

$$\text{population variance} = \sigma^2 = \frac{\sum_{t=1}^T (R_t - \mu)^2}{T}$$

In finance, we typically analyze only a sample of returns, so the sample variance applies instead:

$$\text{sample variance} = S^2 = \frac{\sum_{t=1}^T (R_t - \bar{R})^2}{T - 1}$$

Covariance measures the extent to which two variables move together over time. Positive covariance means the variables (e.g., rates of return on two stocks) tend to move together. Negative covariance means that the two variables tend to move in opposite directions. Covariance of zero means there is no linear relationship between the two variables.

Correlation is a standardized measure of co-movement that is bounded by -1 and $+1$:

$$\rho_{1,2} = \frac{\text{Cov}_{1,2}}{\sigma_1 \sigma_2}$$

LOS 62.g

The standard deviation of returns for a portfolio of two risky assets is calculated as follows:

$$\sigma_{\text{portfolio}} = \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \rho_{1,2} \sigma_1 \sigma_2}$$

LOS 62.h

The greatest portfolio risk will result when the asset returns are perfectly positively correlated. As the correlation decreases from $+1$ to -1 , portfolio risk decreases. The lower the correlation of asset returns, the greater the risk reduction (diversification) benefit of combining assets in a portfolio.

LOS 62.i

For each level of expected portfolio return, the portfolio that has the least risk is known as a minimum-variance portfolio. Taken together, these portfolios form a line called the minimum-variance frontier.

On a risk versus return graph, the one risky portfolio that is farthest to the left (has the least risk) is known as the global minimum-variance portfolio.

Those portfolios that have the greatest expected return for each level of risk make up the efficient frontier. The efficient frontier coincides with the top portion of the minimum variance frontier. Risk-averse investors would only choose a portfolio that lies on the efficient frontier.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 62.1

- C** Using the cash flow functions on your financial calculator, enter $CF_0 = -40$; $CF_1 = -50 + 1 = -49$; $CF_2 = 60 \times 2 + 2 = 122$; CPT IRR = 23.82%. (LOS 62.a)
- A** Small-cap stocks have had the highest annual return and standard deviation of return over time. Large-cap stocks and bonds have historically had lower risk and return than small-cap stocks. (LOS 62.c)

Module Quiz 62.2

- B** Risk-averse investors are generally willing to invest in risky investments, if the expected return of the investment is sufficient to reward the investor for taking on this risk. Participants in securities markets are generally assumed to be risk-averse investors. (LOS 62.d)
- B** An investor's optimal portfolio will lie somewhere on the capital allocation line, which begins at the risk-free asset and runs through the optimal risky portfolio. (LOS 62.e)

Module Quiz 62.3

- B** mean annual return = $(5\% - 3\% - 4\% + 2\% + 6\%) / 5 = 1.2\%$

Squared deviations from the mean:

$$5\% - 1.2\% = 3.8\% \quad 3.8^2 = 14.44$$

$$-3\% - 1.2\% = -4.2\% \quad -4.2^2 = 17.64$$

$$-4\% - 1.2\% = -5.2\% \quad -5.2^2 = 27.04$$

$$2\% - 1.2\% = 0.8\% \quad 0.8^2 = 0.64$$

$$6\% - 1.2\% = 4.8\% \quad 4.8^2 = 23.04$$

$$\text{sum of squared deviations} = 14.44 + 17.64 + 27.04 + 0.64 + 23.04 = 82.8$$

$$\text{sample variance} = 82.8 / (5 - 1) = 20.7$$

$$\text{sample standard deviation} = 20.7^{1/2} = 4.55\%$$

(LOS 62.f)

- B** The covariance is defined as the co-movement of the returns of two assets or how well the returns of two risky assets move together. Range and standard deviation are measures of dispersion and measure risk, not how assets move together. (LOS 62.f)
- B** A zero-variance portfolio can only be constructed if the correlation coefficient between assets is -1. Diversification benefits can be had when correlation is less than +1, and the lower the correlation, the greater the expected benefit. (LOS 62.f)

A $\sqrt{A} = \sqrt{0.09} = 0.30$

A $\sqrt{B} = \sqrt{0.04} = 0.20$

- correlation** = $0.006 / [(0.30)(0.20)] = 0.10$

(LOS 62.f)

A $\sqrt{(0.25)^2(0.15)^2 + (0.75)^2(0.10)^2 + 2(0.25)(0.75)(0.15)(0.10)(0.75)} =$

- $\sqrt{0.001406 + 0.005625 + 0.004219} = \sqrt{0.01125} = 0.106 = 10.6\%$

(LOS 62.g)

C $\sqrt{(0.25)^2(0.15)^2 + (0.75)^2(0.10)^2 + 2(0.25)(0.75)(0.15)(0.10)(-0.75)} =$
6. $\sqrt{0.001406 + 0.005625 - 0.004219} = \sqrt{0.002812} = 0.053 = 5.3\%$
(LOS 62.g)

Module Quiz 62.4

1. **B** If the correlation of returns between the two assets is -1 , the set of possible portfolio risk/return combinations becomes two straight lines (see Figure 62.2). A portfolio of these two assets will have a positive returns variance unless the portfolio weights are those that minimize the portfolio variance. Covariance is equal to the correlation coefficient multiplied by the product of the standard deviations of the returns of the two stocks in a 2-stock portfolio. If covariance is zero, then correlation is also zero, which implies that there is no linear relationship between the two stocks' returns. (LOS 62.h)
2. **B** Portfolio B must be the portfolio that falls below the Markowitz efficient frontier because there is a portfolio (Portfolio C) that offers a higher return and lower risk. (LOS 62.i)

¹ 2018 SBBI Yearbook.

READING 63

PORTFOLIO RISK AND RETURN: PART II

EXAM FOCUS

The concepts developed here are very important to finance theory and are also used extensively in practice. You must know this material completely—not only the formulas and definitions, but the ideas that underlie their use. A model assumption that diversification is costless leads to the conclusion that only systematic risk (which cannot be reduced by further diversification) is priced in equilibrium, so that bearing nonsystematic risk does not increase expected returns.

MODULE 63.1: SYSTEMATIC RISK AND BETA



LOS 63.a: Describe the implications of combining a risk-free asset with a portfolio of risky assets.

Video covering this content is available online.

In the previous reading, we covered the mathematics of calculating the risk and return of a portfolio with a percentage weight of W_A invested in a risky portfolio (P) and a weight of $W_B = 1 - W_A$ invested in a risk-free asset.

$$E(R_P) = W_A E(R_A) + W_B E(R_B)$$

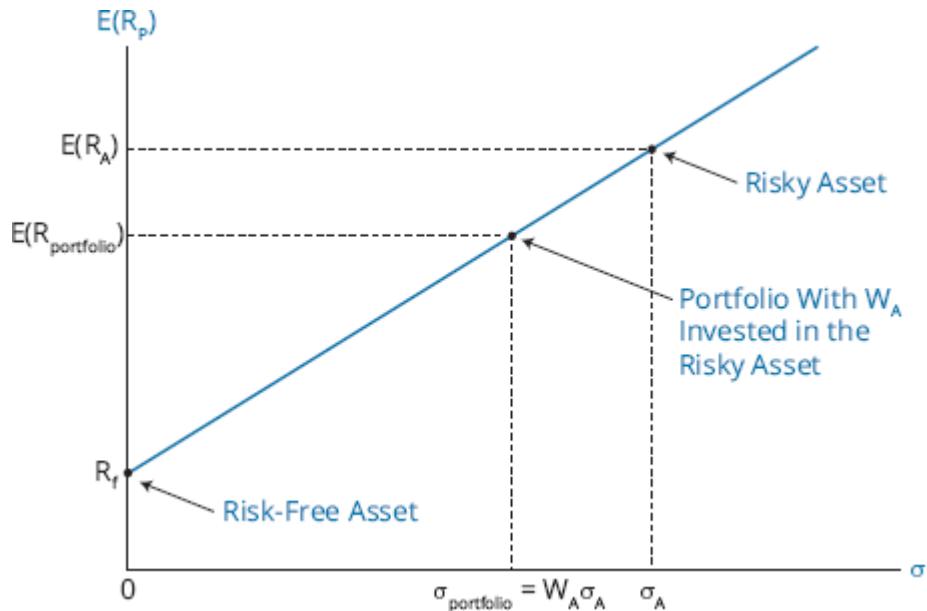
$$\sigma_P = \sqrt{W_A^2 \sigma_A^2 + W_B^2 \sigma_B^2 + 2 W_A W_B \rho_{AB} \sigma_A \sigma_B}$$

Because a risk-free asset has zero standard deviation and zero correlation of returns with a risky portfolio, allowing Asset B to be the risk-free asset and Asset A to be the risky asset portfolio results in the following reduced equation:

$$\sigma_P = \sqrt{W_A^2 \sigma_A^2} = W_A \sigma_A$$

Our result is that the risk (standard deviation of returns) and expected return of portfolios with varying weights in the risk-free asset and a risky portfolio can be plotted as a line that begins at the risk-free rate of return and extends through the risky portfolio. This result is illustrated in Figure 63.1.

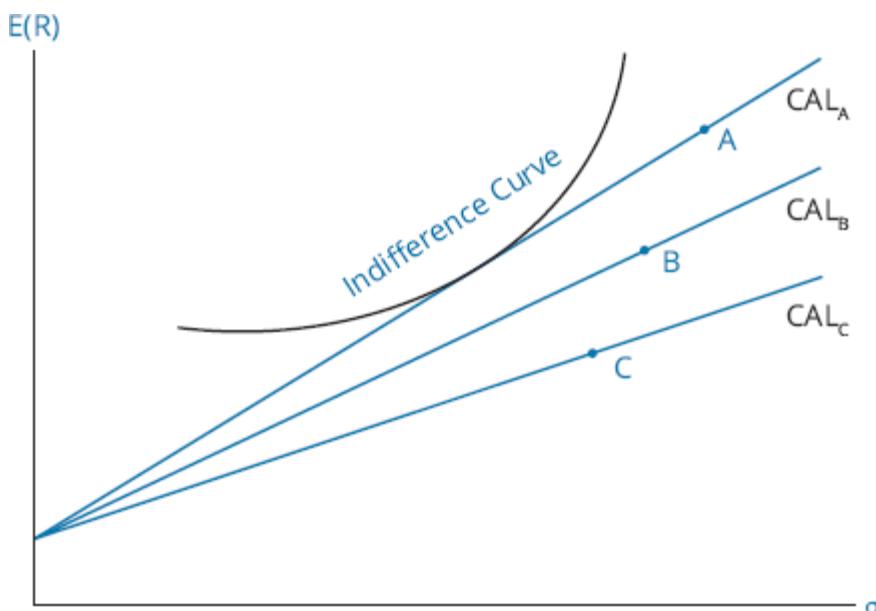
Figure 63.1: Combining a Risk-Free Asset With a Risky Asset



LOS 63.b: Explain the capital allocation line (CAL) and the capital market line (CML).

The line of possible portfolio risk and return combinations given the risk-free rate and the risk and return of a portfolio of risky assets is referred to as the **capital allocation line (CAL)**. For an individual investor, the best CAL is the one that offers the most-preferred set of possible portfolios in terms of their risk and return. Figure 63.2 illustrates three possible investor CALs for three different risky portfolios A, B, and C. The optimal risky portfolio for this investor is Portfolio A because it results in the most preferred set of possible portfolios constructed by combining the risk-free asset with the risky portfolio. Of all the portfolios available to the investor, a combination of the risk-free asset with risky Portfolio A offers the investor the greatest expected utility.

Figure 63.2: Risky Portfolios and Their Associated Capital Allocation Lines

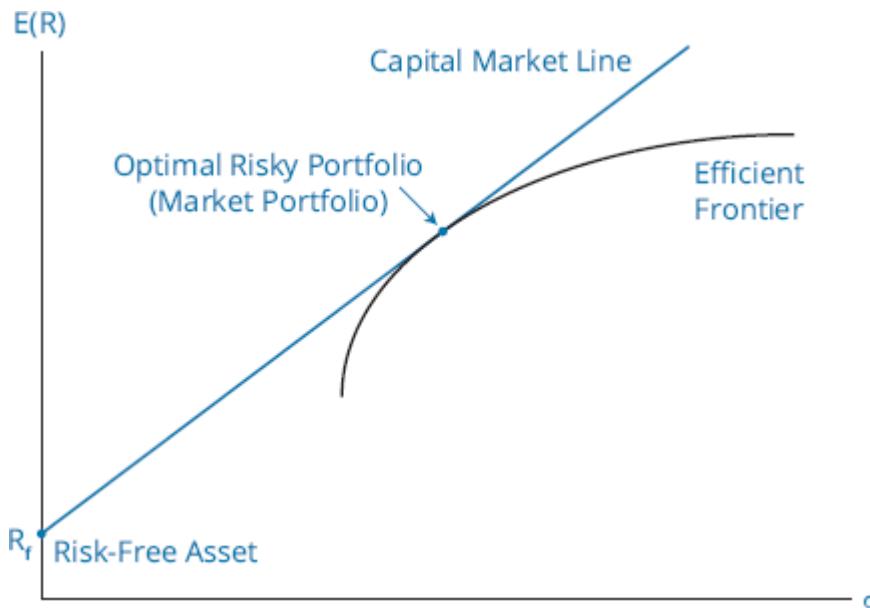


If each investor has different expectations about the expected returns of, standard deviations of, or correlations between risky asset returns, each investor will have a different optimal risky asset portfolio and a different CAL.

A simplifying assumption underlying modern portfolio theory (and the capital asset pricing model, which is introduced later in this reading) is that investors have homogeneous expectations (i.e., they all have the same estimates of risk, return, and correlations with other risky assets for all risky assets). Under this assumption, all investors face the same efficient frontier of risky portfolios and will all have the same optimal risky portfolio and CAL.

Figure 63.3 illustrates the determination of the optimal risky portfolio and optimal CAL for all investors under the assumption of homogeneous expectations. Note that, under this assumption, the optimal CAL for any investor is the one that is just tangent to the efficient frontier. Depending on their preferences for risk and return (their indifference curves), investors may choose different portfolio weights for the risk-free asset and the risky (tangency) portfolio. Every investor, however, will use the same risky portfolio. When this is the case, that portfolio must be the **market portfolio** of all risky assets because all investors that hold any risky assets hold the same portfolio of risky assets.

Figure 63.3: Determining the Optimal Risky Portfolio and Optimal CAL Assuming Homogeneous Expectations



Under the assumption of homogeneous expectations, this optimal CAL for all investors is termed the **capital market line (CML)**. Along this line, expected portfolio return, $E(R_p)$, is a linear function of portfolio risk, σ_p . The equation of this line is as follows:

$$E(R_p) = R_f + \left(\frac{E(R_M) - R_f}{\sigma_M} \right) \sigma_p$$

The y -intercept of this line is R_f and the slope (rise over run) of this line is as follows:

$$\left(\frac{E(R_M) - R_f}{\sigma_M} \right)$$

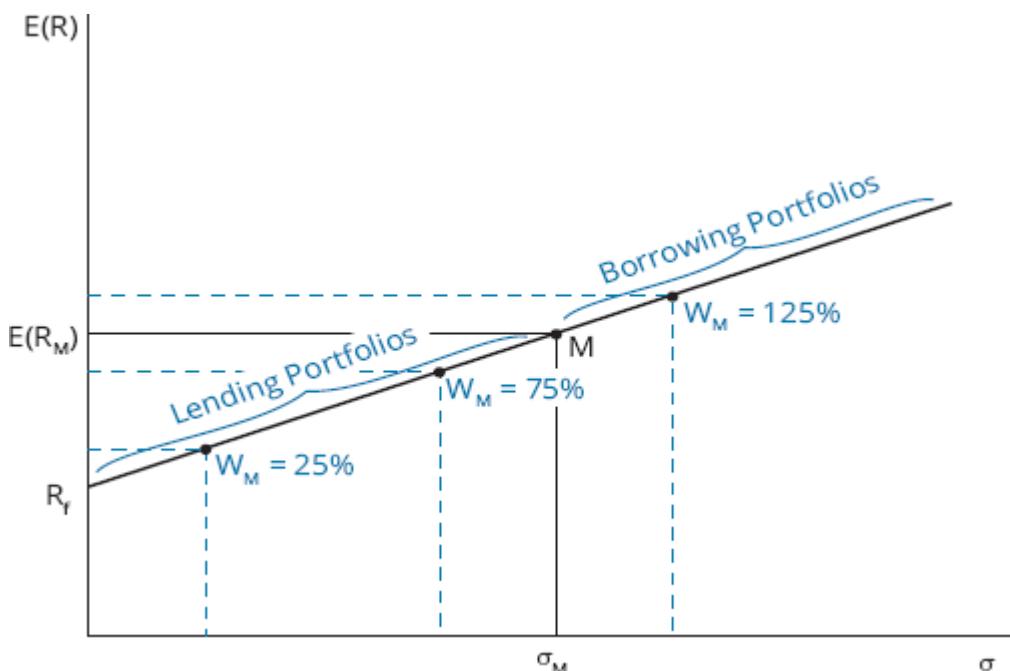
The intuition of this relation is straightforward. An investor who chooses to take on no risk ($\sigma_p = 0$) will earn the risk-free rate, R_f . The difference between the expected return on the market and the risk-free rate is termed the **market risk premium**. If we rewrite the CML equation as

$$E(R_p) = R_f + (E(R_M) - R_f) \left(\frac{\sigma_p}{\sigma_M} \right)$$

we can see that an investor can expect to get one unit of market risk premium in additional return (above the risk-free rate) for every unit of market risk, σ_M , that the investor is willing to accept.

If we assume that investors can both lend (invest in the risk-free asset) at the risk-free rate and borrow (as with a margin account) at the risk-free rate, they can select portfolios to the right of the market portfolio, as illustrated in Figure 63.4.

Figure 63.4: Borrowing and Lending Portfolios



Investors who believe market prices are informationally efficient often follow a **passive investment strategy** (i.e., invest in an index of risky assets that serves as a proxy for the market portfolio and allocate a portion of their investable assets to a risk-free asset, such as short-term government securities). In practice, many investors and portfolio managers believe their estimates of security values are correct and market prices are incorrect. Such investors will not use the weights of the market portfolio but will invest more than the market weights in securities that they believe are undervalued and less than the market weights in securities which they believe are overvalued. This is referred to as **active portfolio management** to differentiate it from a passive investment strategy that utilizes a market index for the optimal risky asset portfolio.

LOS 63.c: Explain systematic and nonsystematic risk, including why an investor should not expect to receive additional return for bearing nonsystematic risk.

When an investor diversifies across assets that are not perfectly correlated, the portfolio's risk is less than the weighted average of the risks of the individual securities in the portfolio. The risk that is eliminated by diversification is called **unsystematic risk** (also called *unique, diversifiable, or firm-specific risk*). Because the market portfolio contains *all* risky assets, it must be a well-diversified portfolio. All the risk that can be diversified away has been. The risk that remains cannot be diversified away and is called the **systematic risk** (also called *nondiversifiable risk or market risk*).

The concept of systematic risk applies to individual securities as well as to portfolios. Some securities' returns are highly correlated with overall market returns. Examples of firms that are highly correlated with market returns are luxury goods manufacturers such as Ferrari automobiles and Harley Davidson motorcycles. These firms have high systematic risk (i.e., they are very responsive to market, or systematic, changes). Other firms, such as utility companies, respond very little to changes in the systematic risk factors. These firms have very little systematic risk. Hence, total risk (as measured by standard deviation) can be broken down into its component parts: unsystematic risk and systematic risk. Mathematically:

$$\text{total risk} = \text{systematic risk} + \text{unsystematic risk}$$

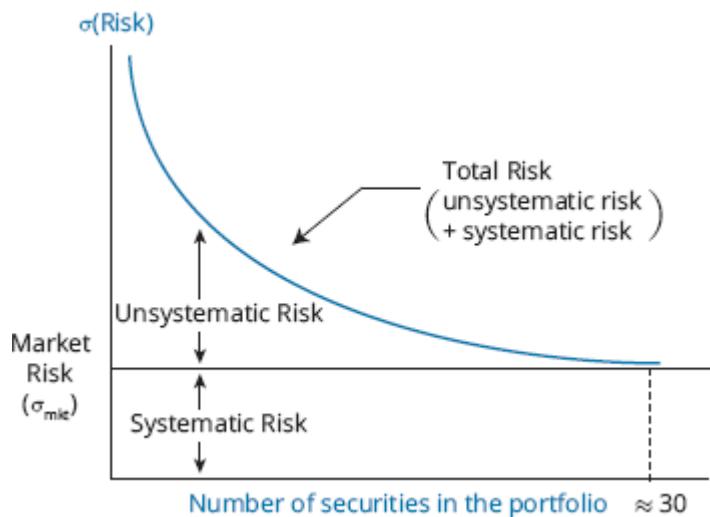


PROFESSOR'S NOTE

Know this concept!

Do you actually have to buy all the securities in the market to diversify away unsystematic risk? No. Academic studies have shown that as you increase the number of stocks in a portfolio, the portfolio's risk falls toward the level of market risk. One study showed that it only took about 12 to 18 stocks in a portfolio to achieve 90% of the maximum diversification possible. Another study indicated it took 30 securities. Whatever the number, it is significantly less than *all* the securities. Figure 63.5 provides a general representation of this concept. Note, in the figure, that once you get to 30 or so securities in a portfolio, the standard deviation remains constant. The remaining risk is systematic, or nondiversifiable, risk. We will develop this concept later when we discuss beta, a measure of systematic risk.

Figure 63.5: Risk vs. Number of Portfolio Assets



Systematic Risk Is Relevant in Portfolios

One important conclusion of capital market theory is that equilibrium security returns depend on a stock's or a portfolio's systematic risk, not its total risk as measured by standard deviation. One of the assumptions of the model is that diversification is free. The reasoning is that investors will not be compensated for bearing risk that can be eliminated at no cost. If you think about the costs of a no-load index fund compared to buying individual stocks, diversification is actually very low cost if not actually free.

The implications of this conclusion are very important to asset pricing (expected returns). The riskiest stock, with risk measured as standard deviation of returns, does not necessarily have the greatest expected return. Consider a biotech stock with one new drug product that is in clinical trials to determine its effectiveness. If it turns out that the drug is effective and safe, stock returns will be quite high. If, on the other hand, the subjects in the clinical trials are killed or otherwise harmed by the drug, the stock will fall to approximately zero and returns will be quite poor. This describes a stock with high standard deviation of returns (i.e., high total risk).

The high risk of our biotech stock, however, is primarily from firm-specific factors, so its unsystematic risk is high. Because market factors such as economic growth rates have little to do with the eventual outcome for this stock, systematic risk is a small proportion of the total risk of the stock. Capital market theory says that the equilibrium return on this stock may be less than that of a stock with much less firm-specific risk but more sensitivity to the factors that drive the return of the overall market. An established manufacturer of machine tools may not be a very risky investment in terms of total risk, but may have a greater sensitivity to market (systematic) risk factors (e.g., GDP growth rates) than our biotech stock. Given this scenario, the stock with more total risk (the biotech stock) has less systematic risk and will therefore have a lower equilibrium rate of return according to capital market theory.

Note that holding many biotech firms in a portfolio will diversify away the firm-specific risk. Some will have blockbuster products and some will fail, but you can imagine that when 50 or 100 such stocks are combined into a portfolio, the uncertainty about the portfolio return is much less than the uncertainty about the return of a single biotech firm stock.

To sum up, unsystematic risk is not compensated in equilibrium because it can be eliminated for free through diversification. Systematic risk is measured by the contribution of a security to the risk of a well-diversified portfolio, and the expected equilibrium return (required return) on an individual security will depend only on its systematic risk.

LOS 63.d: Explain return generating models (including the market model) and their uses.

Return generating models are used to estimate the expected returns on risky securities based on specific factors. For each security, we must estimate the sensitivity of its returns to each specific factor. Factors that explain security returns can be classified as macroeconomic, fundamental, and statistical factors. **Multifactor models** most commonly use macroeconomic factors such as GDP growth, inflation, or consumer confidence, along with fundamental factors such as earnings, earnings growth, firm size, and research expenditures. Statistical factors often have no basis in finance theory and are suspect in that they may represent only relations for a

specific time period which have been identified by data mining (repeated tests on a single dataset).

The general form of a multifactor model with k factors is as follows:

$$E(R_i) - R_f = \beta_{i1} \times E(\text{Factor 1}) + \beta_{i2} \times E(\text{Factor 2}) + \dots + \beta_{ik} \times E(\text{Factor } k)$$

©2022 Kaplan, Inc.

This model states that the expected excess return (above the risk-free rate) for Asset i is the sum of each **factor sensitivity** or **factor loading** (the β s) for Asset i multiplied by the expected value of that factor for the period. The first factor is often the expected excess return on the market, $E(R_m - R_f)$.

One multifactor model that is often used is that of Fama and French. They estimated the sensitivity of security returns to three factors: firm size, firm book value to market value ratio, and the return on the market portfolio minus the risk-free rate (excess return on the market portfolio). Carhart suggests a fourth factor that measures price momentum using prior period returns. Together, these four factors do a relatively good job of explaining returns differences for U.S. equity securities over the period for which the model has been estimated.

The simplest factor model is a single-factor model. A single-factor model with the return on the market, R_m , as its only risk factor can be written (in excess returns form) as:

$$E(R_i) - R_f = \beta_i \times [E(R_m) - R_f]$$

Here, the expected excess return (return above the risk-free rate) is the product of the factor weight or factor sensitivity, Beta i , and the risk factor, which in this model is the excess return on the market portfolio or market index, so that this is also sometimes called a **single-index model**.

A simplified form of a single-index model is the **market model**, which is used to estimate a security's (or portfolio's) beta and to estimate a security's abnormal return (return above its expected return) based on the actual market return.

The form of the market model is as follows:

$$R_i = \alpha_i + \beta_i R_m + e_i$$

where:

R_i = return on Asset i

R_m = market return

β_i = slope coefficient

α_i = intercept

e_i = abnormal return on Asset i

The intercept α_i and slope coefficient β_i are estimated from historical return data. We can require that α_i is the risk-free rate times $(1 - \beta_i)$ to be consistent with the general form of a single-index model in excess returns form.

The expected return on Asset i is $\alpha_i + \beta_i E(R_m)$. A deviation from the expected return in a given period is the abnormal return on Asset i , e_i , or $R_i - (\alpha_i + \beta_i R_m)$.

In the market model, the factor sensitivity or beta for Asset i is a measure of how sensitive the return on Asset i is to the return on the overall market portfolio (market index).

LOS 63.e: Calculate and interpret beta.

The sensitivity of an asset's return to the return on the market index in the context of the market model is referred to as its **beta**. Beta is a standardized measure of the covariance of the asset's return with the market return. Beta can be calculated as follows:

$$\beta_i = \frac{\text{covariance of Asset } i\text{'s return with the market return}}{\text{variance of the market return}} = \frac{\text{Cov}_{im}}{\sigma_m^2}$$

We can use the definition of the correlation between the returns on Asset i with the returns on the market index:

$$\rho_{im} = \frac{\text{Cov}_{im}}{\sigma_i \sigma_m}$$

to get $\text{Cov}_{im} = \rho_{im} \sigma_i \sigma_m$

Substituting for Cov_{im} in the equation for β_i , we can also calculate beta as:

$$\beta_i = \frac{\rho_{im} \sigma_i \sigma_m}{\sigma_m^2} = \rho_{im} \left(\frac{\sigma_i}{\sigma_m} \right)$$

EXAMPLE: Calculating an asset's beta

The standard deviation of the return on the market index is estimated as 20%.

1. If Asset A's standard deviation is 30% and its correlation of returns with the market index is 0.8, what is Asset A's beta?

Using the formula $\beta_i = \rho_{im} \left(\frac{\sigma_i}{\sigma_m} \right)$, we have: $\beta_i = 0.80 \left(\frac{0.30}{0.20} \right) = 1.2$.

2. If the covariance of Asset A's returns with the returns on the market index is 0.048, what is the beta of Asset A?

Using the formula $\beta_i = \frac{\text{Cov}_{im}}{\sigma_m^2}$, we have $\beta_i = \frac{0.048}{0.2^2} = 1.2$.



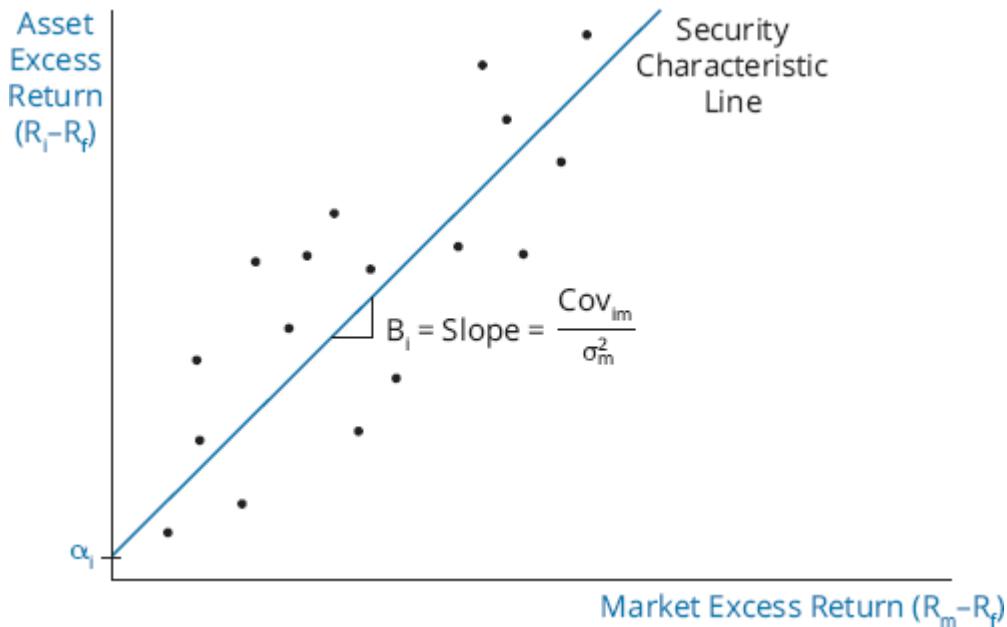
PROFESSOR'S NOTE

Candidates should be prepared to calculate beta in either of the two ways in the example.

In practice, we estimate asset betas by regressing returns on the asset on those of the market index. While regression is a Level II concept, for our purposes, you can think of it as a mathematical estimation procedure that fits a line to a data plot. In Figure 63.5, we represent the excess returns on Asset i as the dependent variable and the excess returns on the market index as the independent variable. The *least squares regression line* is the line that minimizes the

sum of the squared distances of the points plotted from the line (this is what is meant by the line of *best fit*). The slope of this line is our estimate of beta. In Figure 63.6, the line is steeper than 45 degrees, the slope is greater than one, and the asset's estimated beta is greater than one. Our interpretation is that the returns on Asset i are more variable in response to systematic risk factors than is the overall market, which has a beta of one.

Figure 63.6: Regression of Asset Excess Returns Against Market Asset Returns



This regression line is referred to as the asset's **security characteristic line**. Mathematically, the slope of the security characteristic line is $\frac{\text{Cov}_{im}}{\sigma_m^2}$, which is the same formula we used earlier to calculate beta.



MODULE QUIZ 63.1

- An investor put 60% of his portfolio into a risky asset offering a 10% return with a standard deviation of returns of 8% and put the balance of his portfolio in a risk-free asset offering 5%. What is the expected return and standard deviation of his portfolio?

<u>Expected return</u> A. 6.0% B. 8.0% C. 10.0%	<u>Standard deviation</u> 6.8% 4.8% 6.6%
--	---
- What is the risk measure associated with the capital market line (CML)?
 - Beta risk.
 - Unsystematic risk.
 - Total risk.
- A portfolio to the right of the market portfolio on the CML is:
 - a lending portfolio.
 - a borrowing portfolio.
 - an inefficient portfolio.
- As the number of stocks in a portfolio increases, the portfolio's systematic risk:
 - can increase or decrease.

- B. decreases at a decreasing rate.
 - C. decreases at an increasing rate.
5. Total risk equals:
- A. unique plus diversifiable risk.
 - B. market plus nondiversifiable risk.
 - C. systematic plus unsystematic risk.
6. A return generating model is *least likely* to be based on a security's exposure to:
- A. statistical factors.
 - B. macroeconomic factors.
 - C. fundamental factors.
7. The covariance of the market's returns with a stock's returns is 0.005 and the standard deviation of the market's returns is 0.05. What is the stock's beta?
- A. 1.0.
 - B. 1.5.
 - C. 2.0.

MODULE 63.2: THE CAPM AND THE SML



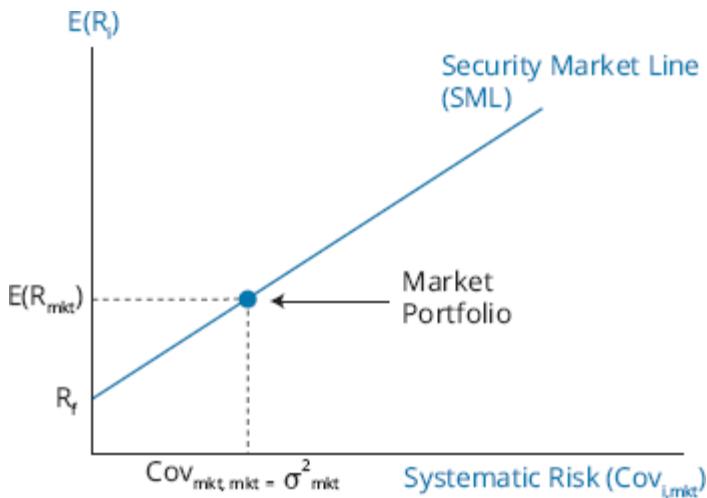
LOS 63.f: Explain the capital asset pricing model (CAPM), including its assumptions, and the security market line (SML).

Video covering this content is available online.

LOS 63.g: Calculate and interpret the expected return of an asset using the CAPM.

Given that the only relevant (priced) risk for an individual Asset i is measured by the covariance between the asset's returns and the returns on the market, $\text{Cov}_{i,\text{mkt}}$, we can plot the relationship between risk and return for individual assets using $\text{Cov}_{i,\text{mkt}}$ as our measure of systematic risk. The resulting line, plotted in Figure 63.7, is one version of what is referred to as the **security market line (SML)**.

Figure 63.7: Security Market Line



The equation of the SML is:

$$E(R_i) = R_f + \frac{E(R_{mkt}) - R_f}{\sigma_{mkt}^2} (\text{Cov}_{i,mkt})$$

which can be rearranged and stated as:

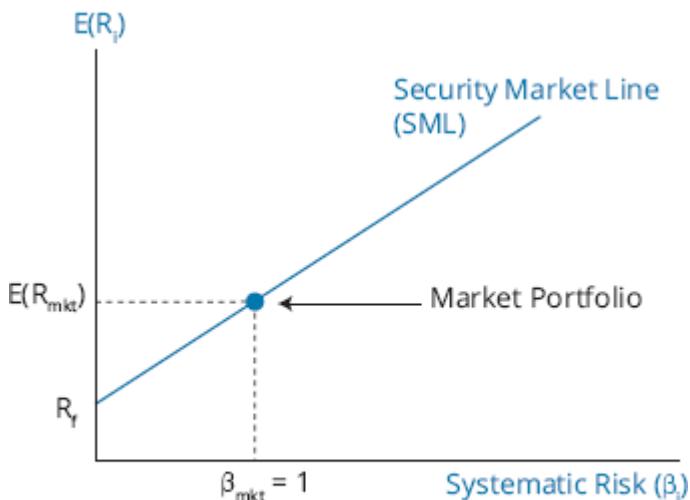
$$E(R_i) = R_f + \frac{\text{Cov}_{i,mkt}}{\sigma_{mkt}^2} [E(R_{mkt}) - R_f]$$

The line described by this last equation is presented in Figure 63.8, where we let the

standardized covariance term, $\frac{\text{Cov}_{i,mkt}}{\sigma_{mkt}^2}$, be defined as beta, β_i .

This is the most common means of describing the SML, and this relation between beta (systematic risk) and expected return is known as the **capital asset pricing model (CAPM)**.

Figure 63.8: The Capital Asset Pricing Model



©2022 Kaplan, Inc.

So, we can define beta, $\beta = \frac{\text{Cov}_{i,mkt}}{\sigma_{mkt}^2}$, as a standardized measure of systematic risk.

Beta measures the relation between a security's excess returns and the excess returns to the market portfolio.

Formally, the CAPM is stated as:

$$E(R_i) = R_f + \beta_i [E(R_{mkt}) - R_f]$$

The CAPM holds that, in equilibrium, the expected return on risky asset $E(R_i)$ is the risk-free rate (R_f) plus a beta-adjusted market risk premium, $\beta_i [E(R_{mkt}) - R_f]$. Beta measures systematic (market or covariance) risk.

EXAMPLE: Capital asset pricing model

The expected return on the market is 8%, the risk-free rate is 2%, and the beta for Stock A is 1.2. Compute the rate of return that would be expected (required) on this stock.

Answer:

$$E(R_A) = 2\% + 1.2(8\% - 2\%) = 9.2\%$$

Note: $\beta_A > 1$, so $E(R_A) > E(R_{mkt})$

The **assumptions of the CAPM** are:

- *Risk aversion.* To accept a greater degree of risk, investors require a higher expected return.
- *Utility maximizing investors.* Investors choose the portfolio, based on their individual preferences, with the risk and return combination that maximizes their (expected) utility.
- *Frictionless markets.* There are no taxes, transaction costs, or other impediments to trading.
- *One-period horizon.* All investors have the same one-period time horizon.
- *Homogeneous expectations.* All investors have the same expectations for assets' expected returns, standard deviation of returns, and returns correlations between assets.
- *Divisible assets.* All investments are infinitely divisible.
- *Competitive markets.* Investors take the market price as given and no investor can influence prices with their trades.

Comparing the CML and the SML

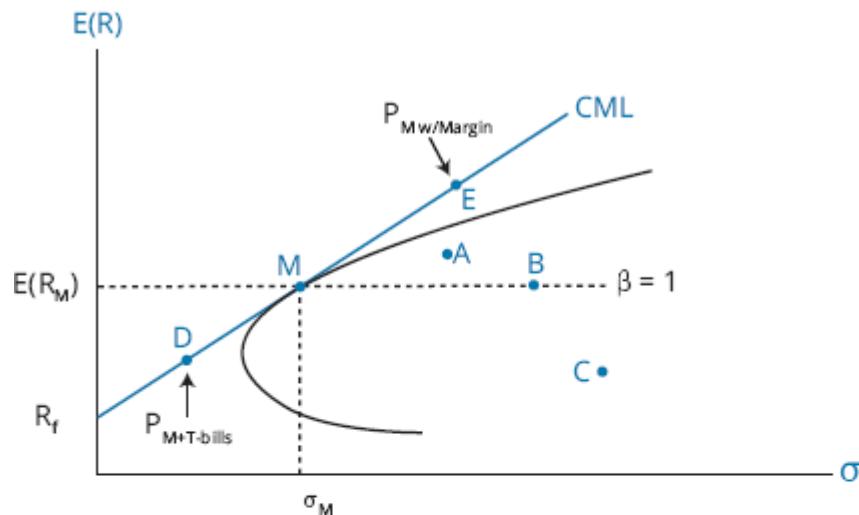
It is important to recognize that the CML and SML are very different. Recall the equation of the CML:

$$E(R_p) = R_f + \sigma_p \left\{ \frac{[E(R_M) - R_f]}{\sigma_M} \right\}$$

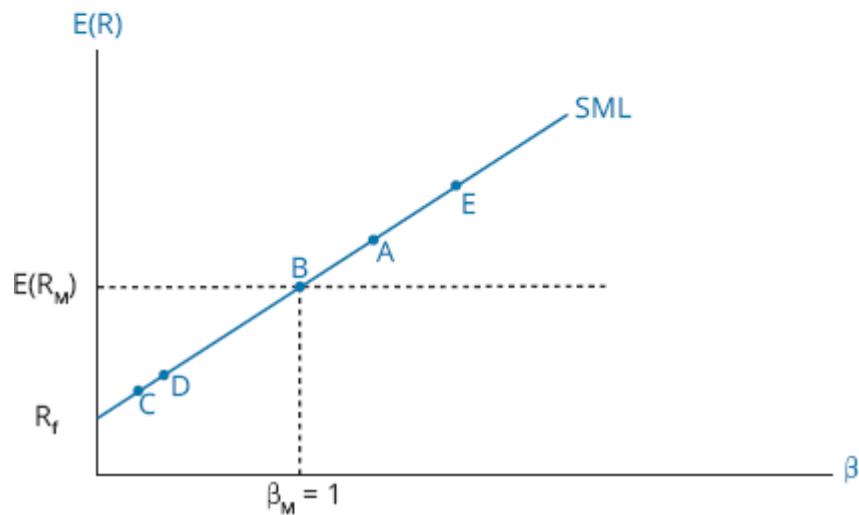
The CML uses total risk = σ_p on the x-axis. Hence, only efficient portfolios will plot on the CML. On the other hand, the SML uses beta (systematic risk) on the x-axis. So in a CAPM world, *all properly priced securities and portfolios of securities will plot on the SML*, as shown in Figure 63.9.

Figure 63.9: Comparing the CML and the SML

(a) Capital Market Line



(b) Security Market Line



Portfolios that are not well diversified (efficient) plot inside the efficient frontier and are represented by risk-return combinations such as points A, B, and C in panel (a) of Figure 63.9. Individual securities are one example of such inefficient portfolios. According to the CAPM, the expected returns on all portfolios, well diversified or not, are determined by their systematic risk. Thus, according to the CAPM, Point A represents a high-beta stock or portfolio, Point B a stock or portfolio with a beta of one, and Point C a low-beta stock or portfolio. We know this because the expected return at Point B is equal to the expected return on the market, and the expected returns at Point A and C are greater and less than the expected return on the market (tangency) portfolio, respectively.

Note that a low-beta stock, such as represented by Point C, is not necessarily low-risk when total risk is considered. While its contribution to the risk of a well-diversified portfolio may be low, its risk when held by itself can be considered quite high. A firm whose only activity is developing a new, but as yet unproven, drug may be quite speculative with highly uncertain returns. It may also have quite low systematic risk if the uncertainty about its future returns depends primarily on firm-specific factors.

All stocks and portfolios that plot along the line labeled $\beta = 1$ in Figure 63.9 have the same expected return as the market portfolio and, thus, according to the CAPM, have the same systematic risk as the market portfolio (i.e., they all have betas of one).

All points on the CML (except the tangency point) represent the risk-return characteristics of portfolios formed by either combining the market portfolio with the risk-free asset or borrowing at the risk-free rate in order to invest more than 100% of the portfolio's net value in the risky market portfolio (investing on margin). Point D in Figure 63.9 represents a portfolio that combines the market portfolio with the risk-free asset, while points above the point of tangency, such as Point E, represent portfolios created by borrowing at the risk-free rate to invest in the market portfolio. Portfolios that do not lie on the CML are not efficient and therefore have risk that will not be rewarded with higher expected returns in equilibrium.

According to the CAPM, all securities and portfolios, diversified or not, will plot on the SML in equilibrium. In fact, all stocks and portfolios along the line labeled $\beta = 1$ in Figure 63.9, including the market portfolio, will plot at the same point on the SML. They will plot at the point on the SML with beta equal to one and expected return equal to the expected return on the market, regardless of their total risk.

LOS 63.h: Describe and demonstrate applications of the CAPM and the SML.

We have used beta to estimate a security's expected return based on our estimate of the risk-free rate and the expected return on the market. In equilibrium, a security's expected return and its required return (by investors) are equal. Therefore, we can use the CAPM to estimate a security's required return.

Because the SML shows the equilibrium (required) return for any security or portfolio based on its beta (systematic risk), analysts often compare their forecast of a security's return to its required return based on its beta risk. The following example illustrates this technique.

EXAMPLE: Identifying mispriced securities

The following figure contains information based on analyst's forecasts for three stocks. Assume a risk-free rate of 7% and a market return of 15%. Compute the expected and required return on each stock, determine whether each stock is undervalued, overvalued, or properly valued, and outline an appropriate trading strategy.

Forecast Data

Stock	Price Today	E(Price) in 1 Year	E(Dividend) in 1 Year	Beta
A	\$25	\$27	\$1.00	1.0
B	40	45	2.00	0.8
C	15	17	0.50	1.2

Answer:

Expected and required returns computations are shown in the following figure.

Forecasts vs. Required Returns

Stock	Forecast Return	Required Return
A	$(\$27 - \$25 + \$1) / \$25 = 12.0\%$	$0.07 + (1.0)(0.15 - 0.07) = 15.0\%$
B	$(\$45 - \$40 + \$2) / \$40 = 17.5\%$	$0.07 + (0.8)(0.15 - 0.07) = 13.4\%$
C	$(\$17 - \$15 + \$0.5) / \$15 = 16.6\%$	$0.07 + (1.2)(0.15 - 0.07) = 16.6\%$

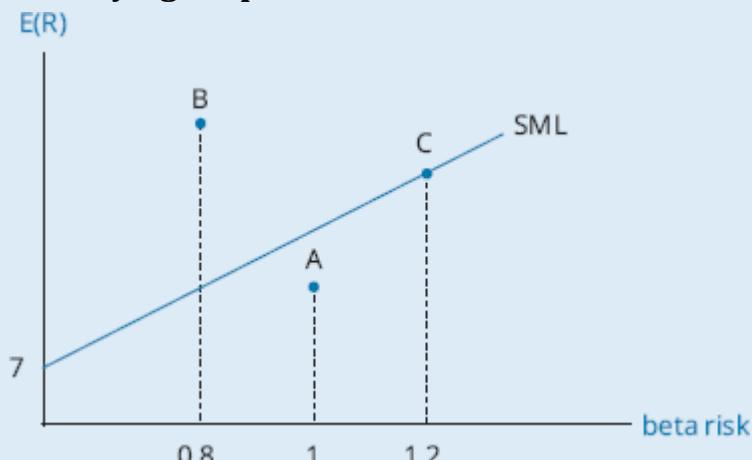
- Stock A is *overvalued*. It is expected to earn 12%, but based on its systematic risk, it should earn 15%. It plots *below* the SML.
- Stock B is *undervalued*. It is expected to earn 17.5%, but based on its systematic risk, it should earn 13.4%. It plots *above* the SML.
- Stock C is *properly valued*. It is expected to earn 16.6%, and based on its systematic risk, it should earn 16.6%. It plots *on* the SML.

The appropriate trading strategy is:

- Short sell Stock A.
- Buy Stock B.
- Buy, sell, or ignore Stock C.

We can do this same analysis graphically. The expected return/beta combinations of all three stocks are graphed in the following figure relative to the SML.

Identifying Mispriced Securities



PROFESSOR'S NOTE

If the estimated return plots “over” the SML, the security is “under” valued. If the estimated return plots “under” the SML, the security is “over” valued.

Remember, all stocks should plot on the SML; any stock not plotting on the SML is mispriced. Notice that Stock A falls below the SML, Stock B lies above the SML, and Stock C is on the SML. If you plot a stock’s expected return and it falls below the SML, the stock is overpriced. That is, the stock’s expected return is too low given its systematic risk. If a stock plots above the SML, it is underpriced and is offering an expected return greater than required for its systematic risk. If it plots on the SML, the stock is properly priced.

Because the equation of the SML is the capital asset pricing model, you can determine if a stock is over- or underpriced graphically or mathematically. Your answers will always be the same.

LOS 63.i: Calculate and interpret the Sharpe ratio, Treynor ratio, M^2 , and Jensen's alpha.

Performance evaluation of an active manager's portfolio choices refers to the analysis of the risk and return of the portfolio. **Attribution analysis**, an analysis of the sources of returns differences between active portfolio returns and those of a passive benchmark portfolio, is part of performance evaluation. Success in active portfolio management cannot be determined simply by comparing portfolio returns to benchmark portfolio returns; the risk taken to achieve returns must also be considered. A portfolio with greater risk than the benchmark portfolio (especially beta risk) is expected to produce higher returns over time than the benchmark portfolio.

When evaluating the performance of a portfolio with risk that differs from that of a benchmark portfolio, we need to adjust the active portfolio return's risk. Of the alternative ways to consider both risk and return in evaluating portfolio performance, the most commonly used is the **Sharpe ratio**. The Sharpe ratio of a portfolio is its excess returns per unit of total portfolio risk. Higher Sharpe ratios indicate better risk-adjusted portfolio performance.

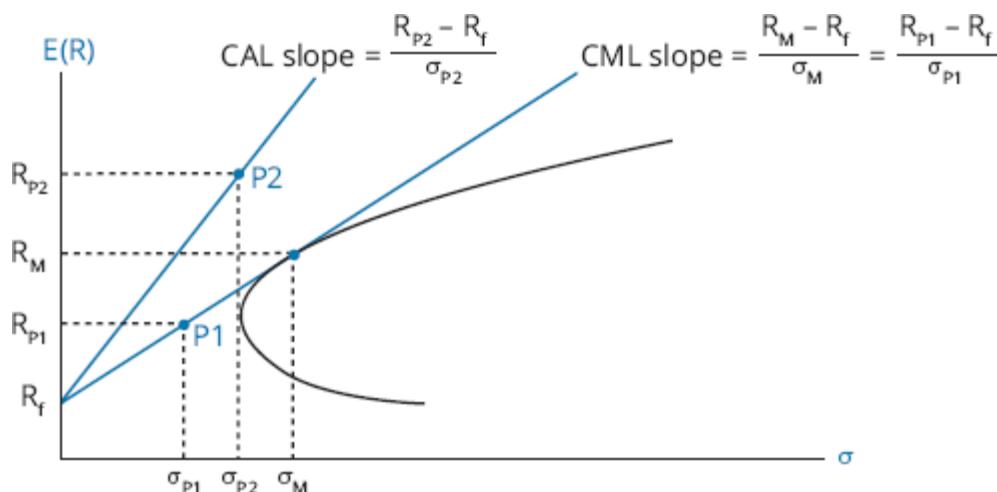
$$\text{Sharpe ratio} = \frac{E[R_{\text{portfolio}}] - R_f}{\sigma_{\text{portfolio}}}$$

We have shown the Sharpe ratio as an ex ante (before the fact) measure, using the expected values of portfolio returns and standard deviation. However, it can also be used as an ex post (after the fact) measure of portfolio performance, using mean returns and sample standard deviation over a period.

The Sharpe ratio is based on total risk (standard deviation of returns), rather than systematic risk (beta). For this reason, the Sharpe ratio can be used to evaluate the performance of concentrated portfolios (those affected by unsystematic risk) as well as well-diversified portfolios (those with only systematic, or beta, risk). Note that the value of the Sharpe ratio is only useful for comparison with the Sharpe ratio of another portfolio.

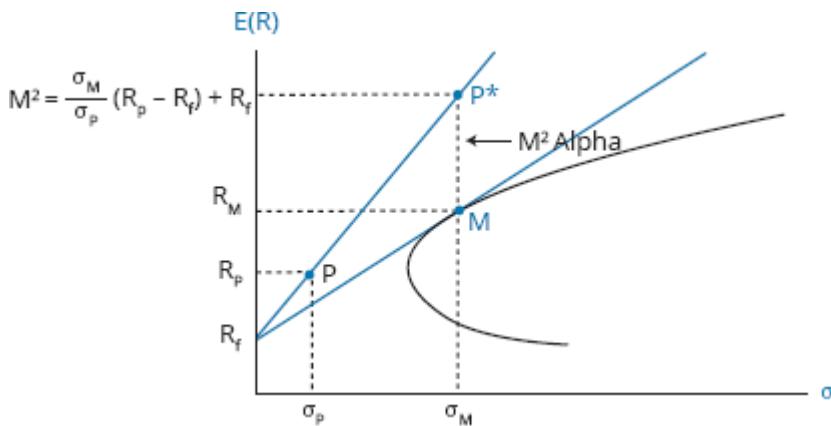
In Figure 63.10, we illustrate that the Sharpe ratio of a portfolio is the slope of the CAL for that portfolio and can be compared to the slope of the CML, which is the Sharpe ratio for portfolios that lie on the CML.

Figure 63.10: Sharpe Ratios as Slopes



For a portfolio of risky assets, **M-squared (M²)** index{M-squared (M2)} is an alternative to the Sharpe ratio as a risk-adjusted rate of return, expressed as a percentage rather than as a slope. Given a Portfolio P, we can calculate the return on a Portfolio P* that is leveraged (when $\sigma_M > \sigma_P$), or deleveraged (when $\sigma_M < \sigma_P$), so that P* has the same risk (standard deviation of returns) as the market portfolio. The return on P* is $R_f + \frac{\sigma_M}{\sigma_P}(R_P - R_f)$ and we refer to that as the M2 measure for Portfolio P. We illustrate the return on the leveraged Portfolio P*, given the standard deviation and return on Portfolio P, in Figure 63.11. The extra return on the Portfolio P* above the return on the market portfolio, ($P^* - R_M$), is referred to as **M² alpha** index{M2 alpha}. Note that in Figure 63.11, P* is created by borrowing at R_f and investing the proceeds in Portfolio P, in an amount so that the standard deviation of P* = σ_M .

Figure 63.11: M-Squared for a Portfolio



The M2 measure produces the same risk-adjusted portfolio rankings as the Sharpe ratio, but is stated in percentage terms. Note that M2 can be derived from the Sharpe ratio (SR) for Portfolio P, $SR = (R_P - R_f)/\sigma_P$, as $SR (\sigma_M) + R_f$, so that if the Sharpe ratio of Portfolio P is greater than the slope of the CML, $M2 > R_m$ and $M2 \text{ alpha} > 0$.

As an example, consider a Portfolio P with return of 10% and standard deviation of returns of 20%, when $R_f = 5\%$, $R_M = 11\%$ and $\sigma_M = 30\%$. The Sharpe ratio of Portfolio P = $(10 - 5)/20 = 0.25$, and $M2 = 0.25(0.30) + 0.05 = 12.5\%$. Comparing that to $R_M = 11\%$, we can see that M2 alpha is 1.5%.

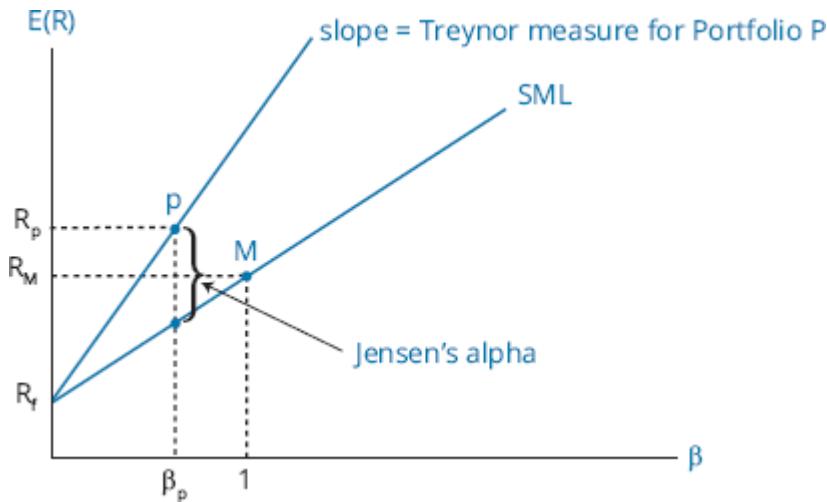
Two measures of portfolio performance based on systematic (beta) risk rather than total risk are the **Treynor measure** and **Jensen's alpha**. They are analogous to the Sharpe ratio and M² in that the Treynor measure is a measure of slope and Jensen's alpha is a measure of percentage returns in excess of those from a portfolio that has the same risk (beta) but lies on the SML.

The Treynor measure is calculated as $\frac{R_P - R_f}{\beta_P}$, interpreted as excess returns per unit of systematic risk, and represented by the slope of a line as illustrated in Figure 63.12. Jensen's alpha for Portfolio P is calculated as

$$\alpha_P = R_p - [R_f + \beta_P(R_M - R_f)]$$

and is the percentage portfolio return above that of a portfolio (or security) with the same beta as the portfolio that lies on the SML, as illustrated in Figure 63.12.

Figure 63.12: Treynor Measure and Jensen's Alpha



Whether risk adjustment should be based on standard deviation of returns or portfolio beta depends on whether a manager's portfolio bears unsystematic risk. If a single manager is used, then the total risk (including any nonsystematic risk) is the relevant measure and risk adjustment using total risk, as with the Sharpe and M^2 measures, is appropriate. If a fund uses multiple managers so that the overall fund portfolio is well diversified (has no unsystematic risk), then performance measures based on systematic (beta) risk, such as the Treynor measure and Jensen's alpha, are appropriate.

These measures of risk-adjusted returns are often used to compare the performance of actively managed funds to passively managed funds. Note in Figure 63.10 and Figure 63.11 that portfolios that lie above the CML have Sharpe ratios greater than those of any portfolios along the CML and have positive M^2 measures. Similarly, in Figure 63.12, we can see that portfolios that lie above the SML have Treynor measures greater than those of any security or portfolio that lies along the SML and also have positive values for Jensen's alpha.

One final note of caution is that estimating the values needed to apply these theoretical models and performance measures is often difficult and is done with error. The expected return on the market, and thus the market risk premium, may not be equal to its average historical value. Estimating security and portfolio betas is done with error as well.



MODULE QUIZ 63.2

- Which of the following statements about the SML and the CML is *least accurate*?
 - Securities that plot above the SML are undervalued.
 - Investors expect to be compensated for systematic risk.
 - Securities that plot on the SML have no value to investors.
- According to the CAPM, what is the expected rate of return for a stock with a beta of 1.2, when the risk-free rate is 6% and the market rate of return is 12%?
 - 7.2%.
 - 12.0%.
 - 13.2%.

3. According to the CAPM, what is the required rate of return for a stock with a beta of 0.7, when the risk-free rate is 7% and the expected market rate of return is 14%?
 - A. 11.9%.
 - B. 14.0%.
 - C. 16.8%.
4. The risk-free rate is 6%, and the expected market return is 15%. A stock with a beta of 1.2 is selling for \$25 and will pay a \$1 dividend at the end of the year. If the stock is priced at \$30 at year-end, it is:
 - A. overpriced, so short it.
 - B. underpriced, so buy it.
 - C. underpriced, so short it.
5. A stock with a beta of 0.7 currently priced at \$50 is expected to increase in price to \$55 by year-end and pay a \$1 dividend. The expected market return is 15%, and the risk-free rate is 8%. The stock is:
 - A. overpriced, so do not buy it.
 - B. underpriced, so buy it.
 - C. properly priced, so buy it.
6. Which of these return metrics is defined as excess return per unit of systematic risk?
 - A. Sharpe ratio.
 - B. Jensen's alpha.
 - C. Treynor measure.

KEY CONCEPTS

LOS 63.a

The availability of a risk-free asset allows investors to build portfolios with superior risk-return properties. By combining a risk-free asset with a portfolio of risky assets, the overall risk and return can be adjusted to appeal to investors with various degrees of risk aversion.

LOS 63.b

On a graph of return versus risk, the various combinations of a risky asset and the risk-free asset form the capital allocation line (CAL). In the specific case where the risky asset is the market portfolio, the combinations of the risky asset and the risk-free asset form the capital market line (CML).

LOS 63.c

Systematic (market) risk is due to factors, such as GDP growth and interest rate changes, that affect the values of all risky securities. Systematic risk cannot be reduced by diversification. Unsystematic (firm-specific) risk can be reduced by portfolio diversification.

Because one of the assumptions underlying the CAPM is that portfolio diversification to eliminate unsystematic risk is costless, investors cannot increase expected equilibrium portfolio returns by taking on unsystematic risk.

LOS 63.d

A return generating model is an equation that estimates the expected return of an investment, based on a security's exposure to one or more macroeconomic, fundamental, or statistical factors.

The simplest return generating model is the market model, which assumes the return on an asset is related to the return on the market portfolio in the following manner:

$$R_i = \alpha_i + \beta_i R_m + e_i$$

LOS 63.e

Beta can be calculated using the following equation:

$$\beta_i = \frac{[\text{Cov}(R_i, R_m)]}{\sigma_m^2} = \rho_{im} \left(\frac{\sigma_i}{\sigma_m} \right)$$

where $[\text{Cov}(R_i, R_m)]$ and ρ_{im} are the covariance and correlation between the asset and the market, and σ_i and σ_m are the standard deviations of asset returns and market returns.

The theoretical average beta of stocks in the market is 1. A beta of zero indicates that a security's return is uncorrelated with the returns of the market.

LOS 63.f

The capital asset pricing model (CAPM) requires several assumptions:

- Investors are risk averse, utility maximizing, and rational.
- Markets are free of frictions like costs and taxes.
- All investors plan using the same time period.
- All investors have the same expectations of security returns.
- Investments are infinitely divisible.
- Prices are unaffected by an investor's trades.

The security market line (SML) is a graphical representation of the CAPM that plots expected return versus beta for any security.

LOS 63.g

The CAPM relates expected return to the market factor (beta) using the following formula:

$$E(R_i) - R_f = \beta_i [E(R_m) - R_f]$$

LOS 63.h

The CAPM and the SML indicate what a security's equilibrium required rate of return should be based on the security's exposure to market risk. An analyst can compare his expected rate of return on a security to the required rate of return indicated by the SML to determine whether the security is overvalued, undervalued, or properly valued.

LOS 63.i

The Sharpe ratio measures excess return per unit of total risk and is useful for comparing portfolios on a risk-adjusted basis.

$$\text{Sharpe ratio} = \left(\frac{R_p - R_f}{\sigma_p} \right)$$

Given a Portfolio P, we can calculate the return on a Portfolio P* that is leveraged or deleveraged, so that P* has the same risk as the market portfolio. The return on P* is the M-squared measure for portfolio P.

$$M^2 = R_f + \frac{\sigma_M}{\sigma_p} (R_p - R_f)$$

M-squared alpha is the extra return on Portfolio P* above the market portfolio.

The Treynor measure measures a portfolio's excess return per unit of systematic risk. Jensen's alpha is the difference between a portfolio's return and the return of a portfolio on the SML that has the same beta:

$$\text{Treynor measure} = \frac{R_p - R_f}{\beta_p}$$

$$\text{Jensen's alpha} = \alpha_p = R_p - [R_f + \beta_p(R_M - R_f)]$$

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 63.1

1. **B** Expected return: $(0.60 \times 0.10) + (0.40 \times 0.05) = 0.08$, or 8.0%
Standard deviation: $0.60 \times 0.08 = 0.048$, or 4.8%
(LOS 63.a)
2. **C** The capital market line (CML) plots return against *total risk*, which is measured by standard deviation of returns. (LOS 63.b)
3. **B** A portfolio to the right of a portfolio on the CML has more risk than the market portfolio. Investors seeking to take on more risk will *borrow* at the risk-free rate to purchase more of the market portfolio. (LOS 63.b)
4. **A** When you increase the number of stocks in a portfolio, *unsystematic risk* will decrease at a decreasing rate. However, the portfolio's *systematic risk* can be increased by adding higher-beta stocks or decreased by adding lower-beta stocks. (LOS 63.c)
5. **C** Total risk equals systematic plus unsystematic risk. Unique risk is diversifiable and is unsystematic. Market (systematic) risk is nondiversifiable risk. (LOS 63.c)
6. **A** Macroeconomic, fundamental, and statistical factor exposures can be included in a return generating model to estimate the expected return of an investment. However, statistical factors may not have any theoretical basis, so analysts prefer macroeconomic and fundamental factor models. (LOS 63.d)
7. **C** $\text{beta} = \text{covariance} / \text{market variance}$
 $\text{market variance} = 0.05^2 = 0.0025$
 $\text{beta} = 0.005 / 0.0025 = 2.0$
(LOS 63.e)

Module Quiz 63.2

1. **C** Securities that plot on the SML are expected to earn their equilibrium rate of return and, therefore, do have value to an investor and may have diversification benefits as well. The other statements are true. (LOS 63.f)
2. **C** $6 + 1.2(12 - 6) = 13.2\%$ (LOS 63.g)
3. **A** $7 + 0.7(14 - 7) = 11.9\%$ (LOS 63.g)
4. **B** required rate = $6 + 1.2(15 - 6) = 16.8\%$
 $\text{return on stock} = (30 - 25 + 1) / 25 = 24\%$
Based on risk, the stock plots above the SML and is underpriced, so buy it. (LOS 63.h)
5. **A** required rate = $8 + 0.7(15 - 8) = 12.9\%$

return on stock = $(55 - 50 + 1) / 50 = 12\%$

The stock falls below the SML, so it is *overpriced*. (LOS 63.h)

6. C The Treynor measure is excess return (return in excess of the risk-free rate) per unit of systematic risk (beta). The Sharpe ratio is excess return per unit of total risk (portfolio standard deviation). Jensen's alpha is the difference between a portfolio's actual rate of return and the equilibrium rate of return for a portfolio with the same level of beta (systematic) risk. (LOS 63.i)

READING 64

BASICS OF PORTFOLIO PLANNING AND CONSTRUCTION

EXAM FOCUS

There is nothing difficult here, but the material is important because it is the foundation for the portfolio construction material at Level II and especially Level III. You should be ready to explain why investment policy statements are created and what their major components are. You should be familiar with the objectives (risk and return) and the constraints: liquidity, legal, time horizon, tax treatment, and unique circumstances. Know the difference between ability and willingness to take risk, the factors that define an asset class, and how asset allocation is used in constructing portfolios.

MODULE 64.1: PORTFOLIO PLANNING AND CONSTRUCTION



Video covering this content is available online.

LOS 64.a: Describe the reasons for a written investment policy statement (IPS).

An investment manager is very unlikely to produce a good result for a client without understanding that client's needs, circumstances, and constraints.

A written **investment policy statement** will typically begin with the investor's goals in terms of risk and return. These should be determined jointly, as the goals of high returns and low risk (while quite popular) are likely to be mutually exclusive in practice. Investor expectations in terms of returns must be compatible with investor's tolerance for risk (uncertainty about portfolio performance).

LOS 64.b: Describe the major components of an IPS.

The major components of an IPS typically address the following:

- *Description of Client* circumstances, situation, and investment objectives.
- *Statement of the Purpose* of the IPS.
- *Statement of Duties and Responsibilities* of investment manager, custodian of assets, and the client.

- *Procedures* to update IPS and to respond to various possible situations.
- *Investment Objectives* derived from communications with the client.
- *Investment Constraints* that must be considered in the plan.
- *Investment Guidelines* such as how the policy will be executed, asset types permitted, and leverage to be used.
- *Evaluation of Performance*, the benchmark portfolio for evaluating investment performance, and other information on evaluation of investment results.
- *Appendices* containing information on strategic (baseline) asset allocation and permitted deviations from policy portfolio allocations, as well as how and when the portfolio allocations should be rebalanced.

In any case, the IPS will, at a minimum, contain a clear statement of client circumstances and constraints, an investment strategy based on these, and some benchmark against which to evaluate the account performance.

LOS 64.c: Describe risk and return objectives and how they may be developed for a client.

The **risk objectives** in an IPS may take several forms. An **absolute risk objective** might be to “have no decrease in portfolio value during any 12-month period” or to “not decrease in value by more than 2% at any point over any 12-month period.” Low absolute percentage risk objectives such as these may result in portfolios made up of securities that offer guaranteed returns (e.g., U.S. Treasury bills).

Absolute risk objectives can also be stated in terms of the probability of specific portfolio results, either percentage losses or dollar losses, rather than strict limits on portfolio results. Examples are as follows:

- “No greater than a 5% probability of returns below -5% in any 12-month period.”
- “No greater than a 4% probability of a loss of more than \$20,000 over any 12-month period.”

An absolute return objective may be stated in nominal terms, such as “an overall return of at least 6% per annum,” or in real returns, such as “a return of 3% more than the annual inflation rate each year.”

Relative risk objectives relate to a specific benchmark and can also be strict, such as, “Returns will not be less than 12-month euro LIBOR over any 12-month period,” or stated in terms of probability, such as, “No greater than a 5% probability of returns more than 4% below the return on the MSCI World Index over any 12-month period.”

Return objectives can be relative to a benchmark portfolio return, such as, “Exceed the return on the S&P 500 Index by 2% per annum.” For a bank, the return objective may be relative to the bank’s cost of funds (deposit rate). While it is possible for an institution to use returns on peer portfolios, such as an endowment with a stated objective to be in the top quartile of endowment fund returns, peer performance benchmarks suffer from not being *investable* portfolios. There is no way to match this investment return by portfolio construction before the fact.

In any event, the account manager must make sure that the stated risk and return objectives are compatible, given the reality of expected investment results and uncertainty over time.

LOS 64.d: Explain the difference between the willingness and the ability (capacity) to take risk in analyzing an investor's financial risk tolerance.

An investor's **ability to bear risk** depends on financial circumstances. Longer investment horizons (20 years rather than 2 years), greater assets versus liabilities (more wealth), more insurance against unexpected occurrences, and a secure job all suggest a greater ability to bear investment risk in terms of uncertainty about periodic investment performance.

An investor's **willingness to bear risk** is based primarily on the investor's attitudes and beliefs about investments (various asset types). The assessment of an investor's attitude about risk is quite subjective and is sometimes done with a short questionnaire that attempts to categorize the investor's risk aversion or risk tolerance.

When the adviser's assessments of an investor's ability and willingness to take investment risk are compatible, there is no real problem selecting an appropriate level of investment risk. If the investor's willingness to take on investment risk is high but the investor's ability to take on risk is low, the low ability to take on investment risk will prevail in the adviser's assessment.

In situations where ability is high but willingness is low, the adviser may attempt to educate the investor about investment risk and correct any misconceptions that may be contributing to the investor's low stated willingness to take on investment risk. However, the adviser's job is not to change the investor's personality characteristics that contribute to a low willingness to take on investment risk. The approach will most likely be to conform to the lower of the investor's ability or willingness to bear risk, as constructing a portfolio with a level of risk that the client is clearly uncomfortable with will not likely lead to a good outcome in the investor's view.

LOS 64.e: Describe the investment constraints of liquidity, time horizon, tax concerns, legal and regulatory factors, and unique circumstances and their implications for the choice of portfolio assets.



PROFESSOR'S NOTE

When I was studying for the CFA exams over 20 years ago, we memorized R-R-T-T-L-L-U as a checklist for addressing the important points of portfolio construction, and it still works today. Then, as now, the important points to cover in an IPS were Risk, Return, Time horizon, Tax situation, Liquidity, Legal restrictions, and the Unique constraints of a specific investor.

Investment constraints include the investor's liquidity needs, time horizon, tax considerations, legal and regulatory constraints, and unique needs and preferences.

Liquidity: Liquidity refers to the ability to turn investment assets into spendable cash in a short period of time without having to make significant price concessions to do so. Investor needs for money to pay tuition, to pay for a parent's assisted living expenses, or to fund other possible

spending needs may all require that some liquid assets be held. As we noted in an earlier reading discussing property and casualty insurance companies, claims arrive unpredictably to some extent and therefore their portfolios must hold a significant proportion of liquid (or maturing) securities in order to be prepared to honor these claims. Illiquid investments in hedge funds and private equity funds, which typically are not traded and have restrictions on redemptions, are not suitable for an investor who may unexpectedly need access to the funds.

Time horizon: In general, the longer an investor's time horizon, the more risk and less liquidity the investor can accept in the portfolio. While the expected returns on a broad equities portfolio may not be too risky for an investor with a 20-year investment horizon, they likely are too risky for an investor who must fund a large purchase at the end of this year. For such an investor, government securities or a bank certificate of deposit may be the most appropriate investments because of their low risk and high liquidity at the time when the funds will be needed.

Tax situation: Besides an individual's overall tax rate, the tax treatment of various types of investment accounts is also a consideration in portfolio construction. For a fully taxable account, investors subject to higher tax rates may prefer tax-free bonds (U.S.) to taxable bonds or prefer equities that are expected to produce capital gains, which are often taxed at a lower rate than other types of income. A focus on expected after-tax returns over time in relation to risk should correctly account for differences in tax treatments as well as investors' overall tax rates.

Some types of investment accounts, such as retirement accounts, may be tax exempt or tax deferred. Investors with such accounts may choose to put securities that generate fully taxed income, such as corporate bond interest, in accounts that are tax deferred, while seeking long-term capital gains, tax-exempt interest income, and dividend income (in jurisdictions where dividends receive preferential tax treatment) in their personal accounts, which have no tax deferral benefit.

Legal and regulatory: In addition to financial market regulations that apply to all investors, more specific legal and regulatory constraints may apply to particular investors. Trust, corporate, and qualified investment accounts may all be restricted by law from investing in particular types of securities and assets. There may also be restrictions on percentage allocations to specific types of investments in such accounts. Corporate officers and directors face legal restrictions on trading in the securities of their firms that the account manager should be aware of.

Unique circumstances: Each investor, whether individual or institutional, may have specific preferences or restrictions on which securities and assets may be purchased for the account. Some of these may be nonfinancial considerations, which are commonly categorized as **responsible investing**. Ethical preferences, such as prohibiting investment in securities issued by tobacco or firearms producers, are not uncommon. Restrictions on investments in companies or countries where human rights abuses are suspected or documented would also fall into this category. Religious preferences may preclude investment in securities that make explicit interest payments.

Unique investor preferences may also be based on diversification needs when the investor's income depends heavily on the prospects for one company or industry. An investor who has

founded or runs a company may not want any investment in securities issued by a competitor to that company.

LOS 64.f: Explain the specification of asset classes in relation to asset allocation.

After having determined the investor objectives and constraints through the exercise of creating an IPS, a **strategic asset allocation** is developed which specifies the percentage allocations to the included asset classes. In choosing which asset classes to consider when developing the strategic asset allocation for the account, the correlations of returns *within* an asset class should be relatively high, indicating that the assets within the class are similar in their investment performance. On the other hand, it is low correlations of returns *between* asset classes that leads to risk reduction through portfolio diversification.

Historically, only the broad categories of equities, bonds, cash, and real estate were considered. More recently, a group of several investable asset classes, referred to collectively as alternative investments, has gained more prominence. Alternative investment asset classes include hedge funds of various types, private equity funds, managed or passively constructed commodity funds, artwork, and intellectual property rights.

We can further divide equities by whether the issuing companies are domestic or foreign, large or small, or whether they are traded in emerging or developed markets. An example of specifying asset classes is world equities. A U.S. investor may want to divide world equities into different regions.

With bonds, we can divide the overall universe of bonds into asset classes based on maturities or on criteria such as whether they are foreign or domestic, government or corporate, or investment grade or speculative (high yield). Overall, the asset classes considered should approximate the universe of permissible investments specified in the IPS.

Once the universe of asset classes has been specified, the investment manager will collect data on the returns, standard deviation of returns, and correlations of returns with those of other asset classes for each asset class.

Figure 64.1 illustrates the strategic asset allocation for a pension fund.

Figure 64.1: Strategic Asset Allocation

The Vermont Pension Investment Committee manages more than \$4 billion in retirement assets for various teachers and state and municipal employees in that state. VPIC's investment policy specifies the following strategic asset allocation:

Asset Class	Target
Growth assets	
Passive global equities	24%
Active global equities	5%
Large cap U.S. equities	4%
Small-/mid-cap U.S. equities	3%
Non-U.S. developed market equities	5%
International small-cap equities	2%
Private equity	10%
Core plus fixed income	6%
Emerging market debt	4%
Private debt	5%
Non-core real estate	3%
Total growth assets	71%
Downturn hedging assets	
Core fixed income	14%
Short-term quality credit	5%
Total downturn hedging	19%
Inflation hedging assets	
Core real estate	5%
U.S. TIPS	3%
Infrastructure/farmland	2%
Total inflation hedging	10%

Source: State of Vermont, Office of the State Treasurer.

Target allocation as of March 25, 2019.

www.vermonttreasurer.gov/content/pension.

LOS 64.g: Describe the principles of portfolio construction and the role of asset allocation in relation to the IPS.

Once the portfolio manager has identified the investable asset classes for the portfolio and the risk, return, and correlation characteristics of each asset class, an *efficient frontier*, analogous to one constructed from individual securities, can be constructed using a computer program. By combining the return and risk objectives from the IPS with the actual risk and return properties of the many portfolios along the efficient frontier, the manager can identify that portfolio which best meets the risk and return requirements of the investor. The asset allocation for the efficient portfolio selected is then the strategic asset allocation for the portfolio.

So far, we have not concerned ourselves with deviations from strategic asset allocations or with selection of individual securities within individual asset classes. These activities are referred to as active (versus passive) portfolio management strategies. A manager who varies from strategic asset allocation weights in order to take advantage of perceived short-term opportunities is adding **tactical asset allocation** to the portfolio strategy. **Security selection** refers to

deviations from index weights on individual securities within an asset class. For example, a portfolio manager might overweight energy stocks and underweight financial stocks, relative to the index weights for U.S. large-cap equities as an asset class. For some asset classes, such as hedge funds, individual real estate properties, and artwork, investable indexes are not available. For these asset classes, selection of individual assets is required by the nature of the asset class.

While each of these active strategies may produce higher returns, they each also increase the risk of the portfolio compared to a passive portfolio of asset class indexes. A practice known as **risk budgeting** sets an overall risk limit for the portfolio and budgets (allocates) a portion of the permitted risk to the systematic risk of the strategic asset allocation, the risk from tactical asset allocation, and the risk from security selection.

Active portfolio management has two specific issues to consider.

1. An investor may have multiple managers actively managing to the same benchmark for the same asset class (or may have significant benchmark overlap). In this case, one manager may overweight an index stock while another may underweight the same stock. Taken together, there is no net active management risk, although each manager has reported active management risk. Overall, the risk budget is underutilized as there is less net active management than gross active management.
2. When all managers are actively managing portfolios relative to an index, trading may be excessive overall. This extra trading could have negative tax consequences, specifically potentially higher capital gains taxes, compared to an overall efficient tax strategy.

One way to address these issues is to use a **core-satellite approach**. The core-satellite approach invests the majority, or core, portion of the portfolio in passively managed indexes and invests a smaller, or satellite, portion in active strategies. This approach reduces the likelihood of excessive trading and offsetting active positions.

Clearly, the success of security selection will depend on the manager's skill and the opportunities (mispricings or inefficiencies) within a particular asset class. Similarly, the success of tactical asset allocation will depend both on the existence of short-term opportunities in specific asset classes and on the manager's ability to identify them.

LOS 64.h: Describe how environmental, social, and governance (ESG) considerations may be integrated into portfolio planning and construction.

In our Corporate Issuers review of "Introduction to Corporate Governance and Other ESG Considerations," we described several approaches to ESG investing. Briefly, these approaches include:

- **Negative screening**, excluding specific companies or industries based on ESG factors.
- **Positive screening**, investing in companies that have positive ESG practices.
- **Thematic investing**, selecting sectors or companies to promote specific ESG-related goals.
- **Impact investing**, selecting investments both to provide a return and to promote positive ESG practices.

- **Engagement/active ownership**, using share ownership as a platform to promote improved ESG practices at a company.
- **ESG integration**, considering ESG factors throughout the asset allocation and security selection process.

Here, we look at some issues these approaches raise for portfolio management.

If a portfolio's investment universe is constrained by negative screening, measuring its performance against a broad market index is unlikely to be appropriate. Indexes excluding companies or industries that investors with ESG concerns commonly avoid are available.

While many investors use positive screening, thematic, or impact investing approaches, choices of which specific companies to invest in and which ESG factors to focus on differ among investors. Thus, portfolios and performance benchmarks must be customized under these approaches, and they may require investment managers who specialize in these styles of investing.

For investment managers with clients who wish to engage in active ownership, it is important to clarify whether the clients intend to vote their shares themselves or direct the managers to vote the shares according to specified ESG factors.

Regardless of the approach chosen, investors should be aware that imposing constraints based on ESG factors will likely affect portfolio performance. How these constraints will affect portfolio performance in practice is uncertain. Both limiting the universe of investment choices and incurring the costs involved in considering ESG factors may decrease returns. On the other hand, investing in companies with good corporate governance practices and avoiding those that face ESG-related risks may increase portfolio returns.



MODULE QUIZ 64.1

1. The investment policy statement is *most accurately* considered:
 - A. the starting point of the portfolio management process.
 - B. the key intermediate step in the portfolio management process.
 - C. the end product of the portfolio management process.
2. The component of an investment policy statement that defines the investment objectives is *most likely* to include information about:
 - A. the investor's risk tolerance.
 - B. unique needs and preferences of the investor.
 - C. permitted asset types and use of leverage in the investment account.
3. When an investment advisor is developing return and risk objectives for a client:
 - A. return objectives should be absolute and risk objectives should be relative.
 - B. risk objectives should be absolute and return objectives should be relative.
 - C. both return and risk objectives may be stated in absolute or relative terms.
4. A client exhibits an above-average willingness to take risk but a below-average ability to take risk. When assigning an overall risk tolerance, the investment adviser is *most likely* to assess the client's overall risk tolerance as:
 - A. above average.
 - B. average.
 - C. below average.
5. Which of the following is *least likely* an example of a portfolio constraint?
 - A. Higher tax rate on dividend income than on capital gains.

- B. Significant spending requirements in the near future.
 - C. Minimum total return requirement of 8%.
6. For asset allocation purposes, asset classes should be specified such that correlations of returns are relatively:
- A. low within each asset class and low among asset classes.
 - B. high within each asset class and low among asset classes.
 - C. low within each asset class and high among asset classes.
7. In determining the appropriate asset allocation for a client's investment account, the manager should:
- A. consider only the investor's risk tolerance.
 - B. incorporate forecasts of future economic conditions.
 - C. consider the investor's risk tolerance and future needs, but not forecasts of market conditions.

KEY CONCEPTS

LOS 64.a

A written investment policy statement, the first step in the portfolio management process, is a plan for achieving investment success. An IPS forces investment discipline and ensures that goals are realistic by requiring investors to articulate their circumstances, objectives, and constraints.

LOS 64.b

Many IPS include the following sections:

- Introduction—Describes the client.
- Statement of Purpose—The intentions of the IPS.
- Statement of Duties and Responsibilities—Of the client, the asset custodian, and the investment managers.
- Procedures—Related to keeping the IPS updated and responding to unforeseen events.
- Investment Objectives—The client's investment needs, specified in terms of required return and risk tolerance.
- Investment Constraints—Factors that may hinder the ability to meet investment objectives; typically categorized as time horizon, taxes, liquidity, legal and regulatory, and unique needs.
- Investment Guidelines—For example, whether leverage, derivatives, or specific kinds of assets are allowed.
- Evaluation and Review—Related to feedback on investment results.
- Appendices—May specify the portfolio's strategic asset allocation (policy portfolio) or the portfolio's rebalancing policy.

LOS 64.c

Risk objectives are specifications for portfolio risk that are developed to embody a client's risk tolerance. Risk objectives can be either absolute (e.g., no losses greater than 10% in any year) or relative (e.g., annual return will be within 2% of FTSE return).

Return objectives are typically based on an investor's desire to meet a future financial goal, such as a particular level of income in retirement. Return objectives can be absolute (e.g., 9% annual return) or relative (e.g., outperform the S&P 500 by 2% per year).

The achievability of an investor's return expectations may be hindered by the investor's risk objectives.

LOS 64.d

Willingness to take financial risk is related to an investor's psychological factors, such as personality type and level of financial knowledge.

Ability or capacity to take risk depends on financial factors, such as wealth relative to liabilities, income stability, and time horizon.

A client's overall risk tolerance depends on both his ability to take risk and his willingness to take risk. A willingness greater than ability, or vice versa, is typically resolved by choosing the more conservative of the two and counseling the client.

LOS 64.e

Investment constraints include:

- Liquidity—The need to draw cash from the portfolio for anticipated or unexpected future spending needs. High liquidity needs often translate to a high portfolio allocation to bonds or cash.
- Time horizon—Often the period over which assets are accumulated and before withdrawals begin. Risky or illiquid investments may be inappropriate for an investor with a short time horizon.
- Tax considerations—Concerns the tax treatments of the investor's various accounts, the relative tax treatment of capital gains and income, and the investor's marginal tax bracket.
- Legal and regulatory—Constraints such as government restrictions on portfolio contents or laws against insider trading.
- Unique circumstances—Restrictions due to investor preferences (religious, ethical, etc.) or other factors not already considered.

LOS 64.f

An asset class is a group of securities with similar risk and performance characteristics. Examples of major asset classes include equity, fixed income, cash, and real estate. Portfolio managers also use more narrowly defined asset classes, such as large-cap U.S. equities or speculative international bonds, and alternative asset classes, such as commodities or investments in hedge funds.

LOS 64.g

Strategic asset allocation is a set of percentage allocations to various asset classes that is designed to meet the investor's objectives. The strategic asset allocation is developed by combining the objectives and constraints in the IPS with the performance expectations of the various asset classes. The strategic asset allocation provides the basic structure of a portfolio.

Tactical asset allocation refers to an allocation that deviates from the baseline (strategic) allocation in order to profit from a forecast of shorter-term opportunities in specific asset classes.

LOS 64.h

Imposing portfolio constraints based on ESG factors may affect performance. Limiting the universe of investment choices may decrease returns, but good corporate governance and low

ESG-related risks may increase returns.

If a portfolio's investment universe is constrained by negative screening, its performance should be measured against a benchmark that excludes companies with negative ESG factors.

Positive screening, best-in-class, or thematic investing approaches typically require portfolio construction to be customized for the investor's choices of which ESG factors to focus on.

For active ownership, it is important to clarify whether investors intend to vote their shares themselves or direct managers to vote the shares according to specified ESG factors.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 64.1

1. **A** An investment policy statement is considered to be the starting point of the portfolio management process. The IPS is a plan for achieving investment success. (LOS 64.a)
2. **A** Investment objectives are defined based on both the investor's return requirements and risk tolerance. Investment constraints include the investor's time horizon, liquidity needs, tax considerations, legal and regulatory requirements, and unique needs and preferences. Policies regarding permitted asset types and the amount of leverage to use are best characterized as investment guidelines. (LOS 64.b)
3. **C** Both risk and return objectives can be defined either in absolute terms or relative to some benchmark. (LOS 64.c)
4. **C** When assigning an overall risk tolerance, the prudent approach is to use the lower of ability to take risk and willingness to take risk. (LOS 64.d)
5. **C** Return objectives are part of a policy statement's objectives, not constraints. (LOS 64.e)
6. **B** Asset classes should be defined such that correlations of returns within the asset class are relatively high (because assets within a class should perform alike over time), while correlations of returns among asset classes are relatively low (to benefit from diversification). (LOS 64.f)
7. **B** An adviser's forecasts of the expected returns and expected volatilities (risk) of different asset classes are an important part of determining an appropriate asset allocation. (LOS 64.g)

READING 65

THE BEHAVIORAL BIASES OF INDIVIDUALS

EXAM FOCUS

Behavioral finance investigates ways in which human behavior differs from the rationality assumed by traditional economic models. Some believe these irrational behaviors or biases lead to predictable deviations of financial markets from the implications of financial models of security price behavior. Here, candidates must learn some of the terminology of behavioral finance and examples that support a belief in irrational investor behavior. Additionally, we cover how investor biases and irrationality may explain certain anomalous results of tests of market efficiency.

MODULE 65.1: COGNITIVE ERRORS VS. EMOTIONAL BIASES



Video covering this content is available online.

Traditional finance assumes that individuals act as perfectly rational economic beings who objectively consider all relevant information to make rational decisions, and that this process results in efficient markets. Research results have called these assumptions into question. Kahneman and Tversky pioneered this work in the early 1970s, initially setting logic tests where individuals' intuitive answers were predictably flawed. This supported the idea that human decision-making has systematic biases that lead to irrational decisions.

This work was extended by Kahneman, Tversky, and others, who have suggested that a better understanding of these biases by clients (and by the professionals who work with those clients) should produce securities prices and returns over time that better match the informationally efficient markets of traditional finance theory.

LOS 65.a: Compare and contrast cognitive errors and emotional biases.

Individuals, when facing complex decision-making, often lack the time or ability to derive the optimal course of action prescribed by traditional finance. Cognitive limitations and emotional responses introduce bias into the decision-making process, leading to decisions that are biased (not perfectly rational). Behavioral finance asserts that certain biases, which are not simply random errors, are widespread and therefore predictable.

Cognitive errors are due primarily to faulty reasoning or irrationality. They can arise from not understanding statistical analysis, information processing errors, illogical reasoning, or memory errors. Such errors can possibly be reduced by increased awareness, better training, or more information.

Emotional biases are not related to conscious thought. Rather, they stem from feelings, impulses, or intuition. As such, they are difficult to overcome and may have to be accommodated.

Despite the distinction in grouping biases as either cognitive or emotional, a bias may have elements of both cognition and emotion. When trying to overcome or mitigate biases that are both emotional and cognitive, success is more likely by focusing on the cognitive issues.

LOS 65.b: Discuss commonly recognized behavioral biases and their implications for financial decision making.

Cognitive errors can be divided into *belief perseverance* biases that reflect an irrational reluctance to change prior conclusions and decisions, and *processing errors* where the information analysis is flawed.

Cognitive Errors: Belief Perseverance

Psychologists use the term **cognitive dissonance** to refer to a situation where an individual holds conflicting beliefs or receives information that causes a current belief to be questioned. Cognitive dissonance causes stress that individuals seek to reduce. They may do so by letting go of prior beliefs in favor of the conflicting belief. On the other hand, they might discount the conflicting information or viewpoints by questioning their truth, source, applicability, or significance. To the extent that it is easier to do the latter than the former, bias in favor of currently held beliefs is the result.

- 1. Conservatism bias** occurs when market participants **rationally form an initial view but then fail to change that view as new information becomes available**. That is, they overweight their prior probabilities and do not adjust them appropriately as new information becomes available. Individuals displaying this bias tend to maintain prior forecasts and securities allocations, ignoring or failing to recognize the significance of new information. Individuals may react slowly to new data or ignore information that is complex to process.

EXAMPLE: Conservatism bias

John Molinari allocates assets based on his observation that over the last 80 years, recessions occurred in 20% of those years. When a coworker informs Molinari that the country's central bank has announced a policy change to a tightening of monetary conditions, Molinari does not adjust his recommended asset allocations. Does this reflect conservatism bias?

Answer:

Molinari should consider that the conditional probability of a recession, given that the central bank is tightening, may differ from the unconditional probability of a recession that

he previously estimated. He is showing conservatism bias by not considering the impact of this new information.

Conservatism bias may result in market participants holding investments too long because they are unwilling or slow to update a view or forecast. They may be avoiding the mental effort or stress of updating prior beliefs by not considering the implications of new information.

2. **Confirmation bias** occurs when market participants focus on or seek information that supports prior beliefs, while avoiding or diminishing the importance of conflicting information or viewpoints. They may distort new information in a way that remains consistent with their prior beliefs.

For example, after buying a car from a given manufacturer, the buyer would exhibit confirmation bias by reading articles about how great cars from that manufacturer are but avoiding news about problems with that particular brand. Because the buyer already purchased the car, information suggesting that it was a bad decision is unwelcome.

Consequences and implications of confirmation bias may include market participants who:

- Consider positive information but ignore negative information.
- Set up a decision process or data screen incorrectly to support a preferred belief.
- Become overconfident about the correctness of a presently held belief.

Market participants can reduce confirmation bias by seeking out contrary views and information—for example, by reading analyses and viewpoints that disagree with a presently held belief, rather than only reading those that reinforce the belief.

3. **Representativeness bias** occurs when certain characteristics are used to put an investment in a category and the individual concludes that it will have the characteristics of investments in that category. Individuals systematically make the error of believing that two things that are similar in some respects are more similar in other respects than they actually are.

Two forms of representativeness bias are base-rate neglect and sample-size neglect.

Base-rate neglect refers to analyzing an individual member of a population without adequately considering the probability of a characteristic in that population (the base rate). Consider this example of base-rate neglect: a group was asked to identify the most likely occupation of a man who was characterized as somewhat shy as a salesperson or a librarian. Most participants chose librarian, thinking that most librarians would tend to be more shy on average than salespeople, who tend to be outgoing. Their mistake was in not considering that there are relatively few male librarians and a great number of male salespeople. Even though a greater percentage of librarians may be characterized as somewhat shy, the absolute number of salespeople who could be characterized as somewhat shy is significantly greater.

Sample-size neglect refers to making a classification based on a small and potentially unrealistic data sample. The error is believing the population reflects the characteristics of the small sample.

For example, a fund manager may show strong performance over a three-year time horizon. This may lead investors to assume this is evidence of superior skill. However, examination of longer sample periods shows a lack of persistence; there will be some three-year “winners” even when annual returns results are actually random. The evidence of the lack of persistence of managers’ relative performance over time does not support investors’ conclusion that this manager will continue to outperform.

EXAMPLE: Representativeness bias

XYZ company has long been recognized as a growth stock, delivering superior earnings growth and stock price appreciation. While earnings have continued to grow, last year’s revenue has not, and neither has the stock price. Under the following two conditions, would an analyst be more likely to buy or sell the stock?

1. The analyst suffers from base-rate and sample-size neglect.
2. The analyst treats the growth classification as representative.

Answer:

If the analyst exhibits sample-size neglect and base-rate neglect biases, the analyst will ignore XYZ’s long record as a growth stock, focus on the short-term disappointing results, and may recommend selling the stock without considering the long-term possibility it will revert to growth behavior.

However, if the analyst over-relies on the initial growth classification, the analyst may assume that the stock will return to growth and recommend buying it, without properly considering the reasons for its recent results or their longer-term implications.

Representativeness bias may lead market participants to attach too much importance to a few characteristics based on a small sample size or make decisions based on simple rules and classifications rather than conducting a more-thorough and complex analysis.

4. **Illusion of control bias** exists when **market participants believe they can control or affect outcomes when they cannot**. It is often associated with emotional biases: illusion of knowledge (belief you know things you do not know), self-attribution (belief you personally caused something to happen), and overconfidence (an unwarranted belief that your beliefs will prove to be correct).

An example of control bias can be seen in humans’ attempts to control the weather using ritual ceremonies. This allowed them to believe they had some control of the environment, when in fact it is highly unlikely that a ceremony can influence the weather.

Illusion of control may cause market participants to overweight securities for which they believe they have control over outcomes, such as a company they work for or are otherwise associated with. This can result in their portfolios being inadequately diversified.

5. **Hindsight bias** is a **selective memory of past events, actions, or what was knowable in the past, resulting in an individual’s tendency to see things as more predictable than they really are**. People tend to remember their correct predictions and forget their incorrect ones. They also overestimate what could have been known. This behavior results from

individuals being able to observe outcomes that did occur but not the outcomes that did not materialize. Hindsight bias is sometimes referred to as the *I-knew-it-all-along phenomenon*.

Hindsight bias is caused by three types of errors:

1. Individuals distort their earlier predictions when looking back. This is the tendency to believe that we knew the outcome of an uncertain event all along.
2. Individuals tend to view events that have occurred as inevitable.
3. Individuals assume they could have foreseen the outcomes of uncertain events.

Hindsight bias can lead to overconfidence in ability to predict outcomes. It may also cause investors to cast aside valid analysis techniques that did not turn out to be correct in favor of poor techniques that turned out well by chance.

Cognitive Errors: Information-Processing Biases

These are related more to the processing of information and less to the decision-making process.

Anchoring and adjustment bias refers to basing expectations on a prior number and overweighting its importance, making adjustments in relation to that number as new information arrives. Examples would be estimating the value of a security relative to its current value or making estimates of earnings per share relative to a previously reported value or relative to a prior estimate. Anchoring leads to underestimating the implications of new information. New data should be considered objectively without regard to any initial anchor point.

Mental accounting bias refers to viewing money in different accounts or from different sources differently when making investment decisions. This conflicts with the idea that security decisions should be made in the context of the investor's overall portfolio of assets based on their financial goals and risk tolerance.

An example of mental accounting is an investor who receives an unexpected bonus at work and chooses to invest it in a very risky biotechnology stock, reasoning that the bonus is "found money" that can acceptably be risked on speculation. In fact, while such a stock may have a place in the investor's portfolio, decisions about whether and how much of it to include should be based on a total portfolio approach. Another example may be an investor who receives an inheritance from a parent and segregates those funds into low-risk bonds out of a desire to "not lose any of the money the parent worked so hard to save." This behavior also conflicts with the principle that investments should be considered in the context of the entire portfolio.

The result of such mental accounting may be that an investor's overall portfolio is not optimal given the investor's circumstances, investment goals, and risk tolerance. It can cause an investor to hold positions that offset each other, rather than considering investments in the context of their correlation of returns. One common form of mental accounting bias is a tendency to view income differently from capital appreciation. This may cause an investor to hold a mix of income-producing and non-income-producing securities that does not match the investor's circumstances.

A study by Camerer et al. (1997)¹ investigated behavior of New York taxi drivers. The study suggests that each new day was a separate account in the minds of the taxi drivers. It appears that New York taxi drivers have a reference rate for daily income and perceive a loss if they fail to meet it (which we may also view as an example of anchoring). On rainy days, demand for taxis is high, while on sunny days, the demand is low. Logically, you would expect taxi drivers to work more hours on rainy days to maximize their incomes. In reality, the opposite was true. Taxi drivers worked longer hours on sunny days as they strove to hit their target income. On rainy days, once the target was achieved, the taxi drivers stopped working.

Framing bias occurs when decisions are affected by the way in which the question or data is “framed.” In other words, the way a question is phrased can influence the answer given.

Tversky and Kahneman (1980)² illustrate framing bias with the following example.

EXAMPLE: Framing bias (framing as a gain)

The United States is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs have been proposed. If Program A is adopted, 200 people will be saved. If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no one will be saved. Which program will people choose?

Answer:

Program A is typically selected. Although the expected value of both Program A and Program B is 200 lives saved, the majority choice is risk averse. The prospect of saving 200 lives with certainty is more attractive than the risky option with the same expected value.

EXAMPLE: Framing bias (framing as a loss)

A different group of individuals is given the same issue, but the two programs are framed differently. If Program A is adopted, 400 people will die. If Program B is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 will die. Which program will people choose?

Answer:

In this situation, Program B is typically selected. The majority choice is now risk-taking, with the certain death of 400 people being less acceptable than a two-thirds chance that 600 people will die.

In these examples, the two programs presented (A and B) are identical for both groups, but the framing of the information results in different choices. For the first group, the information was presented in the context of a gain, while for the second group, it was in the context of a loss. Loss aversion—the idea that we fear losses more than we value gains—is then demonstrated by the choices made.

An example of framing bias in an investment context is overestimating the significance of short-term price volatility (risk) without weighing it against long-term considerations. This may result in overly conservative portfolios.

Investment managers must take care to avoid framing bias when creating questions to assess an investor's risk tolerance. Failing to properly assess risk tolerance may identify investors as more or less risk averse than they actually are, resulting in portfolios that are inconsistent with the investors' needs.

Availability bias refers to putting **undue emphasis on information that is readily available, easy to recall, or based narrowly on personal experience or knowledge**. Availability bias occurs when individuals judge the probability of an event occurring by the ease with which examples and instances come to mind. By the very nature of memory, more-recent events are typically easier to recall than events further in the past, which leads to the bias of attaching too much significance to events that have occurred recently and too little to events that occurred further in the past. People also tend to assume that if something is easily remembered, it must occur with a higher probability.

EXAMPLE: Availability bias

Imagine a word is picked at random from a dictionary. Is it more likely that the word has the letter *r* as the first letter in the word or the third letter?

Answer:

When faced with this problem, most individuals state that it is more likely the letter *r* will be the first letter. In fact, in the English language, there are approximately three times more words with *r* as the third letter than the first. Individuals find it easier to recall words that begin with *r* than words with *r* as the third letter, which distorts their estimation of probabilities.

Availability bias may lead market participants to choose a manager or investment based on advertising or recalling they have heard the name. They may limit their universe of potential investments to familiar firms, resulting in inappropriate asset allocations and lack of diversification. They may also overreact to recent market conditions while ignoring data on historical market performance, or they may place too much emphasis on events that receive a large amount of media attention.

MODULE 65.2: EMOTIONAL BIASES



PROFESSOR'S NOTE

Video covering this content is available online.

Some of the terms we are about to discuss have already come up in the discussion of cognitive biases. In general, if an investor's view is based on unconscious emotion that the holder is unwilling or unable to change, we should regard it as an emotional bias. If a bias can be overcome with a relatively simple change in thought process or information, we should regard it as a cognitive bias.

While there is no formally accepted definition, these six biases generally arise from emotion and feelings rather than through conscious thought:

1. **Loss-aversion bias** arises from **feeling more pain from a loss than pleasure from an equal gain**. Kahneman and Tversky (1979)³ investigated differences between how people feel

when they gain and when they lose and how that affects behavior when faced with risk. They found that individuals' willingness to take a gamble (risk) was very different when facing a loss or a gain.

Consider the following two scenarios:

Scenario 1: An individual is given \$10. The individual is then given the following options:

- Take an additional \$5 with certainty.
- Flip a coin and win an additional \$10 if it lands heads up or nothing if it lands tails up.

Both options represent a gain relative to the original \$10, and the expected value of the gain is \$5 for either option. Option 1 creates a guaranteed outcome of \$15. Option 2 introduces uncertainty, with equal probabilities of an outcome of \$10 or \$20. Most individuals chose the riskless Option 1 over the riskier Option 2.

Scenario 2: An individual is given \$20. The individual is then given the following options:

- Take a \$5 loss with certainty.
- Flip a coin and lose nothing if it lands heads up, but lose \$10 if it lands tails up.

Both options represent a potential loss relative to the original \$20, and the expected loss is \$5 for either option. Most individuals chose risky Option 2 over the riskless certain loss of Option 1.

In both scenarios, the expected value of the individual's wealth is \$15. The options given to the individuals are also identical, with Option 1 resulting in a guaranteed outcome of \$15 and Option 2 providing equal probabilities of either \$10 or \$20. Yet, when faced with gains, people preferred certainty, and when faced with losses, they preferred risk.

The conclusion is that individuals display asymmetrical responses to gains and losses. Kahneman and Tversky suggested that people look at decisions relative to a reference rate. Anything below the reference rate is seen as a loss, and anything above the reference rate is seen as a gain. The reference rate in scenario 1 was the \$10 initially given to the individual, and in scenario 2, it was \$20. The conflicting responses to the scenarios are explained by attitudes toward gains and losses. Crucially, people fear losses far more than they value gains. Thus, in scenario 2, they were willing to take the risk in the hope of avoiding a loss.



PROFESSOR'S NOTE

Be sure to understand the difference between risk aversion and loss aversion. A risk-averse investor is simply an investor who, given two investments with the same expected returns, would select the investment with the lowest risk. A loss-averse investor is one who feels greater pain (decreases in utility) from losses than satisfaction (increase in utility) from gains. As a result, the individual is more likely to take a risk in the hope of avoiding losses than in the hope of achieving gains.

Consequences of loss-aversion bias may include trading too much by selling for small gains, which increases transaction costs and decreases returns, or incurring too much risk by continuing to hold assets that have deteriorated in quality and lost value. If an initial decline in value occurs, loss-averse investors may take excessive risk in the hope of recovering (investment managers may be particularly susceptible to this behavior). A loss-averse investor might view a position inappropriately as a gain or a loss based on the framing of the reference point.

2. **Overconfidence bias** occurs when **market participants overestimate their own intuitive ability or reasoning**. It can show up as illusion of knowledge when they think they do a better job of predicting than they actually do. Combined with self-attribution bias, individuals may give themselves personal credit when things go right (self-enhancing) but blame others or circumstances when things go wrong (self-protecting). Prediction overconfidence leads individuals to underestimate uncertainty and the standard deviation of their predictions, while certainty overconfidence occurs when they overstate the probability they will be right.

While overconfidence is both cognitive and emotional, it is more emotional in nature because it is difficult for most individuals to correct and is rooted in the desire to feel good. Overconfidence bias may cause market participants to underestimate risk, overestimate return, and fail to diversify sufficiently.

3. **Self-control bias** occurs when **individuals lack self-discipline and favor short-term satisfaction over long-term goals**. Often, individuals are not prepared to make short-term sacrifices to meet their long-term goals. They may favor small payoffs now at the expense of larger payoffs in the future, which is known as hyperbolic discounting.

Self-control bias may result in insufficient savings to fund retirement needs, which in turn may cause an investor to take excessive risk to try to compensate for insufficient savings accumulation. It may also result in overemphasis on income-producing assets to meet short-term needs.

Self-control bias might be mitigated by establishing an appropriate investment plan (asset allocation) and a budget to achieve sufficient savings. Both should be reviewed on a regular basis.

4. **Status quo bias** occurs when comfort with an existing situation causes an individual to be resistant to change. If investment choices include the option to maintain existing investments or allocations, or if a choice will happen unless the participant opts out, status quo choices become more likely.

Companies have found that automatically enrolling workers in retirement savings schemes, with the option to opt out, increases participation compared with making it necessary for employees to opt in. Thaler and Sunstein (2008)⁴ argue for framing choices in this way to achieve better participation rates in retirement plans, as well as other choices, such as whether to register as an organ donor.

Consequences of status quo bias may include holding portfolios with inappropriate risk and not considering other, better investment alternatives.

5. **Endowment bias** occurs when an asset is felt to be special and more valuable simply because it is already owned. For example, a spouse may hold on to securities the deceased spouse purchased, for reasons like sentiment that are unrelated to the current merits of the securities. Endowment bias is common with inherited assets and might be detected or mitigated by asking a question such as “Would you make this same investment with new money today?”

In studies, individuals have been asked to state their minimum sales price for an asset they own (say \$25) and the maximum price they would pay now to buy the same asset (say \$23). The price at which they would be willing to sell tended to be higher than the price they would pay. This has been explained as an endowment effect. Once they own an asset, they act as if it is worth more than they would pay.

Market participants who exhibit endowment bias may be failing to sell assets that are no longer appropriate for their investment needs, or they hold assets with which they are familiar because they provide some intangible sense of comfort.

6. **Regret-aversion bias** occurs when market participants **do nothing out of excessive fear that actions could be wrong**. They attach undue weight to errors of commission (doing something that turns out wrong) and not enough weight to errors of omission (not doing something that would have turned out right). Their sense of regret and pain is stronger for acts of commission. This is quite similar to status quo bias.

Herding behavior is a form of regret aversion where participants go with the consensus or popular opinion. Essentially, participants tell themselves they are not to blame if others are wrong too.

Consequences of regret-aversion bias may include excess conservatism in the portfolio because it is easy to see that riskier assets do at times underperform. Therefore, an investor might not buy riskier assets so as not to experience regret when they decline.

LOS 65.c: Describe how behavioral biases of investors can lead to market characteristics that may not be explained by traditional finance.

In our Equity Investments reading on Market Efficiency, we developed the ideas of anomalies (results that do not fit the prevailing model of securities risks and returns) and market inefficiencies (anomalies that present opportunities to earn positive risk-adjusted returns). Many market anomalies have been explained by small sample size, time period bias, or inadequacies in the specification of prevailing models of returns. In the same way, some anomalies once considered evidence of market inefficiency have been explained by the possible misspecification of risk, leading to inaccurate risk adjustment of returns.

Bubbles and subsequent crashes have been with us as long as trading markets have existed. Some have offered explanations for these extremes of valuation based on rational behavior. Investors who leave markets that they determine are exhibiting characteristics of a bubble may miss out on high returns over extended periods. On the other hand, investors who stay invested based on a belief that they will be able to get out at or near the market top often find that that is

not realistic. Either type of mistake, leaving the market too early or staying fully invested too long, can be quite damaging to a fund manager's career. Exploiting the fact that a market is "in a bubble" for an arbitrage profit may not be possible due to the restrictions on and risk of short selling in a rapidly rising market.

While behavioral finance does not provide an overall explanation for bubbles and their aftermath, some have claimed that cognitive and emotional biases exhibited during such periods may have some causal effects. These claims include the following:

- *Overconfidence* may lead to overtrading, underestimation of risk, and lack of diversification.
- Persistently good results combined with *self-attribution* bias can fuel overconfidence, as can *hindsight* bias (as investors give themselves credit for choosing profitable stocks in a bull market).
- *Confirmation bias* may lead investors to ignore or misinterpret new information suggesting that valuations will not continue to rise, or to misinterpret initial decreases in asset values as simply another buying opportunity.
- *Anchoring* may cause investors to believe recent highs are rational prices even after prices begin their eventual decline.
- Fear of *regret* may keep even very skeptical investors in the market.

One anomaly that has been persistent over time is the value/growth anomaly. Value stocks (low market-book, low P/E, high dividend yield stocks) have outperformed growth stocks (high P/E, low dividend yield stocks) over long periods. Fama and French (1992) found that adding risk factors for firm size and book-to-market ratio to a model of stock returns eliminated the evidence of value stock outperformance. These results suggest that the extra returns to value stocks were compensation for additional risks captured by firm size and book-to-market ratios.

Others have suggested behavioral factors as the cause of the seeming outperformance of value stocks. The **halo effect** is a version of representativeness in which a company's good characteristics, such as fast growth and a rising stock price, are extended into a conclusion that it is a good stock to own, leading to overvaluation of growth stocks.

The fact that investors tend to invest heavily in firms in their domestic country in a global portfolio, or more heavily in firms operating in their region of a country, is considered anomalous in that rationality suggests greater diversification. Such **home bias**, it is claimed, may result from a belief that they have better access to information or simply an emotional desire to invest in companies "closer to home." Similarly, investors may underestimate the risk or overestimate the future returns of firms whose products they use or firms for which they are exposed to a great amount of positive marketing messages.



MODULE QUIZ 65.1, 65.2

1. Which of the following would *most likely* be classified as an emotional bias?
 - A. The investor has difficulty interpreting complex new information.
 - B. The investor only partially adjusts forecasts when he receives new information.
 - C. The investor has a tendency to value the same assets higher if he owns them than if he does not own them.
2. Which of the following would *most likely* indicate that an investor is subject to an emotional bias?
 - A. Regularly basing decisions on only a subset of available information.

- B. Reacting spontaneously to a negative earnings announcement by quickly selling a stock.
- C. Remaining invested in a profitable technology stock even though new information indicates its P/E ratio is too high.
3. A cognitive error is *most likely* indicated by which of the following?
- A. A client is the chief executive officer of a public company that she founded and insists she will not diversify her holding of the company stock.
- B. The spouse of a now-deceased company founder becomes upset when it is recommended the portfolio holdings in that company need to be diversified.
- C. A client who initially resists recommendations to diversify the portfolio later thanks the manager for explaining the benefits of diversification.
4. Abby Lane has investments scattered across many different accounts, from bank savings to before- and after-tax retirement accounts to taxable nonretirement accounts. She has multiple investing goals ranging from important short-term goals to longer-term “wish list” goals. She looks at her financial assets and views each holding as designed to meet specific goals. Lane has been very successful in her investment decisions for several decades and believes she can continue to achieve reasonable results. Lane *most likely* exhibits:
- A. framing bias.
- B. mental accounting.
- C. overconfidence bias.
5. Twenty years ago, Jane Ivy set up her initial asset allocation in her defined contribution plan by placing an equal amount in each asset class and never changed it. Over time, she increased her contribution by 1% per year until she reached the maximum amount allowed by law. Due to her steadfastness and good fortune, coupled with matching funds from her employer, she now finds herself in her early 40s with a million-dollar retirement account. Which of the following biases does Ivy *most likely* exhibit?
- A. Representativeness.
- B. Status quo bias.
- C. Availability bias.
6. The halo effect suggests that investors tend to overvalue stocks:
- A. from their own country or region.
- B. with which the investors are most familiar.
- C. that have experienced rapid growth and price appreciation.

KEY CONCEPTS

LOS 65.a

Cognitive errors result from the inability to analyze information or from basing decisions on partial information. Individuals try to process information and make rational decisions, but they may lack the capacity or sufficient information to do so. Cognitive errors can be divided into belief perseverance errors and processing errors.

Emotional biases are caused by the way individuals frame information and decisions, rather than the process used to analyze and interpret information. Emotional bias is more of a spontaneous reaction.

LOS 65.b

Cognitive Errors: Belief Perseverance

- Conservatism bias.
- Confirmation bias.
- Representativeness bias.

- Control bias.
- Hindsight bias.

Cognitive Errors: Information Processing

- Anchoring and adjustment.
- Mental accounting bias.
- Framing bias.
- Availability bias.

Emotional Biases

- Loss-aversion bias.
- Overconfidence bias.
- Self-control bias.
- Status quo bias.
- Endowment bias.
- Regret-aversion bias.

LOS 65.c

Many reported anomalies have been explained by inadequate specification of security risk, but some have attempted to explain the existence of persistent anomalies as the results of cognitive errors and emotional biases.

Behavioral finance has not explained bubbles and crashes, but some cognitive errors and emotional biases may be exhibited during bubbles and crashes.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 65.1, 65.2

- C** This describes the *endowment bias*, where individuals place a higher value on assets they own than if they did not own those same assets. The other two answer choices describe cognitive errors that are due to the inability to analyze all the information. (Modules 65.1, 65.2, LOS 65.a, 65.b)
- B** Emotional biases tend to elicit more of a spontaneous reaction than cognitive errors. Making a decision based only on partial information is indicative of a cognitive error. Ignoring a high P/E ratio could be indicative of the conservatism bias, which is reacting slowly to new information or avoiding analyzing new information. It could also indicate confirmation bias, where the investor focuses on positive information and ignores negative information. Both conservatism and confirmation biases are cognitive errors. (Modules 65.1, 65.2, LOS 65.a, 65.b)
- C** Individuals making cognitive errors are more likely to respond rationally when new information is provided. The client initially resists a rational recommendation but then reverses their thoughts when given more information.

There are rational reasons a CEO may want to hold a large block of her company's stock. Those include legal restrictions on sale or a desire to take concentrated risk in a situation where she has a lot of control. A rational decision is not an error. Alternatively, the "insists" could indicate an emotional bias. Neither interpretation suggests a cognitive error.

The spouse who becomes upset at a rational recommendation to diversify is likely showing an emotional bias. (Module 65.1, LOS 65.a)

4. **B** Viewing each asset in light of meeting a specific goal is mental accounting. There was no indication of framing (the way data is provided overly affects the decision process). An investor with decades-long success who expects to produce reasonable results is acting rationally and is not necessarily overconfident. (Module 65.2, LOS 65.b)
5. **B** Ivy is exhibiting *status quo bias*, where investors leave their asset allocation alone and don't change it according to changing market conditions or changes in their own circumstances. Her actions do not suggest representativeness (placing something in a category and assuming it will have the characteristics associated with that category) or availability (putting undue emphasis on information readily available or easily recalled). (Module 65.2, LOS 65.b)
6. **C** The halo effect suggests investors will view a stock that has experienced rapid growth and price appreciation as a good stock to own, which may result in these stocks being overvalued. Home bias is the tendency for investors to favor stocks from their own country or region because they are more familiar with those stocks. (Module 65.2, LOS 65.c)

¹ Colin Camerer, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor Supply of New York City Cab Drivers: One Day at a Time," *Quarterly Journal of Economics* 112 (1997): 407–442.

² Amos Tversky and Daniel Kahneman, "The Framing of Decisions and the Rationality of Choice," Stanford University Department of Psychology, Defense Technical Information Center – Technical Report (1980).

³ Daniel Kahneman and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica* 47, no. 2 (1979): 263–91.

⁴ Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (New York: Penguin Group, 2008).

READING 66

INTRODUCTION TO RISK MANAGEMENT

EXAM FOCUS

Here we present a framework for risk management that is broad enough to be applied to corporations in general, financial firms, and individuals, as well as to the management of securities portfolios in any context. The main idea is that organizations should estimate the various risks they face and then reduce some risks and accept or increase other risks. The result should be a bundle of risks that simultaneously matches the risk tolerance of the organization and provides the greatest benefits in terms of reaching the organization's goals. Note that risk is not minimized though this process. The concept of risk budgeting, the categorization of types of risks, and the various methods of risk mitigation all offer testable material.

MODULE 66.1: INTRODUCTION TO RISK MANAGEMENT



Video covering
this content is
available online.

LOS 66.a: Define risk management.

The **risk management** process seeks to 1) identify the risk tolerance of the organization, 2) identify and measure the risks that the organization faces, and 3) modify and monitor these risks.

The process does not seek to minimize or eliminate all of these risks. The organization may increase its exposure to risks it decides to take because it is better able to manage and respond to them. The organization may decrease its exposure to risks that it is less well able to manage and respond to by making organizational changes, purchasing insurance, or entering into hedging transactions. Through these choices the firm aligns the risks it takes with its risk tolerances for these various types of risk.

Risk (uncertainty) is not something to be avoided by an organization or in an investment portfolio. Returns above the risk-free rate are earned by taking on risk. While returns for any period are not under the control of managers, the specific risks and overall level of risk the organization takes are under their control. We can think of risk management as determining organizational risks, determining the optimal bundle of risks for the organization, and implementing risk mitigation strategies to achieve that bundle of risks.

We describe the principles of risk management here in a framework that can be applied broadly, not only to firms or organizations in general, but also to the management of investment portfolios and financial firms, and even to individuals deciding how much risk and which

specific risks they will take. Individuals follow a similar approach, selecting a bundle of risks that is optimal for maximizing their expected utility (rather than returns or profit).

LOS 66.b: Describe features of a risk management framework.

An overall **risk management framework** encompasses several activities, including:

- Establishing processes and policies for risk governance.
- Determining the organization's risk tolerance.
- Identifying and measuring existing risks.
- Managing and mitigating risks to achieve the optimal bundle of risks.
- Monitoring risk exposures over time.
- Communicating across the organization.
- Performing strategic risk analysis.

This framework is general, but all of these elements should be addressed in any comprehensive risk management framework. Only by understanding the risks the organization faces, and having the processes and procedures in place to effectively manage and monitor these risks, can an organization align its risk exposures to the goals of the organization.

LOS 66.c: Define risk governance and describe elements of effective risk governance.

Risk governance refers to senior management's determination of the risk tolerance of the organization, the elements of its optimal risk exposure strategy, and the framework for oversight of the risk management function. Risk governance seeks to manage risk in a way that supports the overall goals of the organization so it can achieve the best business outcome consistent with the organization's overall risk tolerance. Risk governance provides organization-wide guidance on the risks that should be pursued in an efficient manner, risks that should be subject to limits, and risks that should be reduced or avoided.

A risk management committee can provide a way for various parts of the organization to bring up issues of risk measurement, integration of risks, and the best ways to mitigate undesirable risks.

LOS 66.d: Explain how risk tolerance affects risk management.

Determining an organization's **risk tolerance** involves setting the overall risk exposure the organization will take by identifying the risks the firm can effectively take and the risks that the organization should reduce or avoid. Some of the factors that determine an organization's risk tolerance are its expertise in its lines of business, its skill at responding to negative outside events, its regulatory environment, and its financial strength and ability to withstand losses.

When analyzing risk tolerance, management should examine risks that may exist within the organization as well as those that may arise from outside. The various risks the firm is exposed

to must each be considered and weighted against the expected benefits of bearing those risks and how these fit the overall goals of the organization.

LOS 66.e: Describe risk budgeting and its role in risk governance.

Risk budgeting is the process of allocating firm resources to assets (or investments) by considering their various risk characteristics and how they combine to meet the organization's risk tolerance. The goal is to allocate the overall amount of acceptable risk to the mix of assets or investments that have the greatest expected returns over time.

The risk budget may be a single metric, such as portfolio beta, value at risk, portfolio duration, or returns variance. A risk budget may be constructed based on categories of investments, such as domestic equities, domestic debt securities, international equities, and international debt securities. Another way to allocate a risk budget is to identify specific risk factors that comprise the overall risk of the portfolio or organization. In this case, specific risk factors that affect asset classes to varying degrees, such as interest rate risk, equity market risk, and foreign exchange rate risk, are estimated and aggregated to determine whether they match the overall risk tolerance of the organization.

LOS 66.f: Identify financial and non-financial sources of risk and describe how they may interact.

Financial risks are those that arise from exposure to financial markets. Examples are:

- **Credit risk.** This is the uncertainty about whether the counterparty to a transaction will fulfill its contractual obligations.
- **Liquidity risk.** This is the risk of loss when selling an asset at a time when market conditions make the sales price less than the underlying fair value of the asset.
- **Market risk.** This is the uncertainty about market prices of assets (stocks, commodities, and currencies) and interest rates.

Non-financial risks arise from the operations of the organization and from sources external to the organization. Examples are:

- **Operational risk.** This is the risk that human error, faulty organizational processes, inadequate security, or business interruptions will result in losses. An example of an operational risk is **cyber risk**, which refers to disruptions of an organization's information technology.
- **Solvency risk.** This is the risk that the organization will be unable to continue to operate because it has run out of cash.
- **Regulatory risk.** This is the risk that the regulatory environment will change, imposing costs on the firm or restricting its activities.
- **Governmental or political risk** (including **tax risk**). This is the risk that political actions outside a specific regulatory framework, such as increases in tax rates, will impose significant costs on an organization.

- **Legal risk.** This is the uncertainty about the organization's exposure to future legal action.
- **Model risk.** This is the risk that asset valuations based on the organization's analytical models are incorrect.
- **Tail risk.** This is the risk that extreme events (those in the tails of the distribution of outcomes) are more likely than the organization's analysis indicates, especially from incorrectly concluding that the distribution of outcomes is normal.
- **Accounting risk.** This is the risk that the organization's accounting policies and estimates are judged to be incorrect.

For individuals, risks, such as risk of death (**mortality risk**) prior to providing for their families' future needs and the risk of living longer than anticipated (**longevity risk**) so that assets run out, are very important in financial planning. Mortality risk is most often addressed with life insurance, and longevity risk can be reduced by purchasing a lifetime annuity. Risk of health care expenses is addressed with health insurance. Although the risks for an individual are in some ways different from those of organizations, the overall approach is the same, choosing which risks to bear (self-insure), which risks to prevent or avoid, and which risks to take in order to maximize the expected outcome in terms of personal utility or satisfaction.

The various risks an organization faces are not independent; they interact in many ways. Consider a firm with market risk that it reduces with option contracts. If markets decline significantly, the firm is owed a payment from the firm on the other side of the option trade, so now there is significant counterparty or credit risk. There also may be legal risks if the counterparty seeks to avoid the payment through loopholes in the contract. Credit losses and legal losses may result in greater liquidity risk as positions must be sold. Additional losses from selling in a declining or less liquid market may increase solvency risk because of the negative impact on the firm's cash position.

Interactions among risks must be considered because such interactions are many and frequent. They can be especially important during periods of stress in financial markets, when risk management is most important to the health and possibly the survival of the organization.

LOS 66.g: Describe methods for measuring and modifying risk exposures and factors to consider in choosing among the methods.

Measures of risk for specific asset types include standard deviation, beta, and duration.

- **Standard deviation** is a measure of the volatility of asset prices and interest rates. Standard deviation may not be the appropriate measure of risk for non-normal probability distributions, especially those with negative skew or positive excess kurtosis (fat tails).
- **Beta** measures the market risk of equity securities and portfolios of equity securities. This measure considers the risk reduction benefits of diversification and is appropriate for securities held in a well-diversified portfolio, whereas standard deviation is a measure of risk on a stand-alone basis.
- **Duration** is a measure of the price sensitivity of debt securities to changes in interest rates.



PROFESSOR'S NOTE

We describe and calculate standard deviation in Quantitative Methods; duration in Fixed Income; and beta in the current topic area, Portfolio Management.

Derivatives risks (sometimes referred to as “the Greeks”) include:

- **Delta.** This is the sensitivity of derivatives values to the price of the underlying asset.
- **Gamma.** This is the sensitivity of delta to changes in the price of the underlying asset.
- **Vega.** This is the sensitivity of derivatives values to the volatility of the price of the underlying asset.
- **Rho.** This is the sensitivity of derivatives values to changes in the risk-free rate.

Tail risk is the uncertainty about the probability of extreme (negative) outcomes. Commonly used measures of tail risk (sometimes referred to as **downside risk**) include Value at Risk and Conditional VaR.

Value at risk (VaR) is the minimum loss over a period that will occur with a specific probability. Consider a bank that has a one-month VaR of \$1 million with a probability of 5%. That means that a one-month loss of at least \$1 million is expected to occur 5% of the time. Note that this is not the maximum one-month loss the bank will experience; it is the minimum loss that will occur 5% of the time. VaR does not provide a maximum loss for a period. VaR has become accepted as a risk measure for banks and is used in establishing minimum capital requirements.

There are various methods of calculating VaR, and both the inputs and models used will affect the calculated value, perhaps significantly. As is always the case with estimates of risk, incorrect inputs or inappropriate distribution assumptions will lead to misleading results. Given these limitations, VaR should be used in conjunction with other risk measures.

Conditional VaR (CVaR) is the expected value of a loss, *given* that the loss exceeds a minimum amount. Relating this to the VaR measure presented previously, the CVaR would be the expected loss, given that the loss was at least \$1 million. It is calculated as the probability-weighted average loss for all losses expected to be at least \$1 million. CVaR is similar to the measure of loss given default that is used in estimating risk for debt securities.

Subjective and Market-Based Estimates of Risk

Two methods of risk assessment that are used to supplement measures such as VaR and CVaR are stress testing and scenario analysis. **Stress testing** examines the effects of a specific (usually extreme) change in a key variable such as an interest rate or exchange rate. **Scenario analysis** refers to a similar what-if analysis of expected loss but incorporates changes in multiple inputs. A given scenario might combine an interest rate change with a significant change in oil prices or exchange rates.

Quantifying the risk to an organization of very infrequent events is quite difficult. The risk of the bankruptcy of a firm that has never experienced significant financial distress is often a subjective estimate rather than a data-driven estimate. Estimates of risk can also be based on the market prices of insurance, derivatives, or other securities that can be used to hedge those risks. These hedging costs provide information on market participants’ aggregate estimate of the expected loss of specific risks.

Operational risks are difficult to quantify for a single organization because they are very difficult to predict and may result in very large costs to the organization. One way to approach this problem is to examine a large sample of firms in order to determine an overall probability of significant losses due to operational risks and the average loss of firms that have experienced such losses.

Unexpected changes in tax laws or the regulatory environment can impose large costs on an organization. The political nature of such changes makes them quite difficult to predict. Subjective estimates, rather than data-driven quantitative estimates, are necessary. As is often the case, even a subjective, non-quantitative estimate of risk probabilities and magnitudes is better than not addressing the risk factor at all.

Modifying Risk Exposures

Risk management does not seek to eliminate all risks. The goal is to retain the optimal mix of risks for the organization. This may mean taking on more of some risks, decreasing others, and eliminating some altogether. Once the risk management team has estimated various risks, management may decide to prevent or avoid a risk, accept a risk, transfer a risk, or shift a risk.

One way to avoid a risk is to not engage in the activity with the uncertain outcome. If political risks in a country are to be avoided, simply not investing in securities of firms based in that country or not expanding a business enterprise to that country would avoid those risks. A decision to avoid certain risks typically would come from top management as a part of establishing the risk tolerance of the organization and would be instituted because the risks are judged to outweigh the potential benefits of specific activities.

Some risks can be prevented. The risk of a data breach can be prevented with a greater level of security for the data and stronger processes. In this case, the benefits of reducing or eliminating the risk are judged to be greater than the cost of doing so.

For risks that management has decided to bear, the organization will seek to bear them efficiently. **Diversification** may offer a way to more efficiently bear a specific risk.



PROFESSOR'S NOTE

We explain how diversification can reduce risk in our review of "Portfolio Risk and Return: Part I."

Sometimes the term **self-insurance** is used to describe a situation where an organization has decided to bear a risk. Note, however, that this simply means that it will bear any associated losses from this risk factor. It is possible that this represents inaction rather than the result of analysis and strategic decision-making. In some cases, the firm will establish a reserve account to cover losses as a way of mitigating the impact of losses on the organization.

For a risk an organization has decided not to bear, risk transfer or risk shifting can be employed. With a **risk transfer**, another party takes on the risk. Insurance is a type of risk transfer. The risk of fire destroying a warehouse complex is shifted to an insurance company by buying an insurance policy and paying the policy premiums. Insurance companies diversify across many risks so the premiums of some insured parties pay the losses of others. Ideally, the various risks the insurance company insures are not highly correlated, as that can reduce or eliminate any diversification benefits. An insurance company with highly correlated risks (or a single very

large risk) may itself shift some of the resulting risk by buying reinsurance from another company.

With a **surety bond**, an insurance company has agreed to make a payment if a third party fails to perform under the terms of a contract or agreement with the organization. For example, a company may be exposed to losses if a key supplier does not deliver on time, slowing a project and resulting in penalty payments by the company. Insurers also issue **fidelity bonds**, which will pay for losses that result from employee theft or misconduct. Managements that purchase insurance, surety bonds, or fidelity bonds have determined that the benefits of risk reduction are greater than the cost of the insurance.

Risk shifting is a way to change the distribution of possible outcomes and is accomplished primarily with derivative contracts. For example, financial firms that do not want to bear currency risk on some foreign currency denominated debt securities can use forward currency contracts, futures contracts, or swaps to reduce or eliminate that risk. A firm with a large position in a specific stock can buy put options that provide a minimum sale price for the securities, altering the distribution of possible outcomes (in this case providing a floor value for the securities). On the other hand, a firm could sell call options on a specific stock, altering the distribution of possible outcomes by giving up some of the upside potential of the stock but decreasing its downside risk by the amount of the premiums received from the sale of the call options.

Choosing Among Risk Modification Methods

Organizations may use multiple methods of risk modification to reduce a single risk. The criterion is always a comparison of the costs and benefits of risk modification. Some risks may be mitigated by diversification, some shifted by insurance where it is available and economical, some shifted through the use of derivatives, and some simply borne or self-insured. The end result is a risk profile that matches the risk tolerance established for the organization and includes the risks that top management has determined match the goals of the organization in terms of cost versus potential returns.



MODULE QUIZ 66.1

1. An investor has the *most* control over her portfolio's:
 - A. risk.
 - B. relative returns.
 - C. risk-adjusted returns.
2. A risk management framework *least likely* includes:
 - A. risk governance, risk mitigation, and strategic risk analysis.
 - B. identifying and measuring risks, risk policies and processes, and risk governance.
 - C. risk mitigation, tracking the organization's risk profile, and establishing position limits.
3. Risk governance should *most appropriately* be addressed within an organization at:
 - A. the enterprise level.
 - B. the business unit level.
 - C. the individual employee level.
4. Effective risk management would *most likely* attempt to:
 - A. maximize expected return for a given level of risk.
 - B. minimize risk for a given level of expected return.
 - C. reduce any significant risks the firm is exposed to.

5. Risk budgeting can *best* be described as:
 - A. setting an annual limit on risk taken.
 - B. selecting assets by their risk characteristics.
 - C. establishing a maximum amount of risk to be taken.
6. Which of the following is *most appropriately* termed a financial risk?
 - A. Credit risk.
 - B. Solvency risk.
 - C. Settlement risk.
7. Risk shifting is *most likely* achieved by:
 - A. risk mitigation.
 - B. using derivative securities.
 - C. transferring risk to an insurance company.

KEY CONCEPTS

LOS 66.a

Risk management is the process of identifying and measuring the risks an organization (or portfolio manager or individual) faces, determining an acceptable level of overall risk (establishing risk tolerance), deciding which risks should be taken and which risks should be reduced or avoided, and putting the structure in place to maintain the bundle of risks that is expected to best achieve the goals of the organization.

LOS 66.b

An overall risk management framework should address the following activities:

- Identifying and measuring existing risks.
- Determining the organization's overall risk tolerance.
- Establishing the processes and policies for risk governance.
- Managing and mitigating risks to achieve the optimal bundle of risks.
- Monitoring risk exposures over time.
- Communicating across the organization.
- Performing strategic risk analysis.

LOS 66.c

Risk governance refers to senior management's determination of the risk tolerance of the organization, the elements of its optimal risk exposure strategy, and the framework for oversight of the risk management function.

LOS 66.d

The risk tolerance for an organization is the overall amount of risk it will take in pursuing its goals and is determined by top management.

LOS 66.e

Risk budgeting is the process of allocating the total risk the firm will take (risk tolerance) to assets or investments by considering the risk characteristics of each and how they can be combined to best meet the organization's goals. The budget can be a single risk measure or the sum of various risk factors.

LOS 66.f

Financial risks are those that arise from exposure to financial markets, including credit risk, liquidity risk, and market risk. Non-financial risks are the risks from the operation of the organization and from sources external to the organization. Individuals face mortality and longevity risk, in addition to financial risks.

Interactions among risks are frequent and can be especially significant during periods of stress in financial markets.

LOS 66.g

Risk of assets is measured by standard deviation, beta, or duration. Derivatives risk measures include delta, gamma, vega, and rho. Tail risk is measured with value at risk (VaR) or conditional VaR. Some risks must be measured subjectively.

An organization may decide to bear a risk (self-insurance), avoid or take steps to prevent a risk, efficiently manage a risk through diversification, transfer a risk with insurance or a surety bond, or shift a risk (change the distribution of uncertain outcomes) with derivatives.

Organizations may use multiple methods of risk modification after considering the costs and benefits of the various methods. The end result is a risk profile that matches the organization's risk tolerance and includes the risks that top management has determined match the organization's goals.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 66.1

1. **A** An investor can select securities to achieve a given level of portfolio risk. Returns cannot be controlled. (LOS 66.a)
2. **C** A risk management framework includes the procedures, analytical tools, and infrastructure to conduct the risk governance process. It includes all of the items listed with the exception of establishing position limits, which is an example of the operational implementation of a system of risk management. (LOS 66.b)
3. **A** Risk governance should be approached from an enterprise view, with senior management determining risk tolerance and a risk management strategy on an organization-wide level. (LOS 66.c)
4. **A** Risk management requires establishment of a risk tolerance (maximum acceptable level of risk) for the organization and will attempt to maximize expected returns for that level of risk. Some significant risks the firm is exposed to may be borne by the firm or even increased as a result of risk management. (LOS 66.d)
5. **B** Risk budgeting refers to selecting assets or securities by their risk characteristics up to the maximum allowable amount of risk. The maximum amount of risk to be taken is established through risk governance. (LOS 66.e)
6. **A** The main sources of financial risk are market risk, credit risk, and liquidity risk. Solvency risk and settlement risk are classified as non-financial risks. (LOS 66.f)
7. **B** Risk shifting changes the distribution of possible outcomes, typically through the use of derivative securities. Risk shifting is one technique for mitigating risk. Transferring risk to an insurance company is termed *risk transfer*. (LOS 66.g)

READING 67

TECHNICAL ANALYSIS

EXAM FOCUS

This reading introduces the “story” that underlies technical analysis, and you should understand how this differs from fundamental analysis. You should learn what the technical indicator names mean and how they are used, but don’t worry about being able to calculate them. The LOS require you to be able to “explain” and “describe.” You are responsible for knowing the assumptions and tools of technical analysis and understanding how technical analysts apply them in practice. You are not required to believe technical analysis can create positive risk-adjusted returns on average over time.

MODULE 67.1: TECHNICAL ANALYSIS



LOS 67.a: Explain principles and assumptions of technical analysis.

Video covering this content is available online.

Technical analysis is the study of collective market sentiment, as expressed in buying and selling of assets. It is based on the idea that prices are determined by the interaction of supply and demand. The market price equates supply and demand at any instant. Only participants who actually trade affect prices, and better-informed participants tend to trade in greater volume. Thus, price and volume reflect the collective behavior of buyers and sellers.

Technical analysis is based on three key principles:

1. *Market prices reflect all known information*, including economic factors, company fundamentals, and investor psychology.
2. *Market prices exhibit trends and countertrends* that tend to persist.
3. *Patterns and cycles repeat themselves in predictable ways*.

LOS 67.b: Describe potential links between technical analysis and behavioral finance.

A key assumption of technical analysis is that market prices reflect both rational and irrational investor behavior. This assumption implies that the *efficient markets hypothesis* does not hold.

Technical analysts believe investor behavior is reflected in trends and patterns that tend to repeat and can be identified and used for forecasting price movements. This behavior may reflect both rational and irrational decisions, as research in behavioral finance has indicated. Technical analysts believe investors respond in similar ways to similar situations over time, so

that patterns in trading prices and volumes will repeat themselves. Further, they believe that recognizing these patterns will provide information about future price changes.

Price and volume data reflect all the reasons why investors buy and sell, whether or not their buying and selling is consistent with any measure of a security's intrinsic value. The data also reflect the **market microstructure**, which consists of specifications such as tick sizes and order types, that may influence how supply and demand manifest themselves in a market.



PROFESSOR'S NOTE

The efficient markets hypothesis is described in the Equity Investments topic area.

LOS 67.c: Compare principles of technical analysis and fundamental analysis.

Technical analysis can be contrasted with fundamental analysis, which attempts to determine the intrinsic value of an asset. While fundamental analysis of a firm uses the company's financial statements and other information to analyze its financial position and determine its value, technical analysis uses only the firm's share price and trading volume data to project a target price. Technical analysis is not concerned with identifying buyers' and sellers' reasons for trading, but only with the trades that have occurred.

An advantage of only using actual price and volume data is that they are observable. Much of the data used in fundamental analysis is subject to assumptions or restatements, and might not be available at all for assets such as currencies or commodities. (This does not, however, imply that technical analysis itself is objective; both kinds of analysis require subjective judgment.) Another advantage of technical analysis is that it can be applied to the prices of assets that do not produce future cash flows (dividends or interest), such as commodities. Technical analysis can also be useful when financial statement fraud occurs. Price and volume may reflect the true value of the company even before the fraud is widely known and before the financial statements are restated.

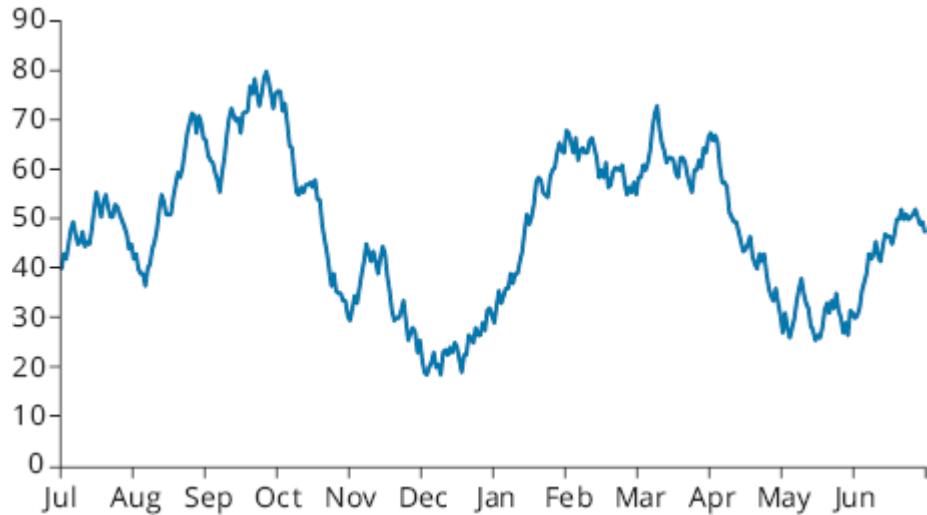
The usefulness of technical analysis is limited in markets where price and volume data might not truly reflect supply and demand. This may be the case in illiquid markets and in markets that are subject to outside manipulation (for example, currency market intervention by central banks). For stocks of bankrupt companies, short covering can create positive technical patterns even when it is known that the stock price will go to zero.

LOS 67.d: Describe and interpret different types of technical analysis charts.

Technical analysts primarily use charts of price and volume to analyze asset prices and overall market movement. Most of these charts have time on the horizontal axis. The time interval chosen (monthly, weekly, daily, or intraday periods) reflects the trading horizon of interest to the analyst. A technical analyst will typically start by observing longer-term trends on monthly and weekly charts, then look at recent activity on daily or intraday charts. If prices have changed exponentially (e.g., a stock index over several decades), an analyst may choose to draw charts on a logarithmic scale instead of the usual linear scale.

Line charts are the simplest technical analysis charts. They show closing prices for each period as a continuous line (see Figure 67.1).

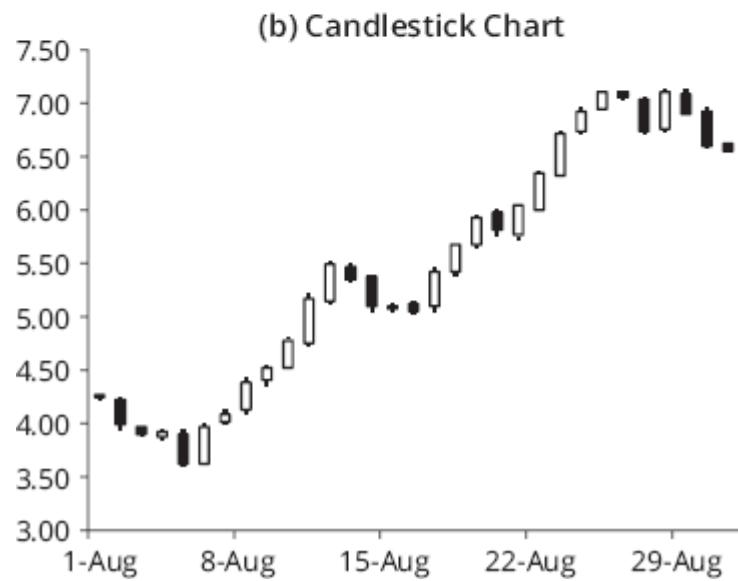
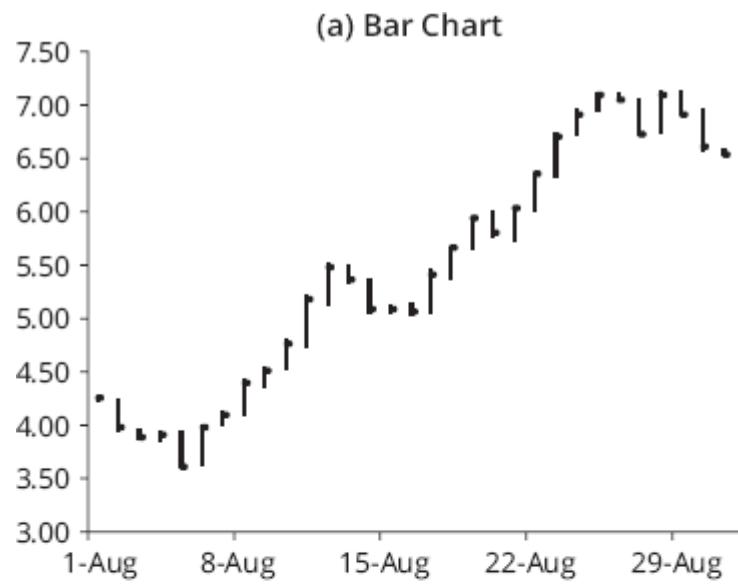
Figure 67.1: Line Chart



Bar charts add the high and low prices for each trading period and often include the opening price as well. Each period is displayed as a vertical line, with the closing price indicated as a point or dash on the right side of the line. If the chart includes opening prices, these are shown on the left side of each vertical line.

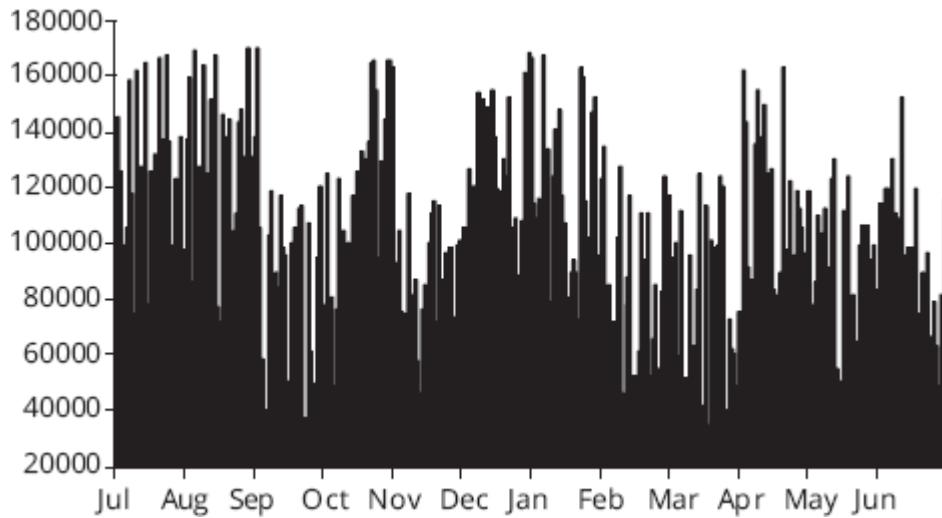
Candlestick charts use the same data as bar charts but display a box bounded by the opening and closing prices. The box is clear if the closing price is higher than the opening price, or filled if the close is lower than the opening price. Candlestick charts can make patterns easier to recognize (see Figure 67.2).

Figure 67.2: Bar and Candlestick Charts



Technical analysts are concerned with volume as well as price. **Volume charts** are usually displayed below price charts with each period's volume shown as a vertical line (see Figure 67.3).

Figure 67.3: Volume Chart



To perform **relative strength analysis**, an analyst calculates the ratios of an asset's closing prices to benchmark values, such as a stock index or comparable asset, and draws a line chart of the ratios. An increasing trend indicates that the asset is outperforming the benchmark (positive relative strength) and a decrease shows that the asset is underperforming the benchmark (negative relative strength).

LOS 67.e: Explain uses of trend, support, and resistance lines.

The most basic concept in technical analysis is the **trend** in prices. A market is said to be in an **uptrend** if prices are consistently reaching higher highs and retracing to higher lows. An uptrend means demand is increasing relative supply.

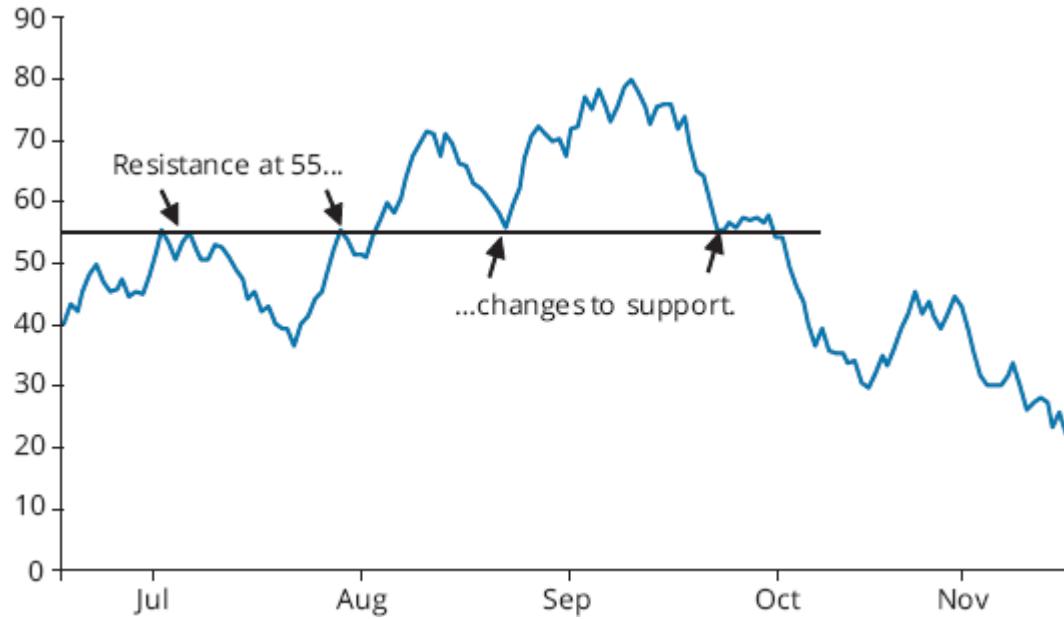
A market is in a **downtrend** if prices are consistently declining to lower lows and retracing to lower highs. A downtrend suggests supply (i.e., selling pressure) is increasing relative to demand. A period when neither an uptrend nor a downtrend is apparent is referred to as a period of **consolidation**.

Drawing a **trendline** on a chart can help to identify whether a trend is continuing or reversing. In an uptrend, a trendline connects the increasing lows in price. In a downtrend, the trendline connects the decreasing highs in price. When the price crosses the trendline by what the analyst considers a significant amount, a **breakout** from a downtrend or a **breakdown** from an uptrend is said to occur. Either a breakout or a breakdown may signal the end of the previous trend.

Breakouts and breakdowns are important because the trendline is thought to represent a level of support or resistance. At a **support level**, buying is expected to emerge that prevents further price decreases. At a **resistance level**, selling is expected to emerge that prevents further price increases. In addition to trendlines, support and resistance levels frequently appear at psychologically important prices such as round-number prices or historical highs and lows.

An important principle in technical analysis is the **change in polarity**. This refers to a belief that breached resistance levels become support levels and that breached support levels become resistance levels. In Figure 67.4, the area of 55 is viewed as a resistance level until the price breaks above that level, then 55 becomes a support level as prices decline from their new highs.

Figure 67.4: Change in Polarity



LOS 67.f: Explain common chart patterns.

Technical analysts look for recurring patterns in charts that suggest the future course of prices. Some patterns tend to appear at the end of trends, while other patterns indicate that a trend is likely to continue.

Reversal patterns occur when a trend approaches a range of prices but fails to continue beyond that range. A well-known example is the **head-and-shoulders pattern**, as shown in Figure 67.6. This pattern suggests the demand that has been driving the uptrend is fading, especially if each of the highs in the pattern occurs on declining volume.

Technical analysts commonly use the size of a head-and-shoulders pattern to project a price target for the ensuing downtrend. The size is the difference in price between the “head,” the highest price reached, and the “neckline,” the support level to which the price retraced after the left “shoulder” and the head have formed. If the price declines beyond the neckline after the right shoulder forms, the downtrend is projected to continue from that breakdown price by about the size of the head-and-shoulders pattern. In Figure 67.5, the top of the head is at 80 and the neckline is at 55, so the size of the pattern is $80 - 55 = 25$. The price target for the ensuing downtrend is $55 - 25 = 30$.

Figure 67.5: Reversal Patterns



Double top and **triple top** patterns are similar to the head-and-shoulders pattern in that they indicate weakening in the buying pressure that has been driving an uptrend. In both cases, the price reaches a resistance level at which selling pressure appears repeatedly, preventing any further increase in the price. As with the head-and-shoulders, the size of a double or triple top pattern can be used to project a price target for the next downtrend.

Reversal patterns for downtrends are called **inverse head-and-shoulders**, **double bottom**, and **triple bottom** patterns and can be analyzed in the same way as the reversal patterns for uptrends.

Continuation patterns suggest a pause in a trend rather than a reversal. **Triangles** form when prices reach lower highs and higher lows over a period of time (see Figure 67.6). Trendlines on the highs and on the lows thus converge when they are projected forward. Triangles can be symmetrical (higher lows and lower highs), ascending (higher lows and a resistance level), or descending (lower highs and a support level).

Figure 67.6: Triangle Continuation Pattern



Triangles suggest buying and selling pressure have become roughly equal temporarily, but they do not imply a change in direction of the trend. The size of a triangle, or the difference between the two trendlines at the time when the pattern begins to form, can be used to set a price target, assuming the price breaks out of the triangle and the previous trend continues.

Rectangles form when trading temporarily forms a range between a support level and a resistance level. As with a triangle, a rectangle suggests the prevailing trend will resume and can be used to set a price target. **Flags** and **pennants** refer to rectangles and triangles that appear on short-term price charts.

LOS 67.g: Explain common technical indicators.

Price-Based Indicators

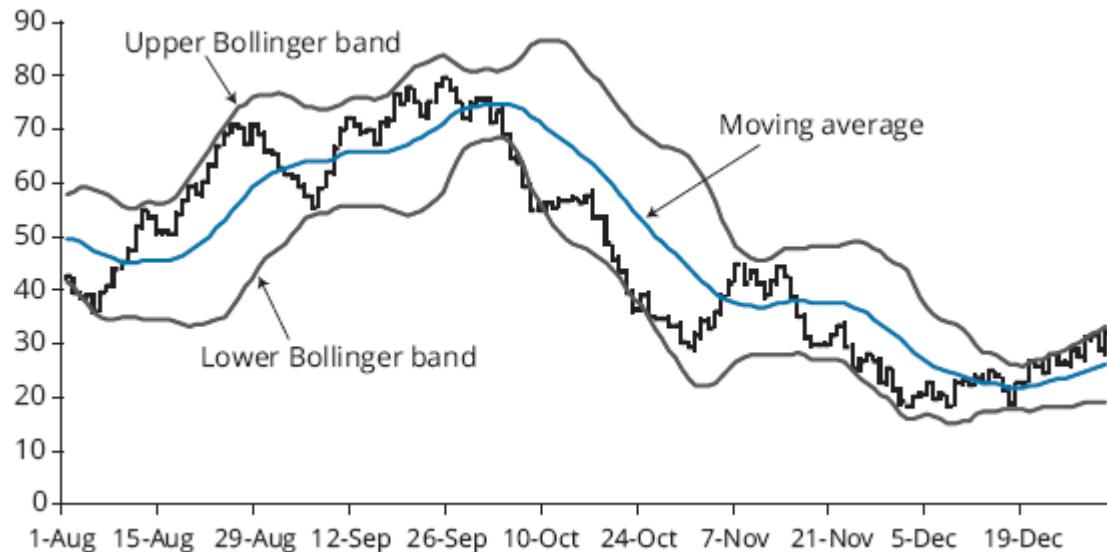
Moving average lines are frequently used to smooth the fluctuations in a price chart (or a chart of any time series). A moving average is simply the mean of the last n closing prices. The larger the chosen value of n , the smoother the moving average line. Analysts often use moving average periods that make intuitive sense, such as 20 days to represent the number of trading days in a month. In an uptrend, the price is higher than the moving average, and in a downtrend, the price is lower than the moving average. Moving average lines are often viewed as support or resistance levels.

Analysts use moving average lines to make changes in the trend easier to see. The longer the period used to calculate the moving average, the more short-term fluctuations are removed from the line. Overly long periods may, in fact, obscure changes in a price trend.

Moving averages for different periods can be used together, such as 20-day and 250-day averages. Points where the short-term average (more volatile) crosses the long-term average (smoother) can indicate changes in the price trend. When the short-term average crosses above long-term average (a “golden cross”), this is often viewed as an indicator of an emerging uptrend or a “buy” signal by technical analysts. The short-term average crossing below the long-term average (a “dead cross”) is often viewed as an indicator of an emerging downtrend or a “sell” signal.

Bollinger bands are constructed based on the standard deviation of closing prices over the last n periods. An analyst can draw high and low bands a chosen number of standard deviations (typically two) above and below the n -period moving average. The bands move away from one another when price volatility increases and move closer together when prices are less volatile (see Figure 67.7).

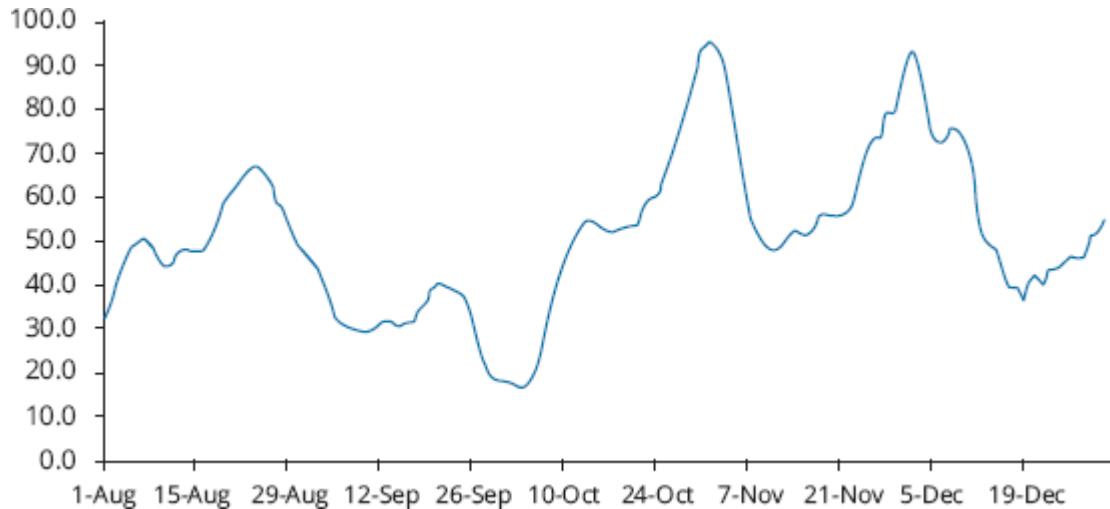
Figure 67.7: Moving Average and Bollinger Bands



Bollinger bands can be used to identify prices that are “extreme” in terms of how many standard deviations they are above or below the moving average price. Prices at or above the upper Bollinger band may be viewed as indicating an **overbought** market, one that is “too high” and likely to decrease in the near term. Likewise, prices at or below the lower Bollinger band may be viewed as indicating an **oversold** market, one that is “too low” and likely to increase in the near term. A possible trading strategy using Bollinger bands is to buy when the price is at the lower band or sell when the price is at the upper band. This is an example of a **contrarian strategy**—one that buys when most traders are selling and sells when most traders are buying. Contrarians believe markets and individual assets become overbought or oversold because most investors tend to buy and sell at the wrong times, and thus it can be profitable to trade in the opposite direction.

The distance between the upper and lower Bollinger bands reflects changes in price volatility (standard deviation) over time. The **Bollinger band width indicator** is the difference between the upper and lower bands as a percentage of the moving average. Figure 67.8 illustrates the Bollinger band width indicator for the market shown in Figure 67.7. A period of low volatility, such as early October in Figure 67.8, is referred to as a **squeeze** and, according to technical analysts, will be followed by a period of increasing volatility.

Figure 67.8: Bollinger Band Width Indicator



Momentum Oscillators

Momentum oscillators are another group of tools technical analysts use to identify overbought or oversold markets. These indicators are based on market prices but scaled so that they “oscillate” around a given value, such as zero, or between two values such as zero and 100. Extreme high values of an oscillator are viewed as indicating that a market is overbought, while extreme low values are viewed as indicating an oversold market. Technical analysts use charts of oscillators over time to highlight periods when they are near their high or low extremes.

Oscillator charts can also be used to identify **convergence** or **divergence** of the oscillator and market prices. Convergence occurs when the oscillator shows the same pattern as prices (e.g., both reaching higher highs), and divergence occurs when the oscillator shows a different pattern than prices (e.g., failing to reach a higher high when the price does). Convergence suggests the price trend is likely to continue, while divergence may indicate a potential change in the price trend.

Examples of oscillators include the rate of change (ROC) oscillator, Relative Strength Index (RSI), moving average convergence/divergence (MACD) lines, and the stochastic oscillator.

- **ROC oscillator.** An ROC is calculated as 100 times the difference between the latest closing price and the closing price n periods earlier. Thus, it oscillates around zero. One way technical analysts use the ROC is to buy when the oscillator changes from negative to positive during an uptrend in prices, and sell when the ROC changes from positive to negative during a downtrend. Some analysts prefer to use the ratio of the current price to the past price in place of the difference. Calculated this way, the ROC oscillates around 100.
- **RSI.** An RSI is based on the ratio of total price increases to total price decreases over a selected number of periods. This ratio is then scaled to oscillate between 0 and 100, with high values (typically those greater than 70) indicating an overbought market and low values (typically those less than 30) indicating an oversold market.
- **MACD.** MACD oscillators are drawn using exponentially smoothed moving averages, which place greater weight on more recent observations. The “MACD line” is the difference between two exponentially smoothed moving averages of the price, and the “signal line” is an exponentially smoothed moving average of the MACD line. The lines oscillate around zero but are not bounded. The MACD oscillator can be used to indicate overbought or oversold conditions or to identify convergence or divergence with the price trend. Points where the

two lines cross can be used as trading signals, much like the use of two different moving averages discussed previously. The MACD line crossing above the smoother signal line is viewed as a buy signal and the MACD line crossing below the signal line is viewed as a sell signal.

- **Stochastic oscillator.** A stochastic oscillator is calculated from the latest closing price and highest and lowest prices reached in a recent period, such as 14 days. In a sustainable uptrend, prices tend to close nearer to the recent high, and in a sustainable downtrend, prices tend to close nearer to the recent low. Stochastic oscillators use two lines that are bounded by 0 and 100. The "%K" line is the difference between the latest price and the recent low as a percentage of the difference between the recent high and low. The "%D" line is a 3-period average of the %K line. Technical analysts typically use stochastic oscillators to identify overbought and oversold markets. Points where the %K line crosses the %D line can also be used as trading signals in the same way as the MACD lines.



PROFESSOR'S NOTE

Is this fun or what? Remember, no calculations are required. Just know the terms as best you can.

Non-Price-Based Indicators

While the technical indicators mentioned so far assume investor sentiment is reflected in price and volume data, technical analysts can also look at indicators of investor sentiment and capital flows to gain insight into potential emerging trends. **Sentiment indicators** can be used to discern the views of potential buyers and sellers. Market sentiment is said to be "bullish" when investors expect increasing prices and "bearish" when they expect decreasing prices.

Indicators can include opinion polls that try to measure investor sentiment directly, as well as several measures that are based on market data:

- **Put/call ratio.** Put options increase in value when the price of an underlying asset decreases, while call options increase in value if the price of the underlying asset increases. For financial assets that have actively traded options, the volume of put and call options reflects activity by investors with negative and positive outlooks, respectively, about the asset. The put/call ratio is put volume divided by call volume. Increases in the put/call ratio indicate a more negative outlook for the price of the asset. This ratio is generally viewed as a contrarian indicator. Extremely high ratios indicate strongly bearish investor sentiment and possibly an oversold market, while extremely low ratios indicate strongly bullish sentiment and perhaps an overbought market.
- **Volatility index (VIX).** The Chicago Board Options Exchange calculates the VIX, which measures the volatility of options on the S&P 500 stock index. High levels of the VIX suggest investors fear declines in the stock market. Technical analysts most often interpret the VIX in a contrarian way, viewing a predominantly bearish investor outlook as a bullish sign.
- **Margin debt.** The amount of margin debt is a readily available indicator because brokers are required to report this data. Increases in total margin debt outstanding suggest aggressive buying by bullish margin investors. As margin investors reach their limits of margin credit, their ability to continue buying decreases, which can cause prices to begin declining. As prices decrease, investors may need to sell securities to meet margin calls, and these sales drive

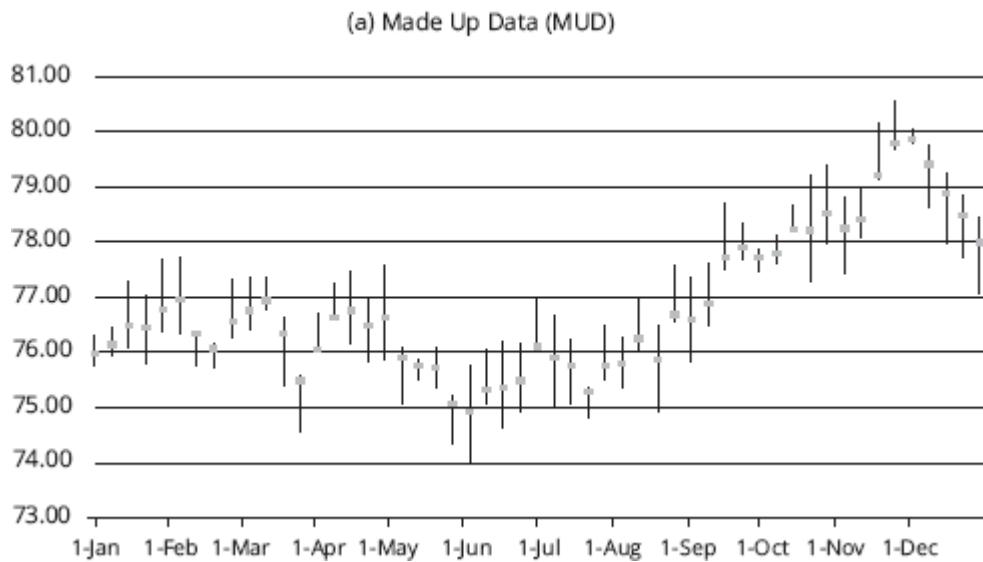
prices lower still. Increasing margin debt tends to coincide with increasing market prices and decreasing margin debt tends to coincide with decreasing prices.

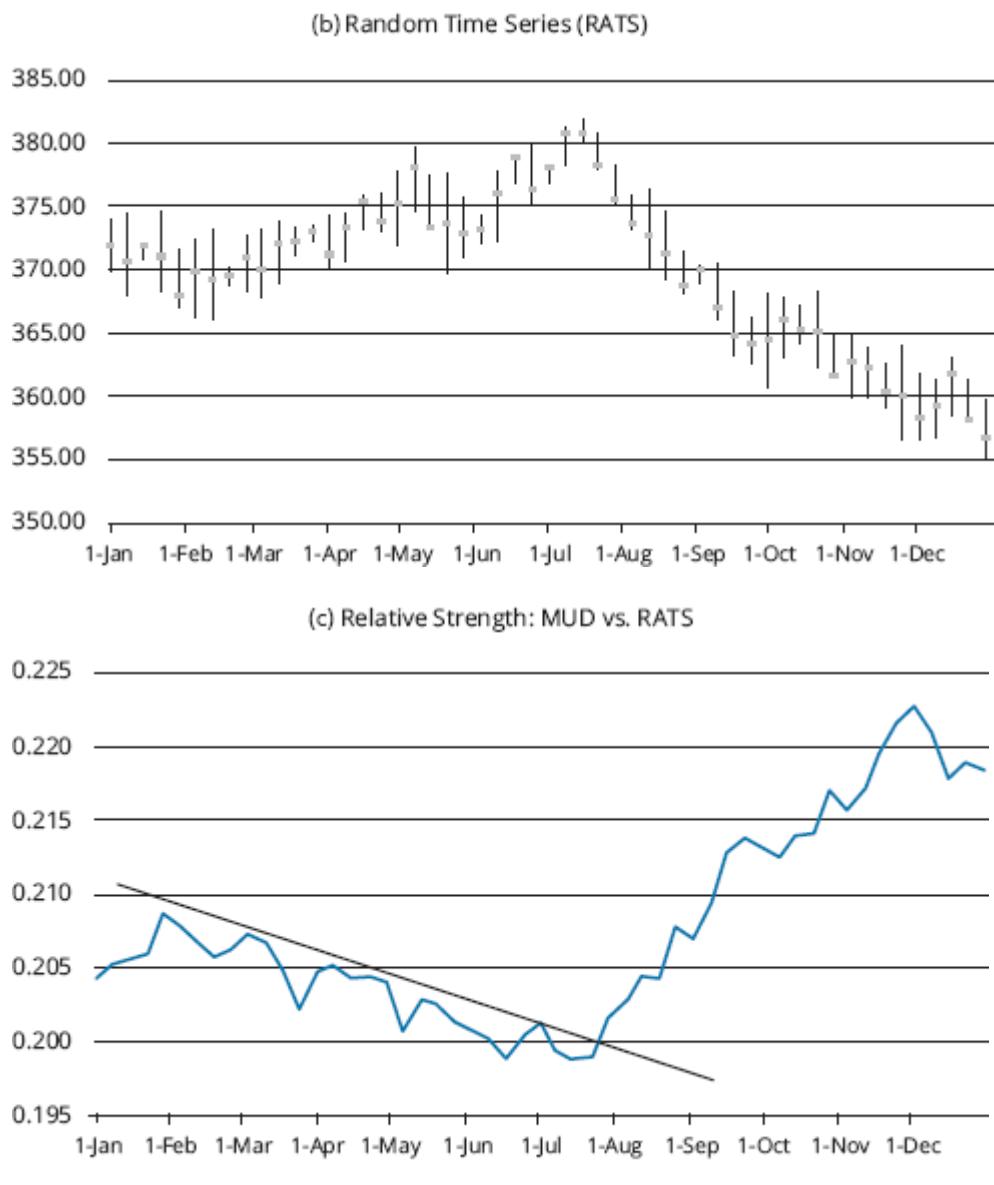
LOS 67.h: Describe principles of intermarket analysis.

Intermarket analysis is an analysis of the interrelationships among the market values of major asset classes, such as stocks, bonds, commodities, and currencies. After identifying attractive asset classes, an analyst can apply relative strength analysis to identify which assets within these classes are outperforming others. This approach is also useful for comparing the relative performance of equity market sectors or industries, and of various international markets.

Relative strength charts are used to determine which assets, asset classes, or markets are outperforming others. Figure 67.9 shows a relative strength analysis of two hypothetical stocks, Made Up Data (MUD) and Random Time Series (RATS). Panels a and b are bar charts for each stock and panel c is the closing price of MUD as a ratio of the closing price of RATS. The relative strength chart indicates a clear breakout from a downtrend in early August. A technical analyst might have interpreted this as a signal to reduce the portfolio investment in RATS and increase the allocation to MUD.

Figure 67.9: Relative Strength Analysis





©2022 Kaplan, Inc.

LOS 67.i: Explain technical analysis applications to portfolio management.

Recall that the execution step of a portfolio management process may involve top-down or bottom-up approaches to asset allocation and security selection. Technical analysis may be used to complement fundamental analysis using either approach.

A top-down approach begins by examining a country's or international region's economic growth, then analyzing how it is affecting the performance of market sectors, industry groups, and securities, and then using that analysis to identify attractive areas for investment.

Intermarket analysis using relative strength charts can identify markets and sectors that have outperformed relative to others. Portfolio managers may use this analysis to make tactical asset allocation decisions.

Bottom-up analysis uses the opposite approach, taking economic and industry performance as given. Instead, this approach begins by identifying an investment universe or opportunity set,

then narrowing it using criteria chosen by the analyst. These criteria may include technical analysis indicators or chart patterns.

At times, a bottom-up approach can complement a top-down approach. For example, if several stocks in an industry exhibit technical patterns indicating selling pressure, it may be that this industry as a whole is likely to underperform.



MODULE QUIZ 67.1

1. An assumption of technical analysis is that market prices:
 - A. exhibit identifiable trends and patterns that persist and repeat.
 - B. are the only information necessary to analyze a freely trading market.
 - C. reflect supply and demand conditions because actual transactions reflect rational decisions by buyers and sellers.
2. Which type of price chart requires data on the opening price?
 - A. Bar chart.
 - B. Line chart.
 - C. Candlestick chart.
3. A stock has been in a downtrend for several days. When its price decreases to near \$30, many investors enter orders to buy the stock and the price increases to \$31. This is *most likely* an example of:
 - A. a support level.
 - B. a resistance level.
 - C. a change in polarity.
4. An analyst who uses the standard deviation of recent prices to identify overbought or oversold conditions is *most likely* to employ:
 - A. Bollinger bands.
 - B. a Relative Strength Index (RSI).
 - C. a rate of change (ROC) oscillator.
5. If a stock's relative strength chart shows an uptrend, the stock is:
 - A. increasing in price.
 - B. outperforming a benchmark.
 - C. increasing on high volume or decreasing on low volume.
6. A technical analyst begins her evaluation of the U.S. equity market by using a computer algorithm to identify which stocks in the S&P 500 Index have shown uptrends on increasing volume over the last month. This analyst is *most accurately* said to be using:
 - A. a top-down approach.
 - B. a center-out approach.
 - C. a bottom-up approach.

KEY CONCEPTS

LOS 67.a

Technical analysis is based on three key principles:

1. Market prices reflect all known information.
2. Prices exhibit trends that persist.
3. Investor behavior repeats itself, producing patterns and cycles that recur.

LOS 67.b

Technical analysts believe investor behavior is reflected in trends and patterns that repeat and can be used for forecasting. This behavior may reflect both rational and irrational motivations, as the theory of behavioral finance suggests.

LOS 67.c

Technical analysis can be contrasted with fundamental analysis, which attempts to determine the intrinsic value of an asset. While fundamental analysis uses a company's financial statements and other information to analyze its financial position and determine its value, technical analysis uses only the firm's share price and trading volume data to project a target price.

LOS 67.d

Technical analysts use charts to identify trends and patterns in prices over time. A line chart is a continuous line that connects closing prices for each period. Bar charts and candlestick charts show the open, high, low, and close for each period. Volume charts often accompany price charts.

LOS 67.e

In an uptrend, prices are reaching higher highs and higher lows. An uptrend line is drawn below the prices on a chart by connecting the increasing lows with a straight line. In a downtrend, prices are reaching lower lows and lower highs. A downtrend line is drawn above the prices on a chart by connecting the decreasing highs with a straight line.

Support and resistance are price levels or ranges at which buying or selling pressure is expected to limit price movement. Commonly identified support and resistance levels include trendlines and previous high and low prices. The change in polarity principle is the idea that breached resistance levels become support levels and breached support levels become resistance levels.

LOS 67.f

Technical analysts look for recurring patterns in price charts. Head-and-shoulders patterns, double tops, and triple tops are thought to be reversal patterns at the ends of uptrends. Inverse head-and-shoulders patterns, double bottoms, and triple bottoms are thought to be reversal patterns at the ends of downtrends. Triangles, rectangles, flags, and pennants are thought to be continuation patterns, which indicate that the trend in which they appear is likely to go further in the same direction.

LOS 67.g

Price-based indicators include moving averages, Bollinger bands, and momentum oscillators such as the Relative Strength Index (RSI), moving average convergence/divergence (MACD) lines, rate of change (ROC) oscillators, and stochastic oscillators. These indicators are commonly used to identify changes in price trends, as well as "overbought" markets that are likely to decrease in the near term and "oversold" markets that are likely to increase in the near term. Sentiment indicators include opinion polls, the put/call ratio, the volatility index, and margin debt.

LOS 67.h

Intermarket analysis refers to analysis of the relationships among the price behavior of major asset classes, such as stocks, bonds, commodities, and currencies. It uses relative strength charts to identify assets, asset classes, or markets that have outperformed others.

LOS 67.i

Technical analysis may be used in either a top-down or bottom-up approach to asset class allocation and security selection. In a top-down approach, intermarket analysis can be used to identify markets and sectors that have outperformed others. In a bottom-up approach, criteria for narrowing down an investment opportunity set may include technical indicators and chart patterns.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 67.1

1. **A** Technical analysis assumes persistent trends and repeating patterns in market prices can be used to forecast price behavior. Technical analysts believe prices reflect supply and demand, but that buying and selling can be motivated by both rational and irrational causes. Volume, along with price, is important information to a technical analyst. (LOS 67.a)
2. **C** Constructing a candlestick chart requires opening, high, low, and closing prices for each trading period. A line chart requires only the closing prices. A bar chart can be constructed using high, low, and closing prices. (LOS 67.c)
3. **A** The downtrend reached a support level where buying demand sustained the price. A resistance level is a price at which selling pressure emerges that stops an uptrend. The change in polarity principle holds that breached support levels become resistance and breached resistance levels become support. With no information given on the stock's history, we cannot determine whether \$30 had once been a resistance level. (LOS 67.e)
4. **A** Bollinger bands are based on the standard deviation of prices over some number of the most recent periods. An RSI is based on the sums of positive and negative price changes during a period. An ROC oscillator is based on the difference between the most recent closing price and the closing price a given number of periods earlier. (LOS 67.g)
5. **B** If a relative strength chart shows an uptrend, the stock is outperforming the benchmark stock or index against which it is being measured. This does not imply that the stock is increasing in price; if the stock price is decreasing, but the benchmark is decreasing by a larger percentage, the relative strength chart will still trend upward. Volume is not an input into a relative strength ratio. (LOS 67.h)
6. **C** Using selected criteria to narrow an investment universe reflects a bottom-up approach. (LOS 67.i)

READING 68

FINTECH IN INVESTMENT MANAGEMENT

EXAM FOCUS

Fintech (financial technology) is increasing in its importance to the financial services industry. As terms like *Big Data*, *blockchain*, and *algorithmic trading* come into common use, CFA exam candidates are expected to be familiar with them and how they relate to investment management. That being said, we do not believe the exam writers expect finance professionals to become data scientists. The Learning Outcome Statements for this topic only ask candidates to describe these fintech concepts. Focus on their applications, such as cryptocurrencies and robo-advisors, the advantages of their use in finance, and the challenges their adoption may present.

MODULE 68.1: FINTECH IN INVESTMENT MANAGEMENT



Video covering this content is available online.

LOS 68.a: Describe “fintech”.

The term **fintech** refers to developments in technology that can be applied to the financial services industry. Companies that are in the business of developing technologies for the finance industry are often referred to as fintech companies.

Some of the primary areas where fintech is developing include:

- Increasing functionality to handle large sets of data that may come from many sources and exist in a variety of forms.
- Tools and techniques such as artificial intelligence for analyzing very large datasets.
- Automation of financial functions such as executing trades and providing investment advice.
- Emerging technologies for financial recordkeeping that may reduce the need for intermediaries.

LOS 68.b: Describe Big Data, artificial intelligence, and machine learning.

Big Data is a widely used expression that refers to all the potentially useful information that is generated in the economy. This includes not only data from traditional sources, such as financial

markets, company financial reports, and government economic statistics, but also **alternative data** from non-traditional sources. Some of these non-traditional sources are:

- Individuals who generate usable data such as social media posts, online reviews, email, and website visits.
- Businesses that generate potentially useful information such as bank records and retail scanner data. These kinds of data are referred to as **corporate exhaust**.
- Sensors, such as radio frequency identification chips, are embedded in numerous devices such as smart phones and smart buildings. The broad network of such devices is referred to as the **Internet of Things**.

Characteristics of Big Data include its volume, velocity, and variety.

The volume of data continues to grow by orders of magnitude. The units in which data can be measured have increased from megabytes and gigabytes to terabytes (1,000 gigabytes) and even petabytes (1,000 terabytes).

Velocity refers to how quickly data are communicated. Real-time data such as stock market price feeds are said to have low **latency**. Data that are only communicated periodically or with a lag are said to have high latency.

The variety of data refers to the varying degrees of structure in which data may exist. These range from structured forms such as spreadsheets and databases, to semistructured forms such as photos and web page code, to unstructured forms such as video.

The field of **data science** concerns how we extract information from Big Data. Data science describes methods for processing and visualizing data. Processing methods include:

- *Capture*—collecting data and transforming it into usable forms.
- *Curation*—assuring data quality by adjusting for bad or missing data.
- *Storage*—archiving and accessing data.
- *Search*—examining stored data to find needed information.
- *Transfer*—moving data from their source or a storage medium to where they are needed.

Visualization techniques include the familiar charts and graphs that display structured data. To visualize less-structured data requires other methods. Some examples of these are word clouds that illustrate the frequency that words appear in a sample of text, or mind maps that display logical relations among concepts.

Taking advantage of Big Data presents a number of challenges. Analysts must ensure that the data they use are of high quality, accounting for the possibilities of outliers, bad or missing data, or sampling biases. The volume of data collected must be sufficient and appropriate for its intended use.

The need to process and organize data before using it can be especially problematic with qualitative and unstructured data. This is a process to which **artificial intelligence**, or computer systems that can be programmed to simulate human cognition, may be applied usefully. **Neural networks** are an example of artificial intelligence in that they are programmed to process information in a way similar to the human brain.

An important development in the field of artificial intelligence is **machine learning**. In machine learning, a computer algorithm is given inputs of source data, with no assumptions about their probability distributions, and may be given outputs of target data. The algorithm is designed to learn, without human assistance, how to model the output data based on the input data or to learn how to detect and recognize patterns in the input data.

Machine learning typically requires vast amounts of data. A typical process begins with a *training* dataset in which the algorithm looks for relationships. A *validation* dataset is then used to refine these relationship models, which can then be applied to a *test* dataset to analyze their predictive ability.

In **supervised learning**, the input and output data are labelled, the machine learns to model the outputs from the inputs, and then the machine is given new data on which to use the model. In **unsupervised learning**, the input data are not labelled and the machine learns to describe the structure of the data. **Deep learning** is a technique that uses layers of neural networks to identify patterns, beginning with simple patterns and advancing to more complex ones. Deep learning may employ supervised or unsupervised learning. Some of the applications of deep learning include image and speech recognition.

Machine learning can produce models that overfit or underfit the data. **Overfitting** occurs when the machine learns the input and output data too exactly, treats noise as true parameters, and identifies spurious patterns and relationships. In effect, the machine creates a model that is too complex. **Underfitting** occurs when the machine fails to identify actual patterns and relationships, treating true parameters as noise. This means the model is not complex enough to describe the data. A further challenge with machine learning is that its results can be a “black box,” producing outcomes based on relationships that are not readily explainable.

LOS 68.c: Describe fintech applications to investment management.

Applications of fintech that are relevant to investment management include text analytics, natural language processing, risk analysis, algorithmic trading, and robo-advisory services.

Text analytics refers to the analysis of unstructured data in text or voice forms. An example of text analytics is analyzing the frequency of words and phrases. In the finance industry, text analytics have the potential to partially automate specific tasks such as evaluating company regulatory filings.

Natural language processing refers to the use of computers and artificial intelligence to interpret human language. Speech recognition and language translation are among the uses of natural language processing. Possible applications in finance could be to check for regulatory compliance in an examination of employee communications, or to evaluate large volumes of research reports to detect more subtle changes in sentiment than can be discerned from analysts' recommendations alone.

As we saw in our reading on Risk Management: An Introduction, risk governance requires an understanding of a firm's exposure to a wide variety of risks. Financial regulators require firms to perform risk assessments and stress testing. The simulations, scenario analysis, and other techniques used for risk analysis require large amounts of quantitative data along with a great

deal of qualitative information. Machine learning and other techniques related to Big Data can be useful in modeling and testing risk, particularly if firms use real-time data to monitor risk exposures.

Algorithmic trading refers to computerized securities trading based on a predetermined set of rules. For example, algorithms may be designed to enter the optimal execution instructions for any given trade based on real-time price and volume data. Algorithmic trading can also be useful for executing large orders by determining the best way to divide the orders across exchanges. Another application of algorithmic trading is **high-frequency trading** that identifies and takes advantage of intraday securities mispricings.

Robo-advisors are online platforms that provide automated investment advice based on a customer's answers to survey questions. The survey questions are designed to elicit an investor's financial position, return objectives, risk tolerance, and constraints such as time horizon and liquidity needs. Robo-advisor services may be fully automated or assisted by a human investment advisor.

Robo-advisory services tend to offer passively managed investments with low fees, low minimum account sizes, traditional asset classes, and conservative recommendations. The primary advantage of robo-advisors is their low cost to customers, which may make advice more accessible to a larger number of investors.

A disadvantage of robo-advisors is that the reasoning behind their recommendations might not be apparent. Without a human investment advisor to explain the reasoning, customers may hesitate to trust the appropriateness of a robo-advisor's recommendations, particularly in crisis periods.

Regulation of robo-advisors is still emerging. However, in many countries robo-advisory services are subject to the same regulations and registration requirements as any other investment advisor.

LOS 68.d: Describe financial applications of distributed ledger technology.

A **distributed ledger** is a database that is shared on a network so that each participant has an identical copy. A distributed ledger must have a consensus mechanism to validate new entries into the ledger. Distributed ledger technology uses cryptography to ensure only authorized network participants can use the data.

A **blockchain** is a distributed ledger that records transactions sequentially in blocks and links these blocks in a chain. Each block has a cryptographically secured "hash" that links it to the previous block. The consensus mechanism in a blockchain requires some of the computers on the network to solve a cryptographic problem. These computers are referred to as **miners**. Mining requires vast resources of computing power and electricity. This imposes substantial costs on any attempt to manipulate a blockchain's historical record. To do so would also require one party to control a majority of the network. For this reason, a blockchain is more likely to succeed with a large number of participants in its network.

Distributed ledgers can take the form of permissionless or permissioned networks. In **permissionless networks**, all network participants can view all transactions. These networks

have no central authority, which gives them the advantage of having no single point of failure. The ledger becomes a permanent record visible to all, and its history cannot be altered (short of the manipulation described previously). This removes the need for trust between the parties to a transaction.

In **permissioned networks**, users have different levels of access. For example, a permissioned network might allow network participants to enter transactions while giving government regulators permission to view the transaction history. A distributed ledger that allowed regulators to view records that firms are required to make available would increase transparency and decrease compliance costs.

Financial Applications of Distributed Ledger Technology

Cryptocurrencies are a current example of distributed ledger technology in finance. A cryptocurrency is an electronic medium of exchange that allows participants to engage in real-time transactions without a financial intermediary. Cryptocurrencies typically reside on permissionless networks.

Demonstrating the impact cryptocurrencies are already having in finance, companies have raised capital through **initial coin offerings**, in which they sell cryptocurrency for money or another cryptocurrency. This reduces the cost and time frame compared to carrying out a regulated IPO, and initial coin offerings typically do not come with voting rights. Investors should note that fraud has occurred with initial coin offerings and they may become subject to securities regulations.

Post-trade clearing and settlement is an area of finance to which distributed ledger technology might be applied productively. Distributed ledgers could automate many of the processes currently carried out by custodians and other third parties. The technology has the potential to bring about real-time trade verification and settlement, which (as we will see in Equity Investments) currently takes one or more days for many securities. This would reduce trading costs and counterparty risk. On the other hand, the inability to alter past transactions on a distributed ledger is problematic when cancelling a trade is required.

Other potential applications of distributed ledger technology in finance include smart contracts and tokenization. **Smart contracts** are electronic contracts that could be programmed to self-execute based on terms agreed to by the counterparties. For example, an options contract could be set up to be exercised automatically if certain defined conditions exist in the market. **Tokenization** refers to electronic proof of ownership of physical assets, which could be maintained on a distributed ledger. For example, such a ledger could potentially replace the paper real estate deeds currently filed at government offices.



MODULE QUIZ 68.1

1. Fintech is *most accurately* described as:
 - A. the application of technology to the financial services industry.
 - B. the replacement of government-issued money with electronic currencies.
 - C. the clearing and settling securities trades through distributed ledger technology.
2. Which of the following technological developments is *most likely* to be useful for analyzing Big Data?
 - A. Machine learning.

- B. High-latency capture.
 - C. The Internet of Things.
3. A key criticism of robo-advisory services is that:
- A. they are costly for investors to use.
 - B. the reasoning behind their recommendations can be unclear.
 - C. they tend to produce overly aggressive investment recommendations.
4. Which of the following statements about distributed ledger technology is *most accurate*?
- A. A disadvantage of blockchain is that past records are vulnerable to manipulation.
 - B. Tokenization can potentially streamline transactions involving high-value physical assets.
 - C. Only parties who trust each other should carry out transactions on a permissionless network.

KEY CONCEPTS

LOS 68.a

Fintech refers to developments in technology that can be applied to the financial services industry. Companies that develop technologies for the finance industry are referred to as fintech companies.

LOS 68.b

Big Data refers to the potentially useful information that is generated in the economy, including data from traditional and non-traditional sources. Characteristics of Big Data include its volume, velocity, and variety.

Artificial intelligence refers to computer systems that can be programmed to simulate human cognition. Neural networks are an example of artificial intelligence.

Machine learning is programming that gives a computer system the ability to improve its performance of a task over time and is often used to detect patterns in large sets of data.

LOS 68.c

Applications of fintech to investment management include text analytics, natural language processing, risk analysis, algorithmic trading, and robo-advisory services.

Text analytics refers to analyzing unstructured data in text or voice forms. Natural language processing is the use of computers and artificial intelligence to interpret human language. Algorithmic trading refers to computerized securities trading based on predetermined rules.

Robo-advisors are online platforms that provide automated investment advice based on a customer's answers to survey questions. The primary advantage of robo-advisors is their low cost to customers. A disadvantage is that the reasoning behind their recommendations might not be apparent.

LOS 68.d

A distributed ledger is a database that is shared on a network, with a consensus mechanism so that each participant has an identical copy of the ledger.

A cryptocurrency is an electronic medium of exchange that allows network participants in a distributed ledger to engage in real-time transactions without a financial intermediary.

Potential financial applications of distributed ledger technology include smart contracts, tokenization, and more efficient post-trade clearing and settlement.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 68.1

1. **A** Fintech refers to the application of technology to the financial services industry and to companies that are involved in developing and applying technology for financial services. Cryptocurrencies and distributed ledger technology are examples of fintech-related developments. (LOS 68.a)
2. **A** Machine learning is a computer programming technique useful for identifying and modeling patterns in large volumes of data. The Internet of Things refers to the network of devices that is one of the sources of Big Data. Capture is one aspect of processing data. Latency refers to the lag between when data is generated and when it is needed. (LOS 68.b)
3. **B** One criticism of robo-advisory services is that the reasoning behind their recommendations might not be readily apparent to customers. Recommendations from robo-advisors tend to be conservative rather than aggressive. Low cost is a primary advantage of robo-advisors. (LOS 68.c)
4. **B** By enabling electronic proof of ownership, tokenization has the potential to streamline transfers of physical assets such as real estate. The high cost and difficulty of manipulating past records is a strength of blockchain technology. Permissionless networks do not require trust between the parties to a transaction because the record of a transaction is unchangeable and visible to all network participants. (LOS 68.d)

TOPIC QUIZ: PORTFOLIO MANAGEMENT

You have now finished the Portfolio Management topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Portfolio Management. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares to the scores of others who entered their answers.

READING 69

ETHICS AND TRUST IN THE INVESTMENT PROFESSION

EXAM FOCUS

From this reading, candidates should learn the definitions of ethics and ethical behavior presented by the authors and the arguments presented for having a code of ethics and following ethical principles. Additionally, the arguments for integrating ethics into the decision-making process include testable material.

MODULE 69.1: ETHICS AND TRUST



LOS 69.a: Explain ethics.

Video covering this content is available online.

Ethics can be described as a set of shared beliefs about what is good or acceptable behavior and what is bad or unacceptable behavior. Ethical conduct has been described as behavior that follows moral principles and is consistent with society's ethical expectations.

Ethical conduct has also been described as conduct that improves outcomes for stakeholders, who are people directly or indirectly affected by the conduct. Examples of stakeholders in the case of investment professionals include their clients, coworkers, employers, and the investment profession as a whole. Some decisions may bring positive results for you, but negative consequences for a stakeholder, such as a coworker. Ethical conduct is behavior that balances your self-interest with the impact on others.

LOS 69.b: Describe the role of a code of ethics in defining a profession.

A code of ethics is a written set of moral principles that can guide behavior by describing what is considered acceptable behavior. Having a code of ethics is a way to communicate the values, principles, and expectations of an organization or other group of people and provides a general guide to what constitutes acceptable behavior. Some codes of ethics include a set of rules or standards that require some minimum level of ethical behavior.

A profession refers to a group of people with specialized skills and knowledge who serve others and agree to behave in accordance with a code of ethics. A professional code of ethics is a way

for a profession to communicate to the public that its members will use their knowledge and skills to serve their clients in an honest and ethical manner.

LOS 69.c: Describe professions and how they establish trust.

A profession is an occupational group (e.g., doctors or lawyers) that has requirements of specialized expert knowledge, and often a focus on ethical behavior and service to the larger community or society. Additionally, a profession may have the following characteristics:

- A code and standards for professional behavior.
- A regulatory body to enforce rules concerning professional behavior and monitor the ethical behavior of members.
- A focus on the needs of their clients (e.g., students, patients).
- A focus on service to society.
- A requirement to put client interests first.
- A focus on or requirement for continuing education.

Ways that professions establish trust include:

- Requiring high standards of expertise, knowledge, and skill.
 - Establishing standards of ethical behavior.
 - Monitoring professional conduct.
 - Encouraging continuing education to maintain and increase competence.
 - Being focused on clients' needs.
 - Mentoring and inspiring others in the profession.
-

LOS 69.d: Describe the need for high ethical standards in investment management.

Investment professionals have a special responsibility because they are entrusted with their clients' wealth. The responsibility to use their specialized knowledge and skills to both protect and grow client assets makes high ethical standards all the more important. Investment advice and management are intangible products, making quality and value received more difficult to evaluate than for tangible products such as a laptop computer or a restaurant meal. For this reason, trust in investment professionals takes on an even greater importance than in many other businesses.

Failure to act in a highly ethical manner can damage not only client wealth but also impede the success of investment firms and investment professionals because potential investors will be less likely to use their services.

Unethical behavior by financial services professionals can have negative effects for society as a whole. The financial services industry serves as an intermediary between savers and those seeking financing for their business activities. A lack of trust in financial advisors will reduce the funds entrusted to them and increase the cost of raising capital for business investment and

growth. When investors cannot rely on the information they receive from financial services professionals, this adds another layer of risk on top of the investment risks that investors face. Even the perception of additional risk will reduce the amounts invested and increase the returns required to attract investor capital.

In addition to reducing the amount of investment overall, unethical behavior—such as providing incomplete, misleading, or false information to investors—can affect the allocation of the capital that is raised. Misallocation of capital to businesses other than those with the most potential for growth and societal benefit reduces the growth of an economy and the well-being of its people. When the allocation of investment capital is constrained or inefficient, the negative consequences extend to all the participants in an economy.

LOS 69.e: Explain professionalism in investment management.

Because clients of investment professionals rely on their expertise, judgment, and ethical principles, many of the characteristics of a profession we have described apply.

Ethical principles are of great importance because clients often do not have significant knowledge about financial securities, fee structures, or sources of potential bias in investment recommendations. Currently, some financial professionals are held to a suitability standard, while others are held to a fiduciary standard. *Suitability* refers to the match between client return requirements and risk tolerances and the characteristics of the securities recommended. A fiduciary standard is stronger, requiring professionals to use their knowledge and expertise to act in the best interests of the client.

LOS 69.f: Identify challenges to ethical behavior.

One challenge to ethical behavior is that individuals tend to overrate the ethical quality of their behavior on a relative basis and overemphasize the importance of their own personal traits in determining the ethical quality of their behavior.

It is claimed that external or situational influences are a more important determinant of the ethical quality of behavior than internal (personal) traits that influence behavior. One situational influence is social pressure from others. Loyalty to an employer, supervisor, organization, or coworkers can cause individuals to act in unethical ways as they place more importance on their self-interest and short-term results than on longer-term results and the ethical quality of their decisions and behavior. The prospect of acquiring more money or greater prestige can cause individuals to engage in unethical behavior.

Firms with strict rules-based compliance procedures run the risk of fostering a culture that is so focused on adhering to compliance rules that individuals only ask themselves what they *can* do. The question of what behavior they *should* engage in, based on ethical principles and longer-term results, is often not addressed in such situations.

LOS 69.g: Compare and contrast ethical standards with legal standards.

Not all unethical actions are illegal, and not all illegal actions are unethical. In some places it may be illegal to report one's employer's actions against the best interests of clients by sharing what is considered private company information with authorities, but doing so may be considered ethical "whistle-blowing" behavior by some. Acts of civil disobedience that are illegal are also considered by many to be ethical behavior. On the other hand, recommending investment in a relative's firm without disclosure may not be illegal, but would be considered unethical by many.

Ethical principles often set a higher standard of behavior than laws and regulations. New laws and regulations often result from recent instances of what is perceived to be unethical behavior. Just as the Securities Act of 1933, the Glass-Steagall Act, and the Securities Exchange Act of 1934 followed the perceived bad behavior by investment professionals and bankers leading to the 1929 market crash, the Sarbanes-Oxley laws followed the accounting scandals at Enron and Worldcom, and the Dodd-Frank Act followed the 2008 financial crisis. New laws and regulations can create opportunities for different unethical behavior. In general, ethical decisions require more judgment and consideration of the impact of behavior on many stakeholders compared to legal decisions.

LOS 69.h: Describe a framework for ethical decision making.

Ethical decisions will be improved when ethics are integrated into a firm's decision-making process. This will allow decision makers and teams to consider alternative actions as well as shorter- and longer-term consequences from various perspectives, improving the ethical aspects of their decisions. To do this it is first necessary that the firm adopt a code of ethics to guide the process.

Such integration provides an opportunity to teach, practice, and reinforce ethical decision-making. This is an important part of developing an ethical culture. The support of senior management for integrating ethics into the decision-making process is also very important in developing a culture and processes that will result in ethical decision-making.

Using a framework for ethical decision-making helps individuals identify the important issues involved, examine these issues from multiple perspectives, develop the necessary judgment and decision-making skills required, and avoid unanticipated ethical consequences.

The following ethical decision-making framework is presented in the Level I CFA curriculum:¹

- Identify: Relevant facts, stakeholders and duties owed, ethical principles, conflicts of interest.
- Consider: Situational influences, additional guidance, alternative actions.
- Decide and act.
- Reflect: Was the outcome as anticipated? Why or why not?

In the first step, decision makers need to identify the facts they have to work with, and the facts they would like to have, before making a decision. Stakeholders—those affected by the decision—must be identified. These stakeholders may include the employer, clients, coworkers, self, family, and others in the industry, and the duties to each stakeholder should be identified. This part of the process will also help in explicitly identifying potential conflicts of interest among the various stakeholders. At this point the decision makers should be able to identify the ethical

principles involved in the decision, although greater clarity about those may also be gained throughout the process.

In the second step, the framework suggests situational factors that may influence decision makers should be identified and considered along with any personal biases that may come into play. At this point, decision makers may seek outside guidance which can come from a mentor, colleagues, or friends who have shown good judgment in the past. Guidance may also be sought from the firm's legal and compliance departments. This guidance from alternative sources will help to provide a variety of perspectives from which the decision under consideration can be viewed, as well as help in developing alternatives that should be considered. Finally, the alternative actions that have been identified are all considered, taking into account both the short-term and long-term effects of each alternative action and any potential but unanticipated ethical implications.

In the final step, decision makers should evaluate the outcomes of the actions that were taken. In particular, they should consider whether the decisions had their intended results and whether appropriate consideration was given to ethical principles, situational influences, and duties to clients and other stakeholders.



MODULE QUIZ 69.1

1. Professional standards of practice:
 - A. are a personal view of acceptable behavior.
 - B. encompass current "best practices."
 - C. specify a minimum level of acceptable conduct.
2. A professional code of conduct:
 - A. can increase public trust in the profession.
 - B. guarantees that members will adhere to a minimum level of ethical conduct.
 - C. includes standards that provide guidance for specific behaviors.
3. Situational factors that influence ethical behavior are *least likely* to include:
 - A. social pressure.
 - B. large financial rewards.
 - C. a lack of ethical principles.
4. Compared to complying with laws and regulations, complying with a code of ethics:
 - A. is considered a lower standard.
 - B. often involves more judgment.
 - C. includes compliance with all laws and regulations.
5. Employing a framework for decision-making that includes the ethical aspects of the decision is *most likely* to:
 - A. lead to higher profits.
 - B. avoid any unintended ethical consequences of decisions.
 - C. balance the interests of various stakeholders.

KEY CONCEPTS

LOS 69.a

Ethical behavior is that which conforms to a set of rules and moral principles based on shared beliefs about what behavior is acceptable and what behavior is unacceptable.

LOS 69.b

A professional code of ethics is a way for a profession to communicate to the public that its members will use their knowledge and skills to serve their clients in an honest and ethical manner, and can increase public confidence and trust that members will act ethically.

LOS 69.c

A profession is an occupational group that has requirements of specialized expert knowledge. Professions establish trust by requiring high standards of expertise, setting standards for ethical behavior, and monitoring professional conduct.

LOS 69.d

Investment professionals have a special responsibility to use their specialized knowledge and skills to both protect and grow client assets. The fact that investment management is an intangible product makes high ethical standards all the more important in the financial services profession.

LOS 69.e

Some financial professionals are held to a suitability standard, while others are held to a fiduciary standard. Suitability refers to the match between client return requirements and risk tolerances and the characteristics of the securities recommended. A fiduciary standard requires professionals to act in the best interests of the client.

LOS 69.f

Challenges to ethical behavior include overestimating one's own ethical character, considering only near-term consequences and not longer-term consequences of behavior, and letting situational (external) influences, such as peer pressure, unduly affect one's decisions and behavior.

LOS 69.g

Not all unethical actions are illegal, and not all illegal actions are unethical. Laws are more specific than ethical principles and often address prior unethical behavior. Ethical behavior requires more judgment; acts such as civil disobedience may be considered ethical even when they are illegal.

LOS 69.h

A framework for ethical decision-making is designed to lead to better decisions by identifying the stakeholders affected and the conflicts of interest among them, considering alternative actions and the relevant situational influences on decision makers, seeking out different perspectives, and evaluating decisions to see if they had unintended consequences.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 69.1

1. **C** Professional standards of practice specify a minimum level of acceptable conduct for a group or organization, whereas “best practices” are suggested behavior, not a minimum acceptable level. (LOS 69.b)
2. **A** A professional code of conduct communicates to the public that members have promised to uphold a minimum level of ethical conduct when acting for clients. This is no guarantee that all members will follow the

code at all times. A code of conduct may include specific standards of behavior or only state principles of conduct without specific standards or guidance. (LOS 69.b)

3. **C** Situational factors are those external to the decision makers, such as financial rewards and desire to please coworkers or others. Researchers have found that external factors are often more likely than a lack of personal ethics to lead to poor ethical decisions. (LOS 69.f)
4. **B** A code of ethics is considered a higher standard of behavior as it goes beyond simply legality of behavior. Compliance with the ethical principles of a code of ethics often requires judgment in balancing the interests of various stakeholders and consideration of short-term effects with longer-term effects of decisions. Some behavior that is illegal, such as civil disobedience or "whistle-blowing," is considered to be ethical behavior by many. (LOS 69.g)
5. **C** A decision-making framework that includes the ethical aspects of the decision should consider the conflicts among the interests of various stakeholders so that decision makers can use the company's stated ethical principles and their judgment to balance these interests in an ethical manner. Profit maximization, at least in the short term, does not necessarily follow from sound ethical judgment. While integrating ethics into the decision-making process can consider and reduce unintended ethical consequences of a decision, avoiding them altogether can never be assured. (LOS 69.h)

¹ "Ethics and Trust in the Investment Profession," Bidhan L Parmar, PhD, Dorothy C. Kelly, CFA, and David B. Stevens, CFA, in CFA Program 2023 Level I Curriculum, Volume 1 (CFA Institute, 2022).

READING 70

CODE OF ETHICS AND STANDARDS OF PROFESSIONAL CONDUCT

EXAM FOCUS

In addition to reading this review of the ethics material, we strongly recommend that all candidates for the CFA® examination read the *Standards of Practice Handbook 11th Edition* (2014) multiple times. As a Level I CFA candidate, it is your responsibility to comply with the *Code and Standards*. The complete *Code and Standards* are reprinted in Volume 6 of the CFA Program Curriculum.

MODULE 70.1: CODE AND STANDARDS



LOS 70.a: Describe the structure of the CFA Institute Professional Conduct Program and the process for the enforcement of the Code and Standards.

Video covering this content is available online.

The CFA Institute Professional Conduct Program is covered by the CFA Institute Bylaws and the Rules of Procedure for Proceedings Related to Professional Conduct. The Program is based on the principles of fairness of the process to members and candidates and maintaining the confidentiality of the proceedings. The CFA Institute Board of Governors has overall responsibility for the Professional Conduct Program and its Disciplinary Review Committee is responsible for enforcing of the Code and Standards.

The CFA Institute Professional Conduct staff conducts inquiries related to professional conduct. Several circumstances can prompt such an inquiry:

1. Self-disclosure by members or candidates on their annual Professional Conduct Statements of involvement in civil litigation or a criminal investigation, or that the member or candidate is the subject of a written complaint.
2. Written complaints about a member or candidate's professional conduct that are received by the Professional Conduct staff.
3. Evidence of misconduct by a member or candidate that the Professional Conduct staff received through public sources, such as a media article or broadcast.
4. A report by a CFA exam proctor of a possible violation during the examination.
5. Analysis of exam materials and monitoring of social media by CFA Institute.

Once an inquiry has begun, the Professional Conduct staff may request (in writing) an explanation from the subject member or candidate and may: (1) interview the subject member or candidate, (2) interview the complainant or other third parties, and/or (3) collect documents and records relevant to the investigation.

The Professional Conduct staff may decide: (1) that no disciplinary sanctions are appropriate, (2) to issue a cautionary letter, or (3) to discipline the member or candidate. In a case where the Professional Conduct staff finds a violation has occurred and proposes a disciplinary sanction, the member or candidate may accept or reject the sanction. If the member or candidate chooses to reject the sanction, the matter will be referred to a disciplinary review panel of CFA Institute members for a hearing. Sanctions imposed may include condemnation by the member's peers or suspension of candidate's continued participation in the CFA Program.

LOS 70.b: Identify the six components of the Code of Ethics and the seven Standards of Professional Conduct.

Code of Ethics

Members of CFA Institute [including Chartered Financial Analyst® (CFA®) charterholders] and candidates for the CFA designation ("Members and Candidates") must:¹

- Act with integrity, competence, diligence, respect, and in an ethical manner with the public, clients, prospective clients, employers, employees, colleagues in the investment profession, and other participants in the global capital markets.
- Place the integrity of the investment profession and the interests of clients above their own personal interests.
- Use reasonable care and exercise independent professional judgment when conducting investment analysis, making investment recommendations, taking investment actions, and engaging in other professional activities.
- Practice and encourage others to practice in a professional and ethical manner that will reflect credit on themselves and the profession.
- Promote the integrity and viability of the global capital markets for the ultimate benefit of society.
- Maintain and improve their professional competence and strive to maintain and improve the competence of other investment professionals.

The Standards of Professional Conduct

- I. Professionalism
- II. Integrity of Capital Markets
- III. Duties to Clients
- IV. Duties to Employers
- V. Investment Analysis, Recommendations, and Actions
- VI. Conflicts of Interest
- VII. Responsibilities as a CFA Institute Member or CFA Candidate

LOS 70.c: Explain the ethical responsibilities required by the Code and Standards, including the sub-sections of each Standard.

Standards of Professional Conduct²

I. PROFESSIONALISM

- A. **Knowledge of the Law.** Members and Candidates must understand and comply with all applicable laws, rules, and regulations (including the CFA Institute *Code of Ethics and Standards of Professional Conduct*) of any government, regulatory organization, licensing agency, or professional association governing their professional activities. In the event of conflict, Members and Candidates must comply with the more strict law, rule, or regulation. Members and Candidates must not knowingly participate or assist in any violation of laws, rules, or regulations and must dissociate themselves from any such violation.
- B. **Independence and Objectivity.** Members and Candidates must use reasonable care and judgment to achieve and maintain independence and objectivity in their professional activities. Members and Candidates must not offer, solicit, or accept any gift, benefit, compensation, or consideration that reasonably could be expected to compromise their own or another's independence and objectivity.
- C. **Misrepresentation.** Members and Candidates must not knowingly make any misrepresentations relating to investment analysis, recommendations, actions, or other professional activities.
- D. **Misconduct.** Members and Candidates must not engage in any professional conduct involving dishonesty, fraud, or deceit or commit any act that reflects adversely on their professional reputation, integrity, or competence.

II. INTEGRITY OF CAPITAL MARKETS

- A. **Material Nonpublic Information.** Members and Candidates who possess material nonpublic information that could affect the value of an investment must not act or cause others to act on the information.
- B. **Market Manipulation.** Members and Candidates must not engage in practices that distort prices or artificially inflate trading volume with the intent to mislead market participants.

III. DUTIES TO CLIENTS

- A. **Loyalty, Prudence, and Care.** Members and Candidates have a duty of loyalty to their clients and must act with reasonable care and exercise prudent judgment. Members and Candidates must act for the benefit of their clients and place their clients' interests before their employer's or their own interests.
- B. **Fair Dealing.** Members and Candidates must deal fairly and objectively with all clients when providing investment analysis, making investment recommendations, taking investment action, or engaging in other professional activities.
- C. **Suitability.**
 - 1. When Members and Candidates are in an advisory relationship with a client, they must:
 - a. Make a reasonable inquiry into a client's or prospective clients' investment experience, risk and return objectives, and financial constraints prior to making any investment recommendation or taking investment action and must reassess and update this information regularly.
 - b. Determine that an investment is suitable to the client's financial situation and consistent with the client's written objectives, mandates, and constraints before making an investment recommendation or taking investment action.
 - c. Judge the suitability of investments in the context of the client's total portfolio.
 - 2. When Members and Candidates are responsible for managing a portfolio to a specific mandate, strategy, or style, they must make only investment recommendations or take only investment actions that are consistent with the stated objectives and constraints of the portfolio.

- D. Performance Presentation.** When communicating investment performance information, Members or Candidates must make reasonable efforts to ensure that it is fair, accurate, and complete.
- E. Preservation of Confidentiality.** Members and Candidates must keep information about current, former, and prospective clients confidential unless:
1. The information concerns illegal activities on the part of the client or prospective client,
 2. Disclosure is required by law, or
 3. The client or prospective client permits disclosure of the information.

IV. DUTIES TO EMPLOYERS

- A. Loyalty.** In matters related to their employment, Members and Candidates must act for the benefit of their employer and not deprive their employer of the advantage of their skills and abilities, divulge confidential information, or otherwise cause harm to their employer.
- B. Additional Compensation Arrangements.** Members and Candidates must not accept gifts, benefits, compensation, or consideration that competes with or might reasonably be expected to create a conflict of interest with their employer's interest unless they obtain written consent from all parties involved.
- C. Responsibilities of Supervisors.** Members and Candidates must make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards.

V. INVESTMENT ANALYSIS, RECOMMENDATIONS, AND ACTIONS

- A. Diligence and Reasonable Basis.** Members and Candidates must:
 1. Exercise diligence, independence, and thoroughness in analyzing investments, making investment recommendations, and taking investment actions.
 2. Have a reasonable and adequate basis, supported by appropriate research and investigation, for any investment analysis, recommendation, or action.
- B. Communication with Clients and Prospective Clients.** Members and Candidates must:
 1. Disclose to clients and prospective clients the basic format and general principles of the investment processes they use to analyze investments, select securities, and construct portfolios and must promptly disclose any changes that might materially affect those processes.
 2. Disclose to clients and prospective clients significant limitations and risks associated with the investment process.
 3. Use reasonable judgment in identifying which factors are important to their investment analyses, recommendations, or actions and include those factors in communications with clients and prospective clients.
 4. Distinguish between fact and opinion in the presentation of investment analysis and recommendations.
- C. Record Retention.** Members and Candidates must develop and maintain appropriate records to support their investment analysis, recommendations, actions, and other investment-related communications with clients and prospective clients.

VI. CONFLICTS OF INTEREST

- A. Disclosure of Conflicts.** Members and Candidates must make full and fair disclosure of all matters that could reasonably be expected to impair their independence and objectivity or interfere with respective duties to their clients, prospective clients, and employer. Members and Candidates must ensure that such disclosures are prominent, are delivered in plain language, and communicate the relevant information effectively.
- B. Priority of Transactions.** Investment transactions for clients and employers must have priority over investment transactions in which a Member or Candidate is the beneficial owner.

- C. Referral Fees.** Members and Candidates must disclose to their employer, clients, and prospective clients, as appropriate, any compensation, consideration, or benefit received by, or paid to, others for the recommendation of products or services.

VII. RESPONSIBILITIES AS A CFA INSTITUTE MEMBER OR CFA CANDIDATE

- A. Conduct as Participants in CFA Institute Programs.** Members and Candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of CFA Institute programs.
- B. Reference to CFA Institute, the CFA Designation, and the CFA Program.** When referring to CFA Institute, CFA Institute membership, the CFA designation, or candidacy in the CFA Program, Members and Candidates must not misrepresent or exaggerate the meaning or implications of membership in CFA Institute, holding the CFA designation, or candidacy in the CFA Program.

MODULE QUIZ 70.1



1. In the case of a complaint about a member's professional conduct, CFA Institute Professional Conduct Program staff are *least likely* to:
 - A. review documents and records related to the complaint.
 - B. request an interview with the member or with the party making the complaint.
 - C. suspend the member's right to use the CFA designation while an investigation is in progress.
2. Which of the following requirements for members and candidates is one of the six components of the Code of Ethics?
 - A. Maintain and improve their professional competence.
 - B. Do not act or cause others to act on material nonpublic information.
 - C. Distinguish between fact and opinion when presenting investment analysis.
3. If a member or candidate is offered an additional compensation arrangement by a client, which of the seven Standards of Professional Conduct states the requirements the member or candidate must follow?
 - A. Duties to Clients.
 - B. Conflicts of Interest.
 - C. Duties to Employers.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 70.1

1. **C** The process for enforcing the Code and Standards does not include suspending a member or candidate while an inquiry is in progress. If CFA Institute Professional Conduct staff receive information that prompts an inquiry, the staff may request information from the member or candidate, interview parties who initiated a complaint, or review relevant records and documents. (LOS 70.a)
2. **A** One of the six components of the Code of Ethics requires members and candidates to "maintain and improve their professional competence and strive to maintain and improve the competence of other investment professionals." The other two answer choices are required by the Standards of Professional Conduct but are not components of the Code of Ethics. (LOS 70.b)
3. **C** The standard related to additional compensation arrangements is a subsection of Standard IV Duties to Employers. (LOS 70.c)

¹ Copyright 2014, CFA Institute. Reproduced and republished from "The Code of Ethics," from *Standards of Practice Handbook, 11th Ed.*, 2014, with permission from CFA Institute. All rights reserved.

² Ibid.

READING 71

GUIDANCE FOR STANDARDS I–VII

EXAM FOCUS

The Standards of Professional Conduct comprise seven Standards (I–VII) and a total of 22 subsections. These Standards and their application are described in the *Standards of Practice Handbook, 11th Edition, 2014*, published by CFA Institute. We recommend carefully reading the *Standards of Practice Handbook* multiple times in preparation for your Level I exam (yes, the whole thing, including all examples). Fifteen percent of your exam questions will be based on this book and the two relatively short readings concerning Global Investment Performance Standards (GIPS). Given that much of this material must simply be memorized, we also suggest that your final reading of the *Standards of Practice Handbook* be on the Friday prior to your exam. You probably don't need to read all the examples that day, but if you highlighted some points during an earlier reading, you can revisit those as you go through all the Standards, the guidance, and the recommended best practices.

MODULE 71.1: GUIDANCE FOR STANDARDS I(A) AND I(B)



Video covering
this content is
available online.

LOS 71.a: Demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity.

LOS 71.b: Recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

LOS 71.c: Identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards.

The *Standards of Practice Handbook* is included in its entirety in the CFA curriculum as Readings 2 and 3 of Volume 1. You can also download a PDF of the *Standards of Practice Handbook* at the website for CFA Institute, www.cfainstitute.org. A third alternative is to purchase the *Standards of Practice Handbook* through Amazon (make sure you get the 11th edition) for about \$30 or get the Kindle edition for \$0.99.

In our summary of the Standards of Practice, we focus on describing three things: (1) actions that clearly violate the subsection, (2) the behaviors that each subsection is intended to either encourage or discourage, and (3) recommended best practices for members and their firms.

In many cases the actions that members and candidates must not take are explained using terms open to interpretation, such as “reasonable,” “adequate,” and “token.”

Some examples from the Standards themselves are:

- ...use reasonable care and judgment to achieve...
- ...accept any gift, that reasonably could be expected to compromise...
- ...act with reasonable care and exercise prudent judgment...
- ...deal fairly and objectively with all clients...
- ...make a reasonable inquiry into...
- ...make reasonable efforts to ensure...
- ...might reasonably be expected to create a conflict of interest with...
- ...Have a reasonable and adequate basis...
- ...Use reasonable judgment in...
- ...matters that could be reasonably expected to impair...

The requirement of the LOS is that you know what constitutes a violation, not that you draw a distinction between what is “reasonable” and what is not in a given situation. We believe the exam writers take this into account and that if they intend, for example, to test whether a recommendation has been given without reasonable care and judgment, it will likely be clear either that the care and judgment exhibited by the analyst did not rise to the level of “reasonable,” or that it did.

No monetary value for a “token” gift is given in the Standards, although it is recommended that a firm establish such a monetary value for its employees. Here, again, the correct answer to a question will not likely hinge on candidate’s determination of what is a token gift and what is not. Questions should be clear in this regard. A business dinner is likely a token gift, but a week at a condominium in Aspen or tickets to the Super Bowl are likely not. Always look for clues in the questions that lead you to the question-writer’s preferred answer choice, such as “lavish” entertainment and “luxury” accommodations.

Next, we present a summary of each subsection of the Standards of Professional Conduct. For each one, we first detail actions that violate the Standard and then list actions and behaviors that are recommended within the Standards. We suggest you learn the violations especially well so you understand that the other items are recommended. For the exam, it is not necessary to memorize the Standard number and subsection letter. Knowing that an action violates, for example, Professionalism, rather than Duties to Employers or Duties to Clients, should be sufficient in this regard. Note that some actions may violate more than one Standard.

One way to write questions for this material is to offer a reason that might make one believe a Standard does not apply in a particular situation. In most, if not all, cases the “reason” does not change the requirement of the Standard. If you are prohibited from some action, the motivations for the action or other circumstances simply do not matter. If the Standard says it’s a violation, it’s a violation. An exception is when intent is key to the Standard, such as intending to mislead clients or market participants in general.

STANDARD I: PROFESSIONALISM¹

Standard I(A) Knowledge of the Law

Members and Candidates must understand and comply with all applicable laws, rules, and regulations (including the CFA Institute Code of Ethics and Standards of Professional Conduct) of any government, regulatory organization, licensing agency, or professional association governing their professional activities. In the event of conflict, Members and Candidates must comply with the more strict law, rule, or regulation. Members and Candidates must not knowingly participate or assist in and must dissociate from any violation of such laws, rules, or regulations.

The Standards begin with a straightforward statement: Don't violate any laws, rules, or regulations that apply to your professional activities. This includes the Code and Standards, so any violation of the Code and Standards will also violate this subsection.

A member may be governed by different rules and regulations among the Standards, the country in which the member resides, and the country where the member is doing business. Follow the most strict of these, or, put another way, do not violate any of the three sets of rules and regulations.

If you know that violations of applicable rules or laws are taking place, either by coworkers or clients, CFA Institute strongly encourages members and candidates to report potential violations.² One way to do so is to approach your supervisor or compliance department to remedy the situation. If they will not or cannot, then you must dissociate from the activity (e.g., not working with a trading group you know is not allocating client trades properly according to the Standard on Fair Dealing, or not using marketing materials that you know or should know are misleading or erroneous). If this cannot be accomplished, you may, in an extreme case, have to resign from the firm to be in compliance with this Standard.

Recommendations for Members

- Establish, or encourage employer to establish, procedures to keep employees informed of changes in relevant laws, rules, and regulations.
- Review, or encourage employer to review, the firm's written compliance procedures on a regular basis.
- Maintain, or encourage employer to maintain, copies of current laws, rules, and regulations.
- When in doubt about legality, consult supervisor, compliance personnel, or a lawyer.
- When dissociating from violations, keep records documenting the violations, encourage employer to bring an end to the violations.
- There is no requirement in the Standards to report wrongdoers, but local law may require it; members are "strongly encouraged" to report violations to CFA Institute Professional Conduct Program.

Recommendations for Firms

- Have a code of ethics.
- Provide employees with information on laws, rules, and regulations governing professional activities.

- Have procedures for reporting suspected violations.

Standard I(B) Independence and Objectivity

Members and Candidates must use reasonable care and judgment to achieve and maintain independence and objectivity in their professional activities. Members and Candidates must not offer, solicit, or accept any gift, benefit, compensation, or consideration that reasonably could be expected to compromise their own or another's independence and objectivity.

Analysts may face pressure or receive inducements to give a security a specific rating, to select certain outside managers or vendors, or to produce favorable or unfavorable research and conclusions. Members who allow their investment recommendations or analysis to be influenced by such pressure or inducements will have violated the requirement to use reasonable care and to maintain independence and objectivity in their professional activities. Allocating shares in oversubscribed IPOs to personal accounts is a violation.

Normal business entertainment is permitted. Members who accept, solicit, or offer things of value that could be expected to influence the member's or others' independence or objectivity are violating the Standard. Gifts from clients are considered less likely to compromise independence and objectivity than gifts from other parties. Client gifts must be disclosed to the member's employer prior to acceptance, if possible, but after acceptance, if not.

Members may prepare reports paid for by the subject firm if compensation is a flat rate not tied to the conclusions of the report (and if the fact that the research is issuer-paid is disclosed). Accepting compensation that is dependent on the conclusions, recommendations, or market impact of the report, and failure to disclose that research is issuer-paid, are violations of this Standard.

Recommendations for Members

Members or their firms should pay for their own travel to company events or tours when practicable and limit use of corporate aircraft to trips for which commercial travel is not an alternative.

Recommendations for Firms

- Restrict employee participation in IPOs and private placements, require pre-approval for participation.
- Appoint a compliance officer, have written policies on independence and objectivity and clear procedures for reporting violations.
- Limit gifts, other than from clients, to token items only.

MODULE 71.2: GUIDANCE FOR STANDARDS I(C) AND I(D)



Video covering
this content is
available online.

Standard I(C) Misrepresentation

Members and Candidates must not knowingly make any misrepresentations relating to investment analysis, recommendations, actions, or other professional activities.

Misrepresentation includes knowingly misleading investors, omitting relevant information, presenting selective data to mislead investors, and plagiarism. Plagiarism is using reports, forecasts, models, ideas, charts, graphs, or spreadsheets created by others without crediting the source. Crediting the source is not required when using projections, statistics, and tables from recognized financial and statistical reporting services. When using models developed or research done by other members of the firm, it is permitted to omit the names of those who are no longer with the firm as long as the member does not represent work previously done by others as his alone.

Actions that would violate the Standard include:

- Presenting third-party research as your own, without attribution to the source.
- Guaranteeing a specific return on securities that do not have an explicit guarantee from a government body or financial institution.
- Selecting a valuation service because it puts the highest value on untraded security holdings.
- Selecting a performance benchmark that is not comparable to the investment strategy employed.
- Presenting performance data or attribution analysis that omits accounts or relevant variables.
- Offering false or misleading information about the analyst's or firm's capabilities, expertise, or experience.
- Using marketing materials from a third party (outside advisor) that are misleading.

Recommendations for Members

- Prepare a summary of experience, qualifications, and services a member is able to perform.
- Encourage employers to develop procedures for verifying marketing materials provided by third parties concerning their capabilities, products, and services.
- Cite the source of any summaries of materials provided by others.
- Keep copies of all reports, articles, or other materials used in the preparation of research reports.
- Provide a list, in writing, of the firm's available services and qualifications.
- Periodically review documents and communications of members for any misrepresentation of employee or firm qualifications and capabilities.

Standard I(D) Misconduct

Members and Candidates must not engage in any professional conduct involving dishonesty, fraud, or deceit or commit any act that reflects adversely on their professional reputation, integrity, or competence.

The first part here regarding professional conduct is clear: no dishonesty, fraud, or deceit. The second part, while it applies to all conduct by the member, specifically requires that the act, "reflects adversely on their professional reputation, integrity, or competence." The guidance states, in fact, that members must not try to use enforcement of this Standard against another member to settle personal, political, or other disputes that are not related to professional ethics or competence.

Recommendations for Firms

- Develop and adopt a code of ethics and make clear that unethical behavior will not be tolerated.
- Give employees a list of potential violations and sanctions, including dismissal.
- Check references of potential employees.

MODULE 71.3: GUIDANCE FOR STANDARD II



STANDARD II: INTEGRITY OF CAPITAL MARKETS

Video covering
this content is
available online.

Standard II(A) Material Nonpublic Information

Members and Candidates who possess material nonpublic information that could affect the value of an investment must not act or cause others to act on the information.

Information is “material” if its disclosure would affect the price of a security or if a reasonable investor would want the information before making an investment decision. Information that is ambiguous as to its likely effect on price may not be considered material.

Information is “nonpublic” until it has been made available to the marketplace. An analyst conference call is not public disclosure. Selective disclosure of information by corporations creates the potential for insider-trading violations.

The prohibition against acting on material nonpublic information extends to mutual funds containing the subject securities as well as related swaps and options contracts. It is the member’s responsibility to determine if information she receives has been publicly disseminated prior acting or causing others to act on it.

Some members and candidates may be involved in transactions during which they are provided with material nonpublic information by firms (e.g., investment banking transactions). Members and candidates may use this information for its intended purpose, but must not use the information for any other purpose unless it becomes public information.

Under the so-called **mosaic theory**, reaching an investment conclusion through perceptive analysis of public information combined with non-material nonpublic information is not a violation of the Standard.

Recommendations for Members

- Make reasonable efforts to achieve public dissemination by the firm of information they possess.
- Encourage their firms to adopt procedures to prevent the misuse of material nonpublic information.

Recommendations for Firms

Use a firewall within the firm, with elements including:

- Exercise substantial control of relevant interdepartmental communications through a clearance area, such as the compliance or legal department.
- Review employee trades.
- Maintain “watch,” “restricted,” and “rumor” lists.

Monitor and restrict proprietary trading while a firm is in possession of material nonpublic information. However, prohibiting all proprietary trading while a firm is in possession of material nonpublic information may be inappropriate because it may send a signal to the market. In these cases, firms should only take the opposite side of unsolicited customer trades.

Standard II(B) Market Manipulation

Members and Candidates must not engage in practices that distort prices or artificially inflate trading volume with the intent to mislead market participants.

Member actions may affect security values and trading volumes without violating this Standard. The key point here is that if there is the *intent to mislead*, then the Standard is violated. Of course, spreading false information to affect prices or volume is a violation of this Standard as is making trades intended to mislead market participants.



MODULE QUIZ 71.1, 71.2, 71.3

1. In situations where the laws of a member or candidate’s country of residence, the local laws of regions where the member or candidate does business, and the Code and Standards specify different requirements, the member or candidate must abide by:
 - A. local law or the Code and Standards, whichever is stricter.
 - B. the Code and Standards or his country’s laws, whichever are stricter.
 - C. the strictest of local law, his country’s laws, or the Code and Standards.
2. According to the Standard on independence and objectivity, members and candidates:
 - A. may accept gifts or bonuses from clients.
 - B. may not accept compensation from an issuer of securities in return for producing research on those securities.
 - C. should consider credit ratings issued by recognized agencies to be objective measures of credit quality.
3. Bill Cooper finds a table of historical bond yields on the website of the U.S. Treasury that supports the work he has done in his analysis and includes the table as part of his report without citing the source. Has Cooper violated the Code and Standards?
 - A. Yes, because he did not cite the source of the table.
 - B. Yes, because he did not verify the accuracy of the information.
 - C. No, because the table is from a recognized source of financial or statistical data.
4. Which of the following statements about the Standard on misconduct is *most accurate*?
 - A. Misconduct applies only to a member or candidate’s professional activities.
 - B. Neglecting to perform due diligence when required is an example of misconduct.
 - C. A member or candidate commits misconduct by engaging in any illegal activity.
5. Ed Ingus, CFA, visits the headquarters and main plant of Bullitt Company and observes that inventories of unsold goods appear unusually large. From the CFO, he learns that a recent increase in returned items may result in earnings for the current quarter that are below analysts’ estimates. Based on his visit, Ingus changes his recommendation on Bullitt to “Sell.” Has Ingus violated the Standard concerning material nonpublic information?
 - A. Yes.
 - B. No, because the information he used is not material.
 - C. No, because his actions are consistent with the mosaic theory.

6. Green Brothers, an emerging market fund manager, has two of its subsidiaries simultaneously buy and sell emerging market stocks. In its marketing literature, Green Brothers cites the overall emerging market volume as evidence of the market's liquidity. As a result of its actions, more investors participate in the emerging markets fund. Green Brothers *most likely*:
- A. did not violate the Code and Standards.
 - B. violated the Standard regarding market manipulation.
 - C. violated the Standard regarding performance presentation.

MODULE 71.4: GUIDANCE FOR STANDARDS III(A) AND III(B)



Video covering
this content is
available online.

STANDARD III: DUTIES TO CLIENTS

Standard III(A) Loyalty, Prudence, and Care

Members and Candidates have a duty of loyalty to their clients and must act with reasonable care and exercise prudent judgment. Members and Candidates must act for the benefit of their clients and place their clients' interests before their employer's or their own interests.

Client interests always come first. Although this Standard does not impose a fiduciary duty on members or candidates where one did not already exist, it does require members and candidates to act in their clients' best interests and recommend products that are suitable given their clients' investment objectives and risk tolerances. Members and candidates must:

- Exercise the prudence, care, skill, and diligence under the circumstances that a person acting in a like capacity and familiar with such matters would use.
- Manage pools of client assets in accordance with the terms of the governing documents, such as trust documents or investment management agreements.
- Make investment decisions in the context of the total portfolio.
- Inform clients of any limitations in an advisory relationship (e.g., an advisor who may only recommend her own firm's products).
- Vote proxies in an informed and responsible manner. Due to cost-benefit considerations, it may not be necessary to vote all proxies.
- Client brokerage, or "soft dollars" or "soft commissions," must be used to benefit the client.
- The "client" may be the investing public as a whole rather than a specific entity or person.

Recommendations for Members

Submit to clients, at least quarterly, itemized statements showing all securities in custody and all debits, credits, and transactions.

Encourage firms to address these topics when drafting policies and procedures regarding fiduciary duty:

- Follow applicable rules and laws.
- Establish investment objectives of client.
- Consider suitability of a portfolio relative to the client's needs and circumstances, the investment's basic characteristics, or the basic characteristics of the total portfolio.

- Diversify.
- Deal fairly with all clients in regard to investment actions.
- Disclose conflicts.
- Disclose compensation arrangements.
- Vote proxies in the best interest of clients and ultimate beneficiaries.
- Maintain confidentiality.
- Seek best execution.

Standard III(B) Fair Dealing

Members and Candidates must deal fairly and objectively with all clients when providing investment analysis, making investment recommendations, taking investment action, or engaging in other professional activities.

Do not discriminate against any clients when disseminating recommendations or taking investment action. “Fairly” does not mean “equally.” In the normal course of business, there will be differences in the time emails, faxes, and other communications are received by different clients.

Different service levels are acceptable, but they must not negatively affect or disadvantage any clients. Disclose the different service levels to all clients and prospects, and make premium levels of service available to all those willing to pay for them.

Give all clients a fair opportunity to act on every recommendation. Clients who are unaware of a change in the recommendation for a security should be advised of the change before an order for the security is accepted.

Treat clients fairly in light of their investment objectives and circumstances. Treat both individual and institutional clients in a fair and impartial manner. Members and candidates should not take advantage of their position in the industry to disadvantage clients (e.g., taking shares of an oversubscribed IPO).

Recommendations for Members

- Encourage firms to establish compliance procedures requiring proper dissemination of investment recommendations and fair treatment of all customers and clients.
- Maintain a list of clients and holdings—use to ensure that all holders are treated fairly.

Recommendations for Firms

- Limit the number of people who are aware that a change in recommendation will be made.
- Shorten the time frame between decision and dissemination.
- Publish personnel guidelines for pre-dissemination—have in place guidelines prohibiting personnel who have prior knowledge of a recommendation from discussing it or taking action on the pending recommendation.
- Disseminate new or changed recommendations simultaneously to all clients who have expressed an interest or for whom an investment is suitable.

- Develop written trade allocation procedures—ensure fairness to clients, timely and efficient order execution, and accuracy of client positions.
- Disclose trade allocation procedures.
- Establish systematic account review—ensure that no client is given preferred treatment and that investment actions are consistent with the account's objectives.
- Disclose available levels of service.

MODULE 71.5: GUIDANCE FOR STANDARDS III(C), III(D), AND III(E)



Video covering this content is available online.

Standard III(C) Suitability

1. When Members and Candidates are in an advisory relationship with a client, they must:
 - a. Make a reasonable inquiry into a client's or prospective client's investment experience, risk and return objectives, and financial constraints prior to making any investment recommendation or taking investment action and must reassess and update this information regularly.
 - b. Determine that an investment is suitable to the client's financial situation and consistent with the client's written objectives, mandates, and constraints before making an investment recommendation or taking investment action.
 - c. Judge the suitability of investments in the context of the client's total portfolio.
2. When Members and Candidates are responsible for managing a portfolio to a specific mandate, strategy, or style, they must make only investment recommendations or take only investment actions that are consistent with the stated objectives and constraints of the portfolio.

In advisory relationships, members must gather client information at the beginning of the relationship, in the form of an investment policy statement (IPS). Consider clients' needs and circumstances and, thus, their risk tolerance. Consider whether or not the use of leverage is suitable for the client.

If a member is responsible for managing a fund to an index or other stated mandate, he must select only investments that are consistent with the stated mandate.

Unsolicited Trade Requests

An investment manager may receive a client request to purchase a security that the manager knows is unsuitable, given the client's investment policy statement. The trade may or may not have a material effect on the risk characteristics of the client's total portfolio and the requirements are different for each case. In either case, however, the manager should not make the trade until he has discussed with the client the reasons (based on the IPS) that the trade is unsuitable for the client's account.

If the manager determines that the *effect on the risk/return profile of the client's total portfolio is minimal*, the manager, after discussing with the client how the trade does not fit the IPS goals and constraints, may follow his firm's policy with regard to unsuitable trades. Regardless of firm

policy, the client must acknowledge the discussion and an understanding of why the trade is unsuitable.

If the trade would have a *material impact on the risk/return profile of the client's total portfolio*, one option is to update the IPS so the client accepts a changed risk profile that would permit the trade. If the client will not accept a changed IPS, the manager may follow firm policy, which may allow the trade to be made in a separate client-directed account. In the absence of other options, the manager may need to reconsider whether to maintain the relationship with the client.

Recommendations for Members

- For each client, put the needs, circumstances, and investment objectives into a written IPS.
- Consider the type of client and whether there are separate beneficiaries, investor objectives (return and risk), investor constraints (liquidity needs, expected cash flows, time, tax, and regulatory and legal circumstances), and performance measurement benchmarks.
- Review the investor's objectives and constraints periodically to reflect any changes in client circumstances.

Standard III(D) Performance Presentation

When communicating investment performance information, Members and Candidates must make reasonable efforts to ensure that it is fair, accurate, and complete.

Members must not misstate performance or mislead clients or prospects about their investment performance or their firm's investment performance.

Members must not misrepresent past performance or reasonably expected performance, and must not state or imply the ability to achieve a rate of return similar to that achieved in the past.

For brief presentations, members must make detailed information available on request and indicate that the presentation has offered only limited information.

Recommendations for Members

- Encourage firms to adhere to Global Investment Performance Standards.
- Consider the sophistication of the audience to whom a performance presentation is addressed.
- Present the performance of a weighted composite of similar portfolios rather than the performance of a single account.
- Include terminated accounts as part of historical performance and clearly state when they were terminated.
- Include all appropriate disclosures to fully explain results (e.g., model results included, gross or net of fees, etc.).
- Maintain data and records used to calculate the performance being presented.

Standard III(E) Preservation of Confidentiality

Members and Candidates must keep information about current, former, and prospective clients confidential unless:

1. The information concerns illegal activities on the part of the client;
2. Disclosure is required by law; or
3. The client or prospective client permits disclosure of the information.

If illegal activities by a client are involved, members may have an obligation to report the activities to authorities.

The confidentiality Standard extends to former clients as well.

The requirements of this Standard are not intended to prevent members and candidates from cooperating with a CFA Institute Professional Conduct Program (PCP) investigation.

Recommendations for Members

- Members should avoid disclosing information received from a client except to authorized coworkers who are also working for the client.
- Members should follow firm procedures for storage of electronic data and recommend adoption of such procedures if they are not in place.



MODULE QUIZ 71.4, 71.5

1. Cobb, Inc., has hired Jude Kasten, CFA, to manage its pension fund. The client(s) to whom Kasten owes a duty of loyalty are:
 - A. Cobb's management.
 - B. the shareholders of Cobb, Inc.
 - C. the beneficiaries of the pension fund.
2. Which of the following actions is *most likely* a violation of the Standard on fair dealing?
 - A. A portfolio manager allocates IPO shares to all client accounts, including her brother's fee-based retirement account.
 - B. An investment firm routinely begins trading for its own account immediately after announcing recommendation changes to clients.
 - C. After releasing a general recommendation to all clients, an analyst calls the firm's largest institutional clients to discuss the recommendation in more detail.
3. The Standard regarding suitability *most likely* requires that:
 - A. an advisor must analyze an investment's suitability for the client prior to recommending or acting on the investment.
 - B. a member or candidate must decline to carry out an unsolicited transaction that she believes is unsuitable for the client.
 - C. when managing a fund to an index, a manager who is evaluating potential investments must consider their suitability for the fund's shareholders.
4. Which of the following is *most likely* a recommended procedure for complying with the Standard on performance presentation?
 - A. Exclude terminated accounts from past performance history.
 - B. Present the performance of a representative account to show how a composite has performed.
 - C. Consider the level of financial knowledge of the audience to whom the performance is presented.
5. The CFA Institute Professional Conduct Program (PCP) has begun an investigation into Chris Jones, a Level II CFA candidate, and a number of his CFA charterholder colleagues. Jones has access to confidential client records that could be useful in clearing his name and wishes to share this information with the PCP. Which of the following *most accurately* describes Jones's duties with regard to preservation of confidentiality?

- A. Sharing the confidential information with the PCP would violate the Standards.
- B. The Standards encourage, but do not require, that Jones support the PCP investigation into his colleagues.
- C. Jones may share confidential information about former clients with the PCP but may not share confidential information about current clients.

MODULE 71.6: GUIDANCE FOR STANDARD IV



STANDARD IV: DUTIES TO EMPLOYERS

Video covering this content is available online.

Standard IV(A) Loyalty

In matters related to their employment, Members and Candidates must act for the benefit of their employer and not deprive their employer of the advantage of their skills and abilities, divulge confidential information, or otherwise cause harm to their employer.

This Standard is applicable to employees. If members are independent contractors, rather than employees, they have a duty to abide by the terms of their agreements.

Members must not engage in any activities that would injure the firm, deprive it of profit, or deprive it of the advantage of employees' skills and abilities.

Members should always place client interests above interests of their employer, but consider the effects of their actions on firm integrity and sustainability.

There is no requirement that the employee put employer interests ahead of family and other personal obligations; it is expected that employers and employees will discuss such matters and balance these obligations with work obligations.

There may be isolated cases where a duty to one's employer may be violated in order to protect clients or the integrity of the market, when the actions are not for personal gain.

Independent practice for compensation is allowed if a notification is provided to the employer fully describing all aspects of the services, including compensation, duration, and the nature of the activities and the employer consents to all terms of the proposed independent practice before it begins.

When leaving an employer, members must continue to act in their employer's best interests until their resignation is effective. Activities that may constitute a violation include:

- Misappropriation of trade secrets.
- Misuse of confidential information.
- Soliciting employer's clients prior to leaving.
- Self-dealing.
- Misappropriation of client lists.

Employer records on any medium (e.g., home computer, tablet, cell phone) are the property of the firm.

When an employee has left a firm, simple knowledge of names and existence of former clients is generally not confidential. There is also no prohibition on the use of experience or knowledge

gained while with a former employer. If an agreement exists among employers (e.g., the U.S. “Protocol for Broker Recruiting”) that permits brokers to take certain client information when leaving a firm, a member may act within the terms of the agreement without violating the Standard.

Members and candidates must adhere to their employers’ policies concerning social media. When planning to leave an employer, members and candidates must ensure that their social media use complies with their employers’ policies for notifying clients about employee separations.

Recommendations for Members

Members are encouraged to give their employer a copy of the Code and Standards.

Best practice is to use separate social media accounts for personal and professional communications.

Recommendations for Firms

Employers should not have incentive and compensation systems that encourage unethical behavior.

Standard IV(B) Additional Compensation Arrangements

Members and Candidates must not accept gifts, benefits, compensation, or consideration that competes with or might reasonably be expected to create a conflict of interest with their employer’s interest unless they obtain written consent from all parties involved.

Compensation includes direct and indirect compensation from a client and other benefits received from third parties.

Written consent from a member’s employer includes email communication.

Understand the difference between an additional compensation arrangement and a gift from a client:

- If a client offers a bonus that depends on the *future performance* of her account, this is an additional compensation arrangement that requires written consent in advance.
- If a client offers a bonus to reward a member for her account’s *past performance*, this is a gift that requires disclosure to the member’s employer to comply with Standard I(B) Independence and Objectivity.

Recommendations for Members

Make an immediate written report to the employer detailing any proposed compensation and services, if additional to that provided by the employer.

Members and candidates who are hired to work part time should discuss any arrangements that may compete with their employer’s interest at the time they are hired and abide by any limitations their employer identifies.

Recommendations for Firms

Details of additional compensation, including any performance incentives, should be verified by the offering party.

Standard IV(C) Responsibilities of Supervisors

Members and Candidates must make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards.

Members must make reasonable efforts to prevent employees from violating laws, rules, regulations, or the Code and Standards, as well as make reasonable efforts to detect violations.

An adequate compliance system must meet industry standards, regulatory requirements, and the requirements of the Code and Standards.

Members with supervisory responsibilities have an obligation to bring an inadequate compliance system to the attention of firm's management and recommend corrective action.

A member or candidate faced with no compliance procedures or with procedures he believes are inadequate must decline supervisory responsibility in writing until adequate procedures are adopted by the firm.

If there is a violation, respond promptly and conduct a thorough investigation while increasing supervision or placing limitations on the wrongdoer's activities.

Recommendations for Members

A member should recommend that his employer adopt a code of ethics. Members should encourage employers to provide their codes of ethics to clients.

Once the compliance program is instituted, the supervisor should:

- Distribute it to the proper personnel.
- Update it as needed.
- Continually educate staff regarding procedures.
- Issue reminders as necessary.
- Require professional conduct evaluations.
- Review employee actions to monitor compliance and identify violations.

Recommendations for Firms

Employers should not commingle compliance procedures with the firm's code of ethics—this can dilute the goal of reinforcing one's ethical obligations.

While investigating a possible breach of compliance procedures, it is appropriate to limit the suspected employee's activities.

Adequate compliance procedures should:

- Be clearly written.
- Be easy to understand.

- Designate a compliance officer with authority clearly defined.
- Have a system of checks and balances.
- Outline the scope of procedures.
- Outline what conduct is permitted.
- Contain procedures for reporting violations and sanctions.
- Structure incentives so unethical behavior is not rewarded.

MODULE 71.7: GUIDANCE FOR STANDARD V



STANDARD V: INVESTMENT ANALYSIS, RECOMMENDATIONS, AND ACTIONS

Video covering
this content is
available online.

Standard V(A) Diligence and Reasonable Basis

Members and Candidates must:

1. Exercise diligence, independence, and thoroughness in analyzing investments, making investment recommendations, and taking investment actions.
2. Have a reasonable and adequate basis, supported by appropriate research and investigation, for any investment analysis, recommendation, or action.

The application of this Standard depends on the investment philosophy adhered to, members' and candidates' roles in the investment decision-making process, and the resources and support provided by employers. These factors dictate the degree of diligence, thoroughness of research, and the proper level of investigation required.

The level of research needed to satisfy the requirement for due diligence will differ depending on the product or service offered. A list of things that should be considered prior to making a recommendation or taking investment action includes:

- Global and national economic conditions.
- A firm's financial results and operating history, and the business cycle stage.
- Fees and historical results for a mutual fund.
- Limitations of any quantitative models used.
- A determination of whether peer group comparisons for valuation are appropriate.

Recommendations for Members

Members should encourage their firms to adopt a policy for periodic review of the quality of third-party research, if they have not. Examples of criteria to use in judging quality are:

- Review assumptions used.
- Determine how rigorous the analysis was.
- Identify how timely the research is.
- Evaluate objectivity and independence of the recommendations.

Members should *encourage their firms to consider* these policies and procedures supporting this Standard:

- Have a policy requiring that research reports and recommendations have a basis that can be substantiated as reasonable and adequate.
- Have detailed, written guidance for proper research and due diligence.
- Have measurable criteria for judging the quality of research, and base analyst compensation on such criteria.
- Have written procedures that provide a minimum acceptable level of scenario testing for computer-based models and include standards for the range of scenarios, model accuracy over time, and a measure of the sensitivity of cash flows to model assumptions and inputs.
- Have a policy for evaluating outside providers of information that addresses the reasonableness and accuracy of the information provided and establishes how often the evaluations should be repeated.
- Adopt a set of standards that provides criteria for evaluating external advisers and states how often a review of external advisers will be performed.

Standard V(B) Communication With Clients and Prospective Clients

Members and Candidates must:

1. Disclose to clients and prospective clients the basic format and general principles of the investment processes they use to analyze investments, select securities, and construct portfolios and must promptly disclose any changes that might materially affect those processes.
2. Disclose to clients and prospective clients significant limitations and risks associated with the investment process.
3. Use reasonable judgment in identifying which factors are important to their investment analyses, recommendations, or actions and include those factors in communications with clients and prospective clients.
4. Distinguish between fact and opinion in the presentation of investment analyses and recommendations.

All means and types of communication with clients are covered by this Standard, not just research reports or other written communications.

Members must distinguish between opinions and facts and always include the basic characteristics of the security being analyzed in a research report. Expectations based on statistical modeling and analysis are not facts.

Members must explain to clients and prospects the investment decision-making process used.

In preparing recommendations for structured securities, allocation strategies, or any other nontraditional investment, members must communicate those risk factors specific to such investments. In all cases, members should communicate the potential gains and losses on the investment clearly in terms of total returns.

Members must communicate significant changes in the risk characteristics of an investment or investment strategy.

Members must update clients regularly about any changes in the investment process, including any risks and limitations that have been newly identified.

When using projections from quantitative models and analysis, members may violate the Standard by not explaining the limitations of the model and the assumptions it uses, which provides a context for judging the uncertainty regarding the estimated investment result.

Members and candidates must inform clients about limitations inherent to an investment. Two examples of such limitations are liquidity and capacity. Liquidity refers to the ability to exit an investment readily without experiencing a significant extra cost from doing so. Capacity refers to an investment vehicle's ability to absorb additional investment without reducing the returns it is able to achieve.

Recommendations for Members

Selection of relevant factors in a report can be a judgment call so members should maintain records indicating the nature of the research, and be able to supply additional information if it is requested by the client or other users of the report.

Standard V(C) Record Retention

Members and Candidates must develop and maintain appropriate records to support their investment analyses, recommendations, actions, and other investment-related communications with clients and prospective clients.

Members must maintain research records that support the reasons for the analyst's conclusions and any investment actions taken. Such records are the property of the firm. All communications with clients through any medium, including emails and text messages, are records that must be retained.

A member who changes firms must re-create the analysis documentation supporting her recommendation using publicly available information or information obtained from the company and must not rely on memory or materials created at her previous firm.

Recommendations for Members

If no regulatory standards or firm policies are in place, the Standard recommends a seven-year minimum holding period.

Recommendations for Firms

This recordkeeping requirement generally is the firm's responsibility.



MODULE QUIZ 71.6, 71.7

1. Connie Fletcher, CFA, works for a small money management firm that specializes in pension accounts. Recently, a friend asked her to act as an unpaid volunteer manager for the city's street sweep pension fund. As part of the position, the city would grant Fletcher a free parking space in front of her downtown office. Before Fletcher accepts, she should *most appropriately*:
 - A. do nothing because this is a volunteer position.

- B. inform her current clients in writing and discuss the offer with her employer.
 - C. disclose the details of the volunteer position to her employer and obtain written permission from her employer.
2. Sarah Johnson, a portfolio manager, is offered a bonus directly by a client if Johnson meets certain performance goals. To comply with the Standard that governs additional compensation arrangements, Johnson should:
- A. decline to accept a bonus outside of her compensation from her employer.
 - B. disclose this arrangement to her employer in writing and obtain her employer's permission.
 - C. disclose this arrangement to her employer only if she actually meets the performance goals and receives the bonus.
3. A member or candidate who has supervisory responsibility:
- A. should place particular emphasis on enforcing investment-related compliance policies.
 - B. is responsible for instructing those to whom he has delegated authority about methods to detect and prevent violations of the law and the Code and Standards.
 - C. has complied with the Standards if she reports employee violations to upper management and provides a written warning to the employee to cease such activities.
4. Which of the following actions is a *required*, rather than *recommended*, action under the Standard regarding diligence and a reasonable basis for a firm's research recommendations?
- A. Compensate analysts based on a measure of the quality of their research.
 - B. Review the assumptions used and evaluate the objectivity of third-party research reports.
 - C. Have a policy requiring that research reports and recommendations have a basis that can be substantiated as reasonable and adequate.
5. Claire Marlin, CFA, manages an investment fund specializing in foreign currency trading. Marlin writes a report to investors that describes the basic characteristics of her strategy, which is based on an expected appreciation of the euro relative to other major currencies. Marlin shows the projected returns from the strategy if the euro appreciates less than 5%, between 5% and 10%, or more than 10%, while clearly stating that these forecasts are her opinion. Has Marlin violated the Standard related to communication with clients?
- A. Yes, because she did not include a scenario in which the euro depreciates.
 - B. No, because she disclosed the basic characteristics of the investment.
 - C. No, because she distinguished fact from opinion and discussed how the strategy may perform under a range of scenarios.
6. If regulations do not specify how long to retain the documents that support an analyst's conclusions, the Code and Standards recommend a period of at least:
- A. 5 years.
 - B. 7 years.
 - C. 10 years.

MODULE 71.8: GUIDANCE FOR STANDARD VI



STANDARD VI: CONFLICTS OF INTEREST

Standard VI(A) Disclosure of Conflicts

Members and Candidates must make full and fair disclosure of all matters that could reasonably be expected to impair their independence and objectivity or interfere with respective duties to their clients, prospective clients, and employer. Members and Candidates must ensure that such disclosures are prominent, are delivered in plain language, and communicate the relevant information effectively.

Video covering
this content is
available online.

Members must fully disclose to clients, prospects, and their employers all actual and potential conflicts of interest in order to protect investors and employers. These disclosures must be clearly stated.

The requirement that all potential areas of conflict be disclosed allows clients and prospects to judge motives and potential biases for themselves. Disclosure of broker-dealer market-making activities would be included here. Board service is another area of potential conflict.

The most common conflict that requires disclosure is actual ownership of stock in companies that the member recommends or that clients hold.

Another common source of conflicts of interest is a member's compensation/bonus structure, which can potentially create incentives to take actions that produce immediate gains for the member with little or no concern for longer-term returns for the client. Such conflicts must be disclosed when the member is acting in an advisory capacity and must be updated in the case of significant change in compensation structure.

Members must give their employers enough information to judge the impact of a conflict, take reasonable steps to avoid conflicts, and report them promptly if they occur.

Recommendations for Members

Any special compensation arrangements, bonus programs, commissions, and incentives should be disclosed.

Standard VI(B) Priority of Transactions

Investment transactions for clients and employers must have priority over investment transactions in which a Member or Candidate is the beneficial owner.

Client transactions take priority over personal transactions and over transactions made on behalf of the member's firm. Personal transactions include situations where the member is a beneficial owner.

Personal transactions may be undertaken only after clients and the member's employer have had an adequate opportunity to act on a recommendation. Note that family member accounts that are client accounts should be treated just like any client account; they should not be disadvantaged.

Members must not act on information about pending trades for personal gain. The overriding considerations with respect to personal trades are that they do not disadvantage any clients.

When requested, members must fully disclose to investors their firm's personal trading policies.

Recommendations for Members

Members can avoid conflicts that arise with IPOs by not participating in them.

Members should encourage their firms to adopt the procedures listed in the following recommendations for firms if they have not done so.

Recommendations for Firms

All firms should have basic procedures in place that address conflicts created by personal investing. The following areas should be included:

- Establish limitations on employee participation in equity IPOs.
- Establish restrictions on participation in private placements. Strict limits should be placed on employee acquisition of these securities and proper supervisory procedures should be in place. Participation in these investments raises conflict of interest issues similar to those of IPOs.
- Establish blackout/restricted periods. Employees involved in investment decision-making should have blackout periods prior to trading for clients—no front running (i.e., purchase or sale of securities in advance of anticipated client or employer purchases and sales). The size of the firm and the type of security should help dictate how severe the blackout requirement should be.
- Establish reporting procedures, including duplicate trade confirmations, disclosure of personal holdings and beneficial ownership positions, and preclearance procedures.

Standard VI(C) Referral Fees

Members and Candidates must disclose to their employer, clients, and prospective clients, as appropriate, any compensation, consideration, or benefit received from or paid to others for the recommendation of products or services.

Members must inform employers, clients, and prospects of any benefit received for referrals of customers and clients, allowing them to evaluate the full cost of the service as well as any potential partiality. All types of consideration must be disclosed.

Recommendations for Members

Members should encourage their firms to adopt clear procedures regarding compensation for referrals.

Members should provide their employers with updates at least quarterly.

Recommendations for Firms

Firms that do not prohibit referral fees should have clear procedures for approval and policies regarding the nature and value of referral compensation received.

MODULE 71.9: GUIDANCE FOR STANDARD VII



STANDARD VII: RESPONSIBILITIES AS A CFA INSTITUTE MEMBER OR CFA CANDIDATE

Video covering this content is available online.

Standard VII(A) Conduct as Participants in CFA Institute Programs

Members and Candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of CFA

Institute programs.

Members must not engage in any activity that undermines the integrity of the CFA charter. This Standard applies to conduct that includes:

- Cheating on the CFA exam or any exam.
- Revealing anything about either broad or specific topics tested, content of exam questions, or formulas required or not required on the exam.
- Not following rules and policies of the CFA Program.
- Giving confidential information on the CFA Program to candidates or the public.
- Improperly using the designation to further personal and professional goals.
- Misrepresenting information on the Professional Conduct Statement (PCS) or the CFA Institute Professional Development Program.

Members and candidates are not precluded from expressing their opinions regarding the exam program or CFA Institute but must not reveal confidential information about the CFA Program.

Candidates who violate any of the CFA exam policies (e.g., calculator, personal belongings, Candidate Pledge) have violated Standard VII(A).

Members who volunteer in the CFA Program may not solicit or reveal information about questions considered for or included on a CFA exam, about the grading process, or about scoring of questions.

Standard VII(B) Reference to CFA Institute, the CFA Designation, and the CFA Program

When referring to CFA Institute, CFA Institute membership, the CFA designation, or candidacy in the CFA Program, Members and Candidates must not misrepresent or exaggerate the meaning or implications of membership in CFA Institute, holding the CFA designation, or candidacy in the CFA Program.

Members must not make promotional promises or guarantees tied to the CFA designation, such as over-promising individual competence or over-promising investment results in the future (i.e., higher performance, less risk, etc.).

Members must satisfy these requirements to maintain membership:

- Sign the PCS annually.
- Pay CFA Institute membership dues annually.

If they fail to do this, they are no longer active members.

Do not misrepresent or exaggerate the meaning of the CFA designation.

There is no partial CFA designation. It is acceptable to state that a candidate successfully completed the program in three years if, in fact, he did, but claiming superior ability because of this is not permitted.

Recommendations for Members

Members should be sure that their firms are aware of the proper references to a member's CFA designation or candidacy, as errors in these references are common.



MODULE QUIZ 71.8, 71.9

1. Daniel Lyons, CFA, is an analyst who covers several stocks including Horizon Company. Lyons's aunt owns 30,000 shares of Horizon. She informs Lyons that she has created a trust in his name into which she has placed 2,000 shares of Horizon. The trust is structured so that Lyons will not be able to sell the shares until his aunt dies, but may vote the shares. Lyons is due to update his research coverage of Horizon next week. Lyons should *most appropriately*:
 - A. update the report as usual because he is not a beneficial owner of the stock.
 - B. advise his superiors that he is no longer able to issue research recommendations on Horizon.
 - C. disclose the situation to his employer and, if then asked to prepare a report, also disclose his beneficial ownership of the shares in his report.
2. Kate Wilson, CFA, is an equity analyst. Wilson enters two transactions for her personal account. Wilson sells 500 shares of Tibon, Inc., a stock on which her firm currently has a "Buy" recommendation. Wilson buys 200 shares of Hayfield Co. and the following day issues a research report on Hayfield with a "Buy" recommendation. Has Wilson violated the Code and Standards?
 - A. No.
 - B. Yes, both of her actions violate the Code and Standards.
 - C. Yes, but only one of her actions violates the Code and Standards.
3. Hern Investments provides monthly emerging market research to Baker Brokerage in exchange for prospective client referrals and European equity research from Baker. Clients and prospects of Hern are not made aware of the agreement, but clients unanimously rave about the high quality of the research provided by Baker. As a result of the research, many clients with nondiscretionary accounts have earned substantial returns on their portfolios. Managers at Hern have also used the research to earn outstanding returns for the firm's discretionary accounts. Hern has *most likely*:
 - A. not violated the Code and Standards.
 - B. violated the Code and Standards by using third-party research in discretionary accounts.
 - C. violated the Code and Standards by failing to disclose the referral agreement with Baker.
4. After writing the Level I CFA exam, Cynthia White goes to internet discussion site *CFA Haven* to express her frustration. White writes, "CFA Institute is not doing a competent job of evaluating candidates because none of the questions in the June exam touched on Alternative Investments." White *most likely* violated the Standard related to conduct as a candidate in the CFA program by:
 - A. publicly disputing CFA Institute policies and procedures.
 - B. disclosing subject matter covered or not covered on a CFA exam.
 - C. participating in an internet forum that is directed toward CFA Program participants.
5. After passing all three levels of the CFA exams on her first attempts and being awarded her CFA charter, Paula Osgood is promoting her new money management firm by issuing an advertisement. Which of these statements would *most likely* violate the Standard related to use of the CFA designation?
 - A. "To earn the right to use the CFA designation, Paula passed three exams covering ethics, financial statement analysis, asset valuation, and portfolio management."
 - B. "Paula passed three 6-hour exams on her first attempts and is a member of her local investment analyst society."
 - C. "Because of her extensive training, Paula will be able to achieve better investment results than managers who have not been awarded the CFA designation."

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 71.1, 71.2, 71.3

1. **C** To comply with Standard I(A) Knowledge of the Law, a member must always abide by the strictest applicable law, regulation, or standard. (Module 71.1, LOS 71.a, 71.b, 71.c)
2. **A** Gifts from clients are acceptable under Standard I(B) Independence and Objectivity, but the Standard requires members and candidates to disclose such gifts to their employers. Standard I(B) allows issuer-paid research as long as the analysis is thorough, independent, unbiased, and has a reasonable and adequate basis for its conclusions, and the compensation from the issuer is disclosed. Members and candidates should consider the potential for conflicts of interest inherent in credit ratings and may need to do independent research to evaluate the soundness of these ratings. (Module 71.1, LOS 71.a, 71.b, 71.c)
3. **C** According to Standard I(C) Misrepresentation, members and candidates must cite the sources of the information they use in their analysis, unless the information is factual data (as opposed to analysis or opinion) from a recognized financial or statistical reporting service. The U.S. Treasury is one example of a recognized source of factual data. (Module 71.2, LOS 71.a, 71.b, 71.c)
4. **B** Failing to act when required by one's professional obligations, such as neglecting to perform due diligence related to an investment recommendation, violates Standard I(D) Misconduct. Acts a member commits outside his professional capacity are misconduct if they reflect poorly on the member or candidate's honesty, integrity, or competence (e.g., theft or fraud). Violations of the law that do not reflect on the member or candidate's honesty, integrity, or competence (e.g., an act related to civil disobedience) are not necessarily regarded as misconduct. (Module 71.2, LOS 71.a, 71.b, 71.c)
5. **A** The statement from the CFO about the current quarter's earnings is material nonpublic information. Ingus violated Standard II(A) Material Nonpublic Information by acting or causing others to act on it. (Module 71.3, LOS 71.a, 71.b, 71.c)
6. **B** The intent of Green Brothers' actions is to manipulate the appearance of market liquidity in order to attract investment to its own funds. The increased trading activity was not based on market fundamentals or an actual trading strategy to benefit investors. It was merely an attempt to mislead market participants in order to increase assets under Green Brothers' management. The action violates Standard II(B) Market Manipulation. (Module 71.3, LOS 71.a, 71.b, 71.c)

Module Quiz 71.4, 71.5

1. **C** Standard III(A) Loyalty, Prudence, and Care specifies that for the manager of a pension or trust, the duty of loyalty is owed to the beneficiaries, not to the individuals who hired the manager. (Module 71.4, LOS 71.a, 71.b, 71.c)
2. **B** The firm must give its clients an opportunity to act on recommendation changes. Firms can offer different levels of service to clients as long as this is disclosed to all clients. The largest institutional clients would likely be paying higher fees for a greater level of service. The portfolio manager's brother's account should be treated the same as any other client account. (Module 71.4, LOS 71.a, 71.b, 71.c)
3. **A** According to Standard III(C) Suitability, a member or candidate who is in an advisory relationship with a client is responsible for analyzing the suitability of an investment for the client before taking investment action or making a recommendation. If a member or candidate believes an unsolicited trade is unsuitable for a client, the appropriate action is to discuss the trade with the client. The advisor may follow her firm's policies for obtaining client approval if the requested trade would not affect the risk and return of the client's portfolio materially. If the trade would have a material effect, the advisor should discuss with the client whether the IPS needs to be updated. When managing a fund to an index or stated mandate, the manager is responsible for ensuring that potential investments are consistent with the fund's mandate. Suitability for individuals would be a concern for an advisor who recommends the fund to clients, but not for the manager of the fund. (Module 71.5, LOS 71.a, 71.b, 71.c)
4. **C** Recommendations stated in Standard III(D) Performance Presentation include considering the sophistication and knowledge of the audience when presenting performance data. Other recommendations are to include terminated accounts in past performance history; to present the performance of a composite as a weighted average of the performance of similar portfolios, rather than using a single representative account; and to maintain the records and data that were used to calculate performance. (Module 71.5, LOS 71.a, 71.b, 71.c)

5. **B** Members and candidates are required to cooperate with PCP investigations into their own conduct and encouraged to cooperate with PCP investigations into the conduct of others. Sharing confidential information with the PCP is not a violation of Standard III(E) Preservation of Confidentiality. Any client information shared with the PCP will be kept in strict confidence. Standard III(E) states that members and candidates are required to maintain confidentiality of client records even after the end of the client relationship. (Module 71.5, LOS 71.a, 71.b, 71.c)

Module Quiz 71.6, 71.7

1. **C** According to Standard IV(A) Loyalty, members and candidates are expected to act for the benefit of their employer and not deprive the employer of their skills. Fletcher is performing work similar to the services that her employer provides. Although the position is a volunteer position, Fletcher will receive compensation in the form of a free parking space. In light of the circumstances, Fletcher must disclose the details of the position to her employer and get written permission before accepting the volunteer position. (Module 71.6, LOS 71.a, 71.b, 71.c)
2. **B** Johnson should disclose her additional compensation arrangement in writing to her employer and obtain her employer's written consent before accepting this offer, in accordance with Standard IV(B) Additional Compensation Arrangements. (Module 71.6, LOS 71.a, 71.b, 71.c)
3. **B** Members or candidates may delegate supervisory duties to subordinates but remain responsible for instructing them about how to detect and prevent violations. Reporting the violation and warning the employee are not sufficient to comply with Standard IV(C) Responsibilities of Supervisors. The supervisor must also take steps to prevent further violations while she conducts an investigation, such as limiting the employee's activity or increasing her monitoring of the employee. Supervisors should enforce investment-related and non-investment-related policies equally. (Module 71.6, LOS 71.a, 71.b, 71.c)
4. **B** Standard V(A) Diligence and Reasonable Basis requires analysts who use third-party research to review its assumptions and evaluate the independence and objectivity of the research. The other choices are recommended procedures for compliance with the Standard. (Module 71.7, LOS 71.a, 71.b, 71.c)
5. **A** Standard V(B) Communication with Clients and Prospective Clients requires that members and candidates communicate the risk associated with the investment strategy used and how the strategy is expected to perform in a range of scenarios. These scenarios should include those different from the current trend. Marlin should have discussed how her strategy would perform if the euro depreciates instead of appreciating as she expects. (Module 71.7, LOS 71.a, 71.b, 71.c)
6. **B** When no other regulatory guidance applies, Standard V(C) Record Retention recommends that records be maintained for a minimum of seven years. (Module 71.7, LOS 71.a, 71.b, 71.c)

Module Quiz 71.8, 71.9

1. **C** Even though the shares are held in trust, Lyons is considered a beneficial owner under Standard VI(A) Disclosure of Conflicts because he has a pecuniary interest in the shares and because has the power to vote the shares. Lyons is obligated to inform his employer of the potential conflict. If Lyons's employer permits him to continue issuing investment recommendations on the company, Lyons must disclose the existence of a potential conflict in his reports. (Module 71.8, LOS 71.a, 71.b, 71.c)
2. **C** Only one of these transactions is a violation. Standard VI(B) Priority of Transactions requires members and candidates to give clients an adequate opportunity to act on a recommendation before trading for accounts in which the member or candidate has a beneficial ownership interest. Members and candidates may trade for their own accounts as long as they do not disadvantage clients, benefit personally from client trades, or violate any regulations that apply. The Standard does not prohibit members and candidates from entering personal

transactions that are contrary to what their firms are recommending for clients, as long as the transaction does not violate any of these criteria. (Module 71.8, LOS 71.a, 71.b, 71.c)

3. **C** According to Standard VI(C) Referral Fees, Hern must disclose the referral arrangement between itself and Baker so that potential clients can judge the true cost of Hern's services and assess whether there is any partiality inherent in the recommendation of services. (Module 71.8, LOS 71.a, 71.b, 71.c)
4. **B** Standard VII(A) Conduct as Participants in CFA Institute Programs prohibits candidates from revealing which portions of the Candidate Body of Knowledge were or were not covered on an exam. Members and candidates are free to disagree with the policies, procedures, or positions taken by the CFA Institute. The Standard does not prohibit participating in CFA Program-related internet blogs, forums, or social networks. (Module 71.9, LOS 71.a, 71.b, 71.c)
5. **C** Standard VII(B) Reference to CFA Institute, the CFA Designation, and the CFA Program prohibits members and candidates from implying superior performance as a result of being a CFA charterholder. Concise factual descriptions of the requirements to obtain the CFA charter are acceptable. Osgood's statement that she passed the exams on her first attempts is acceptable because it states a fact. (Module 71.9, LOS 71.a, 71.b, 71.c)

¹ Copyright 2014, CFA Institute. Reproduced and republished from *Standards of Practice Handbook, 11th Ed.*, 2014, with permission from CFA Institute. All rights reserved.

² Ibid.

READING 72

INTRODUCTION TO THE GLOBAL INVESTMENT PERFORMANCE STANDARDS (GIPS)

EXAM FOCUS

This reading covers the key features of the Global Investment Performance Standards (GIPS) as adopted by CFA Institute in 1999 and subsequently updated. Compliance with GIPS is voluntary. For the Level I exam, you are responsible for the GIPS material included in the book of candidate readings for Level I and also for sections of the GIPS standards for firms that are referenced in the Level I material. The GIPS standards for firms are available on the CFA Institute website.

MODULE 72.1: INTRODUCTION TO GIPS



LOS 72.a: Explain why the GIPS standards were created, who can claim compliance, and who benefits from compliance.

Video covering this content is available online.

When firms choose their own methodologies for reporting investment performance, the results are not comparable across firms, and firms have a tendency to choose a methodology that makes their performance look good. Some choices that tend to bias results are:

- Choosing a top-performing portfolio and claiming it represents a firm's overall results for all assets managed.
- Excluding terminated accounts, which may tend to be accounts closed by investors due to sub-par performance.
- Selecting time periods to report that put firm performance in the best possible light.

GIPS present a standardized methodology for performance reporting that makes comparison of performance across firms meaningful, provides specific information that is useful to current clients, prospective clients, and investors, and avoids misrepresentation of performance.

Widespread usage of GIPS can also give oversight bodies a clearer understanding of the returns achieved and the risks taken by the firms they supervise.

GIPS only apply to firms that manage assets for others. Presenting performance information compliant with GIPS is voluntary for such firms, and they may only claim compliance with GIPS if they comply fully and on a firmwide basis. Other firms related to the asset management business, such as software developers, may state that they endorse GIPS but may not claim compliance with GIPS.

LOS 72.b: Describe the key concepts of the GIPS Standards for Firms.

The GIPS standards for firms consist of eight sections:

1. *Fundamentals of Compliance.* The fundamental issues involved in complying with GIPS are (a) defining the firm, (b) providing GIPS-compliant reports to all clients and prospects, (c) complying with applicable regulations and laws, and (d) presenting information that is neither false nor misleading.
2. *Input Data and Calculation Methodology.* Input data should be consistent in order to establish full, fair, and comparable investment performance presentations. Certain methodologies are required for portfolio return calculations, and certain other methodologies are required for composite return calculations. Uniformity in methods across firms is required so that their results are comparable.
3. *Composite and Pooled Fund Maintenance.* Creation of meaningful, asset-weighted composites is important to achieve a fair presentation. Composite performance is based on the performance of one or more portfolios that have the same investment strategy or investment objective. Composite returns are the asset-weighted average (not a simple average) of the returns on the portfolios that are included in each composite. Pooled funds must be included in a composite if they fit its definition.
4. *Composite Time-Weighted Return Report.*
5. *Composite Money-Weighted Return Report.*
6. *Pooled Fund Time-Weighted Return Report.*
7. *Pooled Fund Money-Weighted Return Report.*

Sections 4 through 7 contain required and recommended procedures for reporting the performance of composites and pooled funds, as well as the necessary disclosures. There are some disclosures that all firms must make, but some disclosures may not apply to all firms. If a disclosure is not applicable to a specific firm, the firm is not required to include any statement regarding it. A firm that has met all the requirements of GIPS may include an appropriate claim of compliance.

8. *GIPS Advertising Guidelines.* If an advertisement includes a claim of compliance with GIPS, the advertisement must comply with these guidelines. The guidelines do not apply to advertisements that do not reference the firm's GIPS compliance.
-

LOS 72.c: Explain the purpose of composites in performance reporting.

A **composite** is a grouping of individual discretionary portfolios representing a similar investment strategy, objective, or mandate. Examples of possible composites are large capitalization stocks, investment-grade domestic bonds, and accounts managed to match the performance of a specific securities index. Reporting on the performance of composites gives clients and prospects information about the firm's success in managing various types of securities and results for various investment styles.

A composite, such as international equities, must include all fee-paying, discretionary portfolios (current and past) that the firm has managed in accordance with a particular strategy. The firm should identify which composite each managed portfolio is to be included in before the portfolio's performance is known. This prevents firms from choosing portfolios to include in a composite in order to create composites with superior returns. All discretionary portfolios must be included in one, and only one, composite.

LOS 72.d: Describe the fundamentals of compliance, including the recommendations of the GIPS standards with respect to the definition of the firm and the firm's definition of discretion.

The definition of the firm, for purposes of GIPS compliance, must be the corporation, subsidiary, or division that is held out to clients as a business entity. If a firm has different geographic locations (e.g., all doing business under the name of Bluestone Advisers), then the definition of the firm should include all the various geographic locations and their clients.

The definition of **discretion** refers to how a firm determines which of the portfolios it manages should be included in a composite. A firm may determine that a portfolio is nondiscretionary—and therefore not include it in a composite—if the client places restrictions on it that prevent the manager from carrying out the intended strategy.

LOS 72.e: Describe the concept of independent verification.

Firms are *encouraged* to pursue independent verification of their compliance with GIPS. Verification applies to the entire firm's performance measurement practices and methods, not a selected composite.

If a firm chooses to pursue verification, it must be performed by a third party, not by the firm itself, on a firmwide basis. This third-party verifier must attest that (1) the firm has complied with all GIPS requirements for composite construction on a firmwide basis and (2) the firm's processes and procedures are established to present performance in accordance with the calculation methodology required by GIPS, the data requirements of GIPS, and in the format required by GIPS.

Verified firms should include the following disclosure language:

[Insert name of firm] has been verified for the periods [insert dates] by [name of verifier]. A copy of the verification report is available upon request.



MODULE QUIZ 72.1

1. Which of the following statements *most accurately* describes the parties that GIPS are intended to apply to and serve?
 - A. GIPS apply to consultants who serve their existing and prospective clients.
 - B. GIPS apply to firms that issue securities and serve investment management firms.
 - C. GIPS apply to investment management firms and serve their existing and prospective clients.

2. For a composite to be constructed in compliance with GIPS, the portfolios included in the composite must:
 - A. have been managed by the firm for the full performance-reporting period.
 - B. be selected immediately after the last business day of the period for which the composite's performance will be presented.
 - C. include all fee-paying, discretionary portfolios that are managed according to the same strategy, mandate, or investment objective.
3. Verification of compliance with GIPS:
 - A. may be performed on single composites.
 - B. is required for a firm to claim GIPS compliance.
 - C. requires the verification report to be issued for the entire firm.

KEY CONCEPTS

LOS 72.a

GIPS are principles that investment firms can voluntarily follow. They are designed to avoid misrepresentations of historical investment results to clients and prospects. GIPS allow clients to more easily compare investment performance among investment firms and have more confidence in reported performance.

LOS 72.b

The eight sections of the GIPS standards for firms are:

1. Fundamentals of Compliance.
2. Input Data and Calculation Methodology.
3. Composite and Pooled Fund Maintenance.
4. Composite Time-Weighted Return Report.
5. Composite Money-Weighted Return Report.
6. Pooled Fund Time-Weighted Return Report.
7. Pooled Fund Money-Weighted Return Report.
8. GIPS Advertising Guidelines.

LOS 72.c

A composite is a grouping of discretionary portfolios representing a similar investment strategy, objective, or mandate. A composite must include all portfolios (current and past) that the firm has managed in accordance with this particular strategy. The firm should identify which composite each managed portfolio is to be included in before the portfolio's performance is known.

LOS 72.d

GIPS compliance must be firmwide, where the *firm* is the distinct business entity that is held out to clients and prospects.

A firm must establish criteria to determine which of the portfolios it manages are discretionary or nondiscretionary.

LOS 72.e

Firms are encouraged to pursue independent verification of GIPS compliance. If they seek verification, it must be performed by a third party, not by the firm itself. The third-party verifier must attest that (1) the firm has complied with all GIPS requirements for composite construction on a firmwide basis and (2) the firm's processes and procedures are established to present performance in accordance with the calculation methodology required by GIPS, the data requirements of GIPS, and in the format required by GIPS.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 72.1

1. **C** GIPS apply to investment management firms. They are intended to serve prospective and existing clients of investment management firms and consultants who advise these clients. (LOS 72.a)
2. **C** To comply with GIPS, a composite must include all fee-paying, discretionary portfolios managed according to the same investment objective, strategy, or mandate. The composite or composites in which a portfolio will be included must be determined on an ex ante basis (i.e., before the period for which the composite's performance will be calculated). A GIPS-compliant composite must include terminated accounts. (LOS 72.c)
3. **C** Verification of GIPS compliance is optional, but if a firm chooses to seek third-party verification, the report must be issued with respect to the whole firm. GIPS verification cannot be carried out for a single composite. (LOS 72.e)

READING 73

ETHICS APPLICATION

EXAM FOCUS

In this reading, we cover key points about applying the Standards, drawn from cases illustrating their application included in the *Ethics in Investment Management Casebook* (CFA Institute, 2019). While this provides a suitable review, just as we do for the *Standards of Practice Handbook*, we strongly suggest that you read the complete coverage of these cases in the Level I CFA Curriculum, Volume 6 (at least once). Studying the full cases will help candidates understand how irrelevant points and unfounded justifications for a member's actions can be used to distract from the essence of violations of the Standards.

MODULE 73.1: ETHICS APPLICATION



LOS 73.a: Evaluate practices, policies, and conduct relative to the CFA Institute Code of Ethics and Standards of Professional Conduct.

Video covering this content is available online.

LOS 73.b: Explain how the practices, policies, and conduct do or do not violate the CFA Institute Code of Ethics and Standards of Professional Conduct.

Standard I: Professionalism

Standard I(A) Knowledge of the Law

Case 1: A member reports to his supervisor that their firm is overcharging clients by seeking to recover the cost of expenses that are actually reimbursable. The actions taken only remedy the situation for some clients but not for others. The member must disassociate from the activity and not work with any clients who are still being overcharged.

Case 2: A member violates the Standard by failing to investigate transactions in an account that appear to be at high risk of violating money-laundering laws (because it was a long-standing client who has relationships with board members).

Case 3: A member violates the Standards (and the law) by forging customer signatures for expediency.

Standard I(B) Independence and Objectivity

Case 1: A member violates the Standard by contributing to a politician's campaign, believing that it may lead to preferential treatment with regard to receiving management contracts for government pension fund money.

Standard I(C) Misrepresentation

Case 1: A member violates the Standard by assuring a client that returns on a fund will outweigh the penalties incurred from shifting funds from an existing investment, essentially guaranteeing a specific rate of return on an investment that has no actual guarantee of returns.

Case 2: A proposal to manage pension assets lists all personnel and their qualifications, but while the proposal is under consideration, one of the key personnel leaves the firm. The member is required to inform the potential client of this personnel change (misrepresentation by omission).

Case 3: A member who is the CEO of an electric car company posts on social media that private financing has been secured to take the company private at \$420 per share when that is untrue and the member admits the price “was a joke.” This is clearly a violation of the Standard regarding misrepresentation.

Standard I(D) Misconduct

Case 1: A member is arrested for minor criminal offenses that are part of civil disobedience related to protesting and expressing her personal beliefs. This is not necessarily a violation of the Standard regarding misconduct, which states in part that members are prohibited from “engaging in any professional conduct involving dishonesty, fraud, or deceit” or committing an act that “reflects adversely on their professional reputation, integrity, or competence.”

Case 2: A member violates the Standard by using his firm’s error-correction policy to effectively give his own money to a client account in order to make his management of the account look better.

Standard II: Integrity of Capital Markets

Standard II(A) Material Nonpublic Information

Case 1: A member violates the Standard by using material nonpublic information he acquired by overhearing a friend’s phone conversation and inferring that a takeover offer for a specific company was imminent.

Case 2: A member violates the Standard by sharing information with clients about a regulator’s positive response to trials of a company’s new drug that he learned of during a meeting of analysts with company management. The information is considered nonpublic because sharing it with a specific group of analysts was “selective sharing” of the information and cannot be considered public disclosure.

Standard II(B) Market Manipulation

Case 1: A member violates the Standard by fraudulently including the names of people who do not actually own shares to meet the minimum required number of shareholders for a firm to be listed on an exchange, misleading market participants about the potential liquidity of the shares.

Standard III: Duties to Clients

Standard III(A) Loyalty, Prudence, and Care

Case 1: A member violates the Standard by putting a policy into an investment firm's client agreements that indicates that their representatives are excused in some cases from acting in a client's best interest or failing to adequately investigate suitability of recommendations to clients, and that clients' rights to make claims for violating securities laws are restricted. Members and candidates cannot "opt out" of the Standards.

Case 2: A member executes trades for a client who self-directs their own account and has received the firm's policies regarding margin requirements. In this case, the member's requirements to act in the client's best interests are limited, compared with a relationship in which the member is providing investment advice. There is no intent to limit the member's responsibilities to act in good faith and not misrepresent the services to be provided according to firm policies. The terms regarding the firm's actions in the event of a margin shortfall are open to negotiation at the inception of the relationship.

Case 3: A member allocates expenses to a client that arise from actions designed to reduce total expenses (staying over a weekend to reduce transportation costs but increasing lodging costs). This is not a violation, but charging a client for expenses that benefit other clients or that are for personal activities of the member is.

Standard III(B) Fair Dealing

Case 1: A member treats clients fairly by simultaneously sending emails regarding a change in investment recommendation. Providing individual clients with updates or clarifications regarding the change in recommendation is not a violation. Further, the member's firm offers some clients, for an additional fee, weekly updates about securities that may indicate a change in recommendation may be forthcoming. This is not a violation as long as clients are all informed about the availability of the weekly updates for a fee and as long as no change in recommendation is included in a weekly update that would disadvantage clients who do not subscribe to the additional service.

Standard III(C) Suitability

Case 1: A member violates the Standard by recommending investments that carry more risk than is suitable for some clients, even though there are tax advantages to the recommended portfolios.

Case 2: A member violates the Standard by making a client-requested change to their portfolio without investigating the client circumstances adequately to determine whether the requested investment is suitable for the clients.

Standard III(D) Performance Presentation

Case 1: A member violates the Standard by presenting performance data based on a composite of separately managed accounts the firm managed before creating the fund that is being reported, giving the impression that the fund has been in existence for many years when, in fact, it is relatively new.

Standard III(E) Preservation of Confidentiality

Case 1: Member A downloads clients' personally identifiable information to his personal server at home to make working from home easier and gets hacked. The member has violated the

Standard by not taking adequate steps to protect client information. Member B, who is the firm's head of compliance, likely violated the Standard by not taking the proper steps to protect client information, as evidenced by the fact that Member A was able to download sensitive client information to his personal server.

Standard IV: Duties to Employers

Standard IV(A) Loyalty

Case 1: A member violates the Standard by making harmful statements about Firm A, her current employer, and promoting the firm that she intends to move to while she is still employed by Firm A.

Case 2: A member is pressured by his firm to sell the firm's proprietary investment products that are relatively expensive and haven't performed well. He has complained to management, and they have done nothing. He copies client records and he records conversations with his supervisor to document his conduct. He takes this information to securities regulators. His actions are not a violation of the Standard because he is acting in the interest of his clients.

Case 3: A member violates the Standard by taking client information with her when she leaves her firm; the client list is the property of her current firm. Although she only intends to send her clients thank-you notes, the client list she copied has personal information.

Standard IV(B) Additional Compensation Arrangements

Case 1: A member works for a firm that produces issuer-paid research reports and is involved with the decision of which companies her firm will cover. Some companies seeking coverage by the firm offer her a bonus payment if their firm is selected. Because this creates a possible conflict between her interests and her firm's interests, she must get approval in writing from her employer to accept the bonus arrangement.

Standard IV(C) Responsibilities of Supervisors

Case 1: A member with supervisory responsibility over a branch office violates the Standard by not making reasonable efforts to ensure that those under his supervision are not engaging in misconduct. He also violates the Standard by not having clear written compliance policies and procedures in place, along with employee training. If these are not in place, he must decline to take supervisory responsibility.

Case 2: A member violates the Standard by accepting the title of chief compliance officer even though she has no experience in compliance, is denied permission to contact clients or review client communications, and is not allowed to enforce company policies. She is in violation of the Standard because it requires her to make reasonable efforts to both detect and prevent violations of laws, rules, and regulations by those subject to her supervision. Knowing she could not do this, she should have declined to take on supervisory responsibilities.

Standard V: Investment Analysis, Recommendations, and Actions

Standard V(A) Diligence and Reasonable Basis

Case 1: A member violates the Standard by recommending the purchase of shares of a company for which he has not performed a diligent, thorough, and independent analysis. A second member bases her recommendation on the analysis of the first member and incorporates part of the first member's research report in her own and is thereby also in violation of the Standards.

Standard V(B) Communication With Clients and Prospective Clients

Case 1: A member's firm begins calculating client fees using different methods from those spelled out in the disclosures made to them when they opened their accounts. The member is required by the Standard to inform clients that the fee calculation methods have changed, prior to changing the methods, even if fees are not higher overall using the new processes.

Case 2: A credit rating agency changes its methodology of determining the ratings of commercial mortgage-backed securities without disclosing the change to potential users of the ratings. A member who is responsible for publishing the ratings violates the Standard by publishing ratings without disclosing the change in methodology.

Standard V(C) Record Retention

Case 1: A member violates the Standard by not updating client records in a timely manner. Although he keeps himself updated on changes in client circumstances and adjusts portfolios accordingly, he does not do a good job of keeping written client profiles up to date.

Standard VI: Conflicts of Interest

Standard VI(A) Disclosure of Conflicts

Case 1: A member receives payment from third-party subadvisors that she uses to manage client funds invested in some asset classes. To avoid violating the Standard, she must disclose these arrangements to clients because the payments may influence her choice of subadvisors.

Standard VI(B) Priority of Transactions

Case 1: A member violates the Standard by buying shares and call options in his personal account just prior to purchasing large blocks of the same stocks in client accounts (front-running the trades), in anticipation of an increase in the price of the stocks as a result of the large purchases.

Case 2: A member violates the Standard by telling friends and relatives about large buy orders he is to execute for clients of his employer prior to executing those orders, allowing them to front-run those trades for quick profits.

Case 3: A member enters client trades grouped as block orders and allocates them to specific client accounts after the market closes for the day. The member violates the Standard (and likely the Standard on fair dealing as well) by allocating profitable trades to personal accounts and allocating a disproportionate amount of losing trades to his largest client accounts where they will likely go unnoticed.

Standard VI(C) Referral Fees

Case 1: A member invites existing clients who have referred very profitable accounts to her to lavish parties, rewarding them with discounts on fees and gift cards as well. The member

violates the Standard by not disclosing these “referral fees” to all existing clients and prospective clients.

Standard VII: Responsibilities as a CFA Institute Member or CFA Candidate

Standard VII(A) Conduct as Participants in CFA Institute Programs

Case 1: A member who teaches exam-prep classes hosts a post-exam party for candidates who took the exam. He solicits their general impressions about the difficulty of the exam and their opinions on the most difficult exam questions. He is permitted to share their opinions about the difficulty of the exam with future candidates in his exam-prep classes. However, he is not permitted to solicit or share any information about specific exam questions or about which topics were or were not tested on the exam.

Standard VII(B) Reference to CFA Institute, the CFA Designation, and the CFA Program

Case 1: A previous member who has not paid her dues to CFA Institute violates the Standard by using the CFA designation in marketing materials and on her business cards. Another member violates the Standard by claiming that all senior employees, which includes the member who has not paid her dues, are CFA charterholders.



MODULE QUIZ 73.1

1. Karen Jones, CFA, is an outside director for Valley Manufacturing. At a director’s meeting, Jones finds out that Valley Corp. has made several contributions to foreign politicians that she suspects were illegal. Jones checks with her firm’s legal counsel and determines that the contributions were indeed illegal. At the next board meeting, Jones urges the board to disclose the contributions. The board, however, votes not to make a disclosure. Jones’s *most appropriate* action would be to:
 - A. protest the board’s actions in writing to the executive officer of Valley.
 - B. resign from the board and seek legal counsel as to her legal disclosure requirements.
 - C. inform her supervisor of her discovery and cease attending meetings until the matter is resolved.
2. Beth Bixby, CFA, uses a quantitative model to actively manage a portfolio of stocks with an objective of earning a greater return than the market. Over the last three years, the returns to a portfolio constructed using the model have been greater than the returns to the S&P Index by between 2% and 4%. In promotional materials, Bixby states, “Through our complex quantitative approach, we select a portfolio that has similar risk to the S&P 500 Index but will receive a return between 2% and 4% greater than the index.” This statement is:
 - A. permissible because prior returns to the firm’s model provide a reasonable and adequate basis for the promotional material.
 - B. permissible because the statement describes the basic characteristics of the fund’s risk and return objectives.
 - C. not permissible because Bixby is misrepresenting the investment performance her firm can reasonably expect to achieve.
3. Over the past two days, Lorraine Quigley, CFA, manager of a hedge fund, has been purchasing large quantities of Craeger Industrial Products’ common stock while at the same time shorting put options on the same stock. Quigley did not notify her clients of the trades, although they are aware of the fund’s general strategy to generate returns. Which of the following statements is *most likely* correct?

- A. Quigley did not violate the Code and Standards.
 - B. Quigley violated the Code and Standards by manipulating the prices of publicly traded securities.
 - C. Quigley violated the Code and Standards by failing to disclose the transactions to clients before they occurred.
4. Julia Green, CFA, has friends from her previous employer who have suggested that she receive information from them via an internet chat room. In this way, she receives news about an exciting new product being developed by a firm in Singapore that has the potential to double the firm's revenue. The firm has not revealed any information regarding the product to the public. According to the Code and Standards, this information is:
- A. both material and nonpublic, and Green may not trade on it in Singapore but may trade on it elsewhere.
 - B. both material and nonpublic and Green may not trade on it in any jurisdiction.
 - C. public by virtue of its release in the chat room and Green may trade on it.
5. Melvin Byrne, CFA, manages a portfolio for James Martin, a very wealthy client. Martin's portfolio is well diversified with a slight tilt toward capital appreciation. Martin requires very little income from the portfolio. Recently, Martin's brother Cliff has become a client of Byrne. Byrne proceeds to invest Cliff's portfolio in a similar manner to James's portfolio based on the fact that both brothers have a similar lifestyle and are only two years apart in age. Which of the following statements is *most accurate*?
- A. Byrne violated the Code and Standards by knowingly creating a conflict between the interests of James's and Cliff's portfolios.
 - B. Byrne violated the Code and Standards by failing to determine Cliff's objectives and constraints prior to investing his portfolio.
 - C. Byrne violated the Code and Standards by failing to have a reasonable and adequate basis for Cliff's portfolio allocation.
6. Beth Anderson, CFA, is a portfolio manager for several wealthy clients, including Reuben Carlyle. Anderson manages Carlyle's personal portfolio of stock and bond investments. Carlyle recently told Anderson that he is under investigation by the IRS for tax evasion related to his business, Carlyle Concrete (CC). After learning about the investigation, Anderson proceeds to inform a friend at a local investment bank so that they may withdraw their proposal to take CC public. Which of the following is *most likely* correct?
- A. Anderson violated the Code and Standards by failing to maintain the confidentiality of her client's information.
 - B. Anderson violated the Code and Standards by failing to detect and report the tax evasion to the proper authorities.
 - C. Anderson did not violate the Code and Standards because the information she conveyed pertained to illegal activities on the part of her client.
7. Robert Blair, CFA, director of research, has had an ongoing battle with management about the adequacy of the firm's compliance system. Recently, it has come to Blair's attention that the firm's compliance procedures are inadequate in that they are not being monitored or carefully followed. What should Blair *most appropriately* do?
- A. Resign from the firm unless the compliance system is strengthened and followed.
 - B. Send his superior a memo outlining the problem.
 - C. Decline in writing to continue to accept supervisory responsibility until reasonable compliance procedures are adopted.
8. Eugene Nieder, CFA, has just accepted a new job as a quantitative analyst for Paschal Investments, LLP. Nieder developed a complex model while working for his previous employer and plans to recreate the model for Paschal. Nieder did not make copies of the model or any supporting documents because his employer refused to grant him permission to do so. Nieder will recreate the model from memory. Which of the following statements is *most likely* correct?
- A. Nieder can recreate the model without violating the Code and Standards as long as he also generates supporting documentation.

- B. Nieder can recreate the model without violating the Code and Standards without documentation if the model is modified from its original form.
 - C. Nieder cannot recreate the model without violating the Code and Standards because it is the property of his former employer.
9. Fred Johnson, CFA, a financial analyst and avid windsurfer, has begun an investment survey of the water sports leisure industry. His brother sells windsurfing gear in Tampa and tells him that Swordfish9 is the “hottest windsurfing rig on the market and will be highly profitable for Swordfish Enterprises.” Johnson had never heard of Swordfish9 previously, but after testing the board himself became very excited about the Swordfish9 and issued an investment recommendation of “buy” on Swordfish Enterprises. As a result of issuing the recommendation, Johnson has:
- A. not violated the Code and Standards.
 - B. violated the Code and Standards by failing to establish a reasonable and adequate basis.
 - C. violated the Code and Standards by failing to consider the suitability of the investment for his clients.
10. Neiman Investment Co. receives brokerage business from Pick Asset Management in exchange for referring prospective clients to Pick. Pick advises clients—in writing, at the time the relationship is established—of the nature of its arrangement with Neiman. With regard to this practice, Pick has:
- A. complied with the Code and Standards.
 - B. violated the Code and Standards by failing to preserve the confidentiality of the agreement with Neiman.
 - C. violated the Code and Standards by inappropriately negotiating an agreement that creates a conflict of interest.
11. After sitting for the Level I CFA exam, Cynthia White visits CFA Haven, an online forum, to express her frustration. White writes, “CFA Institute is not doing a competent job of evaluating candidates, because none of the questions in the June exam touched on Alternative Investments.” Regarding the CFA Institute Standards of Professional Conduct, White *most likely*:
- A. did not violate any Standard, because she was exercising her right to freedom of speech.
 - B. violated the Standards by discussing exam content.
 - C. violated the Standards by impugning the reputation of CFA Institute.

ANSWER KEY FOR MODULE QUIZZES

Module Quiz 73.1

1. **B** According to Standard I(A) Knowledge of the Law, because she has taken steps to stop the illegal activities and the board has ignored her, Jones must dissociate from the board and seek legal advice as to what other actions would be appropriate in this instance. She may need to inform legal or regulatory authorities of the illegal activities. (LOS 73.a, 73.b)
2. **C** There can be no assurance that a premium of 2% to 4% will consistently be obtained. Bixby is in violation of Standard I(C) Misrepresentation, because she has made an implicit guarantee of the fund’s expected performance. (LOS 73.a, 73.b)
3. **A** Quigley’s trades are most likely an attempt to take advantage of an arbitrage opportunity that exists between Craeger’s common stock and its put options. She is not manipulating the prices of securities in an attempt to mislead market participants, which would violate Standard II(B) Market Manipulation. She is pursuing a legitimate investment strategy. Participants in her hedge fund are aware of the fund’s investment strategy, and thus, Quigley did not violate the Code and Standards by not disclosing this specific set of trades in advance of trading. (LOS 73.a, 73.b)

4. **B** The release of such information to a limited circle via an internet chat room does not cause the information to be public. The information is also clearly material. Therefore, Green is not allowed to trade on the information under Standard II(A) Material Nonpublic Information. (LOS 73.a, 73.b)
5. **B** Standard III(C) Suitability requires that before taking investment action, members and candidates must make a reasonable inquiry into a client's or prospect's investment objectives and constraints, as well as their prior investment experience. Byrne cannot assume that because the brothers have similar lifestyles and are close in age that they should have similarly managed portfolios. Byrne should have interviewed Cliff directly before investing his portfolio. (LOS 73.a, 73.b)
6. **A** Standard III(E) Preservation of Confidentiality requires Anderson to maintain the confidentiality of client information. Confidentiality may be broken in instances involving illegal activities on the part of the client, but the client's information may only be relayed to proper authorities. Anderson did not have the right to inform the investment bank of her client's investigation. (LOS 73.a, 73.b)
7. **C** According to Standard IV(C) Responsibilities of Supervisors, because he is aware that the firm's compliance procedures are not being monitored and followed and because he has repeatedly tried to get company management to correct the situation, Blair should decline supervisory responsibility until adequate procedures to detect and prevent violations of laws, regulations, and the Code and Standards are adopted and followed. If he does not do so, he will be in violation of the Code and Standards. (LOS 73.a, 73.b)
8. **A** Nieder must not take models or documents from his previous employer without explicit permission to do so, or he would violate Standard IV(A) Loyalty. He is allowed to reproduce the model from memory but must recreate the supporting documentation to maintain compliance with Standard V(C) Record Retention. (LOS 73.a, 73.b)
9. **B** Johnson has apparently let his recreational passion cloud his judgment. This is not to say that Swordfish Enterprises is not or will not be an excellent investment. However, if he had never heard of the firm previously, issuing an investment recommendation without conducting a thorough financial investigation indicates a failure to exercise diligence and also indicates that he lacks a reasonable and adequate basis for his recommendation. He is in violation of Standard V(A) Diligence and Reasonable Basis. (LOS 73.a, 73.b)
10. **A** There is no violation of the Standards regarding this matter. The referral arrangement is fully disclosed to clients before they agree to do business with Pick. Therefore, clients can fully assess the effect of the agreement on the referral and how the agreement may affect their accounts before hiring Pick as their asset manager. (LOS 73.a, 73.b)
11. **B** Standard VII(A) Conduct as Participants in the CFA Programs prohibits candidates from revealing which portions of the Candidate Body of Knowledge were or were not covered on an exam. Members and candidates are free to disagree with the policies, procedures, or positions taken by the CFA Institute and express their opinion on such policies, procedures, and positions. (LOS 73.a, 73.b)

TOPIC QUIZ: ETHICAL AND PROFESSIONAL STANDARDS

You have now finished the Ethical and Professional Standards topic section. Please log into your Schweser online dashboard and take the Topic Quiz on Ethical and Professional Standards. The Topic Quiz provides immediate feedback on how effective your study has been for this material. The number of questions on this quiz is approximately the number of questions for the topic on one-half of the actual Level I CFA exam. Questions are more exam-like than typical Module Quiz or QBank questions; a score of less than 70% indicates that your study likely needs improvement. These tests are best taken timed; allow 1.5 minutes per question.

After you've completed this Topic Quiz, select "Performance Tracker" to view a breakdown of your score. Select "Compare with Others" to display how your score on the Topic Quiz compares with the scores of others who entered their answers.

FORMULAS

$$\text{holding period return} = \frac{\text{end-of-period value}}{\text{beginning-of-period value}} - 1$$

$$= \frac{P_t + \text{Div}_t}{P_0} - 1 = \frac{P_t - P_0 + \text{Div}_t}{P_0}$$

$$\text{arithmetic mean return} = \frac{(R_1 + R_2 + R_3 + \dots + R_n)}{n}$$

$$\text{geometric mean return} = \sqrt[n]{(1 + R_1) \times (1 + R_2) \times (1 + R_3) \times \dots \times (1 + R_n)} - 1$$

$$\text{correlation: } \rho_{1,2} = \frac{\text{Cov}_{1,2}}{\sigma_1 \times \sigma_2}$$

standard deviation for a two-asset portfolio:

$$\sigma_p = \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \sigma_1 \sigma_2 \rho_{1,2}} \text{ or } \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \text{Cov}_{1,2}}$$

$$\text{equation of the CML: } E(R_p) = R_f + \left(\frac{E(R_M) - R_f}{\sigma_M} \right) \sigma_p$$

$$E(R_p) = R_f + (E(R_M) - R_f) \left(\frac{\sigma_p}{\sigma_M} \right)$$

total risk = systematic risk + unsystematic risk

$$\beta_i = \frac{\text{Cov}_{i,\text{mkt}}}{\sigma_{\text{mkt}}^2} = \rho_{i,\text{mkt}} \frac{\sigma_i}{\sigma_{\text{mkt}}}$$

capital asset pricing model (CAPM): $E(R_i) = R_f + \beta_i [E(R_{\text{mkt}}) - R_f]$

$$\text{Sharpe ratio} = \left(\frac{R_p - R_f}{\sigma_p} \right)$$

$$\text{M-squared} = R_f + \frac{\sigma_M}{\sigma_p} (R_p - R_f)$$

$$\text{Treynor measure} = \frac{R_p - R_f}{\beta_p}$$

$$\text{Jensen's alpha} = \alpha_p = R_p - [R_f + \beta_p (R_M - R_f)]$$

INDEX

A

ability to bear risk, 69
accounting risk, 100
actively managed funds, 7
active management, 6
active portfolio management, 45
after-tax nominal return, 17
algorithmic trading, 130
alternative data, 128
anchoring and adjustment bias, 85
arithmetic mean return, 16
artificial intelligence, 129
attribution analysis, 57
availability bias, 87
average convergence/divergence, 119

B

bank, 3
bar charts, 111
base-rate neglect, 83
Belief Perseverance, 82
best-in-class, 74
beta, 53, 101
Big Data, 128
blockchain, 131
Bollinger bands, 117
Bollinger band width indicator, 117
bond mutual funds, 7
breakdown, 113
breakout, 113
buy-side firms, 6

C

candlestick charts, 111
capital allocation line (CAL), 24, 42
capital asset pricing model (CAPM), 52
capital market line (CML), 44
certainty overconfidence, 89
change in polarity, 114
characteristic line, 50
closed-end funds, 7
Code of Ethics, 146
cognitive dissonance, 82
cognitive errors, 82
cognitive errors, 94
composite, 185
confirmation bias, 83
conservatism bias, 82
consolidation, 113
continuation patterns, 115
contrarian strategy, 117
convergence, 118
core-satellite approach , 73
corporate exhaust, 128
correlation, 27
covariance, 27
credit risk, 99
cryptocurrencies, 131
cyber risk, 100

D

data science, 128
deep learning, 129
defined benefit pension plan, 4
defined contribution pension plan, 4
delta, 101
derivatives risks, 101
discretion, 185
distributed ledger, 130
divergence, 118

diversifiable risk, 45
diversification, 103
diversification ratio, 2
double bottom pattern, 115
double top pattern, 115
downside risk, 101
downtrend, 113
duration, 101

E

efficient frontier, 33
emotional biases, 82, 94
endowment, 3
endowment bias, 90
engagement/active ownership investing, 74
ESG integration, 74
exchange-traded funds (ETFs), 8
execution step, 2

F

factor loading, 48
factor sensitivity, 48
feedback step, 3
fidelity bonds, 103
financial risk, 99
fintech, 127
firm-specific risk, 45
flags, 116
foundation, 3
framing bias, 86
full-service asset managers, 6

G

gamma, 101
geometric mean return, 16

Global Investment Performance Standards (GIPS), 183

global minimum-variance portfolio, 33

governmental risk, 100

gross return, 16

H

halo effect, 91

head-and-shoulders pattern, 114

hedge funds, 8

herding behavior, 90

high-frequency trading, 130

hind sight bias, 85

historical data, 27

holding period return (HPR), 15

home bias, 92

hyperbolic discounting, 89

I

illusion of control bias, 84

illusion of knowledge, 84, 89

impact investing, 74

index funds, 7

indifference curve, 22

individual investors, 3

initial coin offerings, 131

institutions, 3

insurance companies, 3

intermarket analysis, 120

Internet of Things, 128

inverse head-and-shoulders pattern, 115

investment companies, 3

investment constraints, 70

investment policy statement (IPS), 2, 67

J

Jensen's alpha, 59

L

latency, 128
legal and regulatory constraints, 71
legal risk, 100
leveraged return, 17
line charts, 111
liquidity, 70
liquidity risk, 99
load funds, 7
longevity risk, 100
loss-aversion bias, 88

M

M^2 alpha, 58
machine learning, 129
margin debt, 120
market model, 48
market portfolio, 43
market risk, 45, 99
market risk premium, 44
mental accounting bias, 85
miners, 131
minimum-variance portfolios, 33
minimum-variance frontier, 33
model risk, 100
momentum oscillators, 118
money market funds, 7
money-weighted rate of return, 17
mortality risk, 100
mosaic theory, 159
moving average lines, 116
 M -squared (M^2), 58
multi-boutique firm, 6
multifactor models, 47
mutual funds, 3, 4

N

natural language processing, 130
negative screening, 74
net asset value (NAV), 7
net return, 16
neural networks, 129
no-load funds, 7
nondiversifiable risk, 45
non-financial risks, 100

O

open-end fund, 7
operational risk, 100
overbought market, 117
overconfidence, 84
overconfidence bias, 89
overfitting, 129
oversold market, 117

P

passive investment strategy, 45
passive management, 6
pennants, 116
performance evaluation, 57
performance measurement, 20
permissioned networks, 131
permissionless networks, 131
planning step, 2
political risk, 100
pooled investments, 7
portfolio perspective, 1
positive screening, 74
prediction overconfidence, 89
pretax nominal return, 17
private equity, 8
profession, 138

professional conduct program, 145
put/call ratio, 119

R

rate of change (ROC) oscillator, 118
real return, 17
rectangles, 116
regret-aversion bias, 90
regulatory risk, 100
relative risk objectives, 69
relative strength analysis, 113
relative strength charts, 120
Relative Strength Index, 118
representativeness bias, 83
resistance level, 113
responsible investing, 71
return generating models, 47
reversal patterns, 114
rho, 101
risk-averse, 22
risk aversion coefficient, 23
risk budgeting, 73, 99
risk governance, 98
risk management, 97
risk management framework, 98
risk-neutral, 22
risk objectives, 68
risk-seeking, 22
risk shifting, 103
risk tolerance, 99
risk transfer, 103
robo-advisors, 6, 130

S

sample-size neglect, 84
scenario analysis, 102
security characteristic line, 50

security market line (SML), 51
self-attribution, 84
self-attribution bias, 89
self-control bias, 89
self-enhancing, 89
self-insurance, 103
self-protecting, 89
sell-side firms, 6
sentiment indicators, 119
separately managed account, 8
Sharpe ratio, 58
single-index model, 48
smart beta, 6
smart contracts, 131
solvency risk, 100
sovereign wealth funds, 4
specialist asset managers, 6
squeeze, 118
standard deviation, 101
Standard I(A) Knowledge of the Law, 190
Standard I(B) Independence and Objectivity, 190
Standard I(C) Misrepresentation, 190
Standard I(D) Misconduct, 190
Standard II(A) Material Nonpublic Information, 191
Standard II(B) Market Manipulation, 191
Standard III(A) Loyalty, Prudence, and Care, 191
Standard III(B) Fair Dealing, 192
Standard III(C) Suitability, 192
Standard III(D) Performance Presentation, 192
Standard III: Duties to Clients, 191
Standard III(E) Preservation of Confidentiality, 192
Standard II: Integrity of Capital Markets, 191
Standard I: Professionalism, 190
Standard IV(A) Loyalty, 192
Standard IV(B) Additional Compensation Arrangements, 193
Standard IV(C) Responsibilities of Supervisors, 193
Standard IV: Duties to Employers, 192
standards of professional conduct, 147
Standard V(A) Diligence and Reasonable Basis, 193
Standard V(B) Communication with Clients and Prospective Clients, 194
Standard V(C) Record Retention, 194

Standard VI(A) Disclosure of Conflicts, 194
Standard VI(B) Priority of Transactions, 194
Standard VI: Conflicts of Interest, 194
Standard VI(C) Referral Fees, 195
Standard VII(A) Conduct as Participants in CFA Institute Programs, 195
Standard VII(B) Reference to CFA Institute, the CFA Designation, and the CFA Program, 195
Standard VII: Responsibilities as a CFA Institute Member or CFA Candidate, 195
Standard V: Investment Analysis, Recommendations, and Actions, 193
status quo bias, 90
stochastic oscillator, 119
stock mutual funds, 7
strategic asset allocation, 71
stress testing, 102
supervised learning, 129
support level, 113
surety bond, 103
systematic risk, 45

T

tactical asset allocation, 73
tail risk, 100, 101
tax risk, 100
tax situation, 70
technical analysis, 109
text analytics, 129
thematic investing, 74
time horizon, 70
time-weighted rate of return, 19, 20
tokenization, 132
trend, 113
trendline, 113
Treynor measure, 59
triangles, 115
triple bottom pattern, 115
triple top pattern, 115

U

underfitting, 129
unique circumstances, 71
unique risk, 45
unsupervised learning, 129
unsystematic risk, 45
uptrend, 113
utility function, 22

V

value at risk (VaR), 101
vega, 101
venture capital, 8
Volatility Index (VIX), 120
volume charts, 113

W

willingness to bear risk, 69