2019 CFA® PROGRAM CURRICULUM LEVEL III VOLUMES 1-6

CFA Institute®
CFA Program

WILEY

© 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006 by CFA Institute. All rights reserved.

This copyright covers material written expressly for this volume by the editor/s as well as the compilation itself. It does not cover the individual selections herein that first appeared elsewhere. Permission to reprint these has been obtained by CFA Institute for this edition only. Further reproductions by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval systems, must be arranged with the individual copyright holders noted.

CFA®, Chartered Financial Analyst®, AIMR-PPS®, and GIPS® are just a few of the trademarks owned by CFA Institute. To view a list of CFA Institute trademarks and the Guide for Use of CFA Institute Marks, please visit our website at www.cfainstitute.org.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional service. If legal advice or other expert assistance is required, the services of a competent professional should be sought.

All trademarks, service marks, registered trademarks, and registered service marks are the property of their respective owners and are used herein for identification purposes only.

ISBN 978-1-946442-19-2 (paper)
ISBN 978-1-946442-43-7 (ebk)

10 9 8 7 6 5 4 3 2 1

Please visit our website at
www.WileyGlobalFinance.com.

WILEY

CFA Institute®
CFA Program

APPLICATIONS OF ECONOMIC ANALYSIS AND ASSET ALLOCATION

CFA® Program Curriculum
2019 • LEVEL III • VOLUME 3

WILEY

CONTENTS

Portfolio Management

STUDY SESSIONS

Study Session 3	The Asset Management Industry and Professionalism
Study Session 4	Behavioral Finance
Study Session 5	Private Wealth Management (1)
Study Session 6	Private Wealth Management (2)
Study Session 7	Portfolio Management for Institutional Investors
Study Session 8	Applications of Economic Analysis to Portfolio Management
Study Session 9	Asset Allocation and Related Decisions in Portfolio Management (1)
Study Session 10	Asset Allocation and Related Decisions in Portfolio Management (2)
Study Session 11	Fixed-Income Portfolio Management (1)
Study Session 12	Fixed-Income Portfolio Management (2)
Study Session 13	Equity Portfolio Management (1)
Study Session 14	Equity Portfolio Management (2)
Study Session 15	Alternative Investments for Portfolio Management
Study Session 16	Risk Management
Study Session 17	Risk Management Applications of Derivatives
Study Session 18	Trading
Study Session 19	Performance Evaluation and Global Investment Performance Standards

This volume includes Study Sessions 8–10.

© 2018 CFA Institute. All rights reserved.

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to prepare an appropriate investment policy statement and asset allocation; formulate strategies for managing, monitoring, and rebalancing investment portfolios; evaluate portfolio performance; and analyze a presentation of investment returns for consistency with Global Investment Performance Standards (GIPS®).

19年CFA-FRM招生简章

一、SVIP资料包括以下内容：【所有课程均有配套课件讲义】

前导课程：考试介绍、数量基础、金融英语、金融市场与产品的介绍，为学员建立考试框架，打牢基础；
基础课程：紧扣考纲要求，对考点全面解析，建立系统知识框架， 帮助学员根据课程进度系统性学习；
强化课程：突出重点，化繁为简。强化对重要知识点的掌握，让学员完全把握知识点与掌握解题技巧；
直播串讲：以习题串讲的形式，针对每门课程进行对应的重点难点讲解，帮助考生更好的掌握对应知识；
冲刺课程：分析历年考试的重点和难点，通过2次模考与讲解，强化对重点的知识的把握，提高应试技巧；
百题预测：考点精准预测85%以上。以题带点，掌握出题思路与解题方法，最终通过考试取得证书；
赠送资料：官方原版教材 高清NOTES 习题 秘籍 考纲解读等电子资料；

二、SVIP价格说明：【一次收费到当期课程完结，绝无二次收费】

2019年课程价格：500元。购买会赠送18/17年本级别全套课程

往期课程价格：288元。

备注：此价格为一个级别全套课程价格，无其他任何隐形收费。

限时优惠：资料逐步涨价，第一阶段（截止5月1日）500元，第二阶段（截止考试结束）600元。

三、付款方式：

联系微信：fcayyh

（支持支付宝、微信等支付方式！）

四、资料发放方式：

主要通过百度网盘提供下载，课程更新同样在百度网盘里进行。

为确保课程不外传，所有课程全程加密，每位VIP会员给两台设备授权观看。

课程支持windows/mac/ios/安卓等系统，支持手机/平板/电脑等设备。

五、2018年部分会员PASS报喜截图：【限于篇幅，随机抽取】

PORTFOLIO MANAGEMENT

STUDY SESSION

8

Applications of Economic Analysis to Portfolio Management

A necessary task in the investment management process is to formulate capital market expectations. These forecasts of risk and return for various asset classes form the basis for constructing portfolios that maximize expected return for given levels of risk.

This study session examines the process of setting capital market expectations and covers major tools of economic analysis. The application of neo-classical growth theory to develop economic forecasts is presented. The discussion includes how economic forecasts can be integrated with equity valuation techniques to value an equity market.

READING ASSIGNMENT

Reading 16	Capital Market Expectations by John P. Calverley, Alan M. Meder, CPA, CFA, Brian D. Singer, CFA, and Renato Staub, PhD
Reading 17	Equity Market Valuation by Peter C. Stimes, CFA, and Stephen E. Wilcox, PhD, CFA

© 2018 CFA Institute. All rights reserved.

READING

16

Capital Market Expectations

by John P. Calverley, Alan M. Meder, CPA, CFA, Brian D. Singer, CFA, and Renato Staub, PhD

John P. Calverley (Canada). Alan M. Meder, CPA, CFA, is at Duff & Phelps Investment Management Co. (USA). Brian D. Singer, CFA, is at William Blair (USA). Renato Staub, PhD, is at William Blair Investment Management (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** discuss the role of, and a framework for, capital market expectations in the portfolio management process;
☐	**b.** discuss challenges in developing capital market forecasts;
☐	**c.** demonstrate the application of formal tools for setting capital market expectations, including statistical tools, discounted cash flow models, the risk premium approach, and financial equilibrium models;
☐	**d.** explain the use of survey and panel methods and judgment in setting capital market expectations;
☐	**e.** discuss the inventory and business cycles and the effects that consumer and business spending and monetary and fiscal policy have on the business cycle;
☐	**f.** discuss the effects that the phases of the business cycle have on short-term/long-term capital market returns;
☐	**g.** explain the relationship of inflation to the business cycle and the implications of inflation for cash, bonds, equity, and real estate returns;
☐	**h.** demonstrate the use of the Taylor rule to predict central bank behavior;
☐	**i.** interpret the shape of the yield curve as an economic predictor and discuss the relationship between the yield curve and fiscal and monetary policy;
☐	**j.** identify and interpret the components of economic growth trends and demonstrate the application of economic growth trend analysis to the formulation of capital market expectations;

(continued)

Managing Investment Portfolios: A Dynamic Process, Third Edition, John L. Maginn, CFA, Donald L. Tuttle, CFA, Jerald E. Pinto, CFA, and Dennis W. McLeavey, CFA, editors. © 2007 CFA Institute. All rights reserved. Revision © 2017 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	The candidate should be able to:
☐	**k.** explain how exogenous shocks may affect economic growth trends;
☐	**l.** identify and interpret macroeconomic, interest rate, and exchange rate linkages between economies;
☐	**m.** discuss the risks faced by investors in emerging-market securities and the country risk analysis techniques used to evaluate emerging market economies;
☐	**n.** compare the major approaches to economic forecasting;
☐	**o.** demonstrate the use of economic information in forecasting asset class returns;
☐	**p.** explain how economic and competitive factors can affect investment markets, sectors, and specific securities;
☐	**q.** discuss the relative advantages and limitations of the major approaches to forecasting exchange rates;
☐	**r.** recommend and justify changes in the component weights of a global investment portfolio based on trends and expected changes in macroeconomic factors.

1 INTRODUCTION

A noted investment authority has written that the "fundamental law of investing is the uncertainty of the future."[1] Yet investors have no choice but to forecast at least elements of the future because nearly all investment decisions look toward it. Specifically, investment decisions incorporate the decision maker's expectations concerning factors and events believed to affect investment values. The decision maker finally integrates these views into expectations about the risk and return prospects of individual assets and groups of assets.

The particular concern of this reading is **capital market expectations** (CME): the investor's expectations concerning the risk and return prospects of asset classes, however broadly or narrowly the investor defines those asset classes. Capital market expectations are an essential input to formulating a strategic asset allocation. For example, if an investor's investment policy statement specifies and defines eight permissible asset classes, the investor will need to have formulated long-term expectations concerning those asset classes to develop a strategic asset allocation. The investor may also act on short-term expectations. Capital market expectations are expectations about classes of assets, or **macro expectations**. By contrast, **micro expectations** are expectations concerning individual assets. Micro expectations are key ingredients in security selection and valuation. Insights into capital markets gleaned during CME setting should help in formulating accurate micro expectations in security selection and valuation.

One theme of this reading is that a disciplined approach to expectations setting will be rewarded. Therefore, much of the reading is devoted to explaining a widely applicable expectations-setting process. A second theme of this reading is that skillful

1 Peter L. Bernstein in the foreword to Rapaport and Mauboussin (2001), p. xiii.

economic analysis can contribute to expectations setting. That theme is supported by the observation that securities markets trade claims on the cash flows of the business sector and that other markets reflect the macro economy too.

The reading is organized as follows: Section 2 presents a general framework for developing capital market expectations and alerts the reader to the range of problems and pitfalls that await the investor or analyst in this arena. Section 3 then turns to describing the range of tools, both formal and judgmental, that an analyst may use in expectations setting. Section 4 covers economic analysis as applied to formulating capital market expectations, and we then summarize the reading.

2 ORGANIZING THE TASK: FRAMEWORK AND CHALLENGES

In this section, we provide a guide to collecting, organizing, combining, and interpreting information. After illustrating the process, we turn to a discussion of typical problems and challenges to formulating the most informed judgments possible.

2.1 A Framework for Developing Capital Market Expectations

The following is a framework for a disciplined approach to setting CME.

1 *Specify the final set of expectations that are needed, including the time horizon to which they apply.* The analyst needs to understand the specific objectives of the analysis in order to work efficiently toward them. To make this step even more concrete, the analyst should write the questions that need to be answered. Accomplishing this step requires the analyst to formulate his or her specific needs in terms of a relevant set of asset classes that are of concern, giving appropriate regard to the constraints of the client. In many cases, the investor's investment policy statement may provide guidance in this task. For example, for a taxable investor with a 10-year time horizon, the portfolio manager would develop long-term after-tax expectations for use in developing a strategic asset allocation.

2 *Research the historical record.* Most forecasts have some connection to the past. For many markets, the historical record contains useful information on the investment characteristics of the asset, suggesting at least some possible ranges for future results. Beyond the raw historical facts, the analyst should seek to identify the factors that affect asset class returns and to understand the what, when, where, why, and how of these return drivers. The analyst will then have a better sense of the information mosaic that he or she will need to piece together to arrive at well-informed conclusions.

3 *Specify the method(s) and/or model(s) that will be used and their information requirements.* The investor, capital market analyst, or unit responsible for developing capital market expectations (as the case may be) should be explicit about the method(s) and/or model(s) that will be used and should be able to justify the selection. Information requirements (economic and financial market data needs, for example) depend on the decision about method(s).

4 *Determine the best sources for information needs.*

5 *Interpret the current investment environment using the selected data and methods, applying experience and judgment.* The analyst should be sure that he or she is working from a common set of assumptions in interpreting different

elements of the investment and economic scene so that the analyst's conclusions are mutually consistent. The analyst often needs to apply judgment and experience to interpret apparently conflicting signals within the data.

6 *Provide the set of expectations that are needed, documenting conclusions.* These are the analyst's answers to the questions set out in Step 1. The answers should be accompanied by the reasoning and assumptions behind them.

7 *Monitor actual outcomes and compare them to expectations, providing feedback to improve the expectations-setting process.*

Disciplined capital market expectations setting requires experience and expertise in investments and economics. Large asset managers may have a research unit—for example, an economics unit—with responsibility for developing capital market expectations. Through superior forecasts, such asset managers seek to better control risk and improve the results of actively managed accounts in particular. The development of capital market expectations is **beta research** (research related to systematic risk and returns to systematic risk). As such, it is usually centralized so that the CME inputs used across all equity and fixed-income products are consistent. On the other hand, **alpha research** (research related to capturing excess risk-adjusted returns by a particular strategy) is typically conducted within particular product groups with the requisite investment-specific expertise. For institutional investors, professional consultants are a resource for systematically developed capital market expectations. Consultants' assistance may be given in the course of asset allocation reviews or asset–liability planning studies. Institutional investors may develop capital market expectations in-house, although they will usually be aware of the perspectives of professional consultants and peers. Most individual investors rely on their investment adviser or other external source for guidance in setting capital market expectations, as they often do not have expertise in this area. Yet an adviser may incorporate the client's perspectives on capital markets prospects, as the portfolio is run on the client's behalf and the client must be comfortable with the inputs to constructing the portfolio.

The first step in the framework for developing CME requires that analysts set boundaries to focus their attention on the expectations most relevant for their investment situation. Otherwise, effort is wasted. Even pared down to the minimum needs, the scope of the expectations-setting process can be quite challenging. As Example 1 illustrates, there is a direct relationship between the number and variety of permissible asset class alternatives and the scope of the expectations-setting task facing the manager.

EXAMPLE 1

Capital Market Expectations Setting: Information Requirements (1)

Consider the tasks facing two investment managers, John Pearson and Michael Wu.

Pearson runs US balanced separately managed accounts (SMAs) for high-net-worth individuals within a bank trust department. The mandates of these accounts restrict investments to US equities, US investment-grade fixed-income instruments, and prime US money market instruments. These balanced accounts have an investment objective of long-term capital growth and income. In contrast, Wu is the chief investment officer of a large Hong Kong–based, internationally focused asset manager that uses the following types of assets within its investment process:

Equities	Fixed Income	Alternative Investments
Hong Kong-listed equities	Eurozone sovereign debt	Eastern Europe venture capital
Eurozone equities*	US government debt	New Zealand timber assets
US large-cap equities		US apartment properties
US small-cap equities		
Canadian large-cap equities		

*The **Eurozone** is the region using the euro as a currency. As of the end of 2011, the Eurozone consisted of Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia, and Spain.
Note: Venture capital is equity investment in private companies.

Wu runs SMAs with generally long-term time horizons and global tactical asset allocation (GTAA) programs. Compare and contrast the information and knowledge requirements of Pearson and Wu.

Solution:

Pearson's in-depth information requirements relate to US equity and fixed-income markets. By contrast, Wu's information requirements relate not only to US and non-US equity and fixed-income markets, but also to three alternative investment types with nonpublic markets, located on three different continents. Wu's need to be current on political, social, economic, and even trading-oriented operational details worldwide is more urgent than Pearson's. Given their respective investment time horizons, Pearson's focus is on the long term while Wu needs to focus not only on the long term but also on near-term disequilibria among markets (as he runs GTAA programs). One challenge that Pearson has in US fixed-income markets that Wu does not face is the need to cover corporate as well as government debt securities. Nevertheless, Wu's overall information and knowledge requirements are clearly more demanding than Pearson's.

In the next example, the balanced fund manager from Example 1 specifies the final set of expectations needed and the time frame for those expectations.

EXAMPLE 2

Capital Market Expectations Setting: Information Requirements (2)

Following the practice of his employer, Pearson uses the results of constrained mean–variance optimization (MVO) and information from clients' investment policy statements to develop strategic asset allocations for the balanced accounts.

Pearson is now addressing the first step in the framework given in the text for a client whose investment time horizon is five years. What set of final expectational data does Pearson need?

Solution:

Pearson needs the following final set of expectations:

- the expected US broad market annual equity total return over a five-year horizon;

- the expected US investment-grade bond annual total return over a five-year horizon;
- the standard deviation of annual returns of US broad market equities;
- the standard deviation of annual returns of US investment-grade bonds; and
- the correlation of annual US stock and US bond returns.

In total, Pearson needs two expected returns, two standard deviations, and one correlation for the MVO.

Steps 2 and 3 in the expectations-setting process involve understanding the historical performance of the asset classes and researching their return drivers. The analyst can approach these tasks by collecting macroeconomic and market information (e.g., asset returns) by:

- geographic area (e.g., domestic, nondomestic, or some subset—for example, a single international area), or
- broad asset class (e.g., equity, fixed-income, or real estate).

The finer classifications depend on the characteristics of the task and the orientation of the investor. For equities, one approach would be to further classify by economic sector, possibly making style-related (e.g., market-capitalization) distinctions. A fixed-income investor might distinguish between governmental and corporate sectors and make further credit distinctions. For example, an industry rotation equity strategist might formulate expectations on domestic equities as follows:

Economic Sector (e.g., technology manufacturers)

Industry (e.g., computer equipment manufacturers)

Subindustry (e.g., microchip component manufacturers)

EXAMPLE 3

Historical Analysis

As Peter L. Bernstein (2004) has written, forecasters who make predictions without regard to past experience have no benchmarks to distinguish between what is new about their expectations and what may be a continuation of past experience. Dimson, Marsh, and Staunton (2011), in a rigorous study covering the 111-year period from 1900 to 2010, found that equities achieved higher annualized geometric mean real returns than did bonds or bills in seventeen major national markets. It would be appropriate for an analyst forecasting that bonds would outperform equities over some (probably shorter-term) horizon to supply supporting analysis that recognizes the tension between the forecast and past long-term experience.

In Step 3, the analyst also needs to be sensitive to the fact that the effectiveness of forecasting approaches and relationships among variables may be related to the investor's time horizon. As an example, a discounted cash flow approach to developing equity market expectations is usually considered to be most appropriate to long-range forecasting.

The fourth step involves determining the best sources for information needs. Executing this step well requires that the analyst research the quality of alternative data sources. Factors such as data collection principles and definitions, error rates in collection, calculation formulas, and for asset class indexes, qualities such as investability, correction for free float, turnover in index constituents, and biases in the data are relevant. The cost of data may also be relevant. In short, analysts must understand everything they can about the data they will use for analysis. Using flawed or misunderstood data is a recipe for faulty analysis. Furthermore, analysts should constantly be alert to new, superior sources for their data needs.

Besides taking care with data sources, the analyst must select the appropriate data frequency. For instance, long-term data series should not be used for setting short-term trading expectations or evaluating short-term volatility. Daily series are of more use for setting shorter-term capital market expectations. Quarterly or annual data series are useful for setting long-term capital market expectations.

The fifth step involves interpreting the current investment environment using the selected data and methods, applying experience and judgment. In the sixth step, we take all of our analyses of the economic and market environment into forward-looking views on capital markets, developing any required quantitative forecasts. In other words, the questions formulated in Step 1 are answered in Step 6. Economic analysis may work itself into quantitative forecasts in a variety of ways depending on the investor's selection of methodology. Top-down investment approaches often use economic analysis more intensively than bottom-up approaches. Example 4 illustrates several ways an analyst's relative optimism or pessimism concerning a market might be reflected in quantitative forecasts.

EXAMPLE 4

Incorporating Economic Analysis into Expected Return Estimates

Michael Wu has gathered information on consensus expectations in equity and fixed-income markets. On the basis of his economic analysis, Wu is optimistic relative to the consensus on the prospects for Hong Kong equities. On the other hand, Wu is pessimistic relative to the consensus on the prospects for US large-cap equities. Depending on the model chosen, Wu's views might be reflected in his quantitative expectations in several ways, including the following:

- **Historical mean return with adjustments**. If Wu takes a historical mean return as his baseline for each asset class, he may make an upward adjustment to that mean for Hong Kong equities and a downward adjustment for US large-cap equities.
- **Risk premium approach**. Wu may frame his analysis in terms of the equity risk premium (the expected return on equities in excess of the long bond expected return). After translating his views into equity risk premium estimates for Hong Kong and US large-cap equities, his return expectation for each asset class is the expected equity risk premium in each market plus the long bond expected return in each market (which he can estimate directly from the term structure of interest rates).

- **Discounted cash flow (DCF) model estimates**. Wu may use his economic analysis to forecast the growth rates of corporate profits for the United States and Hong Kong and input those forecasts into a DCF model solved for the required return on equities in each market.
- **Implied market estimates of expected return with adjustment**. Making use of a world market benchmark and a methodology known as the Black–Litterman model, Wu may infer the equilibrium expected returns on asset classes as reflected by their values in the allocated world market benchmark. Wu can then incorporate his own views on Hong Kong and US large-cap equities using a procedure specified by Black–Litterman.[2]

For a Hong Kong–based client, Hong Kong dollar returns are relevant, so Wu will also need to make exchange rate forecasts to arrive at his conclusions.

Finally, we want to use experience to improve the expectations-setting process. We measure our previously formed expectations against actual results to assess the level of accuracy that the expectations-setting process is delivering. Generally, good forecasts are:

- unbiased, objective, and well researched;
- efficient, in the sense of reducing the magnitude of forecast errors to a minimum; and
- internally consistent.

Internal inconsistency can take a number of forms. For example, domestic bond and domestic equity expectations developed by different analysts using different inflation projections would not be internally consistent. A restructuring of a portfolio based on those expectations would, at least in part, merely reflect an unresolved difference in assumptions. In some cases, inconsistent forecasts may result in conclusions that are implausible or impossible. Example 5 illustrates inconsistent statistical forecasts.

EXAMPLE 5

Inconsistency of Correlation Estimates: An Illustration

Frequently, the expected correlations between asset classes form part of the final expectational data that an analyst needs. If the number of asset classes is n, the analyst will need to estimate $(n^2 - n)/2$ distinct correlations (or the same number of distinct covariances). In doing so, the analyst must be sure that his or her estimates are consistent. For example, consider the correlation matrix for three assets shown in Exhibit 1.

Exhibit 1 Inconsistent Correlations

	Market 1	Market 2	Market 3
Market 1	1	–1	–1
Market 2	–1	1	–1
Market 3	–1	–1	1

2 The Black–Litterman model is discussed further in the reading on asset allocation.

According to Exhibit 1, the estimated correlation between each asset and each other asset is −1. These estimates are internally inconsistent and, in fact, not possible. If Markets 1 and 2 are perfectly negatively correlated and Markets 2 and 3 are as well, then Markets 1 and 3 should be perfectly positively correlated rather than perfectly negatively correlated.

Other cases of an inconsistent correlation matrix are not so obvious.[3]

As a result of the final feedback step, we may be able to identify and correct weaknesses in our expectations-setting process or methods.

2.2 Challenges in Forecasting

A range of problems can frustrate analysts' expectations-setting efforts. Expectations reflecting faulty analysis or assumptions may cause a portfolio manager to construct a portfolio that is inappropriate for the client. At the least, the portfolio manager may incur the costs of changing portfolio composition without any offsetting benefits. On the principle that forewarned is forearmed, the following sections provide guidance on the points where special caution is warranted. The discussion focuses on problems in the use of data and on analyst mistakes and biases.

2.2.1 *Limitations of Economic Data*

The analyst needs to understand the definition, construction, timeliness, and accuracy of any data used, including any biases. The time lag with which economic data are collected, processed, and disseminated can be an impediment to their use. Although in some highly developed markets some economic data may be reported with a lag as short as one week, other important data may be reported with a lag of more than a quarter. The International Monetary Fund sometimes reports macroeconomic data for developing economies with a lag of two years or more. Older data for a variable increase the uncertainty concerning the current state of the economy with respect to that variable.

Furthermore, one or more official revisions to the initial values are common. In effect, measurements are made with error, but the direction and magnitude of the error are not known at the time the data are initially publicized.

Definitions and calculation methods change too. For example, the sampling procedures and calculation methods for the US Consumer Price Index for All Urban Consumers (CPI-U) used by the Bureau of Labor Statistics (BLS) have changed in significant ways since the series was first published. In 1983, for example, the BLS shifted to a flow-of-services model for pricing owner-occupied housing, based on the costs that would be associated with renting such housing. In 1991, the BLS began the introduction of hedonic or regression-based quality adjustments to prices to reflect any increases in quality and features of various consumption items.

3 What may look like a viable correlation matrix at first inspection is not necessarily feasible. In a three-asset case, it is feasible for all pairwise correlations to be −0.50; however, it can be shown that correlations that are all equal to −0.51 are not feasible (i.e., are inconsistent). Correlations must be consistent for variances to be nonnegative.

EXAMPLE 6

A Change in Focus from GNP to GDP

In the late 1980s, expanding international trade caused economists to favor the use of Gross Domestic Product (GDP) over Gross National Product (GNP). Basically, GDP measures production within national borders regardless of whether the labor and property inputs are domestically or foreign owned. In contrast, GNP makes an adjustment to GDP equal to the receipts of factor income from the rest of the world to the country, less the payments of factor income from a country to the rest of the world. This change in preference reflected the fact that product subcomponents, such as automobile parts, were being created in various regions of the world. Thus, measuring economic activity according to what nation was responsible for activities in various regions of the world was becoming more difficult and less useful than being able to measure what was being made within a nation or particular region. Consistent with this observation, the United Nations System of National Accounts (known as UNSNA, or SNA for short) emphasizes GDP.

An analyst must realize that suppliers of indexes of economic and financial data periodically **re-base** these indexes, meaning that the specific time period used as the base of the index is changed. A re-basing is not a substantive change in the composition of an index. It is more of a mathematical change. Analysts constructing a data series should take care that data relating to different bases are not inadvertently mixed together.

2.2.2 *Data Measurement Errors and Biases*

Analysts need to be aware of possible biases in data measurement of series such as asset class returns. Errors in data series include the following:

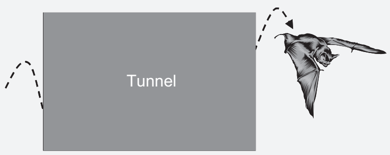
- **Transcription errors**. These are errors in gathering and recording data. Such errors are most serious if they reflect a bias.
- **Survivorship bias**. Survivorship bias arises when a data series reflects only entities that have survived to the end of the period. For example, a share index may be based on companies that trade on an exchange. Such companies are often delisted (removed from trading on the exchange) after events such as bankruptcy filings and mergers. Shares of bankrupt companies may trade elsewhere after delisting. Do reported returns on a share index reflect post-delisting returns? If not, the return series will probably convey an overly optimistic picture of the real-time investment returns from owning all listed shares. Without correction, statistics derived from series subject to survivorship bias can be misleading in the forward-looking context of expectations setting.[4]
- **Appraisal (smoothed) data**. For certain assets without liquid public markets, appraisal data are used in lieu of market price transaction data. Appraised values tend to be less volatile than market-determined values for the identical asset would be. The consequences are 1) the calculated correlations with other assets tend to be smaller in absolute value than the true correlations, and 2) the true standard deviation of the asset is biased downward. This concern has been raised particularly with respect to alternative investments such as real estate.

4 See Brown, Goetzmann, and Ross (1995).

EXAMPLE 7

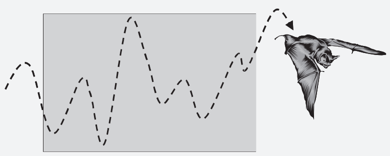
Smoothed Data: The Case of Alternative Investments (1)

The perception of alternative investments is that they yield high returns with low risk and that they barely correlate with traditional asset classes. At least in some cases, this perception results from the uncritical use of flawed historical statistics because alternative assets are not traded on exchanges with continuously observable markets. First, risk is underestimated. Consider the following analogy: A bat is flying through a dark tunnel. While it is in the tunnel, you cannot see it. The bat may exit from the tunnel at about the same height it entered the tunnel, as shown.



Source: UBS Asset Management White Paper Series – "Global Investment Solutions – Capital Market Assumptions," February 2005 © UBS 2005. All rights reserved. Reproduced with permission.

However, the bat's flight within the tunnel, if it could be viewed, would be seen to go up and down:



Source: UBS Asset Management White Paper Series – "Global Investment Solutions – Capital Market Assumptions," February 2005 © UBS 2005. All rights reserved. Reproduced with permission.

In this analogy, the time in the tunnel corresponds to the time between trades (or fund valuations) and the bat's height of flight corresponds to the true price of the asset. In measuring the bat's height only at the points of entry and exit from the tunnel, we would underestimate the real volatility of price. Asset liquidity corresponds to the end of the tunnel, when the true price is first clearly visible. In the context of venture capital, for instance, the end of the tunnel is analogous to the initial public offering date.

Data for alternative investments tend to overly smooth return variation because they are often appraisal-based rather than transaction-based. Many indexes, such as those for real estate, private equity, and natural resources, were created with a focus on measuring return rather than risk. Unfortunately, these indexes have been used to derive risk and correlation estimates that are biased downward. For alternative investments, the issue is not only whether the past is a good indicator of the future, but also whether the past is even correctly recorded.

As an illustration, consider the quarterly returns on the S&P 500 between 1981 and 1999, which include the crash of 1987. The period contains 18 negative quarters and has an annual standard deviation of returns of 16.1 percent. Venture capital also represents equity claims, but on less seasoned and riskier companies. Nevertheless, based on venture economics data, the index-based quarterly venture capital returns over the same period are considerably smoother. Venture capital also seems unaffected by the crash, with a reported 5.2 percent return in the fourth quarter of 1987. Only six negative quarters are reported. The reported annual standard deviation of returns is 9.1 percent, and correlation with the S&P 500 is 0.28.

The analyst can attempt to correct for the biases in datasets (when a bias-free dataset is not available). For example, one heuristic approach to correcting for smoothed data is to rescale the data in such a way that their dispersion is increased but the mean of the data is unchanged. Example 8 illustrates this idea.

EXAMPLE 8

Smoothed Data: The Case of Alternative Investments (2)

How might an analyst address the biases resulting from smoothed data? To continue with the case of venture capital return data, one approach would be to rescale the reported data so that dispersion is increased but the mean is unchanged. The point is that the larger the rescaling, the larger the number of negative quarterly returns, because the frequency distribution is centered in the same place but there is more probability in the tails as dispersion is larger. For example:

- The venture returns rescaled by a factor of 1.4 provide 18 negative quarters—that is, as many as the S&P 500. The estimated standard deviation of the rescaled data is 13 percent.
- The venture returns rescaled by a factor of 4.1 provide 36 negative quarters, which is twice as many as the S&P 500. The estimated standard deviation of the rescaled data is 37 percent.
- The venture returns rescaled by a factor of 4.4 provide 38 negative quarters, 2.1 times as many as the S&P 500. The estimated standard deviation of the rescaled data is 40 percent.

Using these data in conjunction with other analyses, one might propose risks of 43 percent for early-stage venture capital, 34 percent for late-stage venture capital, 29 percent for leveraged buyouts (largely debt-financed purchases of established companies), and 20 percent for distressed debt (the debt of companies that are under financial distress or in or near bankruptcy).[5]

The key is to model the risks of alternative investments as if they were frequently traded, focusing not on statistical observations but on the underlying fundamental and economic drivers of returns.

2.2.3 *The Limitations of Historical Estimates*

With justification, analysts frequently look to history for information in developing capital market forecasts. But although history is usually a guide to what we may expect in the future, the past cannot be simply extrapolated to produce future results uncritically. A historical estimate should be considered a starting point for analysis. The analysis should include a discussion of what may be different from past average results going forward. If we do not see any such differences, we may want to project the historical estimates into the future (perhaps after making certain technical adjustments). However, making such projections without raising the question of differences is questionable.

5 See the reading on alternative investments portfolio management for a discussion of these alternative investments.

Changes in the technological, political, legal, and regulatory environments, as well as disruptions such as wars and other calamities, can alter risk–return relationships. Such shifts are known as changes in **regime** (the governing set of relationships) and give rise to the statistical problem of **nonstationarity** (meaning, informally, that different parts of a data series reflect different underlying statistical properties). For example, the shifts in US central bank policy in 1980 began a period of declining and subsequently stable inflation that is widely recognized as representing a break with the past. Also, disruptive events in a particular time period may boost volatilities in a manner that is simply not relevant for the future. However, extending a dataset to the distant past increases the chance of including irrelevant data. The well-informed analyst tracks the range of events that can indicate an important change in a time series. Statistical tools are available to help identify such changes or turning points.[6]

When many variables are considered, a long data series may be a statistical necessity. (For example, to calculate a historical covariance matrix, the number of observations must exceed the number of variables.) If we could be assured of stationarity, going back farther in time to capture a larger sample should increase the precision with which population parameters of a return distribution are estimated.[7] Related to that point, using larger samples may reduce the sensitivity of parameter estimates to the starting and ending dates of the sample. In practice, using a long data series may involve a variety of problems. For instance:

- The risk that the data cover multiple regimes increases.
- Time series of the required length may not be available.
- In order to get data series of the required length, the temptation is to use high-frequency data (weekly or even daily). Data of high frequency are more sensitive to asynchronism across variables.[8] As a result, high-frequency data tend to produce lower correlation estimates.

Researchers have concluded that the underlying mean returns on volatile asset classes such as equities are particularly difficult to estimate from historical data.[9] Using high-frequency data is of no help in increasing the accuracy of mean return estimates.

A practical approach to deciding whether one should use the whole of a long data series is to answer two questions. The first question is: Is there any fundamental reason to believe that the entirety of the series' time period is no longer relevant? If there is, the next question to answer is: Do the data support that hypothesis? Texts on time-series and regression analysis offer a variety of means to assess objectively whether there is a break in a time series. If the answers to both questions are yes, one should use only that part of the time series that appears to be relevant to the present.

EXAMPLE 9

Using Regression Analysis to Identify a Change in Regime

The effects of specific events on a time series (e.g., the announcement by a central bank of a new monetary policy) can be most simply modeled in a regression framework using a dummy explanatory variable $z(t)$, where $z(t) = 0$ for t before

6 See Hamilton (1994).

7 According to sampling theory, the precision of the estimate of the population mean is proportional to $1/\sqrt{(\text{number of observations})}$.

8 **Asynchronism** is a discrepancy in the dating of observations that occurs because stale (out-of-date) data may be used in the absence of current data.

9 See Luenberger (1998).

the intervention (change) date and $z(t) = 1$ for t at and subsequent to the intervention date. This dummy variable approach models a simple shift in the mean of the dependent variable.

2.2.4 Ex Post *Risk Can Be a Biased Measure of* Ex Ante *Risk*

In interpreting historical prices and returns over a given sample period for their relevance to current decision making, we need to evaluate whether asset prices in the period reflected the possibility of a very negative event that did not materialize during the period. Looking backward, we are likely to underestimate *ex ante* risk and overestimate *ex ante* anticipated returns.[10] For example, suppose that bond prices reflect an anticipation of a small chance of a central bank policy change that would be very negative for inflation and bond returns. When investors become aware that the risk has passed, bond prices should show strong gains. *Ex post* realized bond returns are high although *ex ante* they were lower. Because the bank policy change did not occur, it may be overlooked as a risk that was faced by bond investors at the time. An analyst reviewing the record might conclude that bonds earn high returns in excess of the short-term interest rate.[11] Similarly, a high *ex post* US equity risk premium may reflect fears of adverse events that did not materialize and may be a poor estimate of the *ex ante* risk premium.[12] Only the *ex ante* risk premium is important in decision making.

2.2.5 *Biases in Analysts' Methods*

Analysts naturally search for relationships that will help in developing better capital market expectations. Among the preventable biases that the analyst may introduce in such work are the following:

- **Data-mining bias**. Data-mining bias is introduced by repeatedly "drilling" or searching a dataset until the analyst finds some statistically significant pattern. Such patterns cannot be expected to be of predictive value. It is almost inevitable that the analyst will find some statistically significant relationship by mining the data: Using a given sample, if we examine 50 different variables as predictors of the equity risk premium and set a 10 percent significance level in our tests, we would expect 5 variables to appear significant by random chance alone. The absence of an explicit economic rationale for a variable's usefulness is one warning sign of a data-mining problem: no story, no future.[13]
- **Time-period bias**. Time-period bias relates to results that are time period specific. Research findings are often found to be sensitive to the selection of starting and/or ending dates. As one example, the small-cap stock effect in US stock returns has been found to be largely concentrated in the nine-year period 1975 to 1983, when as a group, small-cap stocks outperformed large-cap equities by 19.6 percent per year. Excluding the 1975–1983 period, a given investment in large-cap equities in 1926 would have grown by the end of 2001 to an amount that was 20 percent greater than the amount resulting from an equal initial investment in small-cap equities.[14]

10 That situation of biased measurement has been called the "generalized peso problem" or the "peso problem." The name comes from an explanation for the fact that forward markets for the Mexican peso in the mid-1970s consistently underpredicted the US dollar/peso exchange rate. The explanation is that traders feared that the Mexican government would devalue the peso from its peg.

11 This explanation has been offered by Bekaert, Hodrick, and Marshall (2001) along with time-varying term premiums for anomalies in the term structure of interest rates noted by Campbell and Shiller (1991).

12 See Goetzmann and Jorion (1999) and references therein.

13 See McQueen and Thorley (1999).

14 See Siegel (2002), pp. 134–135.

How might the analyst avoid the mistake of using a variable in a forecasting mode that historical analysis has suggested as useful but which is actually irrelevant? The analyst should scrutinize the variable selection process for data-mining bias and be able to provide an economic rationale for the variable's usefulness in a forecasting mode. A further practical check is to examine the forecasting relationship out of sample (i.e., on data other than those used to estimate the relationship). For example, the available data period could be split into two subperiods. If the forecasting relationship estimated from the first period does not hold similarly when tested using data from the second subperiod, the variable may not be useful as a forecaster.

2.2.6 *The Failure to Account for Conditioning Information*

We observed above that the analyst should ask whether there are relevant new facts in the present when forecasting the future. Where such information exists, the analyst should condition his or her expectations on it.

We can take the case of estimating mean returns. Long-run mean returns and risk involve an averaging over many different economic and market conditions. Prospective returns and risk for an asset as of today are conditional on the specific characteristics of the current marketplace and prospects looking forward. That fact explains the role of economic analysis in expectations setting: We should not ignore any relevant information or analysis in formulating expectations. Indeed, the use of unconditional expectations can lead to misperceptions of risk, return, and risk-adjusted return.

Consider an asset class that has a beta of 0.80 in economic expansions and 1.2 in recessions (with respect to a world market portfolio). If we make the assumptions in Exhibit 2 on the market return and the risk-free rate, the asset class's expected return is 10 percent in an expansion versus 4.4 percent in a recession and its true unconditional expected return is 0.5(10%) + 0.5(4.4%) = 7.2%. The asset class fairly rewards risk in both expansions and recessions [alpha = 0.5(0%) + 0.5(0%) = 0%]. The asset class would appear to have a beta of 0.5(0.8) + 0.5(1.2) = 1.0 in a regression. Given this unconditional beta, the expected return according to the CAPM would be 8 percent.

Exhibit 2 Misvaluation from Using an Unconditional Benchmark

	Expansion	Recession	Unconditional Expectation
Risk-free rate	2%	2%	2%
Exp. return on market	12%	4%	8%
β_i	0.80	1.20	1.0
$E(R_i)$	2% + 0.8 (12% − 2%) = 10%	2% + 1.2 (4% − 2%) = 4.4%	True: 7.2% Using β_i = 1.0: 2% + 1.0 (8% − 2%) = 8%
α_i	0%	0%	7.2% − 8% = −0.8%

Note: An expansion and a recession are assumed to be equally likely.

Comparing the unconditional expected return using the unconditional beta of 1.0, the asset class appears to inadequately reward risk (alpha = −0.8 percent) although we know from the analysis presented that the asset class fairly rewards risk.[15] How would an analyst avoid drawing the wrong conclusion? The analyst would need to uncover

15 Note that Ferson and Schadt (1996) developed a method to estimate conditional alpha.

through research that the asset class's systematic risk varies with the business cycle. The analyst would then condition his or her forecasts on the state of the economy to formulate the most accurate expectations.

2.2.7 *Misinterpretation of Correlations*

In financial and economic research, the analyst should take care in interpreting correlations. When a variable A is found to be significantly correlated with a variable B, there are at least three possible explanations:

- A predicts B;
- B predicts A;
- a third variable C predicts A and B.

Without the investigation and modeling of underlying linkages, relationships of correlation cannot be used in a predictive model. For example, suppose A relates to natural disasters in quarter t and B represents property insurer claims in quarter t + 1. One can discern on the basis of simple economic reasoning a cause-and-effect link from A to B. Supporting that conclusion, no plausible feedback linkage exists from B to A: A is truly an exogenous variable (an **exogenous variable** is determined outside the system, in contrast to an **endogenous variable**, which is determined within the system). One might consider using A as one predictor of B, but the reverse—using claims in $t + 1$ to predict natural disasters in $t + 2$—would not be fruitful (although the observed correlation by itself would not tell you that).

As the third bullet point represents, there may be no predictive relationship between A and B; the relationship between A and B is conditional on the presence of the variable C, and the correlation between A and B is spurious.

Another surprise that might be in store: A and B could have a strong but *nonlinear* relationship but have a low or zero correlation.[16]

Suppose that one has a plausible model of an underlying causal link to support the use of a variable as a predictor. Are there any more powerful tools to apply to establish the variable's usefulness than simple correlation? Multiple-regression analysis may be one such tool. For example, suppose we have a model that suggests B predicts A but we need to eliminate C as mediating the relationship between A and B (as in the third bullet point above). We can estimate the following regression:

$$A = \beta_0 + \beta_1 B + \beta_2 C + \varepsilon$$

The variable C in this regression is a control variable. The coefficient β_1 represents the effect of B on A after accounting for the effect of the control variable C on A. The coefficient β_1 reflects the **partial correlation** between A and B. If the estimated value of β_1 is significantly different from 0 but β_2 is not significantly different from 0 (based on t-tests), we have a piece of evidence in support of the proposition that B predicts A. The multiple-regression framework supports the introduction of multiple control variables. The analyst may also use time-series analysis. For example, with sufficiently long time series, we can regress A on lagged values of itself, lagged values of B, and lagged values of control variables, and test the null hypothesis that all the coefficients on the lagged values of B jointly equal 0. If we can reject the null hypothesis, the variable B may be useful in predicting A.[17]

---

16 For example, consider $B = A^2$ (A raised to the second power indicates a nonlinear association). The variable B increases with increasing values of A when A is above 0. But consider negative values of A. As A increases from −100 to 0, B decreases. The correlation between A and B is zero, although the relationship between them is so precise that it can be expressed in a mathematical equation.

17 This would be a test of "predictive causality," known as Granger causation. See Granger (1969) and Diebold (2006).

EXAMPLE 10

Causality Relationships

That one event follows another is not sufficient to show that the first event caused the second. For example, a decrease in the number of new accountants following an increase in tax rates would be association without cause-and-effect relationship. But seasonal incoming tax receipts probably bear a direct cause-and-effect relationship to the needs of governments to borrow funds in some months versus others. If an increase in income tax rates causes individuals to be more concerned with minimizing taxes, one might discern an indirect cause-and-effect relationship between the tax rate increase and a subsequent jump in sales of tax preparation software.

2.2.8 *Psychological Traps*

Hammond, Keeney, and Raiffa (1998) formulated several psychological traps that are relevant to our discussion because they can undermine the analyst's ability to make accurate and unbiased forecasts.

The **anchoring trap** is the tendency of the mind to give disproportionate weight to the first information it receives on a topic. In other words, initial impressions, estimates, or data anchor subsequent thoughts and judgments. For instance, in an investment committee in which several different perspectives on capital market returns are presented, the first presentation may tend to function as an anchor for discussion and its lead-off position might give it an edge in being adopted. The analyst can try to address this trap by consciously attempting to avoid premature conclusions.

The **status quo trap** is the tendency for forecasts to perpetuate recent observations—that is, to predict no change from the recent past. If inflation has been rising at a double-digit rate for several recent periods, it is a natural tendency to forecast a similar increase in the next period. In a decision-making context, because doing something other than maintaining the status quo (risking an error of commission) may lead to increased work and regret if the decision is wrong, doing nothing (risking an error of omission) becomes the easy and oft-preferred alternative. The status quo trap may be overcome with rational analysis used within a decision-making process.

The **confirming evidence trap** is the bias that leads individuals to give greater weight to information that supports an existing or preferred point of view than to evidence that contradicts it. The tendency to seek out information that supports an existing point of view also reflects this bias. Several steps may be taken to help ensure objectivity:

- Examine all evidence with equal rigor.
- Enlist an independent-minded person to argue against your preferred conclusion or decision.
- Be honest about your motives.

The **overconfidence trap** is the tendency of individuals to overestimate the accuracy of their forecasts. Many people do not admit or attempt to measure the possibility of failure in predicting uncertain events. In similar fashion, we tend to believe that most people share our particular views. The overconfidence trap would be reflected in admitting too narrow a range of possibilities or scenarios in forecasting. A good practice to prevent this trap from undermining the forecasting endeavor is to widen the range of possibilities around the primary target forecast.

The **prudence trap** is the tendency to temper forecasts so that they do not appear extreme, or the tendency to be overly cautious in forecasting. In a decision-making context, it is the tendency to be cautious when making decisions that could be potentially expensive or damaging to the decision maker's career. To avoid the prudence trap, an analyst is again wise to widen the range of possibilities around the target forecast. In addition, the most sensitive estimates affecting a forecast should be carefully reviewed in light of the supporting analysis.

The **recallability trap** is the tendency of forecasts to be overly influenced by events that have left a strong impression on a person's memory. Often, forecasts are overly influenced by the memory of catastrophic or dramatic past events. For example, investors' memory of the stock market crash of 1929 has sometimes been cited as a depressing influence on equity valuation levels for as long as three decades following the crash. To minimize the distortions of the recallability trap, analysts should ground their conclusions on objective data and procedures rather than on personal emotions and memories.

EXAMPLE 11

Traps in Forecasting

Cynthia Casey is a Canada-based investment adviser with a clientele of ultra-high-net-worth individuals. The Canadian equity allocation of client Philip Lasky's portfolio had favorable risk-adjusted performance from 2006 to 2008 but nevertheless lost 20 percent of its year-end 2007 value by the end of 2008. In a phone call prior to a quarterly portfolio review at the end of 2008, Lasky expressed the thought that the pain of the recent and continuing bear market had made him very cautious about investing in the stock market. Although his equity allocation results with Casey showed healthy appreciation over the entire period he had invested with her, his conversation dwelled mostly on the experience of the past year. Lasky told Casey that he had read a variety of financial reports containing predictions by investment notables on the equity risk premium ranging from near zero to 6 percent. During the call, he repeated to Casey, sometimes inaccurately, the arguments of the most bearish prognosticator. At the time of the call, Casey was preparing to share with clients relatively optimistic forecasts for Canadian equities, developed with an assistant who was well grounded in capital market analysis. Perceiving that Lasky and many of her other clients held more pessimistic viewpoints and that she might lose their trust if her own viewpoint turned out to be wrong, after the phone call, Casey decided to revise downward some of the economic growth assumptions she had previously made.

Critique the forecasts of A) Lasky and B) Casey with respect to psychological traps in forecasting.

Solution to A:

In focusing on the most recent period only and predicting a continuation of the most recent trend, Lasky may have fallen into the status quo trap. The pain of the bear market may have overly influenced his thinking about the present (recallability trap). Furthermore, in sharing the viewpoint of the most bearish prognosticator, Lasky may be falling into the confirming evidence trap.

Solution to B:

By trimming her assumptions to be more conservative without real supporting analysis, Casey may have fallen into the prudence trap.

2.2.9 *Model Uncertainty*

The analyst usually encounters at least two kinds of uncertainty in conducting an analysis: **model uncertainty** (uncertainty concerning whether a selected model is correct) and **input uncertainty** (uncertainty concerning whether the inputs are correct). For example, suppose an analyst takes the equity risk premium of UK equities to be the realized value of the return of UK equities over UK bonds over the past 50 years. The analyst's model might be described as follows: "The *ex ante* UK equity risk premium was, is, and will be equal to some constant number μ." If the model is far off the true state of affairs, the analyst's forecast will also be off. (The sampling error in the estimate of μ using 50 years of data would constitute the input error in this approach.) To take another example, if the analyst uses a monetarist model for forecasting future inflation, the analyst faces uncertainty concerning whether that model is correct. In some cases, the analyst may gauge model uncertainty by observing the variation in results that comes from shifting between the several most promising models.

Input uncertainty and model uncertainty in particular often make it hard to confirm the existence of capital market anomalies (inefficiencies); some valuation model usually underlies the identification of an inefficiency. **Behavioral finance** (the theory that psychological variables affect and often distort individuals' investment decision making) has offered explanations for many perceived capital market anomalies. Kurz, Jin, and Motolese (2005) argue that many of these apparent anomalies could represent equilibria resulting from the actions of investors who use competing models but process and act on information rationally.[18]

3 TOOLS FOR FORMULATING CAPITAL MARKET EXPECTATIONS

The following sections introduce a range of tools that have been used in professional forecasting of capital market returns. Although an analyst may have distinct preferences among these approaches, familiarity with all these major tools will be helpful in addressing the widest variety of forecasting problems according to their particular characteristics.

3.1 Formal Tools

Formal tools are established research methods amenable to precise definition and independent replication of results. The information provided by well-chosen formal tools applied to sound data can help the analyst produce accurate forecasts.

3.1.1 *Statistical Methods*

Statistical methods relevant to expectations setting include **descriptive statistics** (methods for effectively summarizing data to describe important aspects of a dataset) and **inferential statistics** (methods for making estimates or forecasts about a larger group from a smaller group actually observed).

The simplest approach to forecasting is to use past data to directly forecast future outcomes of a variable of interest.

18 That is, apparent anomalies could represent a *rational belief equilibrium* in the sense of Kurz (1994).

3.1.1.1 Historical Statistical Approach: Sample Estimators[19] Suppose an investor uses the FTSE 100 as his benchmark for UK large-cap equity allocations. The investor could use the mean return on the FTSE 100 over some selected sample period as his forecast of the long-run expected return on UK large-cap equities. If future returns over the selected time horizon reflect the same probability distribution as past returns (because the time series is **stationary**—that is, the parameters that describe the return-generating process are unchanged), the resulting estimate will be useful. For example, in a mean–variance framework, the analyst might use:

- the sample arithmetic mean total return or sample geometric mean total return as an estimate of the expected return;
- the sample variance as an estimate of the variance; and
- sample correlations as estimates of correlations.

One decision point relates to the choice between an arithmetic mean and a geometric mean. The arithmetic mean return (which is always used in the calculation of the sample standard deviation) best represents the mean return in a single period. The geometric mean return of a sample represents the compound rate of growth that equates the beginning value to the ending value of a data series. The geometric mean return represents multiperiod growth more accurately than the arithmetic mean return. The geometric mean return is always lower than the arithmetic mean return for a risky variable. The differences between the arithmetic mean and the geometric mean in historical estimates of the equity risk premium can be substantial.[20] Both approaches are used in current practice.

Dimson, Marsh, and Staunton (2011) presented authoritative evidence on asset returns in 17 countries for the 111 years 1900–2010. Exhibit 3 excerpts their findings.

An analyst using a historical statistical approach would use historical data such as those given in Exhibit 3 as the basis for forecasts. Alternatively, using a historical statistical approach for the equity risk premium and a current term-structure estimate for the expected return on bonds (e.g., a yield to maturity on a zero-coupon government bond), the expected return on equities could be estimated as their sum.

Exhibit 3 Real (Inflation-Adjusted) Equity and Bond Returns: Seventeen Major Markets, 1900–2010

	Equities		Bonds	
Country	Arithmetic Mean Equity Return	Standard Deviation of Return	Arithmetic Mean Bond Return	Standard Deviation
Australia	9.1%	18.2%	2.3%	13.2%
Belgium	5.1	23.61	1.1	12.0
Canada	7.3	17.2	2.6	10.4
Denmark	6.9	20.9	3.7	11.7
France	5.7	23.5	0.8	13.0

19 A **sample estimator** is a formula for assigning a unique value (a **point estimate**) to a population parameter.

20 Looking forward to later discussion, the arithmetic historical equity risk premium would be calculated as the difference between the arithmetic mean return on the proxy for equities and the arithmetic mean return on long-term bonds. The geometric mean can be approximated as the difference between the geometric mean return on the proxy on equities ($R_{G,e}$) and the geometric mean return on long-term bonds ($R_{G,b}$), or more precisely, as $(1 + R_{G,e})/(1 + R_{G,b})$. In practice, the geometric mean calculation produces a lower estimate for the equity risk premium than does the arithmetic mean calculation.

Exhibit 3 (Continued)

Country	Equities: Arithmetic Mean Equity Return	Equities: Standard Deviation of Return	Bonds: Arithmetic Mean Bond Return	Bonds: Standard Deviation
Germany	8.1	32.2	0.8	15.7
Ireland	6.4	23.21	1.9	14.9
Italy	6.1	29.0	−0.4	14.1
Japan	8.5	29.8	1.6	20.1
Netherlands	7.1	21.8	1.8	9.4
Norway	7.2	27.4	2.4	12.2
South Africa	9.5	22.6	2.3	10.4
Spain	5.8	22.3	1.9	12.4
Sweden	8.7	22.9	3.2	11.8
Switzerland	6.1	19.8	2.5	9.3
United Kingdom	7.2	20.0	2.2	13.7
United States	8.3	20.3	2.3	10.2

Source: Dimson, Marsh, and Staunton (2011), Tables 2 and 5. German data exclude 1922–1923.

3.1.1.2 Shrinkage Estimators **Shrinkage estimation** involves taking a weighted average of a historical estimate of a parameter and some other parameter estimate, where the weights reflect the analyst's relative belief in the estimates. This "two-estimates-are-better-than-one" approach has desirable statistical properties that have given it a place in professional investment practice. The term "shrinkage" refers to the approach's ability to reduce the impact of extreme values in historical estimates. The procedure has been applied to covariances and mean returns.

A **shrinkage estimator** of the covariance matrix is a weighted average of the historical covariance matrix and another, alternative estimator of the covariance matrix, where the analyst places the larger weight on the covariance matrix he or she believes more strongly in.[21] Why are analysts often not satisfied with using the historical sample covariance matrix? Basically, because investment data series are relatively short and samples often reflect the nonrecurring peculiarities of a historical period. The sample covariance matrix is perfectly well suited for *summarizing* an observed dataset and has the desirable (large-sample) property of unbiasedness. Nevertheless, a shrinkage estimator is a superior approach for estimating the population covariance matrix for the medium- and smaller-size datasets that are typical in finance.

A shrinkage estimator approach involves selecting an alternative estimator of the covariance matrix, called a **target covariance matrix**. For example, an analyst might believe that a particular model relating asset class returns to a particular set of return drivers or systematic risk factors has some validity. The asset classes' estimated betas or factor sensitivities in such a model can be used to estimate the asset classes' covariances. To consider one number in the covariance matrix, suppose that the estimated covariance between domestic shares and bonds is 48 using the factor model and 80 using a historical estimate, and assume further that the optimal weights on the model and historical estimates are 0.75 and 0.25, respectively. The shrinkage

21 This method is usually presented in terms of covariances rather than correlations for technical reasons. Either covariance or correlation can be used in MVO. Stein (1956) introduced shrinkage estimates.

estimate of the covariance would be 0.75(48) + 0.25(80) = 56. There is a systematic way to determine the optimal weights on the two estimates that the analyst can obtain from the investment literature on this topic.[22]

A surprising fact concerning the shrinkage estimator approach is that any choice for the target covariance matrix will lead to an increase (or at least not a decrease) in the efficiency of the covariance estimates versus the historical estimate. The improvement will be greater if a plausible target covariance matrix is selected. If the target covariance matrix is useless in improving the accuracy of the estimate of covariance, the optimal weight on the historical estimate would be calculated as 1. One reasonable choice for the target covariance matrix would be a factor-model-based estimate of the covariance matrix, following the lead of Ledoit and Wolf (2003). Another choice for the target covariance matrix would be a covariance matrix based on assuming each pairwise covariance is equal to the overall average covariance.[23]

EXAMPLE 12

Adjusting a Historical Covariance

Cynthia Casey has estimated the covariance between Canadian equities and US equities as 230 using historical data. Using a factor model approach based on a proxy for the world market portfolio, she estimates the covariance as 190. Casey takes a shrinkage estimator approach to estimating covariances and determines that the optimal weight on the historical estimate is 0.30.

1 Calculate the shrinkage estimate of the covariance between US and Canadian equities.
2 Describe the theoretical advantage of a shrinkage estimate of covariance compared to a raw historical estimate.

Solution to 1:

0.30(230) + 0.70(190) = 202.

Solution to 2:

The shrinkage estimate should be more accurate, given that the weights are chosen appropriately.

A shrinkage estimator of mean returns involves taking a weighted average of each historical mean return and some other target constant—such as the overall (grand) mean historical return across assets. For example, five assets have historical mean returns of 4 percent, 6 percent, 7 percent, 8 percent, and 10 percent, respectively. The overall mean return is 7 percent. Given a weight of 80 percent on an asset's historical mean, we would calculate the shrinkage estimate of the first asset's return as 0.8(4%) + 0.2(7%) = 4.6%.

3.1.1.3 Time-Series Estimators **Time-series estimators** involve forecasting a variable on the basis of lagged values of the variable being forecast and often lagged values of other selected variables.

22 Ledoit and Wolf (2003) give a simple formula for the optimal weights. The criterion their formula satisfies is that the weights minimize the mean square error in the resulting estimate.

23 An identity matrix or a scalar multiple of it is considered a serviceable choice when the researcher has no insight into an intuitive target for the covariance matrix.

Time-series methods have been found useful in developing particularly short-term forecasts for financial and economic variables. Time-series methods have been notably applied to estimating near-term volatility, given persuasive evidence of variance clustering (particularly at high frequencies, such as daily and weekly) in a number of different markets, including equity, currency, and futures markets.[24] **Volatility clustering** is the tendency for large (small) swings in prices to be followed by large (small) swings of random direction. Volatility clustering captures the idea that some markets represent periods of notably high or low volatility. Robert F. Engle shared the 2003 Nobel Prize in Economics in part for the development of time-series models that can accurately capture the property of volatility clustering.[25]

One of the simplest specifications in this broad class of models was developed within a division at JPMorgan that was later established as the RiskMetrics Group. This model specifies that the volatility in period t, σ_t^2, is a weighted average of the volatility in the previous period, σ_{t-1}^2, and the squared value of a random "noise" term, ε_t^2. The expression is

$$\sigma_t^2 = \beta\sigma_{t-1}^2 + (1-\beta)\varepsilon_t^2 \tag{1}$$

with $0 < \beta < 1$. The coefficient β measures the rate of decay of the influence of the value of volatility in one period on future volatility, and the rate of decay is exponential. The higher β is, the more volatility in one period "remembers" what happened in the past and the more it clusters.

To illustrate using $\beta = 0.94$, we will suppose that the standard deviation of returns in $t = 11$ is 10 percent, so $\sigma_{11}^2 = 0.10^2 = 0.01$. The noise term is $\varepsilon_{12} = 0.05$, so $\varepsilon_{12}^2 = 0.05^2 = 0.0025$. The prediction for $t = 12$ is therefore

$$\begin{aligned}\sigma_{12}^2 &= 0.94\sigma_{11}^2 + 0.06\varepsilon_{12}^2\\ &= 0.94(0.01) + 0.06(0.0025)\\ &= 0.00955\end{aligned}$$

implying that $\sigma_{12} = \sqrt{0.00955} = 0.0977$, or 9.77 percent. Intuitively, the high weight on σ_{11}^2 means that it had a strong effect on σ_{12}^2. However, on occasion, the noise term will take on an extreme value and cause volatility to shift quite a bit. In a similar vein to this approach to variance estimation, the correlation matrix has also been estimated with some success using exponentially weighted historical observations.[26]

3.1.1.4 Multifactor Models A **multifactor model** is a model that explains the returns to an asset in terms of the values of a set of return drivers or risk factors.

The structure of a multifactor model, if the analyst believes that K factors drive asset returns, is as follows:

$$R_i = a_i + b_{i1}F_1 + b_{i2}F_2 + \ldots + b_{iK}F_K + \varepsilon_i \tag{2}$$

24 See Drost and Nijman (1993) and references therein.
25 Such models are called autoregressive conditional heteroskedasticity (ARCH) time-series models. The Nobel Prize was shared with Clive W. J. Granger, who developed methods for analyzing cointegrated time series (informally, time series with common trends).
26 For more details on models of volatility clustering, see Bollerslev, Engle, and Nelson (1994).

where

R_i = the return to asset i

a_i = an intercept term in the equation for asset i

F_k = the return to factor k, $k = 1, 2, \ldots, K$

b_{ik} = the sensitivity of the return to asset i to the return to factor k, $k = 1, 2, \ldots, K$

ε_i = an error term with a zero mean that represents the portion of the return to asset i not explained by the factor model. The error term is assumed to be uncorrelated with each of the K factors and to be uncorrelated with the error terms in the equations for other assets.

This structure has been found useful for modeling asset returns and covariances among asset returns. Multifactor models are useful for estimating covariances for the following reasons:

- By relating the returns on all assets to a common set of return drivers, a multifactor model simplifies the task of estimating covariances: estimates of covariances between asset returns can be derived using the assets' factor sensitivities.
- When the factors are well chosen, a multifactor model approach may filter out noise (i.e., random variation in the data specific to the sample period).
- Such models make it relatively easy to verify the consistency of the covariance matrix, because if the smaller factor covariance matrix is consistent, so are any covariances computed on the basis of it.

In the balance of this section, we illustrate a top-down structured approach to using factor models in estimating the covariance matrix. In this approach, we model factors as portfolios of securities and start from a simple two-factor model at the most aggregated level.

Assume that two factors, a global equity factor and a global bonds factor, drive the returns of all assets in the investable universe. In this case, we start the modeling process with a covariance matrix for global equity and global bonds (we will refer to it in this discussion as the *equity–bonds covariance matrix*). A standard deviation of 14 percent for global equity and 4 percent for global bonds and a correlation between them of 0.30 imply the covariance matrix shown in Exhibit 4. In Exhibit 4, 0.0196 is the variance for global equity, 0.0016 is the variance for global bonds, and 0.0017 is the covariance between global equity and global bonds. The covariance between global equity and global bonds is the product of their standard deviations times the correlation between them, or (0.14)(0.04)(0.30) = 0.0017 to four decimal places; for global equity variance, 0.0196 = $(0.14)^2$; for global bonds variance, 0.0016 = $(0.04)^2$.

Exhibit 4 Factor Covariance Matrix

	Global Equity	Global Bonds
Global Equity	0.0196	0.0017
Global Bonds	0.0017	0.0016

This is a **factor covariance matrix**, as it contains the covariances for the factors assumed to drive returns. In order to derive the **asset covariance matrix** (the covariance matrix for the asset classes or markets under consideration), we need to know how each of the markets responds to factor movements. We measure the responsiveness of markets to factor movements by the markets' **factor sensitivities** (also known as factor betas or factor loadings), represented by the quantities b_{ik} in Equation 2. If Market 1 moves by 110 basis points in response to a 100 basis point move of global equities, the corresponding factor sensitivity is 1.10. In addition, every market has some risk that is not explained by the factors. This is called the market's idiosyncratic or residual risk and is represented by the residual variance, $\text{Var}(\varepsilon_i)$ for market i. It is assumed that the residuals are uncorrelated.

Exhibit 5 shows hypothetical statistics for five securities markets.

Exhibit 5 Hypothetical Statistics for Five Markets

	Sensitivities		
	Global Equity	**Global Bonds**	**Residual Risk (%)**
Market A	1.10	0	10.0
Market B	1.05	0	8.0
Market C	0.90	0	7.0
Market D	0	1.03	1.2
Market E	0	0.99	0.9

Source: Staub (2006).

Judged by its factor sensitivities, Market A is an equity market, with zero sensitivity to global bonds and a positive sensitivity to global equity. The zero sensitivity to global bonds does not mean that Market A is uncorrelated with global bonds, although it does mean that its partial correlation with bonds (the correlation after removing the influence of the other markets) is zero and that global bonds are not one of Market A's return drivers.[27]

In the case we are examining, we are assuming that the return of market i, M_i, is as follows:

$$M_i = a_i + b_{i1}F_1 + b_{i2}F_2 + \varepsilon_i,\ i = 1 \text{ to } 5$$

We compute the markets' variances and covariance using Equations 3a and 3b, respectively:

$$M_{ii} = b_{i1}^2\text{Var}(F_1) + b_{i2}^2\text{Var}(F_2) + 2b_{i1}b_{i2}\text{Cov}(F_1,F_2) + \text{Var}(\varepsilon_\text{i}),\quad \text{for } i = 1 \text{ to } 5 \tag{3a}$$

where M_{ii} is the variance of market i;

$$M_{ij} = b_{i1}b_{j1}\text{Var}(F_1) + b_{i2}b_{j2}\text{Var}(F_2) + (b_{i1}b_{j2} + b_{i2}b_{j1})\text{Cov}(F_1,F_2)\quad \text{for } i = 1 \text{ to } 5, j = 1 \text{ to } 5, \text{ and } i \neq j \tag{3b}$$

where M_{ij} is the covariance of market i with market j.

For example, suppose we want to compute the covariance between Markets A and B. Using Equation 3b (with i = 1 for Market A and j = 2 for Market B), we calculate that

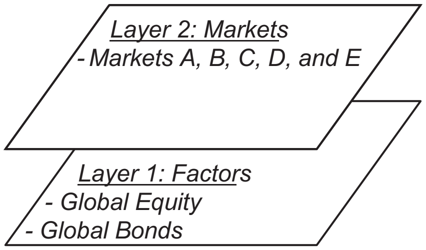
$$M_{12} = (1.1)(1.05)(0.0196) + (0)(0)(0.0016) + [(1.10)(0) + (0)(1.05)](0.0017) = 0.0226$$

27 Through the positive covariance (and correlation) between global equity and global bonds, Market A is still positively correlated with global bonds, although moderately.

Because both Market A and Market B have zero sensitivity to the global bonds factor, their correlation is explained only through their sensitivities to the global equity factor. Equations 3a and 3b are basic formulas for using multifactor models to estimate asset class covariance.

Note that establishing the consistency of the equity–bonds covariance matrix would not be a challenge, because it has only four entries. If the equity–bonds covariance matrix is consistent, then the covariance matrix for the markets calculated using Equations 3a and 3b will be consistent, even if many markets are involved so that consistency might be hard to check directly. The ability to establish consistency efficiently is a significant advantage of a multifactor model approach.

Exhibit 6 A Two-Layer Factor Approach



Source: Staub (2006). © 2006 UBS Global Asset Management (Americas) Inc. All rights reserved.

The above example, illustrated in Exhibit 6, is a two-layer structure with the factors on the first layer and the markets to be modeled on the second and final layer.[28]

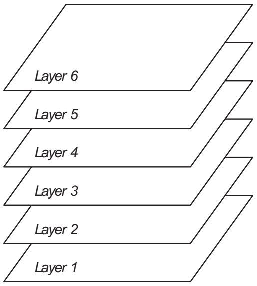
In practice, a two-layer approach is not sufficient to accurately model interrelationships with the level of detail needed. Consider expanding the set of markets from securities markets to real estate markets, including US real estate sectors—apartment, industry, office, and retail. These are mutually fairly highly correlated but have moderate or weak correlations with most other markets.

In the case of these real estate sectors, we would require new factor layers to model co-movements that are unrelated to the movement of factors in the prior layer. To meet our needs, the two-layer approach must be replaced by a multilayer approach, as illustrated in Exhibit 7.

In a layer immediately below the US real estate sectors, we would introduce US real estate as a whole as a factor. The total number of layers would depend on the final set of markets whose covariance structure we wanted to model.[29]

28 The two-layer concept goes back to Grinold and Kahn (1995), who employ it for stock modeling. For more details, see p. 58f.
29 See Staub (2006) for more details.

Exhibit 7 A Multilayer Factor Approach



Source: Staub (2006). © 2006 UBS Global Asset Management (Americas) Inc. All rights reserved.

3.1.2 *Discounted Cash Flow Models*

Discounted cash flow models (DCF models) express the idea that an asset's value is the present value of its (expected) cash flows. Formally, the value of an asset using a DCF approach is as follows:[30]

$$V_0 = \sum_{t=1}^{\infty} \frac{CF_t}{(1+r)^t} \tag{4}$$

where

V_0 = the value of the asset at time t = 0 (today)
CF_t = the cash flow (or the expected cash flow, for risky cash flows) at time t
r = the discount rate or required rate of return

For simplicity, we represent the discount rate in Equation 4 as the same for all time periods—a flat term structure of discount rates.

Analysts use DCF models in expectations setting both for traditional securities markets and for alternative investment markets where the investment (e.g., private equity or real estate) generates cash flows.

DCF models are a basic tool for establishing the intrinsic value of an asset based on fundamentals (e.g., its projected cash flows) and its fair required rate of return. DCF models have the advantage of being forward-looking. They do not address short-run factors such as current supply-and-demand conditions, so practitioners view them as more appropriate for setting long-term rather than short-term expectations. That said, asset prices that are disconnected from fundamentals may reflect conditions of speculative excess that can reverse abruptly.[31]

3.1.2.1 Equity Markets Analysts have frequently used the **Gordon (constant) growth model** form of the dividend discount model, solved for the required rate of return, to formulate the long-term expected return of equity markets. The Gordon growth model

30 If the asset trades in an integrated market, the future cash flows of the asset are to be translated into the home currencies of investors.
31 See Calverley (2004) for a discussion of **bubbles** (episodes in which asset market prices move to extremely high levels in relation to estimated intrinsic value).

assumes that there is a long-term trend in dividends and corporate earnings, which is a reasonable approximation for many developed country economies.[32] The expression for the Gordon growth model solved for $E(R_e)$, the expected rate of return on equity, is

$$E(R_e) = \frac{D_0(1+g)}{P_0} + g = \frac{D_1}{P_0} + g \tag{5}$$

where

D_0 = the most recent annual dividend per share
g = the long-term growth rate in dividends, assumed equal to the long-term earnings growth rate
P_0 = the current share price

According to the Gordon growth model, share price should appreciate at a rate equal to the dividend growth rate. Therefore, in Equation 5, the expected rate of return is composed of two parts: the dividend yield (D_1/P_0) and the capital gains (or appreciation) yield (g).

The quantity g can be estimated most simply as the growth rate in **nominal gross domestic product** (nominal GDP), a money measure of the goods and services produced within a country's borders.[33] Nominal GDP growth can be estimated as the sum of the estimated real growth rate in GDP plus the expected long-run inflation rate. A more advanced analysis can take account of any perceived differences between the expected growth of the overall economy and that of the constituent companies of the particular equity index that the analyst has chosen to represent equities. The analyst can use

Earnings growth rate = GDP growth rate + Excess corporate growth (for the index companies)

where the term *excess corporate growth* may be positive or negative depending on whether the sectoral composition of the index companies is viewed as higher or lower growth than that of the overall economy.[34] If the analyst has chosen a broad-based equity index, the excess corporate growth adjustment, if any, should be small. Exhibit 8 presents the real GDP growth rates for selected countries.

Exhibit 8 Average Annual Real GDP Growth Rates: 1980–2009 (in Percent)

	Time Period		
Country	**1980–89**	**1990–99**	**2000–09**
Australia	3.4%	3.3%	3.2%
Canada	3.0	2.4	2.1
Denmark	2.2	2.3	0.8
France	2.1	1.8	1.3
Germany	1.9	1.3	0.8
Italy	2.4	1.5	0.5
Japan	3.9	1.7	0.6
Netherlands	2.0	3.0	1.6
Sweden	2.4	1.8	1.9

32 See Jagannathan, McGrattan, and Scherbina (2000).
33 See and Singer and Terhaar (1997) for a theoretical analysis of this relationship.
34 See Grinold and Kroner (2002), p. 12.

Exhibit 8 (Continued)

	Time Period		
Country	**1980–89**	**1990–99**	**2000–09**
Switzerland	1.8	1.1	1.7
United Kingdom	2.4	2.1	1.7
United States	3.1	3.1	1.8

Source: OECD, Datastream, Bloomberg, World Bank.

In the United States and other major markets, share repurchases have become an important means for companies to distribute cash to shareholders. Grinold and Kroner (2002) provided a restatement of the Gordon growth model that takes explicit account of repurchases. Their model also provides a means for analysts to incorporate expectations of valuation levels through the familiar P/E ratio. The **Grinold–Kroner model**, which is based on elaborating the expression for the expected single-period return on a share, is[35]

$$E\left(R_e\right) \approx \frac{D}{P} - \Delta S + i + g + \Delta PE \qquad (6)$$

where

$E(R_e)$ = the expected rate of return on equity
D/P = the expected dividend yield
ΔS = the expected percent change in number of shares outstanding
i = the expected inflation rate
g = the expected real total earnings growth rate (not identical to the EPS growth rate in general, with changes in shares outstanding)
ΔPE = the per period percent change in the P/E multiple

The term ΔS is negative in the case of net positive share repurchases, so $-\Delta S$ is a positive **repurchase yield** in such cases.

Equation 6 consists of three components: an expected income return, an expected nominal earnings growth return, and an expected repricing return (from expected P/E ratio expansion or contraction).

- Expected income return: $D/P - \Delta S$.
- Expected nominal earnings growth return: $i + g$.
- Expected repricing return: ΔPE.

The expected nominal earnings growth return and the expected repricing return constitute the expected capital gains.

The Grinold–Kroner model can be used not only in expectations setting, but also as a tool to analyze the sources of historical returns. For example, the S&P 500 achieved a compound growth rate of 10.7 percent per year over the 76-year period 1926–2001 (corresponding to an equity premium of 5.3 percent).[36] Following the Grinold–Kroner analysis, the sources of this return were as follows:

- 4.4 percent from income;

35 See Grinold and Kroner (2002) for a derivation. Ibbotson and Chen (2003) presented a broadly similar analysis of the sources of equity returns but did not model stock repurchases.

36 The equity risk premium was defined by the authors as the mean return on the S&P 500 less the mean 10-year US Treasury bond return over this period.

- 4.8 percent from nominal earnings growth (consisting of 1.7 percent real earnings growth and 3.1 percent annual inflation); and
- 1.5 percent from repricing.

As a check, 4.4% + 4.8% + 1.5% = 10.7%. Repricing return was a volatile contributor to total return. The growth of the P/E ratio from 10.2 in 1926 to 30.6 in 2001 represented a compound annual growth rate of 1.5 percent. However, the P/E of 10.2 in 1926 was actually somewhat above the P/E in 1981: Most of the repricing return was concentrated in the 20 years leading up to the ending date of 2001.

EXAMPLE 13

The Grinold–Kroner Forecast of the US Equity Risk Premium

The details of the Grinold–Kroner (GK) forecast of the US equity risk premium (as of early 2002) are instructive. Their forecast horizon was 10 years.

Expected Income Return

The forecast dividend yield was 1.75 percent (somewhat above the then-current yield of 1.4 percent but below the historical mean of over 4 percent for 1926–2001). The repurchase yield was forecast to be 0.5 percent, down from the 1–2 percent rate of the 1990s, which was viewed as an unusual period. The expected income return was therefore 1.75% + 0.5% = 2.25%.

Expected Nominal Earnings Growth Return

Economic theory suggests that the real GDP growth rate is the sum of labor productivity growth and labor supply growth. GK took the historical 2 percent per year US labor productivity growth rate as their forecast. Using a US population growth forecast of 0.8 percent and assuming a 0.2 percentage point increase in the labor force participation rate, the forecast of the labor supply growth rate was 1 percent per year. The overall real GDP growth estimate of 2% + 1% = 3% was within the 2.7 percent to 3.6 percent range of forecasts by economists. Viewing the S&P 500 companies as having a slightly higher growth profile than the overall economy, GK added a 0.5 percent excess corporate growth return for a 3.5 percent real earnings growth return estimate. GK expected an inflation rate of 2.5 percent, 0.3 percentage points above the contemporaneous consensus estimate of economists (viewed by GK as slightly optimistic). Thus, the expected nominal earnings growth return was 3.5% + 2.5% = 6%.

Expected Repricing Return

This component was perhaps the hardest to forecast. Viewing the contemporaneous P/E of 28 as a slight overreaction to the positive factors of decreased inflation, technological advances (positive productivity shocks), and an expected increase in growth rates from globalization, over a 10-year horizon, GK forecast downward repricing equal to −0.75 percent per year.

The GK forecast of the expected return on US equities was therefore 2.25% + 6% − 0.75% = 7.5%. Subtracting the 10-year government bond yield of 5 percent, the GK forecast of the US equity risk premium was 2.5 percent.

The 2.5 percent estimate put GK in a middle position between the predictions of the "risk premium is dead" and the "rational exuberance" camps.

EXAMPLE 14

Forecasting the Return on Equities Using the Grinold–Kroner Model

Cynthia Casey employs the Grinold–Kroner model in forecasting long-term developed market equity returns. Casey makes the following forecasts:

- a 2.25 percent dividend yield on Canadian equities, based on the S&P/TSE Composite Index;
- a repurchase yield of 1 percent for Canadian equities;
- a long-term inflation rate of 2 percent per year;
- long-term corporate real earnings growth at 4 percent per year, based on a 1 percentage point premium for corporate growth over her expected Canadian GDP growth rate of 3.0 percent; and
- an expansion rate for P/E multiples of 0.25 percent per year.

Based only on the information given, determine the expected rate of return on Canadian equities consistent with Casey's assumptions.

Solution:

Using Casey's assumptions and Equation 6, the expected rate of return on Canadian equities should be 9.5 percent, calculated as

$$E(R_e) \approx 2.25\% - (-1.0\%) + 2.0\% + 4.0\% + 0.25\% = 9.5\%$$

DCF model thinking has provided various methods for evaluating stock market levels. The best known of these is the Fed model, which asserts that the stock market is overvalued if the market's forward earnings yield (earnings divided by price) is less than the 10-year Treasury bond yield.[37] The earnings yield is the required rate of return for no-growth equities and is thus a conservative estimate of the expected return for equities. The intuition of the Fed model is that when the yield of T-bonds is greater than the earnings yield of stocks (a riskier investment than T-bonds), stocks are overvalued.

3.1.2.2 Fixed-Income Markets The DCF model is a standard tool in the pricing of fixed-income instruments. In many such markets, bonds are quoted in terms of the single discount rate (the yield to maturity, or YTM) that equates the present value of the instrument's promised cash flows to its market price. The yield to maturity of a bellwether (reference) instrument for a bond market segment is a readily available first approximation of the market expected return for the asset segment at a time horizon equal to the maturity of the instrument.[38] The YTM calculation makes the strong assumption that as interest payments are received, they can be reinvested at an interest rate that always equals the YTM. Therefore, the YTM of a bond with intermediate cash flows is an estimate of the expected rate of return on the bond that is more or less plausible depending on the level of the YTM. If a representative zero-coupon bond is available at the chosen time horizon, its YTM would be a superior estimate.

37 This model was developed by the US Federal Reserve System (the Fed), the central bank of the United States.

38 If the bond is callable, a downward adjustment would generally need to be made. **Yield to worst** (the yield assuming the bond is called at the earliest opportunity) is sometimes used as a conservative estimate in such cases.

3.1.3 *The Risk Premium Approach*

The **risk premium approach** expresses the expected return on a risky asset as the sum of the risk-free rate of interest and one or more risk premiums that compensate investors for the risky asset's exposure to sources of **priced risk** (risk for which investors demand compensation). Investors would avoid purchasing assets offering inadequate expected compensation for priced risk; the lower demand should lead to lower asset prices until the point is reached at which the compensation for risk is adequate. The risk premium approach (sometimes called the **build-up approach**) is most often applied to estimating the required return in equity and bond markets.[39] In the following discussion, we assume that assets are fairly priced so that an asset's required return is also an investor's expected return.[40]

3.1.3.1 A General Expression Following our verbal definition of the risk premium approach, a formal expression for the expected return on a risky asset is

$$\begin{aligned} E(R_i) = R_F &+ (\text{Risk premium})_1 + (\text{Risk premium})_2 \\ &+ \ldots + (\text{Risk premium})_K \end{aligned} \tag{7}$$

where $E(R_i)$ is the asset's expected return and R_F denotes the risk-free rate of interest.

3.1.3.2 Fixed-Income Premiums The expected bond return, $E(R_b)$, can be built up as the real rate of interest plus a set of premiums:

$$\begin{aligned} E(R_b) = {} & \text{Real risk-free interest rate} + \text{Inflation premium} + \text{Default risk premium} \\ & + \text{Illiquidity premium} + \text{Maturity premium} + \text{Tax premium} \end{aligned}$$

- The **real risk-free interest rate** is the single-period interest rate for a completely risk-free security if no inflation were expected. In economic theory, the real risk-free rate reflects the time preferences of individuals for current versus future real consumption.
- The **inflation premium** compensates investors for expected inflation and reflects the average inflation rate expected over the maturity of the debt plus a premium (or discount) for the probability attached to higher inflation than expected (or greater disinflation). The sum of the real risk-free interest rate and the inflation premium is the **nominal risk-free interest rate**, often represented by a governmental Treasury bill YTM.[41]
- The **default risk premium** compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount. This itself may be analyzed as the sum of the expected default loss in yield terms plus a premium for the nondiversifiable risk of default.[42]
- The **illiquidity premium** compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.[43]

---

39 For more discussion on equity risk premiums, see Arnott and Bernstein (2002), Grinold and Kroner (2002), and Ilmanen (2003).

40 If there is a mispricing, then the expected return would differ from the required return by a capital appreciation or depreciation component reflecting the convergence to fair value over some time frame.

41 Technically, 1 plus the nominal rate equals the product of 1 plus the real rate and 1 plus the inflation rate. As a quick approximation, however, the nominal rate is equal to the real rate plus an inflation premium. In this discussion, we focus on approximate additive relationships to highlight the underlying concepts.

42 See Elton, Gruber, Agrawal, and Mann (2001) for empirical support for such an analysis.

43 Some writers refer to the "illiquidity premium" as the "liquidity" premium (where "lack of liquidity" is understood).

19年CFA-FRM招生简章

一、SVIP资料包括以下内容：【所有课程均有配套课件讲义】

前导课程：考试介绍、数量基础、金融英语、金融市场与产品的介绍，为学员建立考试框架，打牢基础；
基础课程：紧扣考纲要求，对考点全面解析，建立系统知识框架， 帮助学员根据课程进度系统性学习；
强化课程：突出重点，化繁为简。强化对重要知识点的掌握，让学员完全把握知识点与掌握解题技巧；
直播串讲：以习题串讲的形式，针对每门课程进行对应的重点难点讲解，帮助考生更好的掌握对应知识；
冲刺课程：分析历年考试的重点和难点，通过2次模考与讲解，强化对重点的知识的把握，提高应试技巧；
百题预测：考点精准预测85%以上。以题带点，掌握出题思路与解题方法，最终通过考试取得证书；
赠送资料：官方原版教材 高清NOTES 习题 秘籍 考纲解读等电子资料；

二、SVIP价格说明：【一次收费到当期课程完结，绝无二次收费】

2019年课程价格：500元。购买会赠送18/17年本级别全套课程

往期课程价格：288元。

备注：此价格为一个级别全套课程价格，无其他任何隐形收费。

限时优惠：资料逐步涨价，第一阶段（截止5月1日）500元，第二阶段（截止考试结束）600元。

三、付款方式：

联系微信：fcayyh

（支持支付宝、微信等支付方式！）

四、资料发放方式：

主要通过百度网盘提供下载，课程更新同样在百度网盘里进行。

为确保课程不外传，所有课程全程加密，每位VIP会员给两台设备授权观看。

课程支持windows/mac/ios/安卓等系统，支持手机/平板/电脑等设备。

五、2018年部分会员PASS报喜截图：【限于篇幅，随机抽取】

- The **maturity premium** compensates investors for the increased sensitivity, in general, of the market value of debt to a change in market interest rates as maturity is extended, holding all else equal. The difference between the interest rate on longer-maturity, liquid Treasury debt and that on short-term Treasury debt reflects a positive maturity premium for the longer-term debt (and possibly different inflation premiums as well).
- A **tax premium** may also be applicable to certain classes of bonds in some tax jurisdictions.[44]

For example, consider the expected return on a five-year Treasury instrument traded in a developed market when the real risk-free interest rate is 1.5 percent per year, the expected inflation rate over that horizon is 2.5 percent per year, and a one-year Treasury instrument has a yield to maturity of 4 percent per year. Suppose that the five-year Treasury instrument is priced to yield 5 percent. What is the source of the 5 percent to 4 percent spread? As government debt does not have default risk, the longer-term instrument does not bear a default risk premium. Nor does it have an illiquidity premium or differ in taxation from the one-year instrument. The spread would be accounted for as a 1 percent maturity premium.

EXAMPLE 15

The Long-Term Real Risk-Free Rate

The real risk-free rate is compensation for forgoing current consumption in exchange for certain future consumption. Historical real cash rates exhibit high volatility and differ through time and between countries. We distinguish between the current real rate (driven by cyclical factors) and the long-term real rate assumption (based on sustainable equilibrium conditions). In a free economy, the real rate equilibrates the productivity of the economy and society's time preference for consumption. On a forward-looking basis, we can form opinions about the size of the real rate by analyzing societal consumption time preferences and studying the economy's productivity. For developed countries, a range for the long-term real risk-free rate is 2.0 percent to 2.8 percent. Obviously, variation around this estimate has been and is likely to be substantial, but 2.4 percent is an indication of central tendency over the long term.

EXAMPLE 16

The Real Interest Rate and Inflation Premium in Equilibrium

The expected return to any asset or asset class has at least three components: the real risk-free rate, the inflation premium, and the risk premium. In equilibrium and assuming fully integrated markets, the real risk-free rate should be identical for all assets globally. Similarly, from the frame of reference of any individual investor, the inflation premium should be the same for all assets. For

44 For example, in the United States, bonds issued by private corporations are generally tax disadvantaged relative to bonds issued by the federal government, and a tax premium would compensate corporate bondholders. See Elton, Gruber, Agrawal, and Mann (2001) for evidence on the tax premium.

investors with different base currency consumption baskets, different inflation premiums are required to compensate for different rates of depreciation of investment capital.

The inflation premium is the compensation for the depreciation of invested principal because of *expected* price inflation. In equilibrium, we use the inflation rate that each market is using to compensate it for the loss of purchasing power.

EXAMPLE 17

The Risk Premium: Some Facts

The term "risk premium" is often used to refer to the total premium above the nominal default-risk-free interest rate. Some points to keep in mind:

- In comparing risk premium estimates, the analyst should make sure that a common benchmark for the risk-free rate is being used; if not, the estimates should be adjusted to a common risk-free-rate reference point.
- Some analysts do not view illiquidity as a kind of risk and may refer to an illiquidity premium in addition to the risk premium when estimating the required return on an illiquid asset.
- Modeling any risk premium requires an assessment of the degree of capital market integration. Capital market integration will be discussed in Section 3.1.4.

Examples 15, 16, and 17 provide some information on the real risk-free interest rate, including long-term levels. The inflation premium is typically a more volatile element of the yield of bonds. In standard discussions of term-structure theory, the term structure of interest rates for default-free government bonds can provide an estimate of inflation expectations. Furthermore, in markets with active issuance of inflation-indexed bonds, the yield spread at a given maturity of conventional government bonds over inflation-indexed bonds of the same maturity may be able to provide a market-based estimate of the inflation premium at that horizon. The analyst can use market yield data and credit ratings (or other credit models) to estimate the default risk premium by comparing the yields on bonds matched along other dimensions but differing in default risk. An analogous approach may be applied to estimating the other premiums.

3.1.3.3 The Equity Risk Premium The **equity risk premium** is the compensation required by investors for the additional risk of equity compared with debt (debt has a prior claim on the cash flows of the company). An estimate of the equity risk premium directly supplies an estimate of the expected return on equities through the following expression:

$$E(R_e) = \text{YTM on a long-term government bond} + \text{Equity risk premium} \quad (8)$$

where "long-term" has usually been interpreted as 10 or 20 years. (In many markets, bonds with maturities longer than 10 years are not available or actively traded.) As long as one is consistent with the choice of maturity in defining the equity risk premium, either choice is feasible. Equation 8 has been called the **bond-yield-plus-risk-premium method** of estimating the expected return on equity. From Equation 8, we also see that the *equity risk premium in practice is specifically defined as the expected excess return over and above a long-term government bond yield.*

A historical analysis has often been used as a point of departure in estimating the equity risk premium. Exhibit 9 gives the *ex post* data for the 111 years from 1900 to 2010. The standard deviation column represents the volatility in the difference between equity returns and bond returns on a year-by-year basis.

Exhibit 9 Historical Equity Risk Premiums around the World: 1900–2010

	Annual Realized Equity Risk Premium Relative to Long Bond Returns			
Country	**Geometric Mean**	**Arithmetic Mean**	**Standard Error**	**Standard Deviation**
Australia	5.9%	7.8%	1.9%	19.8%
Belgium	2.6	4.9	2.0	21.4
Canada	3.7	5.3	1.7	18.2
Denmark	2.0	3.4	1.6	17.2
France	3.2	5.6	2.2	22.9
Germany (ex-1922/23)	5.4	8.8	2.7	28.4
Ireland	2.9	4.9	1.9	19.8
Italy	3.7	7.2	2.8	29.6
Japan	5.0	9.1	3.1	32.8
Netherlands	3.5	5.8	2.1	22.2
Norway	2.5	5.5	2.7	28.0
South Africa	5.5	7.0	1.9	19.6
Spain	2.3	4.3	2.0	20.8
Sweden	3.8	6.1	2.1	22.3
Switzerland	2.1	3.6	1.7	17.6
United Kingdom	3.9	5.2	1.6	17.0
United States	4.4	6.4	1.9	20.5
World average	3.8	5.0	1.5	15.5

Source: Dimson, Marsh, and Staunton (2011), Table 10. See this source for the details of the equity and bond series used. Note that the world average is based on more markets than are shown in this exhibit.

From Exhibit 9, we can draw the following conclusions:

- The geometric mean historical equity risk premium ranged from a low of 2.0 percent (Denmark) to a high of 5.9 percent (Australia) with an average of 3.8 percent.
- The arithmetic mean historical equity risk premium ranged from a low of 3.4 percent (Denmark) to a high of 9.1 percent (Japan) with an average of 5.0 percent.
- As measured by the standard error (which applies to the arithmetic mean), for most markets there is considerable sampling error in the sample mean estimate of the (presumed unchanging) population mean equity risk premium. Consider the United States, with a standard error of 1.9 percent. Under a normality assumption, one could be only 68 percent confident that the population mean is between 4.5 percent (6.4 percent − 1.9 percent) and 8.3 percent (6.4 percent + 1.9 percent).
- The standard deviation column shows that there is a great amount of variation in the annual return difference between equities and bonds.

The size of *ex ante* equity risk premiums in international markets and the interpretation of the historical record for the purposes of estimating them have been the source of a lively, continuing, and unresolved debate. Grinold and Kroner (2002) and Dimson, Marsh, and Staunton (2002, 2006 and 2011) provide useful observations on the issues raised.

3.1.4 *Financial Market Equilibrium Models*

Financial equilibrium models describe relationships between expected return and risk in which supply and demand are in balance. In that sense, equilibrium prices or equilibrium returns are fair if the equilibrium model is correct.

Equilibrium approaches to setting capital market expectations include the Black–Litterman approach and the international CAPM–based approach presented in Singer and Terhaar (1997). The Black–Litterman approach reverse-engineers the expected returns implicit in a diversified market portfolio, combining them with the investor's own views in a systematic way that takes account of the investor's confidence in his or her views.[45] This approach is discussed at greater length in the reading on asset allocation.

Singer and Terhaar (1997) proposed an equilibrium approach to developing capital market expectations that involves calculating the expected return on each asset class based on the international capital asset pricing model (ICAPM),[46] taking account of market imperfections that are not considered by the ICAPM.[47]

Assuming that the risk premium on any currency equals zero—as it would be if purchasing power parity relationships hold—the ICAPM gives the expected return on any asset as the sum of:

- the (domestic) risk-free rate, and
- a risk premium based on the asset's sensitivity to the world market portfolio and expected return on the world market portfolio in excess of the risk-free rate.

Equation 9 is the formal expression for the ICAPM:

$$E(R_i) = R_F + \beta_i[E(R_M) - R_F] \tag{9}$$

where

$E(R_i)$ = the expected return on asset i given its beta

R_F = the risk-free rate of return

$E(R_M)$ = the expected return on the world market portfolio

β_i = the asset's sensitivity to returns on the world market portfolio, equal to $\mathrm{Cov}(R_i,R_M)/\mathrm{Var}(R_M)$

An important question concerns the identification of an appropriate proxy for the world market portfolio. Based on the criteria of Brinson, Diermeier, and Schlarbaum (1986, p. 17), the analyst can define and use the **global investable market** (GIM). The GIM is a practical proxy for the world market portfolio consisting of traditional and alternative asset classes with sufficient capacity to absorb meaningful investment.[48]

---

45 See Black and Litterman (1991, 1992). The notion of using reverse optimization to infer expected returns was first described in Sharpe (1974).

46 ICAPM has also been used as an acronym for the intertemporal capital asset pricing model developed by Merton (1973). In this reading, ICAPM refers to the international capital asset pricing model.

47 See also Terhaar, Staub, and Singer (2003). The specific suggestion of this approach is to use the factor-model-based estimate of the covariance matrix presented in Section 3.1.1.4.

48 See Brinson, Diermeier, and Schlarbaum (1986) for more details.

Equation 9 implies that an asset class risk premium, RP_i, equal to $E(R_i) - R_F$, is a simple function of the world market risk premium, RP_M, equal to $E(R_M) - R_F$:[49]

$$RP_i = \frac{\sigma_i}{\sigma_m}\rho_{i,M}\left(RP_M\right)$$

Moving the market standard deviation of returns term within the parentheses, we find that an asset class's risk premium equals the product of the Sharpe ratio (RP_M/σ_M) of the world market portfolio, the asset's own volatility, and the asset class's correlation with the world market portfolio:

$$RP_i = \sigma_i\rho_{i,M}\left(\frac{RP_M}{\sigma_M}\right) \tag{10}$$

Equation 10 is one of two key equations in the Singer–Terhaar approach. The Sharpe ratio in Equation 10 (i.e., RP_M/σ_M) is the expected excess return per unit of standard deviation of the world market portfolio. The world market portfolio's standard deviation represents a kind of risk (systematic risk) that cannot be avoided through diversification and that should therefore command a return in excess of the risk-free rate. An asset class's risk premium is therefore the expected excess return accruing to the asset class given its global systematic risk (i.e., its beta relative to the world market portfolio).

Equation 10 requires a market Sharpe ratio estimate. Singer and Terhaar (1997, pp. 44–52) describe a complete analysis for estimating it. As of the date of their analysis, 1997, they recommended a value of 0.30 (a 0.30 percent return per 1 percent of compensated risk). Goodall, Manzini, and Rose (1999, pp. 4–10) revisited this issue on the basis of different macro models and recommended a value of 0.28. For this exposition, we adopt a value of 0.28. In fact, the Sharpe ratio of the global market could change over time with changing global economic fundamentals.

To illustrate Equation 10, suppose that an investor predicts that the standard deviation of Canadian bonds will be 7.0 percent per year and that their correlation with the GIM is 0.54. Then, with our estimate of the market Sharpe ratio, we would estimate the risk premium as

$$7\% \times 0.54 \times 0.28 = 1.06\%$$

For Canadian equities, with a standard deviation of 17 percent and a 0.70 correlation with the GIM, we would estimate the equity risk premium as

$$17\% \times 0.70 \times 0.28 = 3.33\%$$

The Singer–Terhaar approach recognizes the need to account for market imperfections that are not considered by the ICAPM. We will consider two market imperfections: illiquidity and market segmentation.

In the discussion of bonds, we defined the illiquidity premium as compensation for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly. The ICAPM assumes **perfect markets** (markets without any frictional costs, where all assets trade in liquid markets). Thus, we need to add an estimated illiquidity premium to an ICAPM expected return estimate as appropriate. The ICAPM was formulated with developed securities markets such as the Canadian bond and equity markets in mind, and the Singer–Terhaar approach would not add an illiquidity premium to ICAPM expected return estimates for Canadian stocks and bonds.

---

49 The expression is derived as follows: $RP_i = \beta_i\left(RP_M\right) = \left[\mathrm{Cov}(R_i,R_M)/\sigma_M^2\right]\left(RP_M\right) = \left(\sigma_i\sigma_M\rho_{iM}/\sigma_M^2\right)\left(RP_M\right) = \sigma_i\rho_{iM}\left(RP_M/\sigma_M\right)$, where we have used the fact that $\mathrm{Cov}(R_i,R_M) = \sigma_i\sigma_M\rho_{iM}$.

However, the illustrated risk premium estimates for Canadian bonds and equities (1.06 percent and 3.33 percent, respectively) are those that would hold if Canadian bond and equity markets were perfectly integrated with other world asset markets. **Market integration** means that there are no impediments or barriers to capital mobility across markets. Barriers include not only legal barriers, such as restrictions a national emerging market might place on foreign investment, but also cultural predilections and other investor preferences. If markets are perfectly integrated, all investors worldwide participate equally in setting prices in any individual national market. Market integration implies that two assets in different markets with identical risk characteristics must have the same expected return. **Market segmentation** means that there are some meaningful impediments to capital movement across markets. Although many barriers to international capital flows have come down, some do persist and a number of asset markets are in practice at least partially segmented across national borders. The more a market is segmented, the more it is dominated by *local* investors. When markets are segmented, two assets in different markets with identical risk characteristics may have different expected returns. If an asset in a segmented market appears undervalued to a nondomestic investor *not* considering barriers to capital mobility, *after* such barriers are considered, the investor may not actually be able to exploit the opportunity.

Most markets lie between the extremes of perfect market integration and complete market segmentation. A home-biased perspective or partial segmentation is perhaps the best representation of most markets in the world today. We need first to develop an estimate of the risk premium for the case of complete market segmentation. With such an estimate in hand, the estimate of the risk premium for the common case of partial segmentation is just a weighted average of the risk premium assuming perfect market integration and the risk premium assuming complete segmentation, where the weights reflect the analyst's view of the degree of integration of the given asset market.

To address the task of estimating the risk premium for the case of complete market segmentation, we must first recognize that if a market is completely segmented, the market portfolio in Equations 9 and 10 must be identified as the *individual local market*. Because the individual market and the reference market portfolio are identical, $\rho_{i,M}$ in Equation 10 equals 1. (For example, if Canadian equities were a completely segmented market, the reference market portfolio and the individual market portfolio would each be a broad-based index for Canadian equities, and the correlation of such an index with itself would of course be 1.) The value of 1 for correlation is the maximum value, so all else being equal, the risk premium for the completely segmented markets case is higher than that for the perfectly integrated markets case and equal to the amount shown in Equation 11:

$$\mathrm{RP}_i = \sigma_i\left(\frac{\mathrm{RP}_M}{\sigma_M}\right) \qquad (11)$$

This is the second key equation in the Singer–Terhaar approach. Assuming that Canadian bonds and equities trade in completely segmented markets, we would calculate respective risk premiums of[50]

$$7\% \times 0.28 = 1.96\%$$

and

$$17\% \times 0.28 = 4.76\%$$

50 For simplicity, we are assuming that the Sharpe ratios of the GIM and the local market portfolio (used in Equation 11) are the same.

Taking the degree of integration as 0.8 for both Canadian equities and bonds, our final risk premium estimates would be as follows:

- $RP_{\text{Cdn FI}} = (0.8 \times 1.06\%) + (0.2 \times 1.96\%) = 1.24\%$
- $RP_{\text{Cdn equities}} = (0.8 \times 3.33\%) + (0.2 \times 4.76\%) = 3.62\%$

Thus, assuming a risk-free rate of 4 percent, we would estimate the expected returns on Canadian bonds and equities as the sum of the risk-free rate and the relevant risk premium, as follows:

- Canadian bonds: $E(R_{\text{Cdn FI}}) = 4\% + 1.24\% = 5.24\%$
- Canadian equities: $E(R_{\text{Cdn equities}}) = 4\% + 3.62\% = 7.62\%$

To summarize, to arrive at an expected return estimate using the Singer–Terhaar approach, we take the following steps:

- Estimate the perfectly integrated and the completely segmented risk premiums for the asset class using the ICAPM.
- Add the applicable illiquidity premium, if any, to the estimates from the prior step.
- Estimate the degree to which the asset market is perfectly integrated.
- Take a weighted average of the perfectly integrated and the completely segmented risk premiums using the estimate of market integration from the prior step.[51]

The analyst needs to develop estimates of the degree of integration of an asset market, but as a starting point, research has suggested that developed market equities and bonds are approximately 80 percent integrated. To give a flavor of the variation that might be expected, research has also indicated that US and UK real estate is approximately 70 percent integrated; real estate in France, Germany, the Netherlands, and Switzerland is 60 percent integrated; emerging market equities and bonds are about 65 percent integrated; and at the low end of integration are assets such as timber at 50 percent (United States, Australia) or 40 percent (Argentina, Brazil, Chile, Uruguay). Currency and cash markets are 100 percent integrated.[52]

Another task for the analyst is estimating the illiquidity premium for an asset class. Estimating this premium for alternative investments presents a great challenge. Many such investments cannot be traded at all for some time (i.e., are **locked up**, as might be the case for early-stage venture capital). Rebalancing to a target allocation is not feasible during the lockup period and is relatively costly afterward.

EXAMPLE 18

Justifying Capital Market Forecasts

Samuel Breed, CIO of a university endowment, is presenting the capital market expectations shown in Exhibit 10 to the endowment's board of trustees.

51 Alternatively, we can substitute the Sharpe ratio of the "typical" local investor's investment portfolio for the GIM portfolio in Equation 10 and use the correlation of the asset class under consideration with that typical local portfolio.
52 See Staub (2005).

Exhibit 10 Capital Market Projections

Asset Class	Proxy	Projected 5-Year Annual Return (%)	Projected Standard Deviation (%)
Equities			
1. Large-cap US equity	S&P 500	8.8	16.5
2. Small/mid-cap US equity	Russell 2500	9.8	22.0
3. Ex-US equity	MSCI EAFE	9.2	20.0
Fixed Income			
4. Domestic fixed income	Bloomberg Barclays Aggregate	4.7	4.5
5. Non-US fixed income	Citi Non-US Govt.	4.6	9.5
Other Assets			
6. US real estate	NCREIF	7.6	14.0
7. Private equity	VE Post Venture Cap.	12.0	34.0
8. Cash equivalents	90-day T-bill	3.3	1.0
Inflation	CPI-U	2.6	1.4

Correlations:	1	2	3	4	5	6	7	8
1. Large-cap US equity	1.0							
2. Small/mid-cap US equity	0.85	1.0						
3. Ex-US equity	0.74	0.61	1.0					
4. Domestic fixed income	0.27	0.20	0.21	1.0				
5. Non-US fixed income	0.03	−0.03	0.22	0.32	1.0			
6. US real estate	0.64	0.52	0.47	0.20	0.03	1.0		
7. Private equity	0.63	0.57	0.63	0.20	0.10	0.45	1.0	
8. Cash equivalents	−0.10	−0.15	−0.25	0.30	−0.05	−0.06	0.07	1.0

Assume the following:

- The Sharpe ratio of the global investable market portfolio (GIM) is 0.28, and its standard deviation is 7 percent.
- The beta of private equity with respect to the GIM is 3.3, and the beta of small/mid-cap US equity is 2.06.

William Smyth, a trustee, questions various projections for private equity, as follows:

A "I have seen volatility estimates for private equity based on appraisal data that are much smaller than the one you are presenting, in which the volatility of private equity is much larger than that of small/mid-cap US equity. Your volatility estimate for private equity must be wrong."

B "The premium of private equity over small/mid-cap US equity is not justifiable because they both represent ownership interests in US business."

C "Using the ICAPM, the forecast correlation between private equity returns and small/mid-cap US equity returns is lower than your estimate indicates."

1. Evaluate whether Smyth's Comment A is accurate.
2. Evaluate whether Smyth's Comment B is accurate.
3. Evaluate whether Smyth's Comment C is accurate.

Solution to 1:

Smyth's Comment A is not accurate. Although private equity and small-cap stocks both represent ownership interests, private equity is not traded and appraisal data will tend to underestimate volatility.

Solution to 2:

Smyth's Comment B is not accurate. One justification for a higher expected return for private equity is that it has a lockup period and should therefore bear an illiquidity premium.

Solution to 3:

Smyth's Comment C is accurate. According to elementary portfolio theory, the correlation between two assets is given by $\beta_1\beta_2\sigma_M^2 / \sigma_1\sigma_2$. Thus, the correlation between private equity and small/mid-cap US equity is equal to $(3.3)(2.06)(7\%)^2/(34\%)(22\%) = 0.45$, which is lower than the estimate of 0.57 given in Exhibit 10.

The illiquidity premium for an alternative investment should be positively related to the length of the investment's lockup period or illiquidity horizon. How can the amount of the illiquidity premium be estimated? One estimation approach uses the investment's **multiperiod Sharpe ratio** (MPSR), which is based on the investment's multiperiod wealth in excess of the wealth generated by the risk-free investment (i.e., compounded return over compounded cash return). The relevant MPSR is one calculated over a holding period equal to the investment's lockup period. There would be no incentive to invest in an illiquid alternative investment unless its MPSR—its risk-adjusted wealth—were at least as high as the MPSR of the market portfolio at the end of the lockup period. Suppose that an alternative investment has a lockup period of eight years and its ICAPM-given required rate of return is 12 percent but its MPSR is below that of the GIM—say, 0.67—at an eight-year horizon. If increasing its expected return to 20 percent makes the alternative investment's MSPR equal 0.67 at the eight-year horizon, then the estimate of the illiquidity premium is 20% − 12% = 8%.[53]

Example 19 illustrates the Singer–Terhaar approach. In the example, for simplicity's sake, the ICAPM betas are used to develop covariance estimates.

---

53 See Staub and Diermeier (2003) for more details.

EXAMPLE 19

Setting CME Using the Singer–Terhaar Approach

Zimmerman Capital Management (ZCM) is developing a strategic asset allocation for a small US foundation that has approved investment in the following five asset classes: US equities, US fixed income, non-US equities, non-US fixed income, and US real estate. The foundation limits nondomestic assets to no more than 12 percent of invested assets.

- The final set of expectations needed consists of the expected returns, standard deviations, and all distinct pairwise covariances of US equities, US fixed income, non-US equities, non-US fixed income, and US real estate. The investment time horizon is 10 years.
- A risk premium approach will be taken to developing expected return estimates following the methodology of Singer and Terhaar. Historical estimates of standard deviations will be used, and ICAPM betas will be used to develop estimates of covariances.
- Exhibit 11 supplies the standard deviation estimates and gives relevant inputs for other quantities needed. In addition, ZCM has gathered the following facts and estimates:
  - The Sharpe ratio of the GIM is estimated to be 0.28.
  - The standard deviation of the GIM is estimated to be 7 percent.
  - The risk-free rate of interest is 3 percent.
- Equities and bonds are assumed to be 80 percent integrated, and US real estate is assumed to be 70 percent integrated.

Exhibit 11 Equilibrium Approach to Risk Premium Estimation

Asset Class	Standard Deviation	Correlation with GIM	Premium to Equate Sharpe Ratio at Illiquidity Horizon
US equities	15.7%	0.85	0%
US fixed income	3.8	0.75	0
Non-US equities	15.6	0.80	0
Non-US fixed income	9.1	0.70	0
US real estate	11.5	0.50	0.30

Based on the information given, address the following problems:

1 Calculate the expected returns on US equities, US fixed income, non-US equities, non-US fixed income, and US real estate. Make any needed adjustments for illiquidity.

2 Show the calculation of the covariance between US equities and US fixed income.

3 Critique the following statement: “The ZCM risk premium estimates are low, given that the foundation has a very strong home-country bias, reflected in its limitation of nondomestic assets to no more than 12 percent of the portfolio.”

Solution to 1:

To calculate the expected return for an asset class, we take the following steps. First, we calculate the risk premium of the asset class for two distinct cases: full integration and complete segmentation. In the calculation for either case, we take care to add any applicable illiquidity premium. Second, we average the two estimates of the risk premium for an asset class by weighting the full integration estimate by the assumed degree of integration and the complete segmentation estimate by (1 – assumed degree of integration). The result of this step is our informed estimate of the asset class's risk premium. Finally, adding the risk premium estimate to the risk-free rate yields our estimate of the expected return on the asset class.

Step 1

Using Equation 10, we find that in the fully integrated case,

$RP_{\text{US equities}} = 15.7\% \times 0.85 \times 0.28 = 3.74\%$

$RP_{\text{US FI}} = 3.8\% \times 0.75 \times 0.28 = 0.80\%$

$RP_{\text{non-US equities}} = 15.6\% \times 0.80 \times 0.28 = 3.49\%$

$RP_{\text{non-US FI}} = 9.1\% \times 0.70 \times 0.28 = 1.78\%$

$RP_{\text{US RE}} = (11.5\% \times 0.50 \times 0.28) + 0.30\% = 1.61\% + 0.30\% = 1.91\%$

Using Equation 11, we find that in the fully segmented case,

$RP_{\text{US equities}} = 15.7\% \times 0.28 = 4.4\%$

$RP_{\text{US FI}} = 3.8\% \times 0.28 = 1.06\%$

$RP_{\text{non-US equities}} = 15.6\% \times 0.28 = 4.37\%$

$RP_{\text{non-US FI}} = 9.1\% \times 0.28 = 2.55\%$

$RP_{\text{US RE}} = (11.5\% \times 0.28) + 0.30\% = 3.22\% + 0.30\% = 3.52\%$

Note that we added an illiquidity premium of 0.3 percent to the ICAPM derived premium estimates for real estate.

Step 2

We now weight each asset class's fully integrated and segmented premiums according to the assumed degree of integration.

$RP_{\text{US equities}} = (0.8 \times 3.74\%) + (0.2 \times 4.4\%) = 3.87\%$

$RP_{\text{US FI}} = (0.8 \times 0.80\%) + (0.2 \times 1.06\%) = 0.85\%$

$RP_{\text{non-US equities}} = (0.8 \times 3.49\%) + (0.2 \times 4.37\%) = 3.67\%$

$RP_{\text{non-US FI}} = (0.8 \times 1.78\%) + (0.2 \times 2.55\%) = 1.93\%$

$RP_{\text{US RE}} = (0.7 \times 1.91\%) + (0.3 \times 3.52\%) = 2.39\%$

Step 3

The expected return estimates are as follows:

$E(R_{\text{US equities}}) = 3\% + 3.87\% = 6.87\%$

$E(R_{\text{US FI}}) = 3\% + 0.85\% = 3.85\%$

$E(R_{\text{non-US equities}}) = 3\% + 3.67\% = 6.67\%$

$E(R_{\text{non-US FI}}) = 3\% + 1.93\% = 4.93\%$

$E(R_{\text{US RE}}) = 3\% + 2.39\% = 5.39\%$

Solution to 2:

Based on Equation 3b with one factor, the covariance between any two assets in a one-beta model (such as the ICAPM) is equal to the product of each asset's beta with respect to the market times the variance of the market. The needed betas can be calculated as

$$\beta_{\text{US equities}} = (15.7\% \times 0.85)/7\% = 1.91$$

$$\beta_{\text{US FI}} = (3.8\% \times 0.75)/7\% = 0.41$$

and the covariance between US equities and US fixed income returns as

$$\text{Cov(US equities, US FI)} = 1.91 \times 0.41 \times (7\%)^2 = 38.37 \text{ (in units of percent squared)}$$

Solution to 3:

Although the client is correct about the foundation's home-country bias, the point being made is not correct. The equilibrium risk premium is determined by all investors, reflected in the overall degree of integration estimates.

3.2 Survey and Panel Methods

The **survey method** of expectations setting involves asking a group of experts for their expectations and using the responses in capital market formulation. If the group queried and providing responses is fairly stable, the analyst in effect has a panel of experts and the approach can be called a **panel method**. These approaches are based on the straightforward idea that a direct way to uncover a person's expectations is to ask the person what they are.

The oldest continuous survey of expectations is the so-called Livingston Survey, initiated in 1946 by Joseph Livingston, a Philadelphia journalist, and managed since 1990 by the Federal Reserve Bank of Philadelphia, part of the US Federal Reserve System. The survey covers real US GDP growth, CPI and PPI inflation, the unemployment rate, and 3-month T-bill and 10-year T-bond yields. In the United States, Welch surveyed financial economists for their views about the short- and long-term (30-year) equity risk premium in 1998, 2001, and 2009.[54] The results of the three Welch surveys for the 30-year equity risk premium are summarized in Exhibit 12.

Exhibit 12 Consensus Expectations of US Financial Economists of the Long-term (30-Year) US Equity Risk Premium

	2009 Survey	2001 Survey	1998 Survey
Mean	6.2%	5.5%	7.1%
Median	6.0	5.0	7.0
Interquartile range	5%–7%	4%–7%	6%–8.4%

Source: Welch (2000, 2001, 2009).

54 See Welch (2000) and Welch (2001). The 1998 survey had 226 respondents, while the 2001 survey had 510. Graham and Harvey (2001) report a survey of 10-year forecasts of the US equity risk premium by chief financial officers, but their survey question was not specific about whether an arithmetic or geometric mean estimate was sought.

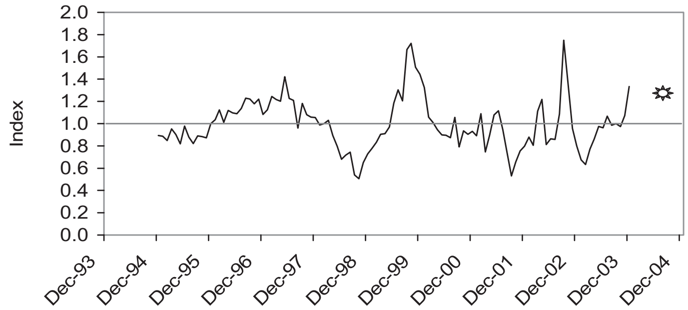
A 2002 survey of global bond investors by Schroder Salomon Smith Barney found an average equity risk premium in the range of 2–2.5 percent, while a Goldman Sachs survey of global clients recorded a mean long-run equity risk premium of 3.9 percent.[55] Such surveys may be sensitive to the professional identity of the respondents. Lally, Roush, and Van Zijl (2004) found the predictions of practitioners for the New Zealand equity risk premium significantly higher than those of academics.[56] Besides direct questions on capital market expectations, for certain equity markets, there are commercial surveys of analysts' forecasts of long-term earnings growth rates that implicitly contain an equity market forecast given a DCF valuation model.

EXAMPLE 20

Short-Term Consumer Spending in the United Kingdom

Bryan Smith is researching the 6- to 12-month expectations for consumer spending in the United Kingdom as of the middle of 2003. One piece of evidence he gathers is changes in consumer sentiment in the United Kingdom as measured by the Economic Optimism Index, shown in Exhibit 13.

Exhibit 13 UK Consumer Economic Optimism Index (Year over Year Change in MORI Survey Data)



Note: Based on an index created from UK Consumer Optimism MORI survey data. The asterisk represents the forecast target for 2004 as of mid-2003.

Interpret Exhibit 13 as it relates to the probable path of consumer spending.

Solution:

Based on the reading at December 2003 of the UK Consumer Optimism Index, it appears that consumers are considerably more optimistic than in December 2002. Rising consumer optimism is a reflection of consumers feeling secure about their income stream and future. Rising consumer optimism suggests that near-term consumer spending will increase.

55 See Ilmanen, Byrne, Gunasekera, and Minikin (2002) and O'Neill, Wilson, and Masih (2002).
56 The median forecast was 7.0 percent for practitioners and 5.5 percent for academics.

3.3 Judgment

In a disciplined expectations-setting process, the analyst should be able to factually explain the basis and rationale for forecasts. Quantitative models such as equilibrium models offer the prospect of providing a non-emotional, objective rationale for a forecast. The expectations-setting process nevertheless can give wide scope to applying judgment—in particular, economic and psychological insight—to improve forecasts. In forecasting, numbers, including those produced by elaborate quantitative models, must be evaluated.

EXAMPLE 21

Judgment Applied to Correlation Estimation

William Chew's firm uses a multifactor model to develop initial correlation forecasts that are then challenged by professionals within the capital markets unit. Based on US historical data including periods of high inflation, Chew finds that the model forecasts a correlation between US equity and US bonds in the range of 0.40 to 0.45. Based on empirical evidence, Chew believes that the correlation between equity and bond returns is higher in high-inflation periods than in low-inflation periods. His firm's chief economist forecasts that in the medium term, US inflation will be low, averaging less than 3 percent per annum. In light of that forecast, Chew has decided that he will recommend a judgmental downward adjustment of the correlation to 0.30.

Other investors who rely on judgment in setting capital market expectations may discipline the process by the use of devices such as checklists. In any case, investment experience, the study of capital markets, and intelligence are requisites for the development of judgment in setting capital market expectations.

4 ECONOMIC ANALYSIS

History has shown that there is a direct yet fluid relationship between actual realized asset returns, expectations for future asset returns, and economic activity. The linkages are consistent with asset-pricing theory, which predicts that the risk premium of an asset is related to the correlation of its payoffs with the marginal utility of consumption in future periods. Assets with low expected payoffs in periods of weak consumption (e.g., business cycle troughs) should bear higher risk premiums than assets with high expected payoffs in such periods. Because investors expect assets of the second type to provide good payoffs when their income may be depressed, they should be willing to pay relatively high prices for them (implying lower risk premiums).[57]

Analysts need to be familiar with the historical relationships that empirical research has uncovered concerning the direction, strength, and lead-lag relationships between economic variables and capital market returns.

The analyst who understands which economic variables may be most relevant to the current economic environment has a competitive advantage, as does the analyst who can discern or forecast a change in trend or point of inflection in economic

57 See Cochrane (1999a, 1999b).

activity. Inflection points often present unique investment opportunities at the same time that they are sources of latent risk. Questions that may help the analyst assess points of inflection include the following:

- What is driving the economy in its current expansion or contraction phase?
- What is helping to maintain economic growth, demand, supply, and/or inflation rates within their current ranges?
- What may trigger the end of a particular trend?

The economic output of many economies has been found to have cyclical and trend growth components. Trend growth is of obvious relevance for setting long-term return expectations for asset classes such as equities. Cyclical variation affects variables such as corporate profits and interest rates, which are directly related to asset class returns and risk. In the following sections, we address business cycles and trend growth.

4.1 Business Cycle Analysis

In business cycle analysis, two cycles are generally recognized: a short-term **inventory cycle**, typically lasting 2–4 years, and a longer-term **business cycle**, usually lasting 9–11 years. Evidence for both these cycles goes back two centuries or more, but they are very far from working like clockwork. In particular, they can be disrupted by major shocks, including wars and shifts in government policy. Also, both the duration and amplitude of each phase of the cycle, as well as the duration of the cycle as a whole, vary considerably and are hard to predict.

Cycles mark variation in economic activity, so we should be clear on how that variation is measured. The chief measurements of economic activity are as follows:

- **Gross domestic product (GDP)**: GDP is a calculation of the total value of final goods and services produced in the economy during a year. The main expenditure components are Consumption, Investment, Change in Inventories, Government Spending, and Exports less Imports. The total value of goods and services can change because the quantities of goods and services change and/or because their prices change. To focus on increases in the quantity (output) of goods and services produced—which are directly associated with increases in the standard of living—rather than on price-driven increases in the value of output, economists focus on real GDP (reflecting an adjustment for changes in prices during the period). For brevity's sake in our discussion, "GDP" is understood as referring to "real GDP" unless otherwise stated.
- **Output gap**: The **output gap** is the difference between the value of GDP estimated as if the economy were on its trend growth path (sometimes referred to as **potential output**) and the actual value of GDP. A positive output gap opens in times of recession or slow growth. When a positive output gap is open, inflation tends to decline. Once the gap closes, inflation tends to rise. When GDP is above its trend value, the economy is under inflationary pressure. Many macroeconomists consider the output gap as the key measure of real activity for policy making because it provides information about future inflationary pressures as well as an output objective. However, because changing demographics and technology affect the economy's trend path, real-time estimates of the output gap can sometimes be quite inaccurate.
- **Recession**: In general terms, a **recession** is a broad-based economic downturn. More formally, a recession occurs when there are two successive quarterly declines in GDP.

The following sections discuss the inventory cycle and the business cycle in more detail.

4.1.1 *The Inventory Cycle*

Economists have found evidence of a short-term inventory cycle, lasting 2–4 years. The **inventory cycle** is a cycle measured in terms of fluctuations in inventories. The inventory cycle is caused by companies trying to keep inventories at desired levels as the expected level of sales changes.

In the up phase of the inventory cycle, businesses are confident about future sales and are increasing production. The increase in production generates more overtime pay and employment, which tends to boost the economy and bring further sales. At some point, there is a disappointment in sales or a change in expectations of future sales, so that businesses start to view inventories as too high. In the recent past, a tightening of monetary policy has often caused this inflection point. It could also be caused by a shock such as higher oil prices. Then, business cuts back production to try to reduce inventories and hires more slowly (or institutes layoffs). The result is a slowdown in growth.

It usually takes a year or two for business to correct inventory levels after an inflection point. A good indicator of the inventory position is the inventory/sales ratio. Exhibit 14 shows the inventory/sales ratio for the United States over the period 1992–2011. When the inventory/sales ratio has moved down, the economy is likely to be strong in the next few quarters as businesses try to rebuild inventory, as in early 2004. Conversely, when the ratio has moved sharply up, as in 2008-09, a period of economic weakness can be expected. Note that while this indicator has been trending down because of improved techniques such as "just in time" inventory management, the 2- to 4-year inventory cycle is still evident.

Exhibit 14 US Inventory/Sales Ratios

Source: FRED, Federal Reserve Economic Data, Federal Reserve Bank of St. Louis: Inventory to Sales Ratio: Total Business (ISRATIO); US Department of Commerce: Census Bureau. http://research.stlouisfed.org/fred2/series/ISRATIO; accessed 21 December 2011.

In the late 1990s, it was argued that improved and computerized techniques of inventory control would make the inventory cycle obsolete. In fact, the 2001 and 2008–09 recessions saw one of the steepest inventory corrections on record. The reason seems to have been that excess inventories were more visible more quickly than in the past, and businesses rapidly cut back production.

4.1.2 *The Business Cycle*

In addition to the inventory cycle, there is evidence of a longer cycle, often lasting 9–11 years, called the business cycle. The **business cycle** represents fluctuations in GDP in relation to long-term trend growth. A typical business cycle has five phases: initial recovery, early upswing, late upswing, slowdown, and recession.

1 **Initial Recovery** This is usually a short phase of a few months in which the economy picks up from its slowdown or recession. Generally, confidence is rising among businesses, although consumer confidence may still be at low levels since unemployment is still high. In the initial recovery phase, there are often stimulatory economic policies from the government in the form of lower interest rates or a budgetary deficit. The business cycle recovery is usually supported by a simultaneous upswing in the inventory cycle, which is sometimes the main cause of the recovery. Inflation will still be falling in the initial recovery phase. The output gap is still large.

   Capital market effects: Government bond yields may continue to come down through this phase in anticipation of a further decline in inflation but are likely to be bottoming. Stock markets may rise strongly at this point because fears of a longer recession (or even a depression) dissipate. Cyclical assets—and riskier assets, such as small stocks, higher-yield corporate bonds, and emerging market equities and bonds—attract investors and perform well.

2 **Early Upswing** After the initial recovery period, confidence is up and the economy is gaining some momentum. This is the healthiest period of the cycle, in a sense, because economic growth can be robust without any signs of overheating or sharply higher inflation. Typically, there is increasing confidence, with consumers prepared to borrow and spend more as unemployment starts to fall. Concurrently, businesses build inventories and step up investment in the face of strong sales and increased capacity use. Higher operating levels allow many businesses to enjoy lower unit costs, so that profits rise rapidly.

   Capital market effects: A key question is how long it will take before inflation starts to become a problem. Short rates are moving up at this time as the central bank starts to withdraw the stimulus put in place during the recession. Longer bond yields are likely to be stable or rising slightly. Stocks are still trending up. This phase usually lasts at least a year and often several years if growth is not too strong and the output gap closes slowly.

3 **Late Upswing** At this stage of the cycle, the output gap has closed and the economy is in danger of overheating. Confidence is high; unemployment is low. The economy may grow rapidly. Inflation starts to pick up, with wages accelerating as shortages of labor develop.

   Capital market effects: Typically, interest rates are rising as the monetary authorities become restrictive. Any heavy borrowing puts pressure on the credit markets. Central banks may aim for a "soft landing," meaning a period of slower growth to cool the economy but not a major downturn. Bond markets anxiously watch this behavior, and bond yields will usually be rising as a result of changed expectations. Stock markets will often rise but may be nervous too, depending on the strength of the boom. Nervous investors mean that equities are volatile.

4 **Slowdown** At this point, the economy is slowing, usually under the impact of rising interest rates. The economy is especially vulnerable at this juncture to a shock, which can turn a "soft landing" into a recession. Business confidence starts to waver. Despite the slowdown, inflation often continues to rise. The slowdown is exacerbated by the inventory correction as companies try to

reduce their inventory levels. This phase may last just a few months, as in the United States in 2000, or it may last a year or more, as in the United States in 1989–1990 and 2009–2011.

Capital market effects: Short-term interest rates are high and rising at first but then may peak. Bonds top out at the first sign of a slowing economy and then rally sharply (yields fall). The yield curve often inverts. The stock market may fall, with interest-sensitive stocks such as utilities and financial services performing best.

5 **Recession** A recession is conventionally defined as two successive quarterly declines in GDP. There is often a large inventory pullback and sometimes a large decline in business investment. Consumer spending on big-ticket items such as cars usually declines (although the US 2001 recession was an exception). Once the recession is confirmed, central banks ease monetary policy, but only cautiously at first. Recessions typically last six months to a year. Both consumer and business confidence decline. Profits drop sharply. In a severe recession, the financial system may be stressed by bad debts, making lenders extremely cautious. Often, recessions are punctuated by major bankruptcies, incidents of uncovered fraud, or a financial crisis. Unemployment can rise quickly, putting downward pressure on inflation.

Capital market effects: Short-term interest rates drop during this phase, as do bond yields. The stock market usually starts to rise in the later stages of the recession, well before the recovery emerges.

Exhibit 15 summarizes the characteristics of the five phases of the business cycle. The description given of business cycles is a stylized one. Each cycle is different because of specific events and trends that fall outside the stylized business cycle framework. Trends that have affected the business cycle from the 1990s through the late 2000s include the growing importance of China in world markets, the aging of populations, and the deregulation of markets. Events such as a petroleum or financial crisis can abruptly take the economy to the next phase of the business cycle or intensify the current phase.

Exhibit 15 Five Phases of the Business Cycle

Phase	Economy	Fiscal and Monetary Policy	Confidence	Capital Markets
1. Initial recovery	Inflation still declining	Stimulatory fiscal policies	Confidence starts to rebound	Short rates low or declining; bond yields bottoming; stock prices strongly rising
2. Early upswing	Healthy economic growth; inflation remains low		Increasing confidence	Short rates moving up; bond yields stable to up slightly; stock prices trending upward
3. Late upswing	Inflation gradually picks up	Policy becomes restrictive	Boom mentality	Short rates rising; bond yields rising; stocks topping out, often volatile
4. Slowdown	Inflation continues to accelerate; inventory correction begins		Confidence drops	Short-term interest rates peaking; bond yields topping out and starting to decline; stocks declining
5. Recession	Production declines; inflation peaks		Confidence weak	Short rates declining; bond yields dropping; stocks bottoming and then starting to rise

EXAMPLE 22

The Yield Curve, Recessions, and Bond Maturity

The yield spread between the 10-year T-bond rate and the 3-month T-bill rate has been found internationally to be a predictor of future growth in output.[58] The observed tendency is for the yield spread to narrow or become negative prior to recessions. Another way of saying the same thing is that the yield curve tends to flatten or become inverted prior to a recession. Effects that may explain a declining yield spread include the following: 1) Future short-term rates are expected to fall, and/or 2) investors' required premium for holding long-term bonds rather than short-term bonds has fallen. At least, the link between an expected decline in short-term rates from expected lower loan demand and declining output growth is economically somewhat intuitive.

When the yield spread is expected to narrow (the yield curve is moving toward inversion), long-duration bonds should outperform short-duration bonds. On the other hand, a widening yield spread (for example, an inverted yield curve moving to an upward-sloping yield curve) favors short-duration bonds.

4.1.3 *Inflation and Deflation in the Business Cycle*

Inflation simply means rising prices, while **deflation** means falling prices. At any given time, some prices are rising and others are falling. Thus, investors look at indexes of prices to discern the overall trend. Consumer price indexes, calculated from a basket of goods and services based on consumers' spending patterns, are commonly watched. Another set of price indexes which are closely watched are the GDP and consumer expenditure deflators, which are inflation indexes used to adjust or deflate the nominal series for inflation.

Inflation is linked to the business cycle, tending to rise in the late stages of a business cycle and to decline during recessions and the early stages of recovery. However, the analyst also needs to note any long-term trends in inflation in formulating capital market expectations.

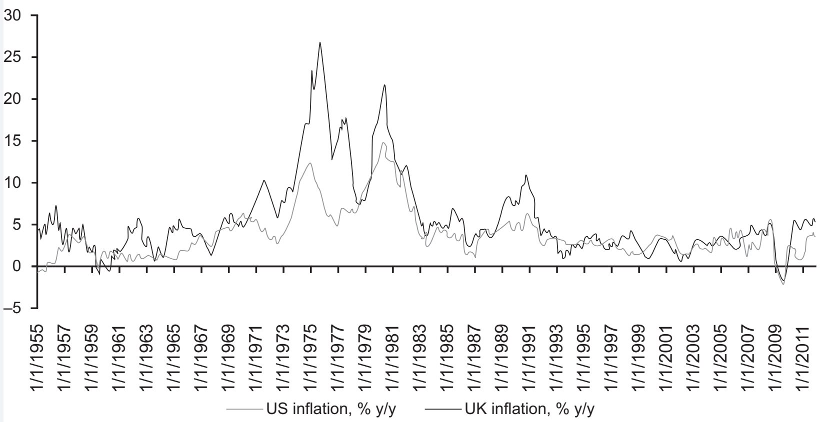
EXAMPLE 23

Inflation, Disinflation, and Deflation

Today, people expect prices of goods and services and of investment assets to trend up over time. However, during most of the 19th century and through the 20th century prior to the 1960s, price inflation was negligible. Indeed, the price level in the United Kingdom fell for a large part of the 19th century. In the United States, the main period of inflation occurred during the Civil War (1861–65). Prices dropped for long periods otherwise. However, from the late 1950s until the late 1970s, inflation gradually accelerated almost everywhere, reaching over 10 percent in the United States and over 30 percent in the United Kingdom for brief periods. Then from about 1979, a period of disinflation set in as inflation gradually retreated back toward zero. Exhibit 16 illustrates the inflation rates in the United States and the United Kingdom since the 1950s.

58 See Estrella and Mishkin (1998).

Exhibit 16 US and UK Inflation over the Long Term



Source: Bloomberg.

Central bank orthodoxy for dealing with inflation rests on three principles:

- Central banks' policy-making decisions must be independent of political influence. If political pressure is brought to bear on central banks, they may be too loose in their monetary policy and allow inflation to gradually accelerate.
- Central banks should have an inflation target, both as a discipline for themselves and as a signal to the markets of their intentions. An inflation target also serves to anchor market expectations.
- Central banks should use monetary policy (primarily interest rates) to control the economy and prevent it from either overheating or languishing in a recession for too long.

By the end of the 20th century, inflation had been defeated almost everywhere. All the major countries enjoyed inflation below 3 percent, and only a handful of emerging countries suffered from double-digit inflation. The challenge is to keep inflation low without succumbing to deflation.

Deflation is a threat to the economy for two main reasons. First, it tends to undermine debt-financed investments. If the price of a debt-financed asset (e.g., new equipment or a house) declines in value, the value of the "equity" in the asset (i.e., the difference between the asset's value and the loan balance) declines at a leveraged rate. For example, if the value of a property financed with a 67 percent loan-to-value mortgage then declines by 5 percent, the value of the equity in the property declines by 15 percent. This phenomenon sometimes leads to panic sales to save some part of the equity and can lead to asset deflation of the kind seen in the United States in the 1930s, the United Kingdom in the early 1990s, and many Asian countries in the late 1990s in the aftermath of the Asian crisis.

Second, deflation undermines the power of central banks. In a deflationary environment, interest rates fall to levels close to zero. When interest rates are already very low, the central bank has less leeway to stimulate the economy by lowering interest rates. A very low interest rate environment during and after the global financial crisis of 2007–2009, accompanied by high unemployment, led central banks including those in the United States, United Kingdom, and Eurozone to engage in a relatively untested policy measure, **quantitative easing** (QE). In QE, a central bank conducts large-scale purchases of high-quality fixed-income securities to inject liquidity into the financial system. Unlike conventional open market operations, QE is designed to provide ongoing and quasi-permanent increases in the level of bank reserves. In addition to government bonds, central banks have purchased mortgage-backed securities and high-quality corporate bonds in their QE operations.

Example 24 looks at some of the considerations that might enter into a short-term inflation forecast.

EXAMPLE 24

An Inflation Forecast for Germany

Early in 2011, Hans Vermaelen, a capital market analyst, has the task of making an inflation forecast for Germany over the next 6 to 12 months. Vermaelen gathers the following inputs and outputs:

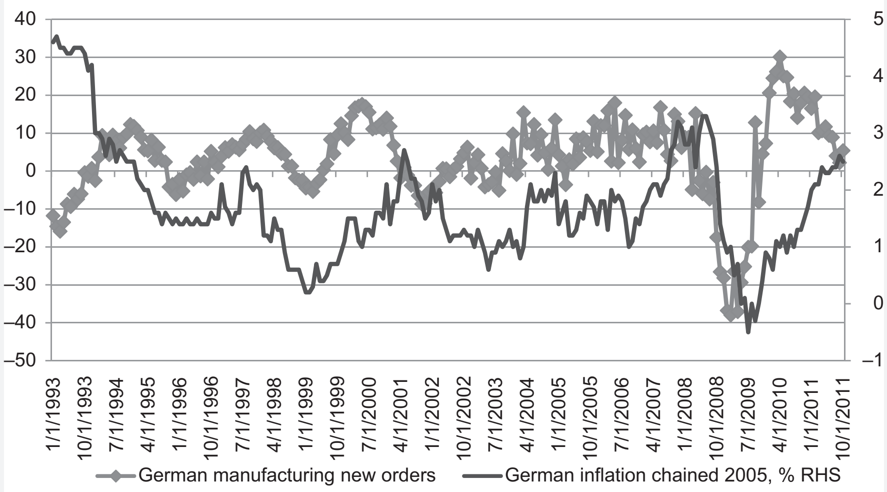
Inputs

1 A survey of manufacturers, asking them whether they expect to see price declines for the products they sell in order to stay globally competitive in light of a then-strengthening euro.
2 Information on German manufacturing orders and consumer price inflation.
3 Data inputs for a multifactor model including the following variables:
   - prices of commodities;
   - prices for labor;
   - wholesale and producer price measures.

Outputs

1 The survey of manufacturers indicates that manufacturers are facing challenges passing some price increases to German customers. After initially lowering export product prices to maintain market share in the face of a rising euro, German manufacturers are now passing price increases on to their international customers, thereby restoring their profit margins.
2 Current year-over-year annual inflation of 2.6 percent is above the average annual rate of 1.6 percent experienced over the past 10 years and over the past 10 quarters. Manufacturing orders have increased at about a 9 percent average rate over the past year. However, over the past quarter, manufacturing orders have decrease at a 4 percent year-over-year rate. Exhibit 17 graphs inflation and manufacturing orders.
3 The multifactor model indicates a positive correlation between the inflation rate and manufacturing orders and a negative correlation between a strengthening local currency (euro) and inflation.

Exhibit 17 Inflation Pressure in Germany (Year over Year Consumer Inflation and Change in Manufacturing Orders)



Based on the above information, Vermaelen forecasts that inflation will decrease to a 1.6 percent rate over the next 6–12 months. Critique Vermaelen's forecast.

Solution:

The manufacturing orders have been decreasing recently at a 4 percent year-over-year rate versus a 9 percent average rate over the past year. This fact suggests that the German economy may be weakening. At the same time, the survey of manufacturers indicates that they are having difficulty passing along price increases to customers. These factors suggest a decrease in inflation from the recent 2.6 percent rate. Overall, it is reasonable to forecast a return to the recent average trend inflation rate of 1.6 percent.

Inflation tends to accelerate in the later stages of the business cycle, when the output gap has been closed. Inflation decelerates when, during a recession or in the early years afterward, there is a large output gap putting downward pressure on prices. As a result, the rate of inflation will decelerate to a low level and **deflation** (an increase in the purchasing power of a unit of currency) becomes possible.[59] Resistance to reduction in wages is a counterweight to deflationary pressures. Except in the worst circumstances, such as in the United States in the early 1930s, the rate of annual deflation is likely to be limited to about 2 percent, with wages holding steady.

During a recession, with falling inflation and interest rates, bonds generally post capital gains (for some bonds, deteriorating credit can offset such gains). In a strong upswing, bond yields will rise as investors fear that central banks will not hold inflation on target, resulting in capital losses to bondholders.

The impact of the inflation cycle on equities is more complex. In theory, as long as inflation stays near its expected or equilibrium level, the inflation rate is not very important. Higher inflation should be reflected in higher profits, so stocks will rise to compensate. However, signs that inflation is moving out of equilibrium indicate

59 The most extreme instances of deflation in the past hundred years occurred in the years surrounding the Great Depression (in particular, in the period 1926–1933).

a potential threat because rising inflation could mean that the central banks need to act to slow the economy. Falling inflation, or possible deflation, is a problem because it threatens a recession and a decline in asset prices.

Exhibit 18 shows how changes in the inflation (deflation) rate affect the relative attractiveness of asset classes.

Exhibit 18 Inflation/Deflation Effects on Asset Classes

	Cash	Bonds	Equity	Real Estate/ Other Real Assets
Inflation at or below expectations	Short-term yields steady or declining. [**Neutral**]	Yield levels maintained; market in equilibrium. [**Neutral**]	Bullish while market in equilibrium state. [**Positive**]	Cash flow steady to rising slightly. Returns equate to long-term average. Market in general equilibrium. [**Neutral**]
Inflation above expectations	Bias toward rising rates. [**Positive**]	Bias toward higher yields due to a higher inflation premium. [**Negative**]	High inflation a negative for financial assets. Less negative for companies/ industries able to pass on inflated costs. [**Negative**]	Asset values increasing; increased cash flows and higher expected returns. [**Positive**]
Deflation	Bias toward 0% short-term rates. [**Negative**]	Purchasing power increasing. Bias toward steady to lower rates (may be offset by increased risk of potential defaults due to falling asset prices). [**Positive**]	Negative wealth effect slows demand. Especially affects asset-intensive, commodity-producing (as opposed to commodity-using), and highly levered companies. [**Negative**]	Cash flows steady to falling. Asset prices face downward pressure. [**Negative**]

4.1.4 *Market Expectations and the Business Cycle*

The description of a typical business cycle may suggest that forming capital market expectations for the short and medium terms is relatively straightforward. If the investor can identify the current phase of the cycle and correctly predict when the next phase will begin, he or she should be able to make money easily. Unfortunately, it is not that simple for several interrelated reasons.

First, the phases of the business cycle vary in length and amplitude. Recessions can be steep, and downturns (such as in the 1930s and, to a lesser extent, the early 1980s) can be frightening. Recessions can be short-lived affairs with only a small decline in output and only a modest rise in unemployment. Sometimes, the weak phase of the cycle does not even involve a recession but merely a period of slower economic growth or a "growth recession." A period of economic growth below trend will open up the output gap. A mild downturn, or growth recession, is most likely if the trend rate of growth of the economy is relatively rapid. For example, China—with a trend rate of annual growth of about 8 percent as of the early 2000s—will see unemployment rise and inflation decline if growth is only 5–6 percent. For the main industrial economies, with trend rates of annual growth of 2–4 percent, a mild downturn is more likely than a recession if some or all of the following conditions hold:

- The upswing was relatively short or mild.
- There was no bubble or severe overheating in the stock market or property market.

- Inflation is relatively low, so the central bank is willing to cut interest rates quickly.
- The world economic and political environments are positive.

EXAMPLE 25

The 1980–1982, 2001, and 2008–09 US Recessions

The US downturn in 1980–1982 was particularly severe. Inflation had reached 12–14 percent in early 1980, partly due to a rise in oil prices. The Board of Governors of the Federal Reserve System under its new chairman, Paul Volcker, was determined to eradicate inflation. The Fed kept interest rates high in 1982. In contrast, the 2001 recession was relatively mild. There had been a stock market bubble, but commercial property prices were not inflated and banks were in good shape. Because inflation was low, the Fed was willing to cut interest rates very rapidly.

The 2001 recession is instructive on the limitations of economic data, which are backward looking and often revised. After the terrorist attack of 11 September 2001 on the World Trade Center, much commentary focused on the risk that it would lead to a recession. In fact, the revised GDP data show that the economy had been in a recession since early in 2001 and began to come out of it starting in October 2001. At the time, it was clear that the economy was weak and therefore growing at less than the trend rate, so bond yields fell and the stock market declined.

In the late 2000s, the collapse of a bubble (speculative run-up) in housing prices, along with a crisis in subprime mortgages, led to a global financial crisis and recessions in most developed countries. The US recession saw a peak-to-trough decline in GDP of 5.1 percent, marking the most severe of post World War II US recessions. A low inflationary environment permitted the Fed to pursue expansionary monetary policy, though with limited impact on economic growth and unemployment.

4.1.5 *Evaluating Factors that Affect the Business Cycle*

For the purposes of setting capital market expectations, we need to focus business cycle analysis on four areas:

- consumers;
- business;
- foreign trade; and
- government activity, both **monetary policy** (concerning interest rates and the money supply) and **fiscal policy** (concerning taxation and governmental spending).

Consumer spending amounts to 60–70 percent of GDP in most large developed economies and is therefore typically the most important business cycle factor.

Business investment has a smaller weight in GDP than consumer spending but is more volatile.

Foreign trade is an important component in many smaller economies, for which trade is often 30–50 percent of GDP. However, for the large economies, such as the United States and Japan, foreign trade is typically only around 10–15 percent of GDP and correspondingly less important. The same range holds true for the European Union (EU) in relation to trade outside the EU (although between countries within the EU, trade generally represents a higher percentage of GDP).

Finally, government policy can influence the business cycle. There are three motivations for government to intervene in the cycle. First, both the government and monetary authorities may try to control the cycle to mitigate severe recessions and also, on occasion, to moderate economic booms. Second, the central bank monetary authorities often have an inflation target and they will consciously try to stimulate or constrain the economy to help meet it. Third, because incumbent politicians prefer to hold elections during economic upswings, they may try to influence fiscal and/or monetary policy to achieve this end.

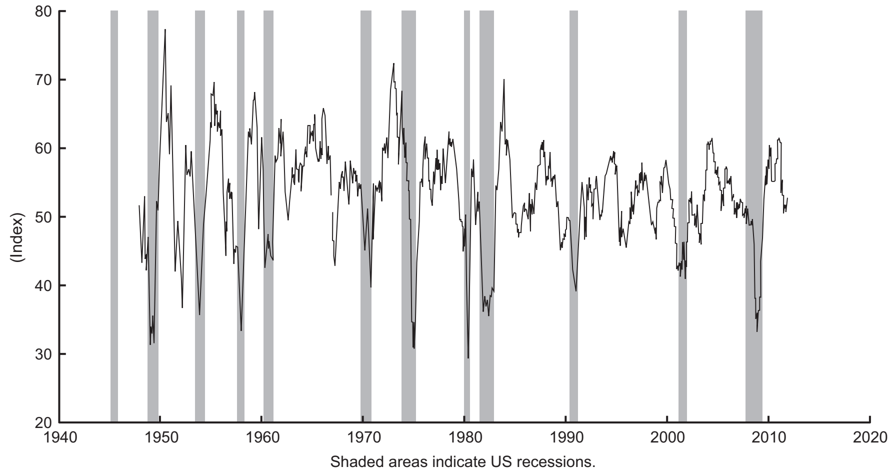
4.1.5.1 Taking the Pulse of Consumers The principal sources of data on consumer spending are retail sales, miscellaneous store sales data, and consumer consumption data. Like most data, consumer spending can be erratic from month to month and can be affected by unusual weather or holidays (such as New Year celebrations).

By far the most important factor affecting consumption is consumer income after tax, which depends on wage settlements, inflation, tax changes, and employment growth. Employment growth is often closely watched because data are usually available on a very timely basis. Most countries have some particular series that analysts scrutinize. In the United Kingdom, besides the unemployment rate, the British Retail Consortium (BRC) retail sales survey is closely watched. In the United States, the monthly non-farm payrolls as well as the weekly new unemployment claims are regular market movers when they diverge from expectations.

If the household savings rate remained constant, then changes in income would exactly predict changes in spending. But the savings rate does change over time, influenced generally by consumer confidence in future jobs and income and also by changes in asset prices. Consumer confidence survey data are also watched closely as indicators of whether consumers are likely to raise or lower their savings rates.

4.1.5.2 Taking the Pulse of Business Data on business investment and spending on inventories reveal recent business activity. As already mentioned, both tend to be relatively volatile so that it is not uncommon for business investment to fall by 10–20 percent or more during a recession and to increase by a similar amount during strong economic upswings. Data for inventories need careful interpretation, however. A report of rising inventories may mean that businesses are very confident of sales and are spending on inventories ahead of expected sales. This would normally be the case in the early stages of an inventory cycle upswing and is bullish for economic growth. But at the late stage of the inventory cycle, a rise in inventories may be involuntary because sales are lower than expected. Such news would be negative.

Exhibit 19 US ISM Manufacturing PMI



Source: FRED, Federal Reserve Economic Data, Federal Reserve Bank of St. Louis: ISM Manufacturing: PMI Composite Index (NAPM); Institute for Supply Management. http://research.stlouisfed.org/fred2/series/NAPM; accessed 21 December 2011.

Some of the most useful data on business are surveys. A particularly useful one is the purchasing managers index (PMI) published for several decades by the Institute of Supply Management (ISM) in the United States (formerly the National Association of Purchasing Managers). The PMI is one of the best indicators of the US economy, and most developed countries have developed PMIs using a similar methodology. In the 1990s, the ISM added a survey of non-manufacturing companies, which is beginning to acquire a useful track record. Exhibit 19 presents US manufacturing PMI data for the years 1948 to 2011. The lower points on the graph coincide predominantly with shaded areas that indicate recessionary periods, when manufacturing activity declines.

The PMI is based on answers to a series of questions about the company's position, including production plans, inventories, prices paid, prices received, and hiring plans. Each component is reported as well as an overall index. The indexes are calibrated so that 50 should be the breakeven point for manufacturing growth. These surveys are particularly useful because of their timeliness.

4.1.5.3 Monetary Policy Monetary policy is sometimes used as a mechanism for intervention in the business cycle. For example, monetary policymakers may use stimulative measures (increasing money supply growth and/or lowering short-term interest rates) when the economy is weak and use restrictive measures (decreasing money supply growth and/or raising short-term interest rates) when the economy is in danger of overheating. If unemployment is relatively high and there is excess productive capacity, then a rate of GDP growth higher than the trend rate will be tolerated for a period of time. This scenario is typical of the recovery and early upswing phases of the business cycle. In the late upswing phase, the economy is threatening to overheat and monetary authorities will restrict money supply to slow growth. If they get it wrong and a recession emerges, then they may cut interest rates to restore growth. Finally, if a major financial crisis threatens the financial system, they may also cut interest rates and increase liquidity and availability of credit, as was seen in the United States in 1987, 1998, 2001, and following the global financial crisis of 2007–2009.

The key variables watched by monetary authorities include:

- the level of unemployment;
- the rate of inflation;

- the pace of economic growth; and
- the amount of excess capacity still available (if any);

A common means for the largest central banks to effect monetary policy has been setting short-term interest rates to levels that are meant to control inflation without inhibiting economic growth. Central banks often see their role as smoothing out the growth rate of the economy to keep it near its long-term sustainable trend rate—in effect, neither too hot nor too cold. A change in short-term interest rates affects the economy through a number of different mechanisms, which vary in their effects at different times.

Lower interest rates encourage more borrowing by consumers and businesses. Lower interest rates also usually result in higher bond and stock prices. These in turn encourage consumers to spend more and encourage businesses to invest more. From an international trade perspective, lower interest rates usually lower the value of a currency and therefore stimulate exports.

The effect of a cut in interest rates also depends on the absolute level of interest rates, not just the direction of change. For example, suppose that interest rates have been raised from 3 to 6 percent to address inflation concerns and then, in response to a recession, are lowered to 4 percent. The lowering of interest rates might stimulate the economy, but interest rates are still higher than where they started. In other words, what matters is not just whether interest rates have most recently been moved up or down but where they stand in relation to their average or "neutral" level. It is common to think of this "neutral" level as a point of interest rate equilibrium within the economy. The concept of the neutral level of interest rates is an important one, though in reality, it is impossible to identify precisely. Conceptually, the argument is that a neutral level of short-term interest rates should include a component to cover inflation and a real rate of return component. For example, in the United States, if inflation is targeted at 2 percent and the economy is growing at 2 percent, many economists argue that the neutral level of interest rates is about 4 percent.

The Taylor Rule One way to assess the central bank's stance and to predict changes is through the so-called **Taylor rule**.[60] In essence, this rule links a central bank's target short-term interest rate to the rate of growth of the economy and inflation. A simple approach to this rule (giving equal weights to GDP growth and inflation) is given by the following Taylor rule equation:

$$R_{\text{optimal}} = R_{\text{neutral}} + \left[0.5 \times \left(\text{GDPg}_{\text{forecast}} - \text{GDPg}_{\text{trend}}\right) + 0.5 \times \left(I_{\text{forecast}} - I_{\text{target}}\right)\right] \tag{12}$$

where

R_{optimal} = the target for the short-term interest rate
R_{neutral} = the short-term interest rate that would be targeted if GDP growth were on trend and inflation on target
$\text{GDPg}_{\text{forecast}}$ = the GDP forecast growth rate
$\text{GDPg}_{\text{trend}}$ = the observed GDP trend growth rate
I_{forecast} = the forecast inflation rate
I_{target} = the target inflation rate

The Taylor rule gives the optimal short-term interest rate as the neutral rate plus an amount that is positively related to the excess of the forecasts of GDP and inflation growth rates above their trend or target values. For example, assume that a current

60 See Taylor (1993).

short-term interest rate of 4 percent is the neutral rate. Thus, if the United States is forecast to achieve its 2 percent trend rate of growth and 2 percent inflation target, then the Fed would be happy with the federal funds rate at the neutral rate of 4 percent. At 4 percent, the Fed would expect that the GDP growth and inflation rates would remain at trend or targeted levels. The Taylor rule then states that if the forecast GDP growth rate and/or the forecast inflation rate are above the trend or target level, short-term interest rates need to be raised by *half* the difference between the forecast and the trend or target. Conversely, GDP growth and/or inflation rates below trend or target would motivate the Fed to lower the fed funds rate. (The **federal funds rate**, or fed funds rate for short, is the interest rate on overnight loans of reserves [deposits] at the Fed between Federal Reserve System member banks.)[61] The belief is that when GDP forecast growth is below trend, lowering the interest rate will stimulate output by lowering corporations' cost of capital. When forecast inflation is below target, lowering the interest rate is expected to help the inflation rate return to target through its stimulative effect on the money supply.

EXAMPLE 26

A Taylor Rule Calculation

Assume the following scenario:

- The neutral value of the short-term interest rate is 3.5 percent.
- The inflation target is 2.5 percent.
- The GDP trend rate of growth is 3 percent.

If the inflation forecast is 4 percent and the forecast for GDP growth is 1 percent, what is the optimal short-term interest rate?

Solution:

According to the Taylor rule,

$$\begin{aligned} R_{\text{optimal}} &= R_{\text{neutral}} + \left[0.5 \times \left(\text{GDP}g_{\text{forecast}} - \text{GDP}g_{\text{trend}}\right) + 0.5 \times \left(I_{\text{forecast}} - I_{\text{target}}\right)\right] \\ &= 3.5\% + \left[0.5(1.0\% - 3.0\%) + 0.5(4.0\% - 2.5\%)\right] \\ &= 3.5\% + (-1.0\% + 0.75\%) \\ &= 3.5\% - 0.25\% \\ &= 3.25\% \end{aligned}$$

The GDP growth forecast by itself implies that the short-term interest rate should be lowered by 1 percentage point, because GDP growth is under trend. Partially offsetting the effect of below-trend GDP growth is the interest rate increase implied by above-target inflation. Net, the Taylor rule implies that the central bank should lower short-term rates by 25 bps to 3.25 percent.

Historically, the Taylor rule has provided a reasonably accurate description of central banks' behavior.

Money Supply Trends Trends in the money supply can be a good indicator of monetary conditions and of the trend of the economy. Over the long run, there has been a reasonably stable relationship between the growth in money supply and the growth

61 According to US law, Federal Reserve System member banks are required to hold reserves at the Fed equal to a fraction of the deposits with the banks. In the Eurozone, banks have broadly similar requirements that relate to holding reserves at national central banks.

in nominal GDP (inflation plus real growth). If money growth is particularly strong in relation to nominal GDP, chances are that growth will accelerate in the near future and that inflation may eventually accelerate.

What Happens When Interest Rates Are Zero or Negative? Prior to the 2007–2009 global financial crisis, it was generally accepted that central banks could not successfully implement negative interest rate policies. The belief in a "zero lower bound" on policy rates was based on an assumption that individuals would choose to hold currency (coins and notes) if faced with earning a negative interest rate on short-term instruments, including deposits. The move towards holding currency would drain deposits and reserves from the banking system, causing bank balance sheets to shrink. The resulting credit contraction would put upward pressure on interest rates, thwarting the central bank's attempt to maintain negative rates. The contraction of credit would likely also put additional downward pressure on economic growth, thereby reinforcing the need for stimulative policies.

This line of reasoning raised questions about the effectiveness of traditional monetary policy when the economy is so weak that economic growth fails to respond to (nominal) interest rates approaching zero. Following the global financial crisis, central banks faced with this situation pursued less conventional measures.

One important measure was quantitative easing (QE) in which the central banks committed to large-scale, ongoing purchases of high-quality domestic fixed-income securities. These purchases were funded by creating an equally large quantity of bank reserves in the form of central bank deposits. As a result of QE, central bank balance sheets and bank reserves grew significantly and sovereign bond yields fell. QE was pursued by (among others) the US Federal Reserve, the European Central Bank, the Bank of Japan, and the Bank of England.

Conventional reasoning suggests that QE should have resulted in the desired growth in nominal spending. In theory, banks could use the increased reserves to extend loans and low interest rates would stimulate businesses and households to borrow. The borrowing was expected to fund capital expenditure by businesses and current consumption and purchases of durables (e.g., houses and cars) by households, thereby stimulating the economy. With interest rates low, investors were expected to bid up the prices of stocks and real estate. Although asset prices did increase and businesses that could issue bonds borrowed heavily, proceeds were more often used to fund dividends and stock buybacks rather than capital expenditures. At the same time, household spending ability was significantly curtailed by the legacy of the global financial crisis.

Whether or not QE was effective remains subject to debate. To achieve desired levels of economic growth, central banks began to try the previously unthinkable: targeting negative interest rates. The central banks of Denmark, Sweden, Japan, Switzerland, and the euro area are among those that adopted negative policy rates. Contrary to the notion of a "zero lower bound," negative policy rates have proven, as of the beginning of 2017, to be sustainable.

The move into currency did not occur as expected because the scale and speed of transactions inherent in modern economies today cannot be supported using cash as the primary method of exchange. Trillions of dollars change hands daily to facilitate trade in goods, services, and financial instruments. These transactions cannot be accomplished using physical cash. Bank deposits and bank reserves held at the central bank, rather than as vault cash, have an implicit yield or convenience value that cash does not. As long as this value exceeds the explicit cost of holding those deposits—in the form of a negative interest rate—there is no incentive to convert deposits into cash. In such circumstances, negative policy rates may be achievable and sustainable.

In theory, using negative nominal rates to stimulate an economy should work similarly to using low but still positive rates. Businesses and consumers are encouraged to hold fewer deposits for transaction purposes; investors are encouraged to seek higher expected returns on other assets; consumers are encouraged to save less and/or borrow more against future income; businesses are encouraged to invest in profitable projects; and banks are encouraged to use their reserves in support of larger loan books. All of this is expected to stimulate economic growth.

However, for consumers, investors, businesses, and banks to behave as above, each must believe they will be adequately rewarded for taking the inherent risks. In a negative interest rate environment, these entities are likely to have greater levels of uncertainty as to whether they will be adequately compensated for risks taken, and therefore may not act as desired by monetary policy makers. As a result, the effectiveness of expansionary monetary policy is more tenuous at low and negative interest rate levels than at higher interest rate levels.

EXAMPLE 27

Monetary Policy in the Eurozone Compared with the United States in 2001 and 2010

Both Europe and the United States saw a sharp economic slowdown in 2001. The Fed responded by cutting the fed funds rate from 6.50 percent to 1.75 percent during 2001. In contrast, the European Central Bank (ECB), the central bank for the Eurozone, cut interest rates from 4.75 percent to 3.50 percent, a much less aggressive move. The reasons for these different responses were twofold:

- In 2001, US CPI inflation stood at 2.6 percent, well within the Fed's likely informal target range of 1–3 percent inflation. Coincidentally, the Eurozone also had inflation of 2.6 percent in mid-2001, but this rate was above the explicit target range of 0–2 percent.
- In the United States, unemployment rose rapidly during 2001 from about 4 percent to nearly 6 percent, opening an output gap in the economy. In contrast, Eurozone unemployment was constant at 8 percent for most of 2001, rising only slightly at the end of the year. Hence, the ECB welcomed the slowdown as a way to put downward pressure on inflation.
- In response to the financial crisis and ensuing recession in the period of 2007–2009, both the ECB and the Fed acted in a coordinated manner by first flooding their respective markets with liquidity and then cutting interest rates to or close to zero. To further address stubborn stagnation in economic growth and high unemployment, the Fed and ECB resorted to such non-traditional measures as implementing bond buying programs which came to be known as Quantitative Easing I and II (QE I and QE II).

Implications of Negative Interest Rates for Capital Market Expectation Long-run capital market expectations typically take the level of the "risk-free rate" as a baseline to which various risk premiums are added to arrive at long-run expected returns for risky assets such as long-term bonds and equities. The implicit assumption is that the risk-free rate is at its long-term equilibrium level. When short-term rates are negative, rather than using the observed negative rate, the long-run equilibrium short-term rate can instead be used as the baseline rate in these models. This rate can be estimated using the neutral policy rate ($R_{neutral}$) in the Taylor rule (or more generally in the central bank's presumed reaction function) adjusted for a modest spread between policy rates and default-free rates available to investors.

In forming capital market expectations for shorter time horizons, consideration must be given to the expected path of interest rates. Paths should be considered that, on average, converge to the long-run equilibrium rate estimate. With negative policy rates in place, this means a negative starting point. In theory, many possible scenarios, each appropriately weighted by its likelihood, should be considered. In practice, it may suffice to consider only a few scenarios. Because shorter horizons provide less opportunity for the impact of events to average out, the shorter the forecast horizon the more important it is to consider deviations from the most likely path.

Negative policy rates are expected to produce asset class returns similar to those occurring in the contraction and early recovery phrases of a "more normal" business/policy cycle. While such historical periods may provide a reasonable starting point in formulating appropriate scenarios, it is important to note that negative rate periods may indicate severe distress in the economy and thus involve greater uncertainty regarding the timing and strength of recovery.

Key considerations when forming capital market expectations in a negative interest rate environment include the following:

- Historical data are less likely to be reliable.
    - Useful data may exist on only a few historical business cycles, which may not include instances of negative rates. In addition, fundamental structural/institutional changes in markets and the economy may have occurred since this data was generated.
    - Quantitative models, especially statistical models, tend to break down in situations which differ from those on which they were estimated/calibrated.
    - Forecasting must account for differences between the current environment and historical averages. Historical averages, which average out differences across phases of the cycle, will be less useful.
- The effects of other monetary policy measures occurring simultaneously (e.g., quantitative easing) may distort market relationships such as the shape of the yield curve or the performance of specific sectors.

Incorporating uncertain dynamics including negative interest rates into capital market expectations over finite horizons is much more difficult than projecting long term average levels. As noted, this is because asset prices depend not only on investor expectations regarding longer term "equilibrium" levels, but also on the path taken to get there.

4.1.5.4 Fiscal Policy **Fiscal policy** means manipulating the budget deficit to influence the economy. Governments increase spending or cut taxes to stimulate the economy and cut spending or raise taxes to slow the economy. In analyzing fiscal policy, or the so-called "fiscal stance," it is crucial to remember two points. First, an analyst should focus on the *changes* in the government budgetary deficit, not its level. For example, although the Japanese budget deficit has been running at around 8 percent of GDP for many years (as of 2005), it has not been a continuous stimulus to the economy. But if the deficit rose to 10 percent, that increase could represent a stimulus. Conversely, a reduction in the deficit would represent a tighter policy.

Second, it is only changes in the deficit due to *deliberate changes* in government fiscal policy that matter. The budget deficit will constantly change in response to the economy without any change in fiscal policy. During recessions, the deficit tends to rise because tax revenues fall and government spending on unemployment benefits increases. In contrast, when the economy grows strongly, the budget deficit naturally falls.

Linkages with Monetary Policy It is useful to consider the overall mix of fiscal and monetary policy. If fiscal and monetary policies are both tight, then the situation is unambiguous and the economy is certain to slow. Similarly, if both monetary policy and fiscal policy are expansionary, then the economy can be expected to grow. However, monetary and fiscal policies are sometimes at odds with one another. These situations create opportunities for investors as well as risks.

The fiscal/monetary mix usually shows up in the shape of the yield curve. Exhibit 20 illustrates the four possibilities. When both fiscal and monetary policies are loose, the yield curve tends to be steeply upward sloping. When fiscal policy is tightened while monetary policy remains loose, bond yields tend to fall and the yield curve comes back to a more moderate upward slope.

Exhibit 20 Policy Mix and the Yield Curve

		Fiscal Policy	
		Loose	**Tight**
Monetary Policy	**Loose**	Yield curve steep	Yield curve moderately steep
	Tight	Yield curve flat	Yield curve inverted

If monetary and fiscal policies are both tight, the yield curve is typically inverted. Finally, when monetary policy is tight but fiscal policy is loose, the yield curve tends to be flat.

4.2 Economic Growth Trends

The economic growth trend is the long-term growth path of GDP. The long-term growth path reflects the average growth rate around which the economy cycles. The differences between economic trends and cycles need to be understood. Economic trends exist independently of the cycle but are related to it. Business cycles take the economy through an alternating sequence of slow and fast growth, often including recessions and economic booms.

Economists are concerned with a variety of trends besides the economic growth trend, because that trend is determined by other economic trends, such as population growth and demographics, business investment and productivity, government structural policies, inflation/deflation, and the health of banking/lending processes.

Trends are more easily forecast than cycles, but there are always uncertainties. In practice, it is often difficult to know which trends are most important. Moreover, some trends or changes in trends are by definition not open to forecasting. These are often called "shocks." Examples include wars that cause market dislocations, abrupt changes in government tax or trade policies, and the sudden collapse in an asset market or in an exchange rate. Often, these abrupt changes in trend affect the overall paradigm of capital market expectations. One example of a paradigm-changing shock was the revelation of accounting irregularities at Enron and other US companies in 2002. Investors' perceptions both of the reliability of companies' earning statements and of corporate leaders' attitudes profoundly shifted. Changes in regulations reinforced these shifts. In contrast, other trends such as demographics are usually very much in the background because they change only very gradually.

While shocks are not forecastable, investors do try to assess the risk that they will occur. Unrest in the Middle East may push up the price of oil as well as "safe haven" investments such as gold, the Swiss franc, and US government bonds. If a particular tax change is being considered, markets may partially anticipate it in the pricing of such assets. Some events do come unexpectedly, and they are the ones likely to have the greatest impact as investors struggle to understand their implications.

EXAMPLE 28

Cycles and Trends: An Example

Consider the following hypothetical passage describing the German economy in late 2010:

- *After a recession in 2009 and stagnation in the first quarter of 2010, the German economy picked up. Starting in the second quarter of 2010, it grew at 2.1 percent annualized. Exports led the way, and business investment picked up. Consumer spending grew strongly in early 2010.*
- *Significant progress on labor market reforms and pension reforms by the government, as well as increased sales to India and China, boosted confidence. R & D spending increases led to significant export growth which was further assisted by a weak Euro. In response, the Bundes Bank increased the economic growth forecast from 1.9% to 3.0%, which was above the projected trend growth rate of 2.5%.*

These two statements contain information about the economy. The first refers to the business cycle, while the second describes other economic trends. The final sentence is a mixture of cycle and trend information. It provides the government forecast for economic growth in the following year (cyclical information) but also implies an estimate for the long-term average rate of growth that the German economy is believed capable of achieving (2.5 percent per year).

The expected trend rate of economic growth is a key input in discounted cash flow models of expected return. First, a country with a higher trend rate of growth may offer equity investors a particularly good return. Second, a higher trend rate of growth in the economy allows actual growth to be faster before there is a danger of inflation.

The trend rate of growth of the economy is usually thought not to change much over time. Indeed, for the United Kingdom, historically the first industrial economy, it would appear that GDP has had a 2–2.5 percent trend growth rate for two hundred years with very little variation. However, most countries have had periods of faster and slower trend growth during their development. Emerging countries naturally can more easily have faster growth as they catch up with the leading industrial countries. But the more developed they become, the more likely it is that their growth will slow. This effect has been very obvious in the case of Japan. After Japan's GDP grew at an annual average rate of 11 percent in the 20 years leading up to 1973, growth in the next 17 years averaged only 3.9 percent and then fell to 1.6 percent between 1990 and 2003.

4.2.1 *Consumer Impacts: Consumption and Demand*

Consumers can be counted upon as the largest source of aggregate economic growth in both developed and developing economies. It is interesting to note that although consumers spend more in response to perceived increases in their wealth due to a "wealth effect," overall consumer consumption is quite stable over the business cycle. Milton Friedman (1957) developed an explanation for this stability in his permanent income hypothesis. The **permanent income hypothesis** asserts that consumers' spending behavior is largely determined by their long-run income expectations. Temporary or unexpected (or one-time) events such as benefiting from an inheritance might temporarily increase an individual's demand for items that might not ordinarily be purchased, but overall spending patterns remain largely determined by long-run expectations. However, if an unexpected event (e.g., winning the lottery) produced an ongoing series of incoming cash flows, it would be expected to permanently alter an individual's spending patterns since the flows would be ongoing and could be depended upon over the long term.

In similar fashion, when temporary events reduce the income flows of consumers, individuals typically reduce the amount they save to maintain their long-term spending patterns. Only when income disruptions occur over the long term may individuals capitulate and reduce their consumption—out of necessity. Thus, consumer trends are usually stable or even countercyclical over a business cycle. When incomes rise the most (during the cyclical expansion phase), spending increases less than income rises. When incomes fall as an economy's growth slows or declines, consumption falls only a fraction and usually only for a relatively short period of time.

4.2.2 *A Decomposition of GDP Growth and Its Use in Forecasting*

The simplest way to analyze an economy's aggregate trend growth is to split it into

- growth from changes in employment (growth from labor inputs), and
- growth from changes in labor productivity.

For longer-term analysis, growth from changes in employment is broken down further into growth in the size of the potential labor force and growth in the actual labor force participation rate (e.g., more or fewer women or older people working; "growth" can be positive or negative).

For example, annual US GDP growth used to be thought of as likely to average 2.5 percent over the long term. This number results from adding a 1 percent growth in the potential labor force, a 0.5 percent growth in labor force participation, and a 1 percent annual growth in labor productivity.

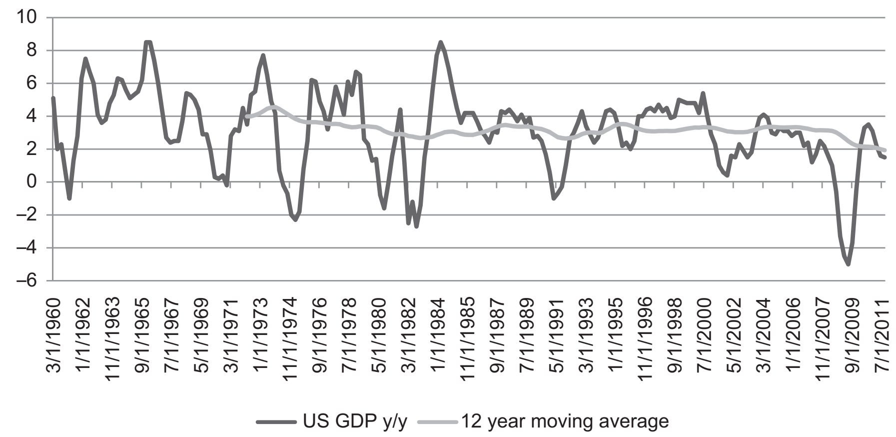
During the late 1990s, there were signs that productivity had risen perhaps to around 2.5 percent annually, so estimates of trend growth were raised to about 4 percent or somewhat less. In contrast, the figures for many developing countries would be closer to 2 percent potential labor force growth, 1 percent growth in labor force participation, and 3–4 percent growth in labor productivity, suggesting that annual growth could average 6–7 percent over a long period. Exhibit 21 shows the US economic trend growth since 1960.

A more sophisticated approach to economic trend growth estimation breaks down the growth in labor productivity into two components, just as growth from labor inputs was analyzed into two components. Productivity increases come from investment in equipment or new machines (growth from capital inputs) and from **growth in total factor productivity** (TFP growth), known also as technical progress and resulting from increased efficiency in using capital inputs.

To summarize, with this approach, the trend growth in GDP is approximately the sum of the following:

- growth from labor inputs, comprising
  - growth in potential labor force size and
  - growth in actual labor force participation, plus
- growth from labor productivity, comprising
  - growth from capital inputs and
  - TFP growth (i.e., growth from increase in the productivity in using capital inputs).

Exhibit 21 US Business Cycles and Economic Trend Growth



Source: Bloomberg.

The potential sources of TFP growth include technological shocks and shifts in government policies. In historical analyses, TFP is often taken simply as a "residual"—that is, output growth that is not accounted for by the other factors.[62]

Many fast-growing emerging countries are successful because they invest heavily and therefore quickly build capital. In Singapore and China, for example, between 30 percent and 40 percent of GDP is invested annually. Slower-growing countries in South America have typically been able to manage capital investment rates of only 15–20 percent of GDP. It is likely, therefore, that the relatively fast rates of economic growth in Asia owe much to the higher rates of capital investment.

The rapid rate of investment helps explain why stock market returns may not be strong despite the rapid growth in the economy. Ultimately, stock market returns depend on the rate of return on invested capital. If capital is growing fast, returns on invested capital are driven down.

Future economic growth trends can be forecast using the model just given. For example, economic trend growth rates in Japan and many parts of Europe are forecast to be relatively low over the next few decades because labor force growth will be constrained by slow population growth. In contrast, the United States—with its relatively young population and high immigration—should enjoy faster labor force growth. Europe and Japan could, however, change the outlook if they could achieve a higher labor force participation rate. Changes to employment laws, pension entitlements, and child care facilities could encourage more women and older people to enter the workforce.

Trend growth will also be boosted if investment is stronger. For example, the surge in economic growth in the United States in the 1990s was partly linked to higher investment. The combination of an economic boom, higher stock market valuations, and a high level of investment in new opportunities in computers and networking in turn boosted overall productivity. Opinions still differ as to how much of the increase in productivity was due to "more machines" and how much was due to greater TFP. Part of the problem is how to value the machines, since computer prices have been falling sharply while their power increases.

62 This component is known as the Solow residual estimate of TFP growth.

By 2005, Europe had not seen a comparable surge in productivity. Investment has not been as strong, and various rigidities seem to be in the way of raising TFP. For example, in the United Kingdom, planning (zoning) restrictions on new large shops on the scale of Wal-Mart in the United States may have limited the scope of retailing efficiencies. Continental European labor laws, which restrict redundancies, or layoffs, also may have made companies move relatively slowly in "de-layering" bureaucracies by using the advantages of networking.

EXAMPLE 29

Forecasting GDP Trend Growth

Cynthia Casey is reviewing a consultant's forecast that Canadian GDP will grow at a long-term 3.5 percent annual rate. According to Casey's own research, a 3.2 percent growth rate is more realistic. Casey and the consultant agree on the following assumptions:

- The size of the Canadian labor force will grow at 1 percent per year based on population projections.
- Labor force participation will grow at 0.25 percent per year.
- Growth from capital inputs will be 1.5 percent per year.

Determine the reason for the discrepancy between Casey's forecast and the consultant's forecast.

Solution:

Casey and the consultant agree on three of the four components of GDP growth, so the reason for the discrepancy in their GDP growth forecasts must be disagreement about the value of the fourth component, total factor productivity growth. For Casey to arrive at a 3.2 percent growth rate estimate, she must be assuming that total factor productivity growth will be 3.2% – (1% + 0.25% + 1.5%) = 0.45%. By contrast, the consultant is predicting that total factor productivity growth will be 3.5% – (1% + 0.25% + 1.5%) = 0.75%. Thus, the consultant is more optimistic than Casey about GDP growth from increases in the productivity in using capital inputs.

4.2.3 *Government Structural Policies*

Government structural policies refer to government policies that affect the limits of economic growth and incentives within the private sector. Government policies affect economic growth trends in profound ways. In the first three-quarters of the 20th century, governments increasingly intervened in the economy in most countries. This intervention often took the form of outright government ownership of large enterprises, combined with labor market and product market regulation. Starting in the 1980s, the trend to privatization led by former Prime Minister Thatcher in Britain substantially reduced the amount of government ownership in most economies. However, the trend toward heavy government regulation of the economy, other than through direct ownership, remains a powerful one.

The following are elements of a pro-growth government structural policy:

1 **Fiscal policy is sound**. Fiscal policy is sometimes used to influence the business cycle and can play a useful role. For example, decreasing a budget surplus (or increasing a budget deficit) may be a justifiable economic stimulus during a recession. But countries that regularly run large deficits tend to have one or more of three potential problems. First, a government budget deficit often

brings a current account deficit (the so-called "twin deficits" problem), which means that the country must borrow abroad. Eventually, when the level of foreign debt becomes too high, that borrowing must be scaled back. This usually requires a large and potentially destabilizing devaluation of the currency. Second, if the deficit is not financed by borrowing, it will ultimately be financed by printing money, which means higher inflation. Third, the financing of the deficit takes resources away from private sector investment, and private sector investment is usually more productive for the country as a whole. It is for all these reasons that investors prefer to see governments hold the budget deficit close to zero over the long term.

2 **The public sector intrudes minimally on the private sector**. According to economic theory, a completely unfettered competitive market would probably supply too little in the way of public goods, such as national defense, and too much in the way of goods with negative externalities, such as goods whose manufacture pollutes the environment.[63] However, the thrust of economic theory is that the marketplace usually provides the right incentives to individuals and businesses and leads to an efficient allocation of scarce resources. Recognizing this, many countries have privatized government-owned businesses over the last few decades and reduced regulations affecting business. The most damaging regulations for business tend to be labor market rules (e.g., restricting hiring and firing) because such regulations tend to raise the **structural level of unemployment** (the level of unemployment resulting from scarcity of a factor of production); however, such regulations are also the most difficult to lift.

3 **Competition within the private sector is encouraged**. Competition is important for trend growth because it drives companies to be more efficient and therefore boosts productivity growth. In the last several decades, the reduction of trade tariffs and barriers has been very important in increasing competition in the goods sector. Advances in networking technologies have spread that competition into the service sector. Another positive government policy is openness to foreign investment. However, note that competition makes it more difficult for companies to earn high returns on capital and thus can work against high stock market valuations.

4 **Infrastructure and human capital development are supported**. Projects supporting these goals may be in partnership with the private sector. Building health and education infrastructure has important economic benefits.

5 **Tax policies are sound**. Governments provide a range of goods, including defense, schools, hospitals, and the legal system. They also engage in a certain amount of redistribution of income directly, through pensions and welfare programs. As a result, developed country governments typically collect between 30 percent and 50 percent of GDP in taxes. According to economic theory, taxes distort economic activity by reducing the equilibrium quantities of goods and services exchanged. A decrease in total societal income and efficiency is the cost of redistributing wealth to the least well-off. As a result, investors often look with skepticism on governments that impose high overall tax burdens. Sound tax policy involves simple, transparent, and rarely altered tax rates; low marginal tax rates; and a very broad tax base.

63 A **public good** is a good that is not divisible and not excludable (a consumer cannot be denied it); because of these properties, a public good cannot be priced or traded. An **externality** is a result of a transaction or process that spills over to the public.

4.3 Exogenous Shocks

Exogenous shocks are events from outside the economic system that affect its course. These could be short-lived political events, changes in government policy, or natural disasters, for example. How do shocks contrast with economic trends? Over time, trends in an economy are likely to stay relatively constant. As such, they should already be discounted in market expectations and prices. Exogenous shocks may have short-lived effects or drive changes in trends. They are typically not built into prices or at most are only partially anticipated.

Most shifts in trends are likely to come from shifts in government policies, which is why investors closely watch both specific measures and the overall direction of government policy (e.g., consumer friendly, business friendly, export friendly).

The biggest impact occurs when there is new government or a major institutional shift. For example, a major fiscal law that prevents the government from borrowing beyond certain limits can be a very effective constraint on excessive spending. A decision to make the central bank more independent or to enter a currency union could have a major impact on the economy. Such government-induced impacts typically are swiftly felt.

Some shocks do not affect trends but are felt in a more immediate or short-term manner. While they are often negative, they are not always so. In 1986, the unexpected breakup of an OPEC meeting without an agreement to cut production led to a sharp decline in oil prices. This event played an important role in keeping inflation low for several years after that. The fall of the Berlin wall triggered German reunification and a "peace dividend" for governments as they cut defense spending.

The creation and assimilation of new products, markets, and technologies provide a positive, longer-term impact on economic trends. Too often, analysts focus on shorter-term benefits, under-appreciating the evolving nature of the technologies and the scope of their effects. For example, the evolution of communication technologies from the telegraph, telephone, phonograph, wireless (cellular and satellite), and internet has been a source of great positive economic impact. These gains show up in TFP growth.

Shocks cannot be forecast in general. But there are two types of economic shock that periodically affect the world economy and often involve a degree of contagion, as problems in one country spread to another. Oil shocks are important because a sharp rise in the price of oil reduces consumer purchasing power and also threatens higher inflation. Financial shocks, which can arise for a variety of reasons, threaten bank lending and therefore economic growth.

4.3.1 *Oil Shocks*

Crises in the Middle East regularly produce spikes in oil prices. Military conflicts that led to declines in world production of oil occurred in 1973–1974, 1979, 1980, 1990, and 2003–2004. Even though oil is a smaller input to the world economy now than it was in the 1970s, a sudden rise in prices affects consumers' income and reduces spending. Inflation rates also rise, though here the effect is ambiguous. Although inflation moves up initially, the contractionary effect of higher oil prices restricts employment and opens up an output gap so that, after a period, inflation slows to below the level where it otherwise might have been.

There have also been episodes of declining oil prices, most notably in 1986 and again in 1999. These tend to have the effect of extending the economic upswing because they contribute to lower inflation. Low oil prices and low inflation boost economic growth that can contribute to overheating, as was seen in the United States in 1987 and again in 1999–2000. Dependence on Middle East oil remains high, and the sources of political instability in the region remain numerous.

4.3.2 *Financial Crises*

Periodic financial crises affect growth rates either directly through bank lending or indirectly through their effect on investor confidence. In the last few decades, events in emerging markets have been the cause of several crises. The Latin American debt crisis of 1982, the Mexican currency crisis of 1994, the Asian financial crisis of 1997, and the Russian crisis of 1998 are examples. The last was particularly important because it threatened both financial markets and investment banks with widespread collapse. The reason was a possible domino effect due to the subsequent collapse of Long-Term Capital Management (LTCM), a large US hedge fund. Most of LTCM's positions had been based on expectations of declining risk spreads. When the Russian crisis sent those spreads upward, it triggered a crisis. Among central banks, the US Federal Reserve's response to these emerging market crises was particularly proactive. The Fed injected liquidity into the system, thereby reducing US interest rates and moderating the impact on financial institutions.

There have been other financial crises. Banks are always potentially vulnerable after a major decline in asset prices, particularly property prices, as in the United States in the early 1990s. In that case, the Fed's response was to keep interest rates low for a prolonged period to provide sufficient liquidity to ensure that the payment system could continue. That action would have been more difficult in a world of low inflation or deflation. Financial crises are therefore potentially more dangerous in a low interest rate environment.

4.4 International Interactions

In general, the dependence of any particular country on international interactions is a function of its relative size and its degree of specialization. Large countries with diverse economies, such as the United States, tend to be less influenced by developments elsewhere than small countries, such as Chile, whose production depends significantly on a few commodities like copper. Increasing globalization of trade, capital flows, and direct investment means that practically all countries are increasingly affected by international interactions.

4.4.1 *Macroeconomic Linkages*

Countries' economies are directly affected by changes in the foreign demand for their exports. This is one way that the business cycle in one country can affect that in others. But there are other international linkages (other than trade) at work, such as those resulting from cross-border direct business investment. The United States has often been the leader in such investment. As a result, an economic slowdown in the United States frequently makes companies worldwide more cautious about investment and hiring.

However, the US economy and the economies of other developed countries are clearly not perfectly integrated. For example, continental Europe did not suffer a recession in 1990 despite the US recession because the stimulative effects of German reunification outweighed the negative effects of the US slowdown. Similarly, while the United States and most other countries suffered a weak economy in the first half of 2003, China's economy boomed under the influence of stimulative monetary and fiscal policies.

4.4.2 *Interest Rate/Exchange Rate Linkages*

One of the linkages of most concern to investors involves interest rates and exchange rates. Sometimes, short-term interest rates are affected by developments in other countries because one central bank pursues a formal or informal exchange rate link with another currency. Some governments *unilaterally* peg their currencies firmly or loosely to one of the major currencies, usually the US dollar or the euro. This strategy is much less common now than it was before the Asian crisis in 1997, but it is still practiced in Asia and by members of the Gulf Cooperation Council (Saudi Arabia, Kuwait, Bahrain, Qatar, the United Arab Emirates, and the Sultanate of Oman). Also, various countries in Africa and Eastern and Central Europe peg to the euro.

The countries that follow this strategy find two benefits. First, domestic business has some reassurance that exchange rates are not going to fluctuate wildly. Second, by pegging the exchange rate, a "pegged" country often hopes to control inflation. This consideration was important in Europe under the Exchange Rate Mechanism (an exchange rate regime established in preparation for the introduction of the euro) and was also the reason for Argentina's convertibility plan, which tied the peso to the dollar in the early 1990s but collapsed amidst a severe economic crisis in 2001.

If a country is following such an exchange rate policy, then the level of interest rates will depend on overall market confidence in the peg. A high degree of confidence in the exchange rate peg means the interest rate differential can converge to near zero. But if the markets see the peg policy as unsustainable, then investors will demand a substantial interest differential.

If a country is known to be linking its currency to another, then bond yields of the weaker currency are nearly always higher. Hence, in Europe, Polish bond yields bear a spread over euro bond yields. If the expectation were that the zloty/euro exchange rate would remain broadly the same as its current level over the long term, then bond yields in Poland would converge with those in the Eurozone for bonds of equal credit risk. An expectation of a stable exchange rate might be justified by a belief in the Polish government's determination to maintain parity or by a perception that inflation and the competitiveness outlook for Poland obviate the need for a devaluation. If, however, markets anticipate a devaluation at some stage before Poland joins the euro, then a bond yield spread should remain in place.

Even if countries are not trying to link their currencies, bond yields can diverge substantially between countries. For example, if one country's exchange rate is severely undervalued and is expected to rise substantially against another country's, then bond yields in the first country will be lower than they would otherwise be in relation to the other country.

Exchange rates can be over- or undervalued, requiring an offset from bond yields, for a number of reasons, such as government action on *short-term* interest rates. For example, the Exchange Rate Mechanism was maintained as long as it was by using high short-term interest rates to limit speculation against currencies.

Misvaluation can also happen when bond yields reflect a particularly strong economy. In 1984, US bond yields averaged 12.5 percent despite an annual inflation of 4 percent. These high real and nominal rates were due to the combination of the increasing US budget deficit, a strong private sector economy, and a tightening monetary policy. In comparison, Germany had bond yields of 8 percent with an annual inflation of 2.5 percent. Hence, investors in the United States could enjoy real yields 3 percent above those seen by investors in German bonds: 8.5% real US yield − 5.5% real German yield = 3% excess real yield difference in favor of the United States. This yield difference was enough to take the dollar up substantially in 1983–1985, leaving the bond markets in some degree of equilibrium, although the US dollar was then viewed as overvalued.

Obviously, *nominal* bond yields vary between countries according to those countries' different inflation outlooks and other factors. It is sometimes thought that *real* bond yields ought to be similar in different countries because international capital flows will equalize them. However, movements in exchange rates to under- or overvalued levels can compensate for different real bond yields. Although real yields can and often do vary among countries, they tend to move together. In the example above from 1984, bond yields in both the United States and Germany were comparatively high in relation to inflation and overall historical experience.

The key factor linking bond yields (especially real bond yields) is world supply and demand for capital or the perceptions of supply and demand. Take the example of the collapse of world bond markets and the sharp rise in bond yields in 1994. These events seem to have been partly due to a perception that synchronized world growth would force short- and long-term interest rates up as the demand for world savings exceeded the supply. Since in the end, the demand has to equal the supply, interest rates everywhere rose to choke off demand and/or stimulate more supply. Similarly, in 2001, bond yields fell everywhere as private demand for capital dropped off in the face of a world slowdown.

4.4.3 *Emerging Markets*

There are some special considerations in setting capital market expectations for emerging countries. Here, we outline some of the key differences from major economies and look briefly at country risk analysis techniques and data sources that analysts use in evaluating emerging markets.

4.4.3.1 Essential Differences between Emerging and Major Economies Emerging countries are engaged in a catch-up process. As a result, they need higher rates of investment than developed countries in physical capital and infrastructure and in human capital. But many emerging countries have inadequate domestic savings and therefore rely heavily on foreign capital. Unfortunately, managing the consequent foreign debt often creates periodic crises, dealing a major blow to investors in emerging stocks and bonds.

Very often, emerging countries have a more volatile political and social environment than developed countries. In comparison with developed countries, emerging countries tend to have a relatively large percentage of people with low income and few assets and a relatively small middle class with its typically major stake in political and economic stability.

Most emerging countries need major structural reform to unlock their potential, which can be difficult to achieve in a volatile political environment. The potential for growth is often blocked by governments protecting vested interests. As a result, the International Monetary Fund (the IMF, an international organization entrusted with fostering global monetary and financial stability) and the World Bank (a group of five international organizations responsible for providing finance to developing countries) have often placed conditions on aid both to manage the risk of crises and to promote growth. For countries within an IMF program, analysts often focus closely on a country's progress in meeting the targets.

Even the largest emerging countries are relatively small in world terms, and their economies are often concentrated in a few areas such as particular commodities or in a narrow range of manufactured goods. Others rely heavily on oil imports and are thus vulnerable to fluctuation in oil prices or are dependent on continuing capital inflows.

4.4.3.2 Country Risk Analysis Techniques Investors in emerging bonds focus on the risk of the country being unable to service its debt (make promised payments of interest and principal). Investors in stocks need to assess the growth prospects of emerging countries as well as their vulnerability to surprises. A common approach is to use a checklist of various economic and financial ratios and a series of qualitative questions. Following are six questions that country risk analysis seeks to answer, with suggestions for data to analyze and points to look for.

1 *How sound is fiscal and monetary policy?*

If there is one single ratio that is most watched in all emerging market analysis, it is the ratio of the fiscal deficit to GDP. Most emerging countries have deficits and are engaged in a perpetual struggle to reduce them. Deficits are a major cause of slow growth and frequently a factor in serious crises. A persistent ratio above 4 percent is regarded with concern. The range of 2–4 percent is acceptable but still damaging. Countries with ratios of 2 percent or less are doing well.

If the fiscal deficit is large for a sustained period, the government is likely to build up significant debt. In developing countries, governments usually borrow in the short term from domestic lenders in local currency and from overseas in foreign currency. The Argentinean crisis in 2001 was essentially the result of too much government debt. For a developing country, the level of debt that would be considered too high is generally lower than for developed countries. Countries with a ratio of debt of more than about 70–80 percent of GDP are extremely vulnerable.

2 *What are the economic growth prospects for the economy?*

The most successful countries in Asia have been able to grow at annual rates of 6–8 percent on a sustained basis. Others have achieved a respectable rate of 4–6 percent per annum. Annual growth rates of less than 4 percent generally mean that the country is catching up with the industrial countries slowly, if at all. It also means that, given some population growth, per capita income is growing very slowly or even falling, which is likely to bring political stresses.

Investors usually welcome a wave of reform because it will typically boost economic growth and the stock market. But growth may slow down after that unless there is further reform or new opportunities are opened up. One of the best indicators of the structural health of an economy is the Economic Freedom Index, published by a consortium of research institutes around the world (www.freetheworld.com). This index consists of a range of indicators of the freedoms enjoyed by the private sector, including tax rates, tariff rates, and the cost of setting up companies. Economies such as the United States and Singapore have scored well. The index has been found to have a broad positive correlation with economic growth.

3 *Is the currency competitive, and are the external accounts under control?*

Managing the currency has proven to be one of the most difficult areas for governments. If the exchange rate swings from heavily undervalued to seriously overvalued, there are negative effects on business confidence and investment. Moreover, if the currency is overvalued for a prolonged period, the country is likely to be borrowing too much, creating a large current account deficit and a growing external debt.

The size of the current account deficit is a key measure of competitiveness and the sustainability of the external accounts.[64] Any country with a deficit persistently greater than 4 percent of GDP is probably uncompetitive to some degree. Current account deficits need to be financed. If the deficits are financed through debt, servicing the debt may become difficult. A combination of currency depreciation and economic slowdown will likely follow. The slowdown will also usually cut the current account deficit by reducing imports. Note,

64 To briefly review accounting for cross-border flows, the **balance of payments** (an accounting of all cash flows between residents and nonresidents of a country) consists of the current account, dominated by the trade balance (reflecting exports and imports), and the financial account, consisting of portfolio flows (from security purchases and sales—e.g., bonds and equities) and foreign direct investment (FDI) by companies (e.g., Toyota Motor Corporation building an automobile assembly plant in the United States), as well as flows such as borrowing from and deposits with foreign banks. The sum of the current account and the financial account, or the **overall trade balance**, should be zero.

however, that a small current account deficit on the order of 1–3 percent of GDP is probably sustainable, provided that the economy is growing. A current account deficit is also more sustainable to the extent that it is financed through foreign direct investment rather than debt, because foreign direct investment creates productive assets.

4 *Is external debt under control?*

External debt means foreign currency debt owed to foreigners by both the government and the private sector. It is perfectly sensible for countries to borrow overseas because such borrowing serves to augment domestic savings. But borrowing needs to be kept within reasonable bounds or lenders will begin to question its sustainability. The resulting reluctance to lend new money may lead to an exodus of capital as money invested in local bonds and stocks flows out.

Analysts watch several measures of debt burden. The ratio of foreign debt to GDP is one of the best measures. Above 50 percent is dangerous territory, while 25–50 percent is the ambiguous area. Another important ratio is debt to current account receipts. A reading above 200 percent for this ratio puts the country into the danger zone, while a reading below 100 percent does not.

5 *Is liquidity plentiful?*

By liquidity, we mean foreign exchange reserves in relation to trade flows and short-term debt. Traditionally, reserves were judged adequate when they were equal in value to three months' worth of imports. However, with the vastly greater importance of debt and capital flows, we now relate reserves to other measures. An important ratio is reserves divided by short-term debt (debt maturing in less than 12 months). A safe level is over 200 percent, while a risky level is under 100 percent.

Excess short-term borrowing is present in most emerging market crises. This is partly a result of the fact that countries often find it more difficult to borrow longer-term in the period leading up to the crisis. But if the country borrows too much, even short-term lending eventually stops and the country is typically in crisis very quickly thereafter.

6 *Is the political situation supportive of the required policies?*

If the economy of the country is healthy, with fast growth, rapid policy liberalization, low debt, and high reserves, then the answer to this question matters less. Poor political leadership is unlikely to create a crisis. However, if the economic indicators and policy are flashing warning signals, the key issue becomes whether the government will implement the necessary adjustment policies. Cutting the budget deficit, which usually requires some combination of higher taxes and lower spending, is always painfully difficult, especially if the economy is weak already. Other key policy changes are reforms such as privatization and the ending of monopolies.

In summary, the evaluation of emerging economies uses many of the same tools as the evaluation of developed countries but places a greater emphasis on the balance of payments, debt, liquidity, and politics. The analysis pays off because despite serious risks caused by political instability and periodic crises, many emerging countries grow faster than developed countries and offer good investment opportunities. Since the Asian crisis in 1997, investors are much more conscious of the potential risks, which include market declines, fixed or quasi fixed exchange rates, major recessions, and contagion. The worst losses have been suffered in the countries that turned out to be the weakest politically.

4.5 Economic Forecasting

Having reviewed some practical basics of macroeconomics for the investment analyst with many real-world illustrations, we can now indicate some of the disciplines that the analyst can apply to economic forecasting. Often, analysts consider the implications of a variety of approaches, which will often raise questions that lead to productive analysis and insight. We may distinguish at least three distinct approaches:

- Econometric models, the most formal and mathematical approach to economic forecasting.
- Leading indicators: variables that have been found to lead (precede) turns in the economy.
- Checklists, requiring the subjective integration of the answers to a set of relevant questions.

In the following, we address each of these approaches in turn.

4.5.1 *Econometric Modeling*

Econometrics is the application of quantitative modeling and analysis grounded in economic theory to the analysis of economic data.[65] Whereas generic data analysis can involve variables of all descriptions (possibly including economic, security characteristic, demographic, and statistical variables), econometric analysis focuses on economic variables, using economic theory to model their relationships.

Econometric models vary from small models with just one equation or perhaps a handful of equations to large, complex models with hundreds of equations. However, they all work in essentially the same way. A model is created of the economy based on variables suggested by economic theory. Optimization (frequently the least-squares criterion from regression analysis) using historical data is used to estimate the parameters of the equations. The estimated system of equations is used to forecast the future values of economic variables, with the forecaster supplying values for the exogenous variables. For example, such a model may require the forecaster to enter an exchange rate, an interest rate, and commodity prices. But the model then uses the estimated past relationships to forecast the future.

A very simple model is presented in the following series of equations:

1 GDP growth = function of (Consumer spending growth and investment growth)
2 Consumer spending growth = function of (Consumer income growth lagged one period and Interest rate*)
3 Investment growth = function of (GDP growth lagged one period and Interest rate*)
4 Consumer income growth lagged one period = Consumer spending growth lagged one period

Here, the asterisk (*) denotes an exogenous variable. So, with this four-equation model estimated on past data and with the actual data for the variables lagged one period, together with the modeler's exogenous forecasts for the interest rate, the model will solve for GDP growth in time in the current period. Note that the final equation asserting that consumer income growth is always identical to consumer spending growth assumes a static relationship between these two variables.

65 There are also quantitative approaches reflecting **unstructured modeling**—that is, modeling without a theory on the underlying structure—such as vector autoregression (VAR), that may be appropriate for certain types of economic forecasting. See Diebold (2006) for an introduction to VAR.

Additional variables make a model more complex, more realistic, but often more difficult to construct, estimate, and interpret. Most models will incorporate variables such as government spending, employment, the savings rate, money supply, exports, and imports. However, it is by no means certain that larger models are superior to smaller models. Also, different models have different structures, and these structures reflect the views of the modeler both in what variables are included and in how they interrelate with one another. Monetarist models, for example, rely heavily on money-supply-related variables and relationships.

Econometric models are widely regarded as very useful for simulating the effects of changes in certain variables. For example, they can be useful for assessing the impact of a 10 percent rise in oil prices or a rise in income tax rates or a faster growth rate in trading partners on consumer demand.[66] Econometric models have several limitations. First, econometric models require the user to find adequate measures for the real-world activities and relationships to be modeled, and these measures may not be available. Variables may also be measured with error. Relationships among the variables may change over time because of changes in the structure of the economy; as a result, the econometric model of the economy may be misspecified.

In practice, therefore, skillful econometric modelers use a great deal of personal judgment in arriving at forecasts. Very often, the first run of the model will generate a forecast that the modelers do not believe. So, they will go back and change some of the exogenous variables to arrive at a forecast they do believe. The great merit of the econometric approach, however, is that it constrains the forecaster to a certain degree of consistency and also challenges the modeler to reassess prior views based on what the model concludes.

In practice, model-based forecasts rarely forecast recessions well, although they have a better record in anticipating upturns. For example, the US economic upswing that gathered pace in the second half of 2003 was well forecast by econometric models.

4.5.2 *Economic Indicators*

Economic indicators are economic statistics provided by government and established private organizations that contain information on an economy's recent past activity or its current or future position in the business cycle. **Lagging economic indicators** and coincident indicators are indicators of recent past and current economic activity, respectively. A **leading economic indicator** (LEI) is a variable that varies with the business cycle but at a fairly consistent time interval before a turn in the business cycle. Most analysts' greatest interest is leading indicators because they may provide information about upcoming changes in economic activity, inflation, interest rates, and security prices.

Leading indicator–based analysis is the simplest forecasting approach to use because it requires following only a limited number of variables. The indicators are best thought of as early signs of probable events to come.

Many private sector forecasters try to gain an edge in identifying the factors that best predict the path of the economy and use their own proprietary indicators. Nevertheless, a good place to start for most investors is the leading indicators published by national governments or, in some countries, by established private organizations such as the Conference Board in the United States.

Analysts may use both individual LEIs and composite LEIs, reflecting a collection of economic data releases that are combined to give an overall reading. Composite LEIs combine these releases using weights based on an analysis of their forecasting

---

66 Mehra and Peterson (2005) found that in the United States a 10 percent increase in oil prices is associated with the level of consumer spending at the end of six quarters being 0.80 percent to 1.60 percent lower than it otherwise would be.

usefulness in past cycles. They can also be combined in a so-called **diffusion index**, which measures how many indicators are pointing up and how many down. For example, if 7 out of 10 are pointing upward, then the odds are that the economy is accelerating.

We review a selection of LEIs, both individual and composite, by geographic region:

Worldwide[67]

OECD Composite Leading Indicators (CLI) (www.oecd.org). The Organization for Economic Cooperation and Development (OECD) is a Paris-headquartered organization comprising 35 member countries as of December 2016. The OECD publishes CLI on Friday of the first full week of each month (these releases relate to information on production two months earlier). The indexes are based on a range of variables (5 to 11 for each country), such as share prices, industrial production data, building permits, and monetary data that may have predictive value for the course of the business cycle. A Total OECD Composite Leading Indicator is published, as well as indexes for the 35 member countries and for regions including:

- Big Four European countries (France, Germany, Italy, and United Kingdom);
- Euro area (European Monetary Union [EMU] countries);
- G–7 (Canada, France, Germany, Italy, Japan, United Kingdom, and United States);
- OECD-Europe (the EMU countries plus others);
- Total OECD countries (OECD-Europe plus other non-European members); and
- NAFTA (Canada, Mexico, and the United States).

The OECD publishes the six-month rate of change (annualized) in the monthly index numbers. Some analysts follow the six-month rate of change series more closely than the monthly index numbers because it may filter out the meaningless variation (noise) in the monthly numbers.[68]

Europe

Selected indicators include the following:

- **Eurozone Harmonized Index of Consumer Prices** (HICP) (epp.eurostat.ec.eu.int). Eurostat, the statistical office of the EU, publishes a composite index of inflation in the Eurozone. The European Central Bank (ECB) developed this index for use in inflation targeting, and indexes standardized on the ECB methodology have been developed for individual EU countries. Inflation indexes are not in general viewed as leading indicators. However, because of the ECB's strong focus on inflation containment, an unexpected increase in this index may presage events such as an interest rate increase.
- **German Industrial Production** (www.destatis.de). The Statistisches Bundesamt Deutschland (German Federal Office of Statistics) publishes in the second week of each month an index of German industrial production relating to production data two months earlier. Germany is Europe's largest economy and has frequently ranked among the top two exporters in the world. (As of 2009, its largest export partners were France, 10.1 percent; United States, 6.7 percent; United Kingdom, 6.6 percent; Netherlands,

67 See *Guide to Economic Indicators*, 5th ed. (The Economist, 2003).
68 See Baumohl (2005), p. 328.

6.6 percent; Italy, 6.3 percent; Austria, 5.7 percent; Belgium, 5.2 percent; China, 4.7 percent; and Switzerland, 4.5 percent.)[69] This series is closely watched as a leading indicator for the German and Eurozone economies.

- The **German IFO Business Survey** (www.ifo.de) and the **French Monthly Business Survey** (www.insee.fr), both released during the fourth week of the month being surveyed, are influential surveys of German and French business executives, respectively. Analysts focus on the answers to the forward-looking component of these series (a six-month-ahead time frame for the German series and a three-month-ahead time frame for the French series) for indications of Eurozone industrial production over the next several months. France is the Eurozone's second-largest economy.

Asia Pacific

Selected indicators include the following:

- The **Tankan Survey** of the Bank of Japan (www.boj.or.jp). Japan's central bank's detailed quarterly survey of business, published at the start of April, July, and October and in mid-December, is a rich source for information on Japanese business conditions and the expectations of business executives.
- **China Industrial Production** (www.stats.gov.cn). China's National Bureau of Statistics (NBS) releases monthly data on industrial production usually four weeks after the month being surveyed. These data are measures of the value added by light industry (mainly producing consumer goods) and heavy industry (producing durable goods such as factory equipment and automobiles).

South America

- **Brazil Industrial Production** (www.ibge.gov.br). The Brazilian Institute for Geography and Statistics releases monthly data on industrial production approximately 40 days after the end of the month surveyed. This series is probably the closest watched by analysts. As of the end of 2004, Brazil represented more than 40 percent of South America's total GDP. With an estimated population of over 185 million as of 2005, Brazil is also by far the most populous South American country. As of 2010, Brazil's major export partners were China, 15.2 percent; United States, 9.6 percent; Argentina, 9.2 percent; Netherlands, 5.4 percent; and Germany, 4.0 percent. Its major import partners were the United States, 15.0 percent; China, 14.1 percent; Argentina, 7.9 percent; Germany, 6.9 percent; and South Korea, 4.6 percent.

North America[70]

- The Conference Board's **Index of Leading Economic Indicators**. The Conference Board is a private, nonprofit, New York City-headquartered research organization that took over the management of this series from the US Commerce Department in 1995. The Conference Board releases this monthly series three weeks after the end of the month reported upon, as well as coincident and lagging indexes. Exhibit 22 shows the 10 components of the US leading indicator index, which is normally quoted as a weighted index but is also available as a diffusion index. The Conference Board also publishes indexes for Australia, France, Germany, Japan, Korea, Mexico, Spain, and the United Kingdom, researching the best indicators to use for each country. The United States is the world's largest economy. As of 2010, the United States' largest exporting partners were Canada, 19.4 percent;

69 See "The World Factbook" at www.cia.gov.
70 This example draws on Baumohl (2005).

Mexico, 12.8 percent; China, 7.2 percent; and Japan, 4.7 percent. Its largest importing partners were China, 19.5 percent; Canada, 14.2 percent; Mexico, 11.8 percent; Japan, 6.3 percent; and Germany, 4.3 percent.

Exhibit 22 US Composite Indexes: Components and Standardization Factors

Leading Index	Factor
1. Average weekly hours, manufacturing	0.2552
2. Average weekly initial claims for unemployment insurance	0.0307
3. Manufacturers' new orders, consumer goods and materials	0.0773
4. Vendor performance, slower deliveries diffusion index	0.0668
5. Manufacturers' new orders, non-defense capital goods	0.0183
6. Building permits, new private housing units	0.0271
7. Stock prices, 500 common stocks	0.0391
8. Money supply, M2	0.3550
9. Interest rate spread, 10-year Treasury bonds less federal funds	0.1021
10. Index of consumer expectations	0.0284
Coincident Index	
1. Employees on nonagricultural payrolls	0.5426
2. Personal income less transfer payments	0.1890
3. Industrial production	0.1493
4. Manufacturing and trade sales	0.1191
Lagging Index	
1. Average duration of unemployment	0.0373
2. Inventory/sales ratio, manufacturing and trade	0.1239
3. Labor cost per unit of output, manufacturing	0.0615
4. Average prime rate	0.2822
5. Commercial and industrial loans	0.1112
6. Consumer installment credit to personal income ratio	0.1880
7. Consumer price index for services	0.1959

Source: Conference Board, www.globalindicators.org.

In a given index, the component factors are inversely related to the standard deviation of the month-to-month changes in each component. They are used to equalize the volatility of the contribution from each component and are "normalized" to sum to 1. When one or more of the components are missing, the other factors are adjusted proportionately to ensure that the total continues to sum to 1.

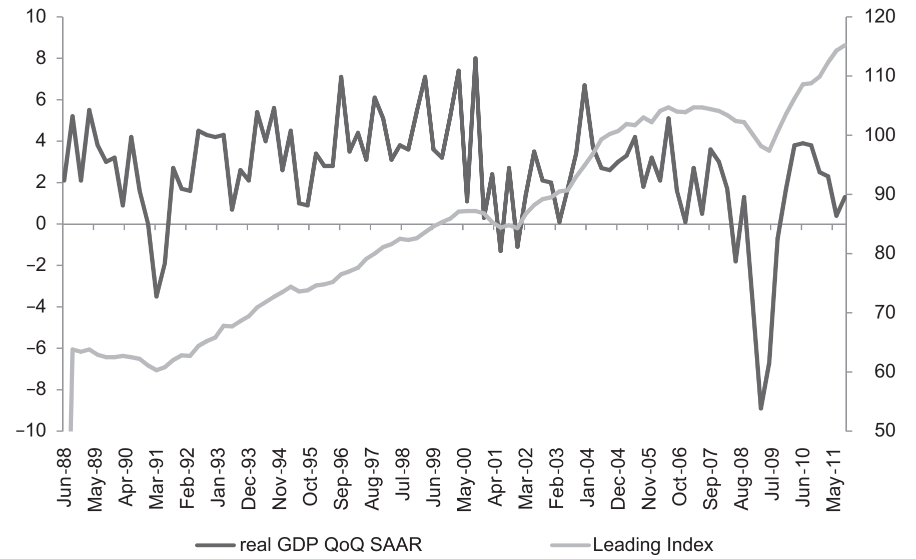
As Exhibit 22 shows, the Conference Board's LEI index consists of seven nonfinancial and three financial components (#7, #8, and #9). Of particular note is #5, provided by the Institute of Supply Management (ISM), which is a professional organization of purchasing managers. This release comes out on the first business day of each month and so is one of the earliest pieces of information on the business cycle available in a given month. This release receives more attention than #3, also from the ISM. Because capital goods orders are more sensitive to the business cycle than consumer goods, the release of #5 is often a market-moving event. Another interesting component is #7, the S&P 500 Index, which historically has been a good leading indicator of the stock market.

In contrast with the release of some of its individual components, the release of the LEI index is rarely a market-moving event because some of its components are already public.

Traditionally, the general rule was that three consecutive months of increases, or three consecutive months of decreases, signaled an upturn or downturn in the economy within three to six months. The Conference Board and others have also indicated other rules for interpreting changes in the index.

Exhibit 23 plots the composite LEI against quarterly GDP changes (expressed as a seasonally adjusted annual rate, or SAAR). The left vertical axis represents percentage changes and applies to real GDP changes; the right vertical axis shows index levels and applies to LEI levels.

Exhibit 23 US: Leading Indicators and GDP



Source: Bloomberg.

The US LEI index gave a somewhat ambiguous signal ahead of the 1990 recession but performed much better in the most recent cycle. In the late 1990s, it correctly showed that the Asian crisis was not threatening a US downturn. Then, in 2000, it peaked in January. By midyear, it was clearly falling, well before there was general agreement that the economy was slowing. In 2003, it picked up strongly in May and continued to rise rapidly, correctly signaling a strong recovery. The strong uptick after the recession of 2008–09 in the face of a tepid US recovery (as indicated by the declining quarter-over-quarter GDP change) is once again an ambiguous signal.

4.5.3 *Checklist Approach*

Formally or informally, many forecasters consider a whole range of economic data to assess the future position of the economy. Checklist assessments are straightforward but time-consuming because they require looking at the widest possible range of data. The data may then be extrapolated into forecasts via objective statistical methods, such as time-series analysis, or via more subjective or judgmental means. An analyst may then assess whether the measures are in an equilibrium state or nearer to an extreme reading.

Inflation reports provided by many central banks or through the minutes of central bank meetings give an idea of the range of indicators that may be included in checklists related to preparing general economic forecasts.

Exhibit 24 is an example of a checklist for evaluating economic growth. In effect, the forecaster asks a series of questions about likely components of spending and then, aggregating the information gathered, reaches a conclusion about the outlook for the economy. Such an approach involves a substantial amount of subjective judgment as to what is important in the economy.

Exhibit 24 Checklist for Economic Growth

Spending Components	Focus
1 Where in the cycle is the economy now?	Aggregate activity
■ Review previous data on GDP growth and its components (consumer spending, business investment, inventories, net exports, and government spending).	Aggregate activity
■ How high is unemployment relative to estimates of "full employment"?	Consumer
■ Has unemployment been falling?	Consumer
■ How large is the output gap?	Business
■ What is the inventory position?	Business
■ Where is inflation relative to target, and is it threatening to rise?	Inflation
2 How strong will consumer spending be?	Consumer
■ Review wage/income patterns.	Consumer
■ How fast will employment grow?	Consumer
■ How confident are consumers? Consumer confidence indexes.	Consumer
3 How strong will business spending be?	Business
■ Review survey data (e.g., purchasing managers indexes).	Business
■ Review recent capital goods orders.	Business
■ Assess balance sheet health of companies.	Business
■ Assess cash flow and earnings growth trends.	Business
■ Has the stock market been rising?	Business
■ What is the inventory position? Low inventory/sales ratio implies GDP strength.	Business
4 How strong will import growth be?	Government
■ Exchange rate competitiveness and recent movements.	Government
■ Strength of economic growth elsewhere.	Government
5 What is the government's fiscal stance?	Government
6 What is the monetary stance?	Central bank
■ Review recent changes in interest rates.	Central bank
■ What do real interest rates tell us?	Central bank
■ What does the Taylor rule tell us?	Central bank
■ Monetary conditions indexes (i.e., trends in asset prices and exchange rate).	Central bank
■ Money supply indicators.	Central bank

Exhibit 24 (Continued)

Spending Components	Focus
7 Inflation	Inflation
■ How fast is inflation rising, or are prices falling?	Inflation

Example 30 is a simple illustration of the checklist approach.

EXAMPLE 30

An Analyst's Forecasts

As a capital market analyst at a large money management firm, Charles Johnson has developed a list of six broad questions for evaluating the economy. The questions are given below with his responses for his own national market. The current inflation rate is 2 percent per year.

A Is consumer spending increasing or decreasing? *Johnson*: Consumer spending is increasing at a lackluster rate of 0.75 percent per annum.

B Are business conditions and fundamentals growing stronger or weakening? *Johnson*: Based on recent values of manufacturers' new orders for consumer goods and materials and non-defense capital goods, business demand is weakening.

C What is the consensus forecast for the GDP growth rate over the next year? *Johnson*: The median forecast from a survey of economists is that GDP will decline from a 3.5 percent to a 3.0 percent annual growth rate.

D Are government programs and fiscal policy becoming more restrictive or expansive? *Johnson*: Political support for a stimulative fiscal policy is absent; fiscal policy will be neutral.

E Is monetary policy neutral, tightening, or loosening? *Johnson*: Monetary policy is neutral.

F Is inflation in a steady state (state of equilibrium)? *Johnson*: The current inflation rate of 2 percent is close to a steady state value.

Based on the information given, what conclusions will Johnson reach concerning:

1 inflation over the next six months?

2 short-term interest rates?

Solution to 1:

Based on the expected slow growth in consumer demand and weakening business demand, inflation should remain muted over the next six months.

Solution to 2:

Reduced aggregate economic activity and stable inflation should allow for stable to falling interest rates.

The subjectivity of the checklist approach is perhaps its main weakness. The checklist's strength is its flexibility. It allows the forecaster to take changes in the structure of the economy into account quickly by changing the variables or the weights assigned to variables within the analysis. The next section summarizes the three chief approaches.

4.5.4 *Economic Forecasting Approaches: Summary of Strengths and Weaknesses*

Exhibit 25 summarizes the advantages and disadvantages of forecasting using econometric models, leading indicators, and checklists.

Exhibit 25 Advantages and Disadvantages of Three Approaches to Economic Forecasting

Advantages	Disadvantages
Econometric Models Approach	
■ Models can be quite robust with many factors used that can approximate reality. ■ Once models are built, new data may be collected and consistently used within models to quickly generate output. ■ Provides quantitative estimates of the effects on the economy of changes in exogenous variables.	■ Most complex and time-consuming to formulate. ■ Data inputs and relationships not easy to forecast and not static. ■ Requires careful analysis of output. ■ Rarely forecasts recessions well.
Leading Indicator–Based Approach	
■ Usually intuitive and simple in construction. ■ May be available from third parties. ■ May be tailored for individual needs. ■ A literature exists on the effective use of various third-party indicators.	■ Historically, has not consistently worked, as relationships between inputs are not static. ■ Can provide false signals.
Checklist Approach	
■ Limited complexity. ■ Flexible: allows structural changes to be easily incorporated.	■ Subjective. ■ Time-consuming. ■ Complexity has to be limited due to the manual nature of the process.

4.6 Using Economic Information in Forecasting Asset Class Returns

Movements in economic variables play a key role in forming investors' expectations. Although some investors, such as pure bottom-up stock pickers or fully hedged arbitrage specialists, do not care much about the way that economic developments move markets, it is important for most investors. In this section, we look at how the principal asset classes are moved by different economic variables.

4.6.1 *Cash and Equivalents*

Cash managers make money through selection of the maturity of the paper in their portfolio or, if permitted by investment policy, by taking credit risk. Longer maturities and lower credit ratings reward the extra risk with higher expected returns. Managers lengthen or shorten maturities according to their expectations of where interest rates will go next. Normally, longer-maturity paper will pay a higher interest rate than shorter-maturity paper, even if overnight interest rates are expected to remain the same, because the risk of loss is greater for the longer-term paper if this expectation is not fulfilled. If further rises in rates are expected over time, then 6- and 12-month paper should offer even higher rates than shorter-term paper.

The overnight interest rate is targeted by the central bank and will normally vary only slightly from the target set. For example, in the United States, the Federal Reserve's target for the fed funds rate along with open market operations usually ensures the overnight interest rate is close to the target. (**Open market operations** are the purchase or sale by a central bank of government securities, which are settled using reserves, to influence interest rates and the supply of credit by banks.) Occasional variations are due normally to liquidity factors, especially close to year-end or during unusual market turbulence. In other countries, the target rate may be the repo rate (repurchase rate), as in the Eurozone, where the European Central Bank conducts open market operations.

At any given time, the yield curve of interest rates of a particular security (e.g., Treasury bills or interest rate futures) reflects the market's expectations of rates over that period. The money manager is trying to be ahead of others in correctly forecasting those levels. In practice, this means forecasting both the behavior of the economy and the reaction of the central bank to that behavior. It also means understanding what the markets currently anticipate and distinguishing between future data surprises and what is already factored into expectations (i.e., likely to have no effect on the market). Money managers therefore spend a great deal of time in so-called "central bank watching."

EXAMPLE 31

Central Bank Watching and Short-Term Interest Rate Expectations

At the beginning of 2000, the US stock market bubble peaked and the economy was strong. At this point, one-month US interest rates stood at about 5.7 percent per annum, with the six-month yield about 40 basis points higher at 6.1 percent per annum. Interest rates had already moved up in 1999, but the market was expecting the Federal Reserve to announce additional small increases in rates to help slow the economy and avoid rising inflation. A money manager might have been tempted to buy the longer-term paper, given its higher yield. However, the Federal Reserve raised interest rates faster than expected, with the fed funds rate moving up from 5.5 percent to 6.5 percent by June 2000. The best place to be was therefore at the short end of the curve, because by May 2000, one-month paper could be bought to yield 6.5 percent per annum and six-month paper could be bought to yield 6.8 percent per annum. During periods of rising short-term rates, keeping maturities short is a good strategy.

Early summer 2000 turned out to be the peak for US interest rates, and rates were cut sharply in 2001 when the US economy went into recession. In November 2000, shortly before the markets began to expect that the Fed would cut interest rates sharply, six-month rates of 6.7 percent per annum were available, just above one-month rates of 6.5 percent per annum. In the first months of 2001, the Fed cut rates rapidly and one-month yields fell to 4 percent per annum by May 2001 with an average yield of only just over 5 percent per annum.

Consider a manager who, during the summer of 2000, correctly anticipated the actions of the Fed in 2001. For such a manager, contrast the appropriateness of the following two strategies:

- An investment strategy of rolling over one-month paper.
- An investment strategy of buying longer-maturity paper (in this case, six-month paper).

Solution:

The second strategy is superior, as it would lock in the higher yields for six months in a declining interest rate environment. By contrast, the first strategy counts on interest rates rising, not declining. The first strategy would produce higher returns only if interest rates rose.

4.6.2 *Nominal Default-Free Bonds*

Nominal default-free bonds are conventional bonds that have no (or minimal) default risk. Conventional government bonds of developed countries are the best example. Thus, our focus is on the government yield curve. One way to think of the yield on a government bond is that it reflects the expected future short-term Treasury bill yields over the same horizon. Another approach, which is more useful for longer-term bonds, is to break down the yield into at least two components. First, the so-called real bond yield is determined by the growth rate of GDP and the supply and demand for capital. Second, yields are affected by forecast inflation over the investment period. For default-risk-free bonds, the credit spread or default risk premium is zero. Investors may thus assess whether bonds are cheap or expensive according to their view on whether the markets are too optimistic or too pessimistic based on real yields and inflation.

Historically, taking the period 1900–2010, the average annual real return on long-term government bonds above inflation was 2.3 percent for the United States and 2.2 percent for the United Kingdom.[71] However, there is some evidence that investors underestimated inflation during several periods, including the world wars and the peacetime inflation of the 1960s and 1970s. Hence, a better estimate of the *ex ante* expectations for annual returns above inflation (i.e., real yields) is 2–4 percent.

The investor then needs to forecast the inflation rate over the long term. For example, if 10-year bonds yield 5 percent and inflation is forecast at 2 percent, then the investor is hoping to receive approximately a 3 percent real return. If his or her judgment is that annual inflation is likely to be only 0.5 percent or perhaps that deflation will occur, then these bonds will be particularly attractive. Conversely, if inflation is thought likely to accelerate—for example, to 6 percent—then the bonds are very unattractive because they will not compensate for this higher inflation rate and will likely fall below par value during their lifetime as market yields rise.

For investors buying and selling long-term bonds over a shorter time period, the emphasis is on how bond yields will respond to developments in the business cycle and changes in short-term interest rates. News of stronger economic growth usually makes bond yields rise (prices fall) because it implies greater demand for capital and perhaps higher inflation too. Changes in short-term rates have less predictable effects on bond yields. More often than not, a rise in short-term rates will lead to a rise in longer-term bond yields. However, a rise in rates will sometimes be expected to slow the economy, and bond yields could fall as a result. If bond markets expect that central

71 See Dimson, Marsh, and Staunton (2011).

banks will exactly achieve their inflation objectives, then bond yields should not change on inflation expectations but nevertheless could go up and down according to changes in short rates (with higher short rates making bonds less attractive in relative terms).

As bond investors look toward the long-term picture, they must carefully assess the future effects of inflation, which erodes the future purchasing power of the yields earned on their fixed-income investments. In the 1970s, bond investors in the United States and most major countries suffered severe losses in real terms because of unexpected inflation. Yields in the 1960s were at single-digit levels, and yet inflation in the following 15 years moved up into double digits. In the 1980s, after this bad experience, investors were apprehensive of a new surge in inflation and therefore demanded higher yields (a higher inflation premium). Hence, yields were often 4–6 percent above recorded inflation—much higher than normal. As of early 2005, with inflation very low, US bond investors have begun to believe not only that inflation will stay low, but also that there is a risk of deflation. Thus, as inflation fears declined during the 1990s, the inflation premium in bond yields likewise declined, reducing overall nominal yields. The inflation premium embedded in bond yields fell further, to unusually low levels in the first half of 2003, when deflation fears reached a high point and bond yields registered as little as 1–2 percent above recorded inflation.

4.6.3 *Defaultable Debt*

Defaultable debt is debt with some meaningful amount of credit risk—in particular, most corporate debt. For corporate debt, such as certificates of deposit and bonds, the spread over Treasuries represents at least in part the market's perception of default risk.[72] Individual securities move in response to particular corporate circumstances, but the market as a whole responds primarily to changes in short-term rates and changes in the business cycle. During a business cycle, spreads tend to rise during a recession because companies are under stress from both weak business conditions and, typically, higher interest rates. Sometimes, borrowing from banks or in the commercial paper market becomes more difficult too, so that companies can be severely squeezed. Default rates typically rise during recessions. Investors demand higher rates to pay for the uncertainties and possible surprises, such as fraud. In contrast, during periods of strong economic growth, spreads narrow as fears of default decline.

4.6.4 *Emerging Market Bonds*

Emerging market debt refers here to the sovereign debt of nondeveloped countries. So far, we have considered only government issues and regarded them as virtually risk-free from a credit point of view. Almost all of the main industrial country governments, the members of the OECD area, are AAA rated by the rating agencies and likely to remain so. Japan is the main exception (rated AA), with its rapidly rising debt/GDP ratio an increasing concern. In practice, even in Japan, rising debt is more likely to lead to a bout of inflation than an outright default, as long as governments can control monetary policy and therefore can ultimately print money to pay off debts. Emerging market bonds, as an asset class, are different in that the country is borrowing in a foreign currency. The government therefore cannot simply inflate its way out of a problem in servicing the debt, and so the risk of default is correspondingly higher. Assessing this risk, using what is known as country risk analysis, involves a large array of economic and political factors. Much of country risk analysis comes down to predicting policy moves and therefore often hinges on politics—that is, whether a government has the

---

72 Other factors, such as differences in taxation or the presence or absence of a call provision, can also account for part of this spread.

power to follow the necessary policies to stabilize the economy. Emerging market bonds are usually analyzed by developed market investors in terms of their spread over domestic Treasuries compared to similarly rated domestic corporate debt.

4.6.5 *Inflation-Indexed Bonds*

Many governments now issue bonds linked to inflation, so in principle, we can directly observe the market's forecast of inflation by comparing the yield of these indexed bonds with the yield on similarly dated conventional bonds. Examples of this important class of bonds are Treasury Inflation-Protected Securities (TIPS) in the United States and Index-Linked Gilts (ILGs) in the United Kingdom. These provide a fixed coupon (the real portion) plus an adjustment equal to the change in consumer prices. In principle, indexed bonds are the perfect risk-free asset because, unlike conventional bonds, they entail no risk from unexpected inflation. However, the yield on indexed bonds still changes over time, and in practice it varies with three economic factors.

First, the yield goes up and down with the real economy and particularly with the level of short-term interest rates. If real yields generally are high because the economy is strong, then real yields on TIPS and ILGs will be higher. Second, yields fall if inflation accelerates because these securities are more attractive when inflation is volatile. In other words, their value in hedging against inflation risk is higher. Finally, as with all assets, the yield can vary according to institutional supply and demand. In practice, tax effects and the limited size of the market (particularly for TIPS in the United States) may also distort the real yield; thus, investors usually find it worthwhile to forecast all three components (real yield, inflation, and supply and demand).

The yield-related relationships that affect indexed fixed-income securities can be shown as in Exhibit 26.

Exhibit 26 The Macroeconomy and Real Yields

Economic Observation	Effect on Real Bond Yields
Economic growth rising (falling)	Rise (fall)
Inflation expectations rising (falling)	Fall (rise)
Investor demand rising (falling)	Fall (rise)

4.6.6 *Common Shares*

To relate economic analysis to common equity valuation, it is useful to think of economic factors, first, in the way that they affect company earnings and, second, in the way that they affect interest rates, bond yields, and liquidity. The two views combined provide a forecast for the equity markets and can lead to new investor ideas and trading activity. Particular economic factors will also affect the outlook for specific companies—for example, the price of oil or the demand for airline travel. Here, we focus on the impact on the overall stock market.

4.6.6.1 Economic Factors Affecting Earnings In the long term, the trend growth in aggregate company earnings is mainly determined by the trend rate of growth of the economy. A faster-growing economy is likely to show faster average earnings growth, while a slower economy is correlated with slower earnings growth. The trend rate of growth of an economy is dependent on labor force growth, the level of investment, and the rate of labor productivity growth. Variations in growth rates among countries are usually due to past overinvestment, government overregulation or political instability, or the bursting of a major asset price bubble.

EXAMPLE 32

Economic Return Drivers: Energy and Transportation

Willem DeVries is researching the macroeconomic return drivers of the energy and transportation industries. He has gathered the information in Exhibit 27 from a US investment manager's research report.

Exhibit 27 Correlations of GDP, Inflation, and Interest Rates with Industry Sales, Earnings, and Dividends[73]

	Energy Industry			Transportation Industry		
	GDP	Inflation	Interest Rates	GDP	Inflation	Interest Rates
Sales	+0.10	+0.77	+0.78	+0.58	+0.75	+0.74
Earnings	+0.13	+0.66	+0.67	+0.81	+0.26	+0.25
Dividends	+0.16	+0.03	+0.05	+0.65	−0.03	−0.08

Using only the information given, compare and contrast the macroeconomic return drivers of the energy and transportation industries.

Solution:

The larger positive correlations between GDP and the transportation industry's sales, earnings, and dividends compared to the corresponding correlations for the energy industry are an indication that transportation companies are more procyclical.

Sales for the energy and transportation industries are approximately equally positively related to inflation and interest rates. However, earnings are less positively correlated with inflation and interest rates for the transportation industry. Transportation companies appear to be less able to pass through to customers the increased costs of higher inflation and interest rates. These observations should be helpful when one is using economic factors to draw inferences on future industry fundamentals.

Over the shorter term, the share of profits in GDP varies with the business cycle and is influenced by a variety of factors, including final sales, wages, capacity utilization, and interest rates.

During a recession, earnings are depressed because of reduced sales and a set amount of fixed costs. Capacity utilization is typically low. In severe recessions, earnings can disappear altogether for many companies. Other companies, less affected by the cycle (e.g., food companies), may see very little change in earnings. Companies that can maintain earnings growth through recessions receive high market valuations from investors.

During the early stages of an economic upswing, earnings recover strongly. One reason is the rise in capacity utilization and increasing employment. Many costs stay the same while volume rises, which brings large increases in profits. Wage awards usually remain modest because of continuing relatively high unemployment, so that

73 Analysis excerpted from a Duff & Phelps Investment Management Co. study of the 1,000 largest US companies over 1990–2003, using annual data.

most of the productivity gains flow straight into profits. A second factor is often the efficiency gains made during the recession that become evident when output rises. A leaner, fitter company emerges from recession as some of the fat built up during the growth years, including both obvious waste and "luxury" projects, is cut out.

Later in the economic upswing, wage growth starts to quickly rise, profits contract, and earnings growth slows. Some companies, generally the ones with large fixed costs and a pronounced sales cycle, are more sensitive to the business cycle than others. These are called **cyclical stocks**. Examples include car manufacturers and chemical producers.

Example 33 shows an analyst methodically organizing the economic analysis to formulate an answer to a client's question on the equity risk premium.

EXAMPLE 33

Researching US Equity Prospects for a Client

In the beginning of 2004, Michael Wu has one of his regular quarterly meetings with an institutional client for whom he manages a US equity portfolio. The economic forecasts of Wu's firm covering the next 12–18 months are consistent with the client's view that short-term interest rates will be increasing from 3.0 percent to 3.5 percent while long-term government bonds will return 5.5 percent. The client views US equities as currently slightly overvalued. The client is not optimistic about long-term prospects for US equities either and states to Wu that the long-term US equity risk premium will be in the range of 1.0 percent to 2.5 percent. The client asks Wu to help him decide, based on economic analysis, whether a 1.0 percent or 2.5 equity risk premium is more likely.

Wu summarizes his firm's research in Exhibit 28.

Exhibit 28 Current and Expected Economic/Market Trends: United States

	Expected Economic Trends/Impact Forecast		
Category	**Short-Term (1 Year)**	**Long-Term (> 1 Year)**	**Comments on Economic Measures and Categories**
Macroeconomy	$E_{(Trend)}$: Slowing GDP growth $E_{(Economic\ Impact)}$: Negative: Growth slowing from 4 percent to a lower rate	Average growth [3.1% annual GDP growth] $E_{(Trend)}$: Stable $E_{(Economic\ Impact)}$: Neutral	High current economic growth rate is due to fiscal and monetary stimuli and is not sustainable. Overall economic growth rate to slow to a lower 3.1% rate beyond 1-year time horizon.
Consumer	$E_{(Trend)}$: Improving consumer spending $E_{(Economic\ Impact)}$: Positive	$E_{(Trend)}$: Stable $E_{(Economic\ Impact)}$: Positive	Looking forward, stabilization in employment patterns and personal income will aid the consumer component of the economy.

Exhibit 28 (Continued)

Category	Expected Economic Trends/Impact Forecast: Short-Term (1 Year)	Expected Economic Trends/Impact Forecast: Long-Term (> 1 Year)	Comments on Economic Measures and Categories
Business	$E_{(Trend)}$: Stable $E_{(Economic\ Impact)}$: Neutral	$E_{(Trend)}$: Stable $E_{(Economic\ Impact)}$: Positive	The low base against which current results are being compared has aided profit and sales growth rate comparisons. Productivity growth has been aided by the weak employment (hiring) practices of the past few years. As employment rises, profit and productivity increases will diminish. Export-oriented businesses will be in the best position over the next few years as the US dollar is expected to decline further. Sales and profits showing signs of strength but are being compared to weak prior year results.
Central bank	Economic strength and fiscal deficits likely to put pressure on short-term interest rates; slightly higher inflation $E_{(Trend)}$: Declining stimulation $E_{(Economic\ Impact)}$: Negative	$E_{(Trend)}$: Stable $E_{(Economic\ Impact)}$: Neutral Short-term rates and inflation rate will stabilize near long-term average rates	Monetary stimulus expected to be reduced in light of the increased economic strength. The stronger economy will place upward pressure on short-term interest rates and on the rate of inflation near-term, but short-term interest rates and inflation are expected to quickly stabilize near their long-term average rates.
Government	$E_{(Trend)}$: Stable measures $E_{(Economic\ Impact)}$: Positive	$E_{(Trend)}$: Weakening $E_{(Economic\ Impact)}$: Negative	Fiscal stimulus (i.e., deficit spending and tax cuts) is giving a current boost to GDP. More work must be done to cut the budget deficit and to deal with a declining dollar. The US government needs to deal with the problem of increasing long-term transfer payment costs.

Note: Expected economic trends are denoted by $E_{(Trend)}$, while expected economic impact is denoted by $E_{(Economic\ Impact)}$.

Using only the information in Exhibit 28, address the following problems:

1 State and justify a long-term expected return for equities within the client's guidelines.

2 Comment on whether the economic data support the client's belief that the equity market is overvalued.

Solution to 1:

The consumer and business sectors are critical for corporate profits, and the long-term forecast strengths of these sectors are a positive for US equities. The central bank appears to be a neutral factor long-term. Although the government sector is a negative, it is not expected to push inflation and interest rates above their long-term averages. Overall, a 2.5 percent equity risk premium, at the upper end of the client's range, appears to be justified by the positive economic outlook, which would lead to a forecast of an 8.0 percent arithmetic average return on equities. (A long government bond expected return of 5.5% + expected equity risk premium of 2.5% = 8.0% expected equity return over the forecast period.)

Solution to 2:

By contrast to the long-term forecasts, the short-term economic forecasts of decelerating growth and increasing interest rates might constitute a negative for short-term equity returns. However, the analyst would need to evaluate whether current market prices incorporate this information before concurring in the client's assessment.

4.6.6.2 The P/E Ratio and the Business Cycle The price-to-earnings ratio of a stock market is the price that the market is willing to pay for the earnings of that market. During the business cycle, the P/E ratio tends to be high and rising when earnings are expected to rise. For example, the P/E would be high in the early stages of an economic recovery, or when interest rates are low and the return on fixed-rate investments such as cash or bonds is less attractive. Conversely, P/Es are likely to be low and falling if the outlook for earnings worsens (e.g., in an economic slump). Nevertheless, P/Es of cyclical companies may be above their own historical means during economic downturns as investors anticipate a sharp future earnings recovery when the economy turns up (a phenomenon known as the Molodovsky effect).

P/E ratios vary over longer periods too. In general, they are lower for an economy stuck on a slower growth path. During the 1990s, P/E ratios were at relatively high levels (e.g., multiples greater than 20 in the United States). At the time, some saw this situation as reflecting the benign economic influences of falling inflation, relatively low interest rates, fast productivity and profits growth, and a relatively stable economy. Another view was that these valuations were too high and would decline in the future, and this view has been borne out since 2000.

High inflation rates tend to depress P/E ratios. Inflation can distort the economic meaning of reported earnings, leading investors to value a given amount of reported earnings less during inflationary periods, which tends to lower observed P/Es during those periods. Consequently, comparisons of current P/E with past average P/E that do not control for differences in inflation rates may be suspect.[74]

4.6.6.3 Emerging Market Equities Empirical evidence points to *ex post* equity risk premiums for emerging markets that are on average higher and more volatile than those in developed markets. *Ex post*, emerging market equity risk premiums in US dollar terms appear to be positively correlated with expansion phases in G–7 economies as proxied by industrial production.[75] Transmission channels for G–7 macroeconomic fluctuations to developing economies include trade (demand for many of the goods produced by emerging countries, such as natural resources, is procyclical), finance, and direct sectoral linkages. In addition to evaluating linkages, the analyst needs to do considerable country and often sector-specific research to appraise the prospects for equity investments in a particular emerging country.

4.6.7 *Real Estate*

Ling and Naranjo (1997, 1998) identify growth in consumption, real interest rates, the term structure of interest rates, and unexpected inflation as systematic determinants of real estate returns. Interest rates are linked with a number of factors that affect the supply and demand for real estate, such as construction financing costs and the costs of mortgage financing. In general, lower interest rates are net positive for real estate valuation, resulting in lower capitalization rates.

74 For more details, see Bodie, Kane, and Marcus (2017).
75 See Salomons and Grootveld (2002).

In Example 34, an analyst with a one-year horizon applies a checklist approach to economic forecasting in modifying baseline historical capital market forecasts. The set of asset classes includes real estate.

EXAMPLE 34

Modifying Historical Capital Market Expectations

Cortney Young is an investment analyst in a firm serving an international clientele. Young's firm has developed the baseline forecasts shown in Exhibit 29 for six asset classes that are particularly relevant for UK-focused portfolios. The forecasts are based on a recent part of the historical record of the asset classes. Young is currently working on establishing capital market expectations for mean return and standard deviation of returns for these six asset classes based on a one-year horizon; she focuses first on the UK equities, UK intermediate-term bonds, UK long-term bonds, and UK real estate.

Exhibit 29 Baseline Forecasts

Asset Class	Mean Annual Returns (%)	Standard Deviation of Returns (%)
UK equities	9.72	15.3
Non-UK equities	8.94	11.6
UK intermediate-term bonds	3.60	6.5
UK long-term bonds	4.42	7.7
International bonds	4.81	8.3
UK real estate	12.63	8.7

Young's economic analysis leads her to the conclusions on the UK economy shown in Exhibit 30.

Exhibit 30 Economic Conclusions

Economic Category	Economic Conclusion
Consumers	Consumer spending is expected to be stronger over the next year with very positive effects for the UK economy.
Business	Business spending, revenues, and profits are expected to show solid growth in year-over-year results over the next 12 months.
Government	Tax policies are stable. Government is currently a source of moderate economic stimulation.
Central bank	It is anticipated that the Bank of England (the central bank) will want to hold short-term interest rates steady over the next year. The inflation target is 2 percent.

(continued)

Exhibit 30 (Continued)

Economic Category	Economic Conclusion
Inflation rates	The inflation rate is expected to increase to 2.2 percent per year over the next year.
Other/Unique	The UK economy is expected to outperform other major economies over the next 12 months. The growth of the real estate sector will moderate.

1 Explain the expected impact on UK asset classes of each of Young's economic conclusions.
2 Demonstrate and justify the direction of judgmental modifications that Young might make to the baseline forecasts of her firm.

Solution to 1:

Young reaches the conclusions shown in Exhibit 31.

Exhibit 31 Market/Asset Class Conclusions

Economic Conclusion	Market/Asset Class Impact
Consumers	Consumers are creating a positive investment environment for corporate profits and therefore for UK equities and credit quality. However, if spending rises much more steeply than anticipated, we might expect upward pressure on both short-term and long-term interest rates that would be a negative for bonds and real estate.
Business	The economic conclusion is a positive for UK equities and bonds via improved credit quality. However, predicted business growth may put upward pressure on wages, costs, and inflation rates.
Government	The government sector conclusion is a slight positive for the economy at this time.
Central bank	The expected steady interest rate environment is a positive factor for the UK investment market. If the economy expands too quickly, there may be pressure from the central bank for higher interest rates looking out 12 months.
Inflation rates	The current stable inflation picture should have a positive impact on the economy and on the financial assets we are comparing in this analysis.
Other/Unique	Returns to UK real estate should moderate from unusually high rates.

Solution to 2:

The arrows in Exhibit 32 indicate the direction of adjustment to the baseline forecasts.

Exhibit 32 Modifications to Capital Market Forecasts for UK Asset Classes

Asset Class	Average Annual Returns	Average Annual Standard Deviation
UK equities	↑	↓ or →
Intermediate bonds	→	↑
Long-term bonds	→	↑
Real estate	↓	↑

The growth outlook for consumers and businesses is a strong positive for UK equities. The steady central bank, government, and inflation outlooks suggest below-average or at least unchanged volatility. The outlook of steady interest rates is neutral for intermediate- and long-term bonds, but the uncertainty about the economy overheating suggests an increase in risk. Real estate's returns should decrease from the high baseline forecast. The break with the past trend growth should translate into higher risk for real estate.

4.6.8 *Currencies*

The exchange rate between two countries reflects the balance of buyers and sellers. One major reason for buying and selling foreign currencies is to facilitate trade in goods and services (exports and imports). If a country begins to import more, the currency will tend to depreciate (all else being equal). Hence, considerable attention is usually paid to determining a competitive exchange rate at which the trade balance—or, more broadly, the current account balance (which includes services and transfers)—is zero. Governments and central banks are often concerned with maintaining a competitive exchange rate and may try to do so by buying or selling foreign currencies or by raising or lowering interest rates.

However, trade is only one motive for purchases and sales of foreign currency and it has become relatively less important. The other motive is international flows of capital. Companies wishing to invest in a country are likely to be buyers of the currency as they bring in capital to build assets. Strong domestic economic growth and an opening of new industries to foreign ownership are two possible drivers of a rise in foreign direct investment that will likely push up a currency too. The liberalization of capital flows and the increasing trend toward global diversification mean that there may also be inflows to buy local stocks, bonds, or short-term instruments, including deposits. These flows can be volatile and may quickly reverse. The foreign direct investment is likely to be more stable.

Portfolio flows may be influenced by the growth of the economy or by domestic interest rates. When interest rates are high, inflows are likely to be higher and the currency value rises. Conversely, falling interest rates often weaken a currency. However, the link between interest rates and the currency sometimes works the other way. This is because investors may see higher interest rates as slowing the economy. If a currency departs from the level that equilibrates trade for a long time, the resulting deficits or surpluses may eventually become too large for capital flows to finance. Among the major currencies, there are often prolonged over- and undervaluations around a long-term equilibrium level. For this reason, many governments in emerging countries use some combination of capital controls and currency management (pegs, currency boards, managed floats, etc.) to try to keep the currency competitive. This approach tends to lead to stability for extended periods punctuated by periodic sudden, large movements.

EXAMPLE 35

A Currency Example

Between 1990 and July 1997, the Bank of Thailand managed the Thai baht in a narrow range. Over time, a gradual loss of competitiveness through higher inflation pushed the current account deficit up to 8 percent of GDP, financed by strong capital inflows. In 1996, the economy slowed and capital inflows faltered, prompting speculation that the baht might fall. The central bank intervened heavily to defend the baht in early 1997 but by midyear had exhausted its reserves and was forced to float the currency. Within a few months, the baht halved in value and other currencies in the region were also under pressure. The Asian crisis of 1997 had begun.

Forecasting exchange rate movements is widely viewed as especially difficult. For this reason, some investors try to fully hedge currency exposure. Others see opportunities in currency forecasting because of the volatility of many exchange rates and the highly liquid markets. The following sections review the major approaches to exchange rate forecasting.

4.6.9 *Approaches to Forecasting Exchange Rates*

There are four broad approaches to forecasting exchange rates, and most forecasters probably use a combination of them all.

4.6.9.1 Purchasing Power Parity **Purchasing power parity** (PPP) asserts that movements in an exchange rate should offset any difference in the inflation rates between two countries.[76] PPP reflects the idea that exchange rates should find a level that keeps different countries broadly competitive with each other.

To illustrate PPP, suppose that over a five-year period, Canadian prices are forecast to increase by 10.41 percent (equal to 2 percent annual inflation) while Eurozone prices are forecast to show 15.93 percent growth (equal to 3 percent annual inflation). Over the five-year period, the Canadian–Eurozone inflation differential is 10.41% − 15.93% = −5.52%. PPP would predict that the Canadian dollar will appreciate against the euro by approximately the same percentage. For example, if the exchange rate is C\$1.3843 per euro, PPP would predict an exchange rate of approximately (1 − 0.0552) × (C\$1.3843 per euro) = C\$1.3079 per euro.

PPP in broad terms does seem to be useful in the long run—say, over periods of five years or longer. Furthermore, governments and central banks take PPP very seriously in their approach to exchange rates because periods of under- or overvaluation of a currency may lead to sudden exchange rate instability or be destabilizing for business.

However, with the huge rise in capital flows over the last three decades, exchange rates can depart from PPP levels for long periods of time. PPP is often not a useful guide to the direction of exchange rates in the short or even medium run (up to three years or so). There are also times when factors other than PPP dominate exchange rate movements. This usually happens when a large current account deficit is opening up and the markets question whether a growing deficit can be financed. Markets then focus on what level of the currency is needed to correct the deficit.

4.6.9.2 Relative Economic Strength The **relative economic strength forecasting approach** focuses on investment flows rather than trade flows. It suggests that a strong pace of economic growth in a country creates attractive investment opportunities, increasing the demand for the country's currency and causing it to appreciate. Sometimes, demand comes from a higher short-term deposit rate in that country combined with an expectation that the currency will stay the same or appreciate. More recently, the focus has been on the pace of economic growth and the existence of attractive investment opportunities in general.

When interest rates are relatively high in a country, capital moves into that country and, as a result, the currency strengthens. Even if investors begin to see the exchange rate as overvalued in some long-term sense, they may still be content if they feel the

76 The definition refers to relative PPP, the form of PPP most economists are concerned with. See Solnik and McLeavey (2004) for further details.

extra yield compensates for that overvaluation. However, once the exchange rate reaches an excessive level, they will question whether the high yield is enough to justify the likely exchange rate depreciation.

What is the role of short rates? There is little question that short-term interest rates can influence exchange rates but primarily over a short-term time horizon. The level of short-term interest rates influences the extent to which speculators are willing to bet against a currency. If interest rates in a particular country are especially high, speculators are less likely to short that currency because that currency will probably strengthen as a result of the higher interest rates. Similarly, very low interest rates on Japanese yen have periodically encouraged investors to borrow yen to fund other investments (a so-called carry trade).

It can be helpful to combine the PPP and relative strength approaches. The relative strength approach indicates the response to news on the economy but does not tell us anything about the level of exchange rates. The PPP approach indicates what level of the exchange rate can be regarded as a long-term equilibrium. By combining the two, we can generate a more complete theory.

4.6.9.3 Capital Flows The **capital flows forecasting approach** focuses on expected capital flows, particularly long-term flows such as equity investment and foreign direct investment (FDI). Inflows of FDI into a country increase the demand for the country's currency, all else being equal.

From 1999 onward, there was considerable focus on this approach because of the surprising strength of the dollar versus the euro. This situation coincided with a clear increase in long-term flows from the Eurozone to the United States, especially FDI and US equities. This capital was being attracted into the United States by the boom in the domestic economy and the attractiveness of equity assets, particularly in the internet and technology sectors, until at least 2001.

Note that long-term capital flows may have the effect of reversing the usual relationship between short-term interest rates and the currency. This is explained by the fact that a cut in short-term rates would be expected to boost economic growth and the stock markets, thereby making long-term investments more attractive. In this environment, central banks face a dilemma. Whereas they might want to raise interest rates to respond to a weak currency that is threatening to stimulate the economy too much and boost inflation, the effect may actually be to push the currency lower. Hence, the effectiveness of monetary policy is much reduced.

This appeared to be a problem for the Eurozone at times during 2001. As the economy slowed, the ECB was reluctant to cut interest rates because of rising inflation and a weak currency. The inaction seemed to make the currency weaker. Similarly, the Fed's aggressive cutting in the first half of 2001 pushed the dollar higher, which attracted capital and thus reduced the impact of lower interest rates in stimulating the economy.

4.6.9.4 Savings–Investment Imbalances The **savings–investment imbalances forecasting approach** explains currency movements in terms of the effects of domestic savings–investment imbalances on the exchange rate. Although it is not easy to use for forecasting, this approach can sometimes help with understanding why currencies depart from equilibrium for long periods. It starts from the fact that the current account deficit of a country is the sum of its government deficit and private sector deficit. For example, in the United States in 2004, the current account deficit was estimated at about 6 percent of GDP, with the government deficit at about 5 percent and the private sector deficit at about 1 percent. In contrast, Japan had a current account surplus of 4 percent of GDP, with a government deficit of 8 percent balanced by a private sector surplus of 12 percent of GDP. So, in Japan, the private sector was financing the government deficit as well as an outflow of capital.

However, if the private sector or government currency-related trends change, then the current account position must change too and the exchange rate moves to help achieve that. Suppose that an economy suddenly begins to expand rapidly, driven by a new government budget deficit or bullish entrepreneurs. If domestic savings do not change, there will be excess demand for capital as investment tries to exceed savings. The only way that investment can exceed savings in reality is for foreign savings to be used, since the accounts have to balance. But this solution requires a deficit on the current account of the balance of payments.

So, where does this deficit on the current account come from? Some of it may arise simply because imports are strong due to the buoyant economy or because exports are weak as companies focus on the domestic market. But if that is not enough, the exchange rate needs to rise. If capital flows are attracted to the country, either due to high interest rates or due to attractive expected returns on investments, then the exchange rate will indeed rise as needed.

Because trade takes time to adjust, the exchange rate will frequently depart far from generally accepted equilibrium rates for prolonged periods, typically two to four years. Eventually, the rising currency will widen the current account deficit sufficiently and the domestic currency may start to decline. Of course, it needs to stay reasonably strong as long as domestic investment exceeds savings.

If the economy becomes weak enough at this point and domestic investments no longer exceed domestic savings, then the currency will also weaken. To return to a current account surplus, the exchange rate may need to drop to a level well below its equilibrium rate. Hence, there is a risk that the currency could swing sharply to an undervalued position.

EXAMPLE 36

The USD/Euro Exchange Rate, 1999–2004

The euro was first established as a currency at the beginning of 1999. To the surprise of nearly everyone, it started trading weakly, its value against the US dollar falling from about US$1.17 to a low of US$0.82 in late 2000. But beginning in 2001 and accelerating in 2002–2004, the dollar fell. In 2004, the euro reached US$1.37. On a PPP basis, the euro probably lies in the range of US$1.10-US$1.20. So, in the course of five years, the exchange rate cycled around that level. These swings can be considered according to the three explanations below.

Relative economic strength: This approach explains why the dollar rose strongly in 1999–2000, with faster economic growth and consequent higher interest rates in the United States. In 2001, the fact that the US economy was weaker than that of the Eurozone helps to explain why the dollar peaked and went sideways during that year. The explanation breaks down in 2002–2003, however. Despite the superior performance of the US economy over the Eurozone in 2002 and beyond, the dollar retraced its path all the way back to its starting point.

Capital flows: This approach explains more about the dollar's moves. The dollar's strength in 1999–2000 was matched by massive long-term inflows into the United States in the form of foreign direct investment and equity purchases. In 2001, these flows fell off rapidly, though there were still large inflows into US bonds. The current account deficit had expanded by then as a result of the strong dollar, so the capital flows were no longer large relative to the necessary inflow to finance the deficit. Hence, the dollar's decline.

Savings–investment imbalances: During 1999–2000, the US economy grew very rapidly with pressure to reduce domestic savings and increase investment. Households reduced savings, encouraged by low and falling unemployment and the rising stock prices. Businesses cut savings because they saw major new investment opportunities. The result was a soaring US dollar opening up the current account deficit, further encouraged by the inflow of capital described above. In 2001–2002, the private sector deficit was slashed drastically as companies cut back on borrowing and spending. The government cut taxes and shifted the government accounts from surplus to deficit. But the dollar still fell back against the euro because the current account deficit needed to be in the 4–6 percent range to balance the internal savings balances. Hence, the dollar fell back from its still-overvalued position and nosed into undervalued territory.

This approach, if correct, suggests that the dollar's weakness (at least in 2004–2005) may be limited by the continued large government borrowing requirement and low private savings. Beyond 2005, however, the dollar could reach past lows if either the government makes a major effort to reduce the budget deficit or an economic slowdown prompts increased private savings.

4.6.10 *Government Intervention*

Since the developed world moved to floating rates in the early 1970s, periodic attempts have been made to control exchange rates. However, economists and the markets have been skeptical about whether governments really can control exchange rates with market intervention alone because of three factors. First, the total value of foreign exchange trading, in excess of US$1 trillion *daily*, is large relative to the total foreign exchange reserves of the major central banks combined. Second, many people believe that market prices are determined by fundamentals and that government authorities are just another player. Third, experience with trying to control foreign exchange trends is not encouraging in the absence of capital controls. Unless governments are prepared to move interest rates and other policies, they cannot expect to succeed.

4.7 Information Sources for Economic Data and Forecasts

Having presented economic analysis for capital market expectations setting, we can indicate several fruitful sources for gathering economic data. The sources we present link to many other useful resources.

The main sources of leading indicators are the Conference Board and national sources. Most other economic data also come from national statistical sources, such as central banks and government statistical offices. Some survey data come from other organizations, such as the Institute of Supply Management. Useful international sources include the OECD, IMF, and World Bank. A list of websites is provided below.

Forecasts from econometric models are published by governments. The OECD publishes forecasts twice a year in its *Economic Outlook* reports. The IMF publishes forecasts, and various private sector forecasters also publish forecasts, though these are sometimes proprietary. Exhibit 33 summarizes some sources for researching US markets.

Exhibit 33 A Selection of Data Sources for Researching US Markets

Categories of Economic Interest	Factor Measures	Data Use LT	Data Use ST	Data Source
Economic fundamentals	Measures of economic output/growth (e.g., GDP, industrial production)	✓		www.bea.gov Bloomberg
	General price level stability	✓		

(continued)

Exhibit 33 (Continued)

Categories of Economic Interest	Factor Measures	Data Use LT	Data Use ST	Data Source
Consumers	Employment/unemployment	✓		Bloomberg
	Measures of consumption/income	✓		www.bea.gov
	Measures of savings, investment, and leverage	✓		
	Measures of sentiment	✓		U. of Michigan Survey
Business	Measures of profitability	✓		www.bea.gov
	Measures of productivity	✓		Federal Reserve Bank
	Industry price level stability	✓	✓	Internal or third-party research; Trade pub.
	Capacity utilization rates		✓	
Central bank	Measures of monetary policy	✓	✓	www.stls.frb.org
	General price level stability (inflation)	✓	✓	Bloomberg
	Assessment of central bank independence		✓	Internal analysis
Government	Fiscal policy	✓	✓	Congressional Budget
	Assessment of exchange rate stability/trends	✓	✓	Office; Bloomberg Internal; www.wto.org
	Measures of political stability	✓		Internal analysis
	Assessment of legal system's ability to protect assets (including intangible assets) and ability to settle disputes (due process)	✓		Internal analysis
Economic technical factors	Capital flows		✓	Internal/third-party research;
	Sector/industry supply and demand		✓	
				Trade publications
Market fundamentals	Rates of return	✓	✓	Relative (industry) internal research;
	Valuation trends (e.g., equity P/E multiples)	✓	✓	
	Asset class price volatility			Third-party research; Trade publications
	Large-cap equities	✓		
	Corporate bonds vs. overall market	✓	✓	
	Short sovereign debt		✓	
	Exchange rate movements	✓		
Market technical factors	Ratio of advances/declines in equity market		✓	Reuters;
	Corporate bond issuance (market yield)		✓	Internal research
Other: unique; social; political	Demographic influences	✓	✓	Internal; third-party
	Seasonal patterns of consumption	✓	✓	Third-party; Trade pub.
	Current account trends; net exports versus imports	✓	✓	Bloomberg

Some additional useful resources include the following:

- www.imf.org
- www.worldbank.org
- www.oecd.org
- www.federalreserve.gov
- www.ecb.int

Exhibit 33 (Continued)

- www.bankofengland.co.uk
- www.boj.or.jp/en (the English site for the Bank of Japan)
- www.bis.org
- www.nber.org (the website of the National Bureau of Economic Research, the US organization that dates business cycles; it contains useful data and research on past business cycles)

SUMMARY

In this reading, we have discussed how investment professionals address the setting of capital market expectations.

- Capital market expectations are essential inputs to deciding on a strategic asset allocation. The process of capital market expectations setting involves the following steps:
  1. Specify the final set of expectations that are needed, including the time horizon to which they apply.
  2. Research the historical record.
  3. Specify the method(s) and/or model(s) that will be used and the information needs for developing expectations.
  4. Determine the best sources for information needs.
  5. Interpret the current investment environment using the selected data and methods, applying experience and judgment.
  6. Formulate the set of expectations that are needed, documenting conclusions.
  7. Monitor actual outcomes and compare to expectations, providing feedback to improve the expectations-setting process.
- Among the challenges in setting capital market expectations are *the limitations of economic data* (including lack of timeliness as well as changing definitions and calculations); *data measurement errors and biases* (including transcription errors, survivorship bias, and appraisal [smoothed] data); *the limitations of historical estimates* (including nonstationarity); *ex post risk as a biased risk measure* (historical prices may reflect expectations of a low-probability catastrophe that did not occur); *biases in analysts' methods* (including data-mining bias and time-period bias); *the failure to account for conditioning information*; *the misinterpretation of correlations*; and *psychological traps* (including the anchoring trap, the status quo trap, the confirming evidence trap, the overconfidence trap, the prudence trap, and the recallability trap); and *model uncertainty*.
- The tools for formulating capital market expectations include formal tools, survey and panel methods, and judgment. Formal tools include statistical tools, discounted cash flow models, the risk premium approach, and financial market equilibrium models. Analyst judgment includes economic and psychological insight.

- Economic output has cyclical and growth trend components. The cyclical components include the inventory cycle (measured in terms of fluctuation of inventory) and the business cycle (representing fluctuations in GDP in relation to long-term trend growth). A typical business cycle has five phases: initial recovery, early upswing, late upswing, slowdown, and recession. Each of the two cyclical components has implications for variables such as interest rates and corporate profits, which are important for capital market expectations. The economic trend growth component (the long-term growth path of GDP) is important particularly for setting long-term expectations.
- Consumer spending is typically the most important factor affecting GDP (it often accounts for 60–70 percent of GDP). Retail sales and consumer consumption are closely watched for indications of consumer spending.
- Business investment has a smaller weight in GDP than consumer spending but is more volatile. Data on business investment and spending on inventories reveal recent real business activity.
- Fiscal policy and monetary policy are means by which governments attempt to influence the business cycle.
- Monetary policymakers often target inflation rates and use the central bank's influence over interest rates to achieve policy goals. The Taylor rule gives the optimal short-term interest rate as the neutral rate plus an amount that is positively related to the excess of the GDP and inflation growth rates above their respective trend and target values.
- If monetary and fiscal policies are both tight, the yield curve is typically inverted. When monetary policy is tight but fiscal policy is loose, the yield curve tends to be flat. An inverted yield curve has often preceded a recession.
- In managing cash and equivalents, central bank actions are closely watched.
- For investors buying and selling nominal default-free bonds for the short term, developments in the business cycle and changes in short-term interest rates must be closely watched. News of stronger economic growth usually makes bond yields rise. For long-term investors, inflation expectations are of great importance. For holders of corporate bonds and other defaultable debt, the spread over Treasuries in relation to the business cycle is an important factor.
- Investing in emerging market debt involves special considerations, such as country risk analysis. Emerging market governments borrow in a foreign currency and so cannot simply inflate their way out of a problem in servicing the debt; this limitation increases the risk of default.
- Inflation-indexed bonds are not exposed to the risk of unexpected inflation. Generally, yields on such instruments rise with real economic growth and the level of short-term interest rates.
- Investors in common shares should analyze economic factors, first, in the way that they affect company earnings and, second, in the way that they affect interest rates, bond yields, and liquidity. The trend growth in the aggregate economy largely determines the trend growth in aggregate corporate earnings.
- During the economic cycle, the P/E ratio tends to be high and rising when earnings are expected to rise but low and falling if the outlook for earnings worsens. P/E ratios are usually lower for an economy stuck on a slower growth path. High inflation often tends to depress P/E ratios.

- Among the systematic factors affecting real estate returns are growth in consumption, real interest rates, the term structure of interest rates, and unexpected inflation.
- Among the factors affecting exchange rate movements are purchasing power parity, relative economic strength, capital flows, and savings–investment imbalances.

REFERENCES

Arnott, Robert, and Peter Bernstein. 2002. "What Premium is 'Normal'?" *Financial Analysts Journal* 58 (2): 64–85.

Baumohl, Bernard. 2005. *The Secrets of Economic Indicators*. Upper Saddle River, NJ: Pearson Education, Inc.

Bekaert, Geert, Robert Hodrick, and David Marshall. 2001. "Peso Problem Explanations for Term Structure Anomalies." *Journal of Monetary Economics* 48 (2): 241–70.

Bernstein, Peter L. 2004. "A Do-It-Yourself Forecasting Kit Updated." *Financial Analysts Journal* 60 (6): 27–32.

Black, Fischer, and Robert Litterman. 1991. "Asset Allocation: Combining Investor Views with Market Equilibrium." *Journal of Fixed Income* 1 (2): 7–18.

Black, Fischer, and Robert Litterman. 1992. "Global Portfolio Optimization." *Financial Analysts Journal* 48 (5): 28–43.

Bodie, Zvi, Alex Kane, and Alan Marcus. 2017. *Investments*. 11th ed. New York: McGraw-Hill Education.

Bollerslev, Tim, Robert Engle, and Daniel Nelson. 1994. "ARCH Models." *Handbook of Econometrics*, vol. 4. Robert Engle and Daniel McFadden, eds. Amsterdam: Elsevier.

Brinson, Gary, Jeffrey Diermeier, and Gary Schlarbaum. 1986. "A Composite Portfolio Benchmark for Pension Plans." *Financial Analysts Journal* 42 (2): 15–24.

Brown, Stephen, William Goetzmann, and Stephen Ross. 1995. "Survival." *Journal of Finance* 50:853–73.

Calverley, John. 2004. *Bubbles and How to Survive Them*. London: Nicholas Brealey Publishing.

Campbell, John, and Robert Shiller. 1991. "Yield Spreads and Interest Rate Movements." *Review of Economic Studies* 58 (3): 495–514.

Cochrane, John. 1999a. "New Facts in Finance." *Economic Perspectives*. Federal Reserve Bank of Chicago, vol. 23, no. 3:36–58. (Revision of NBER Working Paper 7169).

Cochrane, John. 1999b. "Portfolio Advice for a Multifactor World." *Economic Perspectives*. Federal Reserve Bank of Chicago, vol. 23, no. 3:59–78. (Revision of NBER Working Paper 7170).

Diebold, Francis. 2006. *Elements of Forecasting*. 4th ed. Mason, OH: South-Western Cengage Leraning.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2002. *Triumphs of the Optimists: 101 Years of Global Investment Returns*. Princeton, NJ: Princeton University Press.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2006. *Global Investment Returns Yearbook 2006*. ABN-AMRO.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2011. *Equity Premiums around the World*. Charlottesville, VA: Research Foundation of CFA Institute.

Drost, Feike, and Theo Nijman. 1993. "Temporal Aggregation of GARCH Processes." *Econometrica* 61 (4): 909–27.

Elton, Edwin, Martin Gruber, Deepak Agrawal, and Christopher Mann. 2001. "Explaining the Rate Spread on Corporate Bonds." *Journal of Finance* 56 (1): 247–77.

Estrella, Arturo, and Frederic Mishkin. 1998. "Predicting U.S. Recessions: Financial Variables as Leading Indicators." *Review of Economics and Statistics* 80 (1): 45–61.

Ferson, Wayne, and Rudi Schadt. 1996. "Measuring Fund Strategy and Performance in Changing Economic Conditions." *Journal of Finance* 51 (2): 425–61.

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press.

Goetzmann, William, and Philippe Jorion. 1999. "Re-Emerging Markets." *Journal of Financial and Quantitative Analysis* 34 (1): 1–32.

Goodall, Thilo, Antonio Manzini, and Thomas Rose. 1999. "Risk Premium Project." Working Paper, UBS Global Asset Management.

Graham, John R., and Campbell R. Harvey. 2001. "The Theory and Practice of Corporate Finance: Evidence from the Field." *Journal of Financial Economics* 60 (2–3): 187–243.

Granger, Clive. 1969. "Investigating Causal Relationships by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (3): 424–38.

Grinold, Richard, and Ronald Kahn. 1995. *Active Portfolio Management*. Chicago, IL: Probus Publications.

Grinold, Richard, and Kenneth Kroner. 2002. "The Equity Risk Premium." *Investment Insights from Barclays Global Investors*, vol. 5, no. 3.

Hamilton, James. 1994. *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hammond, John, Ralph Keeney, and Howard Raiffa. 1998. "The Hidden Traps in Decision Making." *Harvard Business Review* 76 (5): 47–58.

Ibbotson, Roger, and Peng Chen. 2003. "Long-Run Stock Returns: Participating in the Real Economy." *Financial Analysts Journal* 59 (1): 88–98.

Ilmanen, Antti. 2003. "Expected Returns on Stocks and Bonds." *Journal of Portfolio Management* 29 (2): 7–27.

Ilmanen, Antti, Rory Byrne, Heinz Gunasekera, and Robert Minikin. 2002. *Stocks versus Bonds: Balancing Expectations and Reality*. Schroder Salomon Smith Barney.

Jagannathan, Ravi, Ellen McGrattan, and Anna Scherbina. 2000. ""The Declining U.S. Equity Premium." *Quarterly Review*." *Federal Reserve Bank of Minnesota* 24 (4): 3–19.

Kurz, Mordecai. 1994. "On the Structure and Diversity of Rational Beliefs." *Economic Theory* 4 (6): 877–900.

Kurz, Mordecai, Hehui Jin, and Maurizio Motolese. 2005. "Determinants of Stock Market Volatility and Risk Premia." *Annals of Finance* 1 (2): 109–47.

Lally, Martin, Melvin Roush, and Tony Van Zijl. 2004. "The Market Risk Premium in New Zealand—Survey Evidence." *INFINZ Journal* 1:5–12.

Ledoit, Olivier, and Michael Wolf. 2003. "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection." *Journal of Empirical Finance*, vol. 2003, no. 5:603–621.

Ling, David, and Andy Naranjo. 1997. "Economic Risk Factors and Commercial Real Estate Returns." *Journal of Real Estate Finance and Economics* 14 (3): 283–307.

Ling, David, and Andy Naranjo. 1998. "'The Fundamental Determinants of Commercial Real Estate Returns." *Real Estate Finance* 14 (4): 13–24.

Luenberger, David. 1998. *Investment Science*. New York: Oxford University Press.

McQueen, Grant, and Steven Thorley. 1999. "Mining Fools Gold." *Financial Analysts Journal* 55 (2): 61–72.

Mehra, Yash, and Jon Peterson. 2005. ""Oil Prices and Consumer Spending." Federal Reserve Bank of Richmond." *Economic Quarterly* 91 (3): 51–70.

Merton, Robert. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41 (5): 867–87.

O'Neill, Jim, Dominic Wilson, and Rumi Masih. 2002. "The Equity Risk Premium from an Economics Perspective." Goldman Sachs Global Economics Paper No. 84.

Rapaport, Alfred, and Michael Mauboussin. 2001. *Expectations Investing: Reading Stock Prices for Better Returns*. Boston, MA: Harvard Business School Press.

Salomons, Roelof, and Henk Grootveld. 2002. "The Equity Market Risk Premium: Emerging vs. Developed Markets." *University of Groningen, Research Report* 4 (2): 121–44.

Sharpe, William. 1974. "Inputting Expected Security Returns from Portfolio Composition." *Journal of Financial and Quantitative Analysis* 9 (3): 463–72.

Siegel, Jeremy. 2002. *Stocks for the Long Run*. 3rd ed. New York: McGraw-Hill.

Singer, Brian, and Kevin Terhaar. 1997. *Economic Foundations of Capital Market Returns*. Charlottesville, VA: Research Foundation of AIMR.

Solnik, Bruno, and Dennis McLeavey. 2004. *International Investments*. 5th ed. Boston: Pearson Addison-Wesley.

Staub, Renato. 2005. "Capital Market Assumptions." UBS white paper, February.

Staub, Renato. 2006. "Multilayer Modeling of a Market Covariance Matrix." *Journal of Portfolio Management* 32 (3): 33–44.

Staub, Renato, and Jeffrey Diermeier. 2003. "Segmentation, Illiquidity, and Returns." *Journal of Investment Management* 1 (1): 135–51.

Stein, Charles. 1956. "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution." *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1:197–206.

Taylor, John. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39:195–214.

Terhaar, Kevin, Renato Staub, and Brian Singer. 2003. "The Appropriate Policy Allocation for Alternative Investments." *Journal of Portfolio Management* 29 (3): 101–10.

Welch, Ivo. 2000. "Views of Financial Economists on the Equity Premium and on Professional Controversies." *Journal of Business* 73 (4): 501–37.

Welch, Ivo. 2001. "The Equity Premium Consensus Forecast Revisited." Cowles Foundation Discussion Paper no. 1325.

Welch, Ivo. 2009. "The Results of the Equity Premium January 2009 Survey." Available at https://ssrn.com/abstract=1519544.

PRACTICE PROBLEMS

1 An analyst is assembling data for use in her firm's expectations-setting process. Several historical measures have been collected and used to set expectations on inflation and consumer consumption trends. Previously, only the most recent 25 years of historical data concerning these measures had been collected and analyzed. Now, an executive has suggested extending the starting point of the data 25 years further back to make the overall analysis more robust. Discuss why the inclusion of the additional data may present problems for the expectations-setting process despite the request's objective of making the analysis more robust.

2 Seth Bildownes is an analyst who has prepared forecasts regarding the current capital market environment. He recently gave his presentation to the managing directors of his firm. Excerpts of his presentation follow:

> "Noting that year-end holiday sales have been weak over the past several years, I believe that current expectations should be likewise muted. In fact, just last week, I had an occasion to visit Harrods and noticed that the number of shoppers seemed quite low. The last time I saw a retail establishment with so little pedestrian traffic at the beginning of December was in 1990, and that coincided with one of the worst holiday sales periods in the past 50 years. Thus, there will be no overall year-over-year retail sales growth this holiday season."

A Identify any psychological traps that may be interfering with the creation of Bildownes' forecasts.

B Recommend a way to mitigate the bias caused by any trap identified in Part A.

3 An investor is considering adding three new securities to his internationally focused fixed-income portfolio. The securities under consideration are as follows:

- 1-year US Treasury note (noncallable);
- 10-year BBB/Baa rated corporate bond (callable);
- 10-year mortgage-backed security (MBS) (government-backed collateral).

The investor will invest equally in all three securities being analyzed or will invest in none of them at this time. He will only make the added investment provided that the expected spread/premium of the equally weighted investment is at least 0.5 percent (50 bps) over the 10-year Treasury bond. The investor has gathered the following information:

Real risk-free interest rate	1.2%
Current inflation rate	2.2%
Spread of 10-year over 1-year Treasury note	1.0%
Long-term inflation expectation	2.6%
10-year MBS prepayment risk spread (over 10-year Treasuries)[a]	95 bps

(continued)

Practice Problems and Solutions: *Managing Investment Portfolios: A Dynamic Process,* Third Edition, John L. Maginn, CFA, Donald L. Tuttle, CFA, Dennis W. McLeavey, CFA, and Jerald E. Pinto, CFA, editors. © 2007 CFA Institute. All rights reserved.

10-year call risk spread	80 bps
10-year BBB credit risk spread (over 10-year Treasuries)	90 bps

[a] This spread implicitly includes a maturity premium in relation to the 1-year T-note as well as compensation for prepayment risk.

Using only the information given, address the following problems using the risk premium approach:

A Calculate the expected return that an equal-weighted investment in the three securities could provide.

B Calculate the expected total risk premium of the three securities, and determine the investor's probable course of action.

4 An Australian investor currently holds a A$240 million equity portfolio. He is considering rebalancing the portfolio based on an assessment of the risk and return prospects facing the Australian economy. Information pertaining to the Australian investment markets and the economy has been collected in the following table:

10-Year Historical	Current	Capital Market Expectations
10-yr avg govt. bond yield: 6.2%	10-yr govt. bond yield: 3.8 %	
Avg annual equity return: 8.2%	Year-over-year equity return: –9.4%	
Avg annual inflation rate: 2.8%	Year-over-year inflation rate: 2.6%	Expected annual inflation: 3.5%
Equity market P/E (beginning of period): 15.0x	Current equity market P/E: 14.5x	Expected equity market P/E: 14.0x
Avg annual income return: 2.0%		Expected annual income return: 1.5%
Avg annual real earnings growth: 6.0%		Expected annual real earnings growth: 5.0%

Using the information in the table, address the following problems:

A Calculate the historical Australian equity risk premium using the bond-yield-plus-risk-premium method.

B Calculate the expected annual equity return using the Grinold–Kroner model (assume no change in the number of shares outstanding).

C Using your answer to Part B, calculate the expected annual equity risk premium.

5 An analyst is reviewing various asset alternatives and is presented with the following information directly pertaining to the broad equity market in Switzerland and various industries within the Swiss market that are of particular investment interest.

Expected risk premium for overall global investable market (GIM) portfolio	3.5%
Expected standard deviation for the GIM portfolio	8.5%
Expected standard deviation for Swiss Health Care Industry equity investments	12.0%
Expected standard deviation for Swiss Watch Industry equity investments	6.0%
Expected standard deviation for Swiss Consumer Products Industry equity investments	7.5%

- Assume that the Swiss market is perfectly integrated with world markets.
- Swiss Health Care has a correlation of 0.7 with the GIM portfolio.
- Swiss Watch has a correlation of 0.8 with the GIM portfolio.
- Swiss Consumer Products has a correlation of 0.8 with the GIM portfolio.

A Basing your answers only upon the data presented in the table above and using the international capital asset pricing model—in particular, the Singer–Terhaar approach—estimate the expected risk premium for the following:

  i. Swiss Health Care Industry.

  ii. Swiss Watch Industry.

  iii. Swiss Consumer Products Industry.

B Judge which industry is most attractive from a valuation perspective.

6 Consider the information given in the following table:

Eurodollar Short Rates (as of Start of Month)

Month	1-Month Rates	6-Month Rates	Spread: 6-Mo. vs. 1-Mo.
Jan 20X0	5.71%	6.14%	43 bps
Feb 20X0	5.80	6.26	46
Mar 20X0	5.97	6.35	38
Apr 20X0	6.07	6.48	41
May 20X0	6.47	6.93	46
Jun 20X0	6.58	6.90	32
Jul 20X0	6.54	6.84	30
Aug 20X0	6.53	6.75	22
Sept 20X0	6.54	6.68	14
Oct 20X0	6.53	6.64	11
Nov 20X0	6.54	6.61	07
Dec 20X0	6.58	6.28	–30
Jan 20X1	5.48	5.16	–32
Feb 20X1	5.13	4.82	–31
Mar 20X1	4.98	4.58	–40

A Determine the implicit economic forecast in the interest rate data given.

B For a one-year holding period extending from 1 March 20X0 to 1 March 20X1, determine the relative merits of buying a six-month security now and then another one in six months, or purchasing a one-month Eurodollar security and then rolling that security over each month at the then-prevailing yield.

7 **A** How might an analyst use the data reflected below to confirm her suspicion that the country is currently experiencing an output gap?

Variable	3/31/2009	3/31/2010	3/31/2011
GDP (index)	129.0	128.5	128.0
Unemployment rate	10.5%	11.0%	11.5%
Capacity utilization rate	80.5%	80.0%	80.0%
Inflation rate	9.0%	8.5%	7.0%

B Given your response in Part A, would you expect inflation over the next year to accelerate or decline?

8 Based on the trends that may be calculated from the following economic measures, which of the countries below would be expected to achieve higher economic growth rates over the next year if current trends are sustained? Justify your response.

	Croatia (Millions of Kuna)		Czech Republic (Millions of Koruna)	
Economic Measures	**Qtr 1**	**Qtr 2**	**Qtr 1**	**Qtr 2**
Consumer spending	28	30	350	386
Business capital investment	12	13	205	250
Government investment/fiscal spending	10	11	110	140
Other miscellaneous GDP factors	−1	−2	−58	−111
Total GDP	49	52	607	665

9 **A** Based on targets for inflation and overall economic growth rate and on actual observations of inflation and economic growth rates, apply the Taylor rule to estimate what short-term interest rate level should be an appropriate target for monetary authorities.

B Explain why the monetary action suggested by the Taylor rule output may not actually be taken by central bank authorities.

Economic Measures	GDP Trend/Inflation Target (%)	Forecast (%)
Year-over-year increase in GDP	3.2	2.6
Inflation rate	2.0	4.0
Short-term interest rate (neutral value)	4.0	

10 Pharmavest is an investment advisory firm that focuses solely on companies within the health care (HC) sector. The firm conducts research and manages several commingled health sector funds. Pharmavest is conducting an analysis of health sector companies that have business exposures to the economies of Europe and the United States. The following tables show current and historical economic data and Pharmavest's forecasts of the most likely economic outcomes for the next year in Europe and the United States. In the tables, "Y/Y" is short for "year-over-year."

Using the economic categories shown in the tables, compare and contrast European and US economic trends and forecasts. Indicate and justify which economic region is expected to provide a relative advantage for the health care sector (Europe or the United States).

Europe

	3-Year Trend	1-Year Trend	Current Measure	1-Year Forecast
Broad economic output measure	Stable	Improving	1.3% Y/Y GDP growth in HC	3.0% Y/Y GDP growth in HC
Economic impact of consumers	Stable	Improving	8.9% consumer spending HC % of GDP	10.0% consumer spending HC % of GDP
Economic impact of businesses	Weakening	Improving	4.0% HC business profits 3.6% HC sales	5.5% HC business profits 8.0% HC sales
Economic impact of central bank	Improving/ stimulative	Stable	2.7% avg. short interest rates 2.6% inflation	2.8% avg. short interest rates 2.7% inflation
Economic impact of government	Stimulative	Stimulative but stable	5.0% government spending % of GDP (fiscal stimulus)	5.0% government spending % of GDP (fiscal stimulus)
Other unique economic factors, population growth, demographics	Stable	Stable	49 average age of aggregate population	49 average age of aggregate population

United States

	3-Year Trend	1-Year Trend	Current Measure	1-Year Forecast
Broad economic output measure	Stable	Improving	3.8% Y/Y GDP growth in HC	2.8% Y/Y growth in HC
Economic impact of consumers	Stable	Improving	9.6% consumer spending HC % of GDP	9.5% consumer spending HC % of GDP
Economic impact of businesses	Stable	Improving	5.2% HC business profits 9.0% HC sales	5.0% HC business profits 9.0% HC sales
Economic impact of central bank	Stimulative	Stable	1.0% avg short interest rates 2.2% inflation	2.2% avg short interest rates 2.5% Inflation
Economic impact of government	Stimulative	Increasingly stimulative	6.0% government spending % of GDP (fiscal stimulus)	6.5% government spending % of GDP (fiscal stimulus)
Other unique economic factors, population growth, demographics	Stable	Stable	44 average age of aggregate population	44 average age of aggregate population

11 Plim Ltd. is a manufacturing company in Finland that is a defined-benefit pension plan sponsor. Plim intends to increase its overall plan diversification by making an investment in Country X. The table below provides data for Country X indexes representing various economic variables.

Based only on the data presented, from the perspective of year-end 2011, indicate whether the 1-year trend for each of the economic variables is stronger than, weaker than, or the same as its 3-year and 20-year data trend growth rates.

Economic Index Data for Country X

Variable	Year-End 2008	Year-End 2009	Year-End 2010	Year-End 2011	20-Yr. L/T Average Annual % Increase
GDP	118.3	121.3	124.3	127.4	4.2
Consumer spending	1,569.2	1,596.2	1,584.3	1,647.7	2.5
Business spending	650.1	632.0	707.8	726.9	2.6
Inflation	2,749.8	2,901.1	3,133.1	3,446.4	14.3
Government spending (% of GDP)	16.2	16.5	16.0	15.8	3.6

12 **A** List five general elements of a pro-growth government structural policy.

B For each of the variables given below, describe the change or changes in the variable that would be pro-growth and determine the element of a pro-growth government structural policy that would best describe the change or changes.

- **i.** Tax receipts as a percent of GDP.
- **ii.** Government tariff receipts.
- **iii.** Number of publicly funded schools.
- **iv.** Number of state-owned businesses.
- **v.** Long-term average budget deficit as a percent of GDP.

13 Identify four differences between developed economies and emerging market economies.

14 In late 2011, K.C. Sung is planning an asset allocation strategy but would first like to assess Country M's current economic environment, then make a forecast for the economic conditions expected over the succeeding six- to nine-month period. Sung has learned that the leading indicator measures that he has compiled are quite indicative of current economic activity. However, Sung has seen over time that these specific measures impact many parts of the economy and thus are also predictive of potential longer-term (six- to nine-month) economic impacts as the initial economic activities create jobs and other beneficial output throughout Country M's economy.

Leading Indicators: Component Contributions and Total Index	June 2011	July 2011	August 2011 (Most Current Period)
Consumer orders growth	−0.01	0.02	0.07
Business capital goods orders growth	0.04	0.05	0.04
Central bank money supply growth	0.12	0.15	0.16
Total index value	111	115	116

Using the leading indicator approach to forecasting, draw a conclusion for Country M's economy for the next six to nine months using only the above table.

15 Other than changes in the rate of inflation, specify two factors that impact the yields available on inflation-indexed bonds.

16 J. Wolf is an individual investor who intends to make an additional investment in various South Korean–based assets based on the outcome of your capital market expectations-setting framework analysis. Your analysis should use the data provided in the following table. However, each measure should be analyzed independently of the other measures. While examining the forward-looking one-year forecast relative to the data provided for the recent trends, indicate whether the equity market impact and the corporate fixed-income market impact would be positive or negative. Justify your answer considering the likely risk premium impact that would result if the one-year forecast actually occurred.

Index Data (South Korea)	Current Index Measure	Index 1-Year Forecast	South Korean Equity Market Impact	Corporate Fixed-Income Market Impact
GDP	159	173		
Consumer spending	432	430		
Business profits	115	100		
Central bank money supply	396	455		
Government spending relative to tax receipts	1,385	1,600		

17 Discuss four approaches to forecasting exchange rates.

18 Looking independently at each of the economic observations below, indicate the country where an analyst would expect to see a strengthening currency for each observation.

	Country X	Country Y
Expected inflation over next year	2.0%	3.0%
Real (inflation-adjusted) government 10-year bond rate	4.8%	5.1%
Short-term (1-month) government rate	1.9%	5.0%
Expected (forward-looking) GDP growth over next year	2.0%	3.3%
New national laws have been passed that enable foreign direct investment in real estate/financial companies	Yes	No
Current account surplus (deficit)	8.0%	−1.0%

19 Fap is a small country whose currency is the fip. Ten years ago, the exchange rate with the Swiss franc (CHF) was 3 fips per 1 CHF, the inflation indexes were equal to 100 in both Switzerland and Fap, and the exchange rate reflected purchasing power parity (PPP). Now, the exchange rate is 2 fips per 1 CHF. The Swiss inflation index level is at 150, and the Fap inflation index is at 140.

A What should the current exchange rate be if PPP prevails?

B Are fips over- or undervalued, according to PPP, relative to CHF?

SOLUTIONS

1 The chief point is that extending the data series further back in time increases the risk of using data representing more than one regime. The analyst also needs to be aware of adjustments or revisions to the data, which can create inconsistencies in the data and make interpretation of those data difficult. The analyst must be sure that any adjustments to the data were made on a consistent and uniform basis. Furthermore, the analyst needs to be aware that variable definitions and calculation methods may have changed over the original and extended periods.

2 **A-B.** Drawing an inference that current sales expectations might be muted due to the weak sales numbers posted over the past several years is an example of the *status quo trap.* An objective assessment of early sales figures, surveys of shoppers, and collecting data about retail order patterns and shipments represent more unbiased bases for a year-end sales forecast.

The *confirming evidence trap* was evident when Bildownes used the recent observation of the number of customers in a single department store as further support of his forecast for weak year-end retail sales. To help prevent this bias, an analyst could undertake more observations and then honestly and independently assess them with equal rigor before drawing a conclusion.

Bildownes also seemed to be strongly influenced by his memory of previous weak year-end sales periods being associated with low pedestrian traffic during early December 1990. In assuming that a correlation witnessed in the past will repeat again without further analysis, Bildownes has fallen into the *recallability trap.* To counter this forecast-tainting bias, Bildownes should emphasize objective assessments of data, rather than personal memories, in forecasting.

In drawing the conclusion that "there will be no overall year-over-year retail sales growth this holiday season," Bildownes has fallen into the *overconfidence trap.* This trap relates to the natural tendency for individuals to be overconfident about the accuracy of their forecasts. The easiest way to help prevent this trap from biasing a forecast is to admit the possibility that the forecast may be inaccurate and to increase the range of possible outcomes around the primary target outcome.

3 **A**

	Real risk-free rate (%)	+	Expected inflation (%)	+	Spreads or premiums (%)	=	Expected annual fixed-income return (%)
1-year US T-note	1.2	+	2.6	+	0	=	3.8
10-year corp. bond	1.2	+	2.6	+	1.0 + 0.8 + 0.9	=	6.5
10-year MBS	1.2	+	2.6	+	0.95	=	4.75

Note: We assign the 10-year corporate a 1% maturity premium based on the 10-year over 1-year government spread.

Estimate of the expected return of an equal-weighted investment in the three securities: (3.8% + 6.5% + 4.75%)/3 = 5.02%

B The average spread at issue is [0 + (1.0% + 0.8% + 0.9%) + 0.95%]/3 = 1.22%. As 1.22% − 1% = 0.22% is less than 0.5 percent, the investor will not make the investment.

4 **A** The historical equity risk premium is 2.0%, calculated as follows:

Historical equity returns	–	Historical 10-Year government bond yield	=	Historical equity risk premium
8.2%	–	6.2%	=	2.0%

B The Grinold–Kroner model states that the expected return on equity is the sum of the expected income return (1.5%), the expected nominal earnings growth return (8.5% = 3.5% from inflation + 5.0% from real earnings growth), and the expected repricing return (–3.45%). The expected change in market valuation of –3.45% is calculated as the percentage change in the P/E level from the current 14.5x to the expected level of 14.0x: (14 – 14.5)/14.5 = –3.45%.

Thus, the expected return is 1.5% + 8.5% – 3.45% = 6.55%, or approximately 6.6 percent.

C Using the results from Part B, the expected equity risk premium is 2.8 percent.

Expected equity return	–	Current 10-Year government bond yield	=	Expected equity risk premium
6.6%	–	3.8%	=	2.8%

5 A Using the formula $\mathrm{RP}_i = \sigma_i \rho_{iM}\left(\frac{\mathrm{RP}_M}{\sigma_M}\right)$ we can solve for each expected industry risk premium. The term in brackets is the Sharpe ratio for the GIM, computed as 3.5/8.5 = 0.412.

i. $\mathrm{RP}_{\text{Health Care}}$ = (12)(0.7)(0.412) = 3.46%

ii. $\mathrm{RP}_{\text{Watch}}$ = (6)(0.8)(0.412) = 1.98%

iii. $\mathrm{RP}_{\text{Consumer Products}}$ = (7.5)(0.8)(0.412) = 2.47%

B Based on the above analysis, the Swiss Health Care Industry would have the highest expected return. However, that return is expected compensation for systematic risk. We cannot conclude which industry is most attractive from a valuation standpoint.

6 A The yield curve is inverting over the time period specified. This is an implicit forecast of an economic slowdown or a recession.

B Rolling over six-month Eurodollar securities would have provided superior results to rolling over one-month securities during the stated period. Extending the duration of the bond portfolio will be profitable when the yield curve subsequently flattens or inverts.

7 A All of the measures in the table would lead an analyst to conclude that the country is currently experiencing an output gap. The economy, as measured by GDP, has been contracting in each of the past two years. Thus, the economy has produced a higher level of output in the recent past. The unemployment rate has also been increasing steadily over the past two years and is quite high at over 10 percent. Thus, there is an ample supply of labor that could be put to work to increase economic output. Further, while the capacity utilization rate has been holding quite steady near 80 percent, it has declined from the reading posted two years ago. Thus, there is spare capacity that could be used to increase economic output. Finally, the decline in inflation in the latest year confirms that there is an output gap. The decline in inflation is important because otherwise the economic slowdown could be from an extreme overheating position and might not have opened up an output gap as yet.

B Given the conclusion that the country is experiencing an output gap, an analyst would expect a further decline in the rate of inflation in the next year.

8 The changes between Quarter 1 and Quarter 2 will be observed more clearly if we convert the economic measures from currency amounts into absolute percentage changes. The results are shown in the following table.

	Percentage Change: Qtr 1 to Qtr 2	
Economic Variable	**Croatia**	**Czech Republic**
Consumer spending	7.1%	10.3%
Business capital investment	8.3	22.0
Government investment/fiscal spending	10.0	27.3
Other miscellaneous GDP factors	100.0	91.4

Aggregate economic activity (including measures such as the components of GDP) is derived from many of the same factors from country to country. However, the proportion of economic activity derived from the multiple factors varies between economic regions and within a single region over time. Other issues that must be confronted when one analyzes across countries are the varying units of currency used and the differing levels of inflation between countries, which require an adjustment of some sort to be made before direct comparisons may be accomplished. Constant or inflation-adjusted measures may be used, and absolute currency measures may also be translated to a base currency to allow for direct comparisons. Alternatively, an analyst can examine the factors contributing to economic growth in proportion to an aggregate measure of activity such as GDP, as is shown in the following table.

	Absolute Currency Amounts Translated into Percentages of GDP			
	Croatia		**Czech Republic**	
Economic Variable	**Qtr 1**	**Qtr 2**	**Qtr 1**	**Qtr 2**
Consumer spending	57.1%	57.7%	57.7%	58.0%
Business capital investment	24.5	25.0	33.8	37.6
Government investment/fiscal spending	20.4	21.2	18.1	21.1
Other miscellaneous GDP factors	−2.0	−3.8	−9.6	−16.7

Finally, using the information in the two previous tables allows one to analyze more clearly how meaningful the recent changes in economic output are to a specific economy. If we multiply the percentage change in GDP from Quarter 1 to Quarter 2 by the percentage of GDP that that factor represents in Quarter 2 to a particular economy, we can more clearly gauge the importance of the economic trends in the economic output data that we have collected.

For example, the 7.1 percent increase in consumer spending in Croatia in Quarter 2 versus Quarter 1 is weighted quite heavily since the table above reflects that consumer spending represents about 57.7 percent of Croatia's total GDP as of the end of Quarter 2. Multiplying 7.1 percent by the 57.7 percent weighting that consumer spending represented to the Croatian economy produces a positive 4.1 percent weighted-average percentage change to the overall Croatian economy between Quarter 1 and Quarter 2. Here, we see the combined effect of underlying changes in each component of economic output and the relative importance of each component's change to overall economic output.

Economic Variable	Percentage Change: Qtr 1 to Qtr 2 (Weighted by Each Factor's GDP Percentage) Croatia	Czech Republic
Consumer spending	4.1%	6.0%
Business capital investment	2.1	8.3
Government investment/ fiscal spending	2.1	5.8
Other miscellaneous GDP factors	−3.8	−15.3
Sum of weighted percentage change: Qtr 1 to Qtr 2	4.5	4.8

Analyzing the weighted percentage changes in the economic measures reflected in the table above shows that the Czech Republic would be expected to achieve a slightly higher economic growth rate relative to Croatia over the next year if current trends are sustained. However, the large and growing drag (negative absolute values) that other miscellaneous GDP factors are placing on the Czech economy should be explored further. This drag is probably due to factors such as a large amount of imports (relative to exports) of goods and supplies required to maintain the current robust pace of the economy.

9 **A** The Taylor rule can be used to estimate the direction and magnitude of a short-term interest rate adjustment that could be made by central bank authorities. The calculation follows:

$$
\begin{aligned}
R_{\text{optimal}} &= R_{\text{neutral}} + \left[0.5 \times \left(\text{GDPg}_{\text{forecast}} - \text{GDPg}_{\text{trend}}\right) + 0.5 \times \left(I_{\text{forecast}} - I_{\text{target}}\right)\right] \\
&= 4.0\% + \left[0.5 \times \left(2.6\% - 3.2\%\right) + 0.5 \times \left(4.0\% - 2.0\%\right)\right] \\
&= 4.70\%
\end{aligned}
$$

B The Taylor rule is a simple yet useful tool that an analyst can use to determine from macroeconomic factors when it might be appropriate for central bank authorities to push down short-term interest rates in order to monetarily stimulate an economy that is functioning below potential (i.e., has an output gap). Conversely, if an economy is operating above its long-term average rate (substantial growth) and there are signs of accelerating inflation or rising inflation expectations, the Taylor rule can signal the need for central bank authorities to raise the level of short-term interest rates so that economic activity and increasing inflation expectations may be mitigated.

However, many other economic factors also affect aggregate output, inflation, and interest rates. Beyond economic factors, political and social factors come into the decision-making process when central bank authorities are setting interest rate targets. Central banks also typically take account of fiscal policies, wage behavior, asset prices, and developments in other economies. Periods of deflation would also call for more unorthodox central bank actions due to the fact that the central bank's ability to influence the economy is diminished once short-term interest rates approach zero during deflationary periods.

10 Comparing the broad economic output measures for Europe and the United States, the data show that both Europe and the United States have posted similar trends over the past few years and have had improving trends over the past year. The US economy has posted higher absolute levels of GDP health care growth, which is expected to moderate over the next year. However, overall output and health care-related output are expected to continue to increase in Europe over the next year. Advantage: Europe.

Recent consumer trends for Europe and the United States have also been quite correlated. However, the United States recently has registered higher absolute measures of overall consumer health care spending. Whereas this consumer impact is likely to be stable in the United States over the next year, Europe is expected to continue to show increased growth. Advantage: Europe.

Regarding the economic impact of businesses, the United States posted better results than Europe over the past three years. However, over the past year, improving conditions have been seen for both regions. Currently, business profits have rebounded more in the United States. Thus, absolute economic measures look stronger in the United States than in Europe in the near term. However, over the next year, business-related economic impacts should be stronger in Europe. Advantage: Europe.

Monetary authorities have been stimulative in both Europe and the United States over the past several years. However, the stimulation is expected to reverse in the United States over the next year, with short-term rates rising from 1 percent to 2.2 percent. Short-term rates are expected to rise only slightly over the next year in Europe. Thus, monetary policy will be a more neutral factor for both economies over the next year. The United States is expected to have lower overall interest rates and slightly lower inflation. Advantage: United States.

Fiscal policies for Europe and for the United States have been stimulative over the past few years but have been increasingly stimulative in the United States. Over the next year, the fiscal stimulation in the United States is expected to widen further in comparison to Europe. This trend is favorable to the United States as long as inflation is not an unexpected outcome of this fiscal stimulation and government borrowing caused by the fiscal stimulation does not crowd corporate borrowers out of the health care industry. Advantage: United States.

The higher average age of the aggregate general population in Europe relative to the United States is a positive factor for the health care industry from the perspective of Pharmavest. Advantage: Europe.

With more advantages pointing toward Europe's economies over the next year, Europe appears to be the economic region that should provide the stronger economic backdrop for the health care sector over the next year.

Matrix Summarization of Relative Advantages: Europe versus United States

	Recent Trends	Current Measures	1-Year Forecast
Broad economic measures	Similar (in Europe/US)	Similar	Advantage: Europe
Impact of consumers	Similar	Similar	Advantage: Europe
Impact of businesses	Advantage: US	Similar	Advantage: Europe
Impact of central bank	Similar	Similar	Advantage: US
Impact of government	Similar	Advantage: Europe	Advantage: US
Other/unique factors	Similar	Advantage: Europe	Advantage: Europe

11 To address the question, we first convert the data from the table into trend information (1-year, 3-year, and long-term 20-year trend information). The time periods are from the perspective of the end of 2011.

Trend Comparisons	1-Year Trend	3-Year Trend	20-Year Trend
GDP	2.5%	7.7% (2.5% compound average annual growth)	4.2%
Consumer spending	4.0	5.0 (1.6% compound average annual growth)	2.5
Business spending	2.7	11.8 (3.8% compound average annual growth)	2.6
Inflation	10.0	25.3 (7.8% compound average annual growth)	14.3
Gov't. spending (% of GDP)	–1.25	–2.5 (–0.8% compound average annual growth)	3.6

Based on the above comparisons, our conclusions about current trends witnessed over the past year for the economy of Country X relative to longer-term 3-year and 20-year trends are reflected in the table below.

Trend Analysis	1-Year Trend	3-Year Trend	20-Year Trend
GDP	2.5%	1-year trend is the same as the average annual GDP growth 3-year trend of 2.5%.	1-year trend of GDP growth is weaker than the 20-year trends.
Consumer spending	4.0	1-year trend is stronger than average annual consumer spending growth 3-year trend of 1.6%.	1-year trend is stronger than the average annual consumer spending growth 20-year trend of 2.5%.
Business spending	2.7	1-year trend is weaker than the average annual business spending growth 3-year trend of 3.8%.	1-year trend is stronger than the average annual business spending growth 20-year trend of 2.6%.
Inflation	10.0	1-year trend is weaker (higher inflation) than the average annual inflation growth 3-year trend of 7.8%.	1-year trend is stronger (lower inflation) than the average annual inflation growth 20-year trend of 14.3%.
Gov't. spending (% of GDP)	–1.25	1-year trend is weaker than the average annual govt spending growth 3-year trend of –0.8%.	1-year trend is weaker than the average annual govt spending growth 20-year trend of 3.6%.

12 **A** Five elements of a pro-growth government structural policy are as follows:

- Fiscal policy is sound. Fiscal policy is sometimes used to control the business cycle.
- The public sector intrudes minimally on the private sector.
- Competition within the private sector is encouraged.
- Infrastructure and human capital development are supported.
- Tax policies are sound.

B **i.** Declines in government tax receipts as a percent of GDP would be pro-growth because the equilibrium level of goods and services would increase. *Associated structural policy element: Tax policies are sound.*

ii. Declines in government tariff receipts would be pro-growth. Such declines would imply that government is fostering competition. By contrast, increases in tariff receipts would imply that government is protecting domestic businesses from international competition. *Associated structural policy element: Competition within the private sector is encouraged.*

- **iii.** Increases in the number of publicly funded schools would be pro-growth because businesses stand to gain from a well-educated workforce. *Associated structural policy element: Infrastructure and human capital development are supported.*
- **iv.** A negative net change in the number of state-owned businesses (that is, fewer such businesses) would be pro-growth. Such a change would increase the private sector's share of output, which would favor the efficient allocation of scarce resources. *Associated structural policy element: The public sector intrudes minimally on the private sector.*
- **v.** A decrease in the long-term average budget deficit as a percent of GDP is pro-growth, because it would be a positive for controlling the current account deficit. *Associated structural policy element: Fiscal policy is sound.*

13 Four factors usually associated with emerging market economies are as follows:

- Emerging market economies require high rates of investment, which is usually in short supply within the emerging economy itself. This situation creates a reliance on foreign capital. Areas of needed investment usually include both physical assets (capital equipment and infrastructure) and human capital (education and skills building).
- Emerging economies typically have volatile political and social situations. Leaders usually acquire and maintain power using force and other less-than-democratic means. The social environment is usually strained by the fact that a large portion of the population possesses few assets, has little formal education, and is unable to generate income to feed/support family and neighbors.
- To alleviate the first two factors above, organizations such as the IMF and World Bank provide conduits for external sources of investment and a means to push for structural reforms—political, social, educational, pro-growth, etc. The "conditions" usually prescribed by these institutions are often felt to be draconian.
- Emerging countries typically have economies that are relatively small and undiversified. Those emerging economies that are dependent on oil imports are especially vulnerable in periods of rising energy prices and can become dependent on ongoing capital inflows.

14 The consumer-oriented aspect of Country M's economy (as measured by consumer orders) has been consistently strengthening over the past several months. The consumer measure had a negative reading in June, a slightly positive reading in July, and an even stronger August reading. From this information, we can state that the consumer-oriented aspect of Country M's economy appears to be improving.

The business-oriented aspect of Country M's economy (as measured by business capital goods orders) has remained quite flat in the period reviewed. From this information, we can state that the business-oriented aspect of the Country M's economy currently appears to be a positive contributor but is not necessarily showing signs of improvement or weakness. Thus, in recent months, the measure has been a stable but positive contributor.

The central bank-oriented aspect of Country M's economy (as measured by central bank money supply) has also steadily improved over the period being analyzed.

All of the leading indicator components have been positively contributing to the economy, and most have been contributing at an increasingly positive rate (all except business capital goods orders). These components are a sign of current economic activity and also typically create additional economic activity by their very nature. In addition, the total index of leading indicators, which includes all the components, has likewise moved upward steadily over the past few months. Thus, we can conclude from the measures in the table that Country M's economy should show continued growth over the next six to nine months.

15 Two factors that affect the yields available on inflation-indexed bonds (IIBs) are as follows:

- Overall economic growth and its corresponding impact on real interest rates bear a direct impact on IIB yields. A growing economy places upward pressure on all bond yields. Though the impact may be muted due to the nature of the IIB structure, IIBs are not immune to interest rate risk.
- Investor demand for bonds in general and for IIBs in particular has an inverse impact on IIB yields. As with non-IIBs, rising investor demand serves to drive interest rates lower and the lack of investor demand drives up the yields that issuers must pay in order to sell the bonds they need to issue.

16

Index Data (South Korea)	Current Index Measure	Index 1-Year Forecast	South Korean Equity Market Impact	Corporate Fixed-Income Market Impact
GDP	159	173	Positive	Negative
Consumer spending	432	430	Negative	Positive
Business profits	115	100	Negative	Negative
Central bank money supply	396	455	Positive	Negative
Government spending relative to tax receipts	1,385	1,600	Positive	Negative

Justification:

GDP: A large 9 percent annual increase in GDP would give rise to strong corporate profits and would represent a favorable economic environment for equity investors (positive equity impact). However, such a strong economy would be a negative for corporate bond investors in that such economic growth and aggregate demand would place upward pressure on bond yields. In addition, in time, expectations of rising inflation could also hurt corporate bond investors (negative corporate bond impact).

Consumer Spending: A slight decrease in consumer spending represents an economic drag on overall economic output. This drag serves to reduce real earnings growth (negative equity impact). A slight decrease in consumer spending will not place any upward pressure on corporate bond yields or inflation or generally impact a debt issuer's ability to pay back the bondholders (positive corporate bond impact).

Business Profits: A 13 percent decline in business profits is a negative factor for both the equity market and the corporate bond market. Equity returns ultimately depend on businesses being able to earn a profit on the capital being employed. Thus, a steep decline in corporate profitability also increases the credit risk of corporate bondholders. Falling business profits can lead to corporate bond rating downgrades or insolvency (negative corporate bond impact).

Central Bank Money Supply: A 15 percent increase in money supply represents central bank monetary stimulation. Such stimulation should foster stronger economic growth (positive equity impact). However, bond yields could be expected to increase because monetary stimulation may increase expectations for higher aggregate growth and because of the potential higher inflation that monetary stimulation can cause over time (negative corporate bond impact).

Government Spending Relative to Tax Receipts (government budget deficit spending): A large increase in government spending relative to tax receipts (fiscal budget deficit) also represents stimulation to the economy. This stimulation can create an attractive environment for increasing corporate profits (positive equity impact). However, bond yields could be expected to increase because fiscal stimulation may increase expectations for higher aggregate growth and because of the potential higher inflation that monetary stimulation can cause over time (negative corporate bond impact).

17 Four approaches to forecasting exchange rates are as follows:

- PPP (or relative inflation rates), as exchange rate movements should offset inflation differentials.
- Relative economic strength, because a strong pace of economic growth tends to attract investment.
- Capital flows, as net inflows into a country, such as foreign direct investment, increase the demand for that country's currency.
- Savings-investment imbalances, through their ultimate effect on the need for foreign savings.

18

	Country X	Country Y
Expected inflation over next year	2.0% ✓	3.0%
Real (inflation-adjusted) government 10-year bond rate	4.8%	5.1% ✓
Short-term (1-month) government rate	1.9%	5.0% ✓
Expected (forward-looking) GDP growth over next year	2.0%	3.3% ✓
New national laws have been passed that enable foreign direct investment in real estate/financial companies	Yes ✓	No
Current account surplus (deficit)	8.0% ✓	−1.0%

Note: A ✓ represents the comparatively stronger measure, where an analyst could expect to see a strengthening currency based on the factor being independently reviewed.

19 **A** The Swiss price index has increased by (150 − 100)/100 = 50%. The Fap price index has increased by (140 − 100)/100 = 40%. The inflation differential is therefore 10 percent. According to PPP, to offset higher Swiss inflation, the fip should appreciate against the Swiss franc by approximately the same percentage to 0.90 × (3.0 fips per 1 CHF) = 2.70 fips per 1 CHF.

B Contrasting the exchange rate of 2.7 fips per 1 CHF implied by PPP and the actual exchange rate of 2 fips per 1 CHF, we see that the fip is overvalued relative to its PPP value: At actual exchange rates, fewer fips are required to purchase 1 CHF than PPP would predict.

READING

17

Equity Market Valuation

by Peter C. Stimes, CFA, and Stephen E. Wilcox, PhD, CFA

Peter C. Stimes, CFA, is a private investor in Fallbrook, California (USA). Stephen E. Wilcox, PhD, CFA, is at Minnesota State University, Mankato (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain the terms of the Cobb-Douglas production function and demonstrate how the function can be used to model growth in real output under the assumption of constant returns to scale;
☐	**b.** evaluate the relative importance of growth in total factor productivity, in capital stock, and in labor input given relevant historical data;
☐	**c.** demonstrate the use of the Cobb-Douglas production function in obtaining a discounted dividend model estimate of the intrinsic value of an equity market;
☐	**d.** critique the use of discounted dividend models and macroeconomic forecasts to estimate the intrinsic value of an equity market;
☐	**e.** contrast top-down and bottom-up approaches to forecasting the earnings per share of an equity market index;
☐	**f.** discuss the strengths and limitations of relative valuation models;
☐	**g.** judge whether an equity market is under-, fairly, or over-valued using a relative equity valuation model.

INTRODUCTION

1

Economic strength or weakness affects equity prices through its effect on risk-free rates, risk premiums, and corporate earnings. These economic drivers of security prices are often considered fundamental because they will affect security returns throughout most investment horizons. It is widely accepted that equity prices are negatively related to risk-free rates and risk premiums and positively related to earnings growth.

© 2010 CFA Institute. All rights reserved.

There are, of course, other drivers of equity returns and most of these can be considered behavioral. The cognitive and emotional factors experienced by investors can create both positive and negative feedback mechanisms. Market momentum may thus result in both bull market rallies and bear market declines lasting longer than may be justified by fundamental factors. This reading does not deal specifically with such behavioral drivers. Rather, this reading illustrates the application of economic forecasts to the valuation of equity markets. While many factors interact to determine whether equity prices are currently rising or falling, economic fundamentals will ultimately dictate long-term equity market price trends.

Section 2 uses GDP forecasts for a developing country, China, to develop inputs for a discounted cash flow valuation of that country's equity market.[1] Section 3 contrasts the top-down and bottom-up valuation approaches. Section 4 explains and critiques popular earnings- and asset-based models to relative equity market valuation. Section 5 summarizes the reading, and practice problems in the CFA Institute format follow.

2 ESTIMATING A JUSTIFIED P/E RATIO

Investors commonly use the market's price-to-earnings (P/E) ratio or multiple to gauge the prospects for future equity returns. Sections 2.1 through 2.3 develop the **Cobb-Douglas production function** (also called the **Cobb-Douglas model**) for obtaining growth rates for an economy and, thus, the dividend growth rate trajectories for a corresponding equity market. This model is particularly useful in the case of developing markets such as China where the structure of the underlying economy has experienced, and may experience, fundamental changes (as compared with the relatively stable growth rates of more developed economies).

In Section 2.4 we apply a form of the dividend discount model known as the H-model to the complicated dividend growth trajectory because it is well suited to instances where near-term growth rates can diverge significantly from the ultimately sustainable dividend growth rate. We also standardize the results in justified P/E form. This facilitates inter-temporal and cross-border market value comparisons. The difference between prevailing P/Es and justified P/Es is a measure of potential investment attractiveness.

As will be shown, the Cobb-Douglas and dividend discount models may also be applied to developed economies and equity markets.

2.1 Neoclassical Approach to Growth Accounting

Growth accounting is used in economics to measure the contribution of different factors—usually broadly defined as capital and labor—to economic growth and, indirectly, to compute the rate of an economy's technological progress. The neoclassical approach to growth accounting uses the Cobb-Douglas production function.[2] This approach can be useful to financial analysts because it gives insights into the long-term potential economic growth in individual countries, in larger regions, and for the world as a whole. The Cobb-Douglas estimate of the growth of total production can help to estimate corporate profit growth and develop corporate cash flow projections for stock market composites.

1 Forecasts and opinions offered in this reading are those of the authors (or the writers cited) and are not positions of CFA Institute.

2 See Cobb and Douglas (1928).

The basic form of the Cobb-Douglas production function is set forth as Equation 1, where Y represents total real economic output, A is total factor productivity, K is capital stock, α is output elasticity of K, L is labor input, and β is the output elasticity of L. **Total factor productivity (TFP)** is a variable which accounts for that part of Y not directly accounted for by the levels of the production factors (K and L).

$$Y = AK^{\alpha}L^{\beta} \tag{1}$$

If we assume that the production function exhibits **constant returns to scale** (i.e., a given percentage increase in capital stock and labor input results in an equal percentage increase in output), we can substitute $\beta = (1 - \alpha)$ into Equation 1.[3] Taking the natural logarithm of both sides of the equation gives

$$\ln(Y) = \ln(A) + \alpha\ln(K) + (1 - \alpha)\ln(L) \tag{2}$$

Taking first differences of Equation 2 and utilizing the fact that, for small changes in any variable x, $\ln(x + \Delta x) - \ln(x) = \ln\left(\frac{x + \Delta x}{x}\right) \approx \frac{\Delta x}{x}$, we obtain the expression:

$$\frac{\Delta Y}{Y} \approx \frac{\Delta A}{A} + \alpha\frac{\Delta K}{K} + (1 - \alpha)\frac{\Delta L}{L} \tag{3}$$

Equation 3 is the expression that we will employ in our analysis. In Equation 3, the percentage growth in real output (or gross domestic product, GDP) is shown as $\Delta Y/Y$, and it is decomposed into its components: $\Delta A/A$ is growth in TFP; $\Delta K/K$ is the growth in the capital stock; $\Delta L/L$ is the growth in the labor input; α is the output elasticity of capital; and $1 - \alpha$ is the output elasticity of labor where $0 < \alpha < 1$.

In practice, all the variables in Equation 3, with the exception of the growth in TFP, are directly observable or can be derived from national income and product accounts.[4] However, growth in TFP is determined using the other inputs as noted by Equation 3 and is commonly referred to as the **Solow residual**.[5]

TFP growth means that aggregate output (i.e., GDP) can grow at a faster rate than would be predicted simply from growth in accumulated capital stock and the labor force. Interpreting TFP as a measure of the level of technology, growth in TFP is often described as a measure of "technical progress" and linked to innovation. As examples, such technological advances as the introduction of the steam engine, electricity, the internal combustion engine, telecommunications, microchips, penicillin, and the internet are thought to have contributed to growth in TFP. However, growth in TFP, as a residual in the sense described, can be driven by factors other than improvements in technology. These factors could be particularly significant in economies that are experiencing major changes in political and/or regulatory structures. As examples, liberalization of trade policies, abolition of restrictions on the movement and ownership of capital and labor, the establishment of peace and the predictable rule of law, and even the dismantling of punitive taxation policies would be expected to contribute to growth in TFP. Finally, growth in TFP can benefit from improvements in the division of labor that arise from the growth of the economy itself. By contrast, developments such as the depletion and degradation of natural resources would detract from growth in TFP.

The robustness and simplicity of the approach we have presented can be tested against the complex and important case of valuing the equity markets in mainland China.

---

3 As a result, if both capital and labor change by a percentage x, then the total change in output is $\alpha x + (1 - \alpha)x = x$. The use of constant returns to scale is predicated on empirical results from several large economies over various time periods during the 19th and 20th centuries.

4 Capital, α, and labor, $(1 - \alpha)$, output elasticities may differ across national economies.

5 See Solow (1957). The Solow residual is thus simply: $\frac{\Delta A}{A} \approx \frac{\Delta Y}{Y} - \alpha\frac{\Delta K}{K} - (1 - \alpha)\frac{\Delta L}{L}$

2.2 The China Economic Experience

China has been widely regarded as the most influential emerging economy, and its growth performance since reform has been hailed as an economic miracle. Historical growth accounting results, as presented in Zheng, Hu, and Bigsten (2009), are reported in Exhibit 1. Note particularly the comparisons of China's growth in the capital stock, $\Delta K/K$, and growth in the labor input, $\Delta L/L$, to those of the (former) Soviet Union, United States, and European Union. The growth in capital stock stands out particularly for China and is most apparent during the period of economic liberalization that commenced in the early 1990s. According to estimates by the World Bank and other institutions, the gross effective savings in China (loosely defined as investment in plant, property, equipment, and inventories) divided by economic output have been in the neighborhood of 40 percent. This compares with 15 to 20 percent over the comparable periods for the other countries in Exhibit 1.

Exhibit 1 Historical Growth Accounting for China, the (Former) Soviet Union, United States, and European Union

Countries	Time Period	Real GDP Growth $\Delta Y/Y$ (%)	Growth in Total Factor Productivity $\Delta A/A$ (%)	Growth in Capital Stock $\Delta K/K$ (%)	Growth in Labor Input $\Delta L/L$ (%)
China	1978–1995	10.11	3.80	9.12	3.49
	1995–2007	9.25	1.45	12.81	2.78
Soviet Union	1950–1970	5.4	1.6	8.8	1.8
	1970–1985	2.7	−0.4	7.0	1.1
United States	1950–1972	3.9	1.6	2.6	1.4
	1972–1996	3.3	0.6	3.1	1.7
	1996–2004	3.6	1.5	2.6	0.7
European Union	1960–1973	5.1	3.2		4.8
	1973–2003	2.2	1.0	0.5	2.8

Source: Zheng, Hu, and Bigsten (2009). China's output elasticity for capital (α) and output elasticity for labor ($1 - \alpha$) were both estimated to be 0.5.

Zheng, Bigsten, and Hu (2006) studied the Chinese economy and found that reform measures had a significant positive impact on TFP, but this impact should be considered a one-time event. These authors make a case that China should now focus on achieving sustained increases in productivity.

Exhibit 1 also shows that Chinese economic growth has been largely driven by growth in the capital stock. Zheng, Bigsten, and Hu note that government policies in the mid- to late 1990s supported this extraordinary growth in investments. Key input prices were kept low, and a high savings rate allowed for the availability of cheap credit. A huge trade surplus has been another side-effect of both high investment[6] and a low fixed exchange-rate policy designed to support exports. China's foreign reserves are currently the world's largest by a considerable amount and have surpassed \$2 trillion.[7] A necessary, eventual "course correction" in exchange rate and monetary policies would reduce or reverse the forces that contributed to a de facto subsidization of capital formation.

6 In lieu of higher consumption spending, particularly on imported goods.

7 Preston (2009).

In addition to the foregoing structural factors, changes in consumer behavior are also likely to cause the Chinese savings/investment rate to moderate. Altogether, government policy changes and an increased propensity to consume all point to an eventual reduction from the double-digit growth rates of capital stock. At the same time, while the labor force of China has grown at a much more rapid pace than for European and American economies, this has been attributable both to higher population growth rates and to a rise in labor force participation rates. The Chinese population growth rate has slowed to less than 1.0 percent per year, according to the World Bank. Furthermore, major changes in labor force participation rates, largely due to more people leaving rural occupations and household/childcare activities, represent one-time changes rather than sustainable trends. In sum, these considerations suggest that Chinese economic growth will eventually moderate.

Finally, in addition to the factors above, an investment analyst might wish to consider other, more qualitative factors in producing a long-term growth forecast (e.g., China's educational system). Because adjustments for such factors would typically have a large judgmental element, this reading does not address them.

EXAMPLE 1

The Neoclassical Approach to Growth

1. The savings rate for a national economy is comparatively stable. The economy faces a sharp uptick in energy prices and at the same time imposes stringent restrictions on environmental pollution. The combined impact of energy and environmental factors renders a large portion of the existing stock of manufacturing equipment and structures economically obsolescent. What is the impact on the economy according to Equation 3?
2. A country experiences a sharp demographic rise in the divorce rate and single-parent households. Using the framework of Equations 1 and 3, what is likely to happen to total national production, total per capita income, and total income per household?

Solution to 1:

The sudden, unexpected obsolescence of a significant portion of the capital stock means that the percentage growth rate in capital stock in that period will be negative, that is, $\Delta K/K < 0$. All other things being equal, this implies a one-time reduction in economic output. Assuming no change in technological innovation, savings rates, and labor force growth trends, the subsequent long-term growth rates should be relatively close to the previously prevailing growth rates, starting from the lower base value for Y.

Solution to 2:

The change in demographics implies an increase in the aggregate labor force as stay-at-home parents and spouses re-enter the workforce. That is, the labor force will grow, for some period of time, at a pace faster than underlying population growth until a new steady-state labor force participation rate is attained. Total economic production (and income) will thus also rise at an above-trend rate during this adjustment period. Above-trend growth in national income, holding population trends constant, means that per capita income will also grow above trend during this period of demographic adjustment. Per-household income, by contrast, will grow at a below-trend rate (and may even decline) due to an uptick in new household formation during the shift in divorce and separation rates to ultimately prevailing steady-state levels.

2.3 Quantifying China's Future Economic Growth

Now that we have covered a simple model for estimating an economy's growth rate, the next step is to apply the model using our best estimates of the model inputs. As in any forecasting exercise, the specific forecasts must be based on currently available information. Any forecast has an "as of" date associated with it. Comparing the forecasts to actual outcomes subsequently, some inputs or elements of the forecast may appear to be misjudged or dated. With that caution in mind, we can proceed to develop our economic growth projections for China.

Zheng, Hu, and Bigsten (2009) offer the GDP growth projections presented in Exhibit 2 for China, the United States, and the European Union. The forecast of an 8 percent GDP growth rate for China is consistent with the Chinese government's 8 percent GDP growth target as presented by Premier Wen Jiabao. Zheng, Hu, and Bigsten note their own projections rely heavily on two basic assumptions: 1) growth in the capital stock cannot exceed GDP growth and 2) a TFP growth rate of 2 to 3 percent will prevail for the foreseeable future. These authors believe that the potential for China to absorb new technologies from developed nations is double that for the United States and European Union. Given the history of other developing countries and the record of economic recovery of developed countries after World War II, this does not seem unreasonable.

Exhibit 2 Growth Projections (2009–2030)

Country	Real GDP Growth $\Delta Y/Y$ (%)	Growth in Total Factor Productivity $\Delta A/A$ (%)	Output Elasticity of Capital α	Growth in Capital Stock $\Delta K/K$ (%)	Output Elasticity of Labor $1 - \alpha$	Growth in Labor Input $\Delta L/L$ (%)
China	8.0	2.5	0.5	8.0	0.5	3.0
United States	2.75	1.2	0.3	4.0	0.7	0.5
European Union	2.2	1.0	0.4	3.0	0.6	0.0

Source: Zheng, Hu, and Bigsten (2009).

The neoclassical framework we have presented permits analysts to apply their own forecasts of factors of production, with particular emphasis on how such factor trajectories might change over time. Once the analyst has developed a long-term macro forecast, it can then be used in conjunction with traditional valuation models.

In applying the framework, we modify the Zheng, Hu, and Bigsten ("ZHB") projections by using a lower estimate of the growth rate in the labor force, since World Bank data indicate that population growth in China now appears to have declined to below 1.0 percent annually. At the same time, we are inclined to think that savings and investment rates will only decline gradually from over 35 percent of GDP, thereby keeping the growth rate of capital stock much higher than the 8 percent per annum assumed by ZHB. We have no disagreement with the ZHB projection of 2.5 percent per year for TFP growth. If we utilize the labor and capital elasticities from the ZHB study, a reasonable projection for economic growth would therefore be:

Total factor productivity	plus	Effect of growth in capital stock	plus	Effect of labor force growth	
2.5%	+	0.5 × 12%	+	0.5 × 1.5%	= 9.25%

This near-term rate is higher than the ZHB forecast, the official forecast of the Chinese government, and the consensus of many economic forecasters. We note, however, that there are several factors that are consistent with our higher near-term forecast. First, actual real growth has cumulatively exceeded the 8 percent Chinese official growth target of the past several years. Second, and more importantly, our forecast is to be thought of as a normalized forecast of sustainable cash flow growth potential.

While our near-term forecast for economic growth is higher than ZHB and the Chinese government, the reasoning set forth in the preceding section leads us to believe that economic growth will gradually decline to levels lower than the ZHB analysis.[8] This is because, as economies develop and as the stock of accumulated capital per person rises, savings rates tend to decline and TFP trends fall to levels closer to those of more highly developed countries. Finally, although labor force growth can exceed population growth for some time (as labor force participation rates increase), in the long run, labor force growth is constrained by population growth. China appears to be on its way towards zero population growth (much like Western Europe). With this in mind, an ultimately sustainable economic growth rate might be:

Total factor productivity	plus	Effect of growth in capital stock	plus	Effect of labor force growth	
1.25%	+	0.5 × 6%	+	0.5 × 0.0%	= 4.25%

2.4 Equity Market Valuation

In this section we translate macroeconomic forecasts into corporate cash flow forecasts and combine those corporate forecasts with an appropriate discounted cash flow model to estimate the intrinsic value of an equity market in terms of justified P/E ratios.

The growth rate of corporate earnings and dividend cash flow, adjusted for inflation, should bear a close relationship with real GDP growth over the long term. For purposes of this analysis, we assume that earnings and dividend cash flow for the underlying comprehensive stock composite grow at the same rate as the core growth rate of Chinese GDP.[9]

In theory, we would like to be able to forecast, year by year, each of the underlying factors of production and the change in TFP. In practice, however, we recognize that a less complicated cash flow representation might be more suitable, because it lessens the possibility of compounding forecast errors. Fuller and Hsia (1984) developed a valuation model, known as the **H-model**, in which dividend *growth rates* are expected to decline in a linear fashion, over a finite horizon, towards an ultimately sustainable rate from the end of that horizon into perpetuity. It incorporates a growth rate in dividends that is expected to prevail in the initial period g_S, a period of years, N, where the dividend growth rate declines in a linear fashion, and a long-term dividend

---

8 Our forecast is for a 30-year time period and was made in the summer of 2009. As noted in Exhibit 2, the ZHB forecast was for 2009–2030. We believe the choice of a longer time horizon for our forecast is also supportive of the choice of a lower terminal growth rate.

9 In principle, the sector of publicly traded companies could grow somewhat above or below the overall growth rate of GDP, because it is a subset of the overall economy. However, the approach used in the text should serve as a good approximation for analytical purposes.

growth rate g_L that is expected to prevail to perpetuity beginning at the end of period N. With an initial annualized dividend rate at time zero of D_0 and a discount rate to perpetuity of r, the H-model estimate of value, V_0, is given by Equation 4:

$$V_0 = \frac{D_0}{r - g_L}\left[\left(1 + g_L\right) + \frac{N}{2}\left(g_s - g_L\right)\right] \tag{4}$$

The H-model provides a convenient means for modeling initially high ("supernormal") dividend growth rates that gradually transition to a lower, long-run growth at a constant mature-stage growth rate. The H-model involves an approximation to the value estimate that would result from period-by-period discounting of cash flows in the phase prior to the mature or terminal phase when a constant growth rate is assumed. The approximation is generally very good in most practical situations and the gain from using an approximation is an easy to evaluate expression.[10] In the case of valuation of mature developed equity markets, the Gordon (constant) growth dividend discount model would be more commonly used than the H-model because supernormal growth would not generally need to be modeled in such cases.

In our valuation analysis, we express the discount rate and both growth factors in real, that is, inflation-adjusted terms. A key to valuation is consistency: stating variables consistently on a nominal basis or consistently on a real basis are both feasible approaches. Economists, however, typically prefer to use real variables as they tend to be more stable and, therefore, easier to predict than their nominal counterparts.

We use our growth rate trajectory and apply the H-model to the S&P China BMI Index. This index underlies the SPDR S&P China ETF, which is an exchange-traded fund designed to track the investment performance of the mainland China stock markets. The index underlying the ETF and the information provided by the ETF's sponsor (State Street Global Advisors) provide up-to-date information that can enter traditional valuation models.

In evaluating the investment attractiveness of a market index, we utilize a price–earnings ratio or P/E approach. Because of the behavioral factors mentioned in the introduction, prices of equities and equity market composites tend to vary more than underlying normalized earnings and growth prospects. P/E analysis permits us to make useful inter-temporal valuation comparisons and has the additional benefit of providing intuition when making comparisons across international borders. As of 15 July 2009, the forward or prospective P/E ratio for the underlying S&P China BMI Index was 19.1 (this P/E is the level of the S&P China BMI Index divided by year-ahead expected earnings for that index). In the following analysis, we estimate what justified P/E ratios should be under differing inflation-adjusted equity discount rates and for different estimates of the ultimately prevailing terminal inflation-adjusted dividend growth rate to perpetuity.

The (forward) **justified P/E** is the estimated intrinsic value divided by year-ahead expected earnings; in this case we are estimating intrinsic value using the H-model. Reflecting the meaning of *justified* here as *warranted by fundamentals*, price in the justified P/E ratio is assumed in this discussion to equal intrinsic value as estimated by the valuation model, i.e., P_0 (or P) = V_0.[11] In all instances, we assume that core

10 Valuation differences between the H-Model and a period-by-period approach should be minimal so long as the differences between long-term and interim growth rates are of single-digit magnitude and the "interim" period length is not much longer than 30–40 years. In any event, the possible valuation error from adopting a simpler model is reasonable in comparison with the incremental valuation error that can arise from introducing an excessive number of valuation parameters, i.e., year-by-year cash flows.

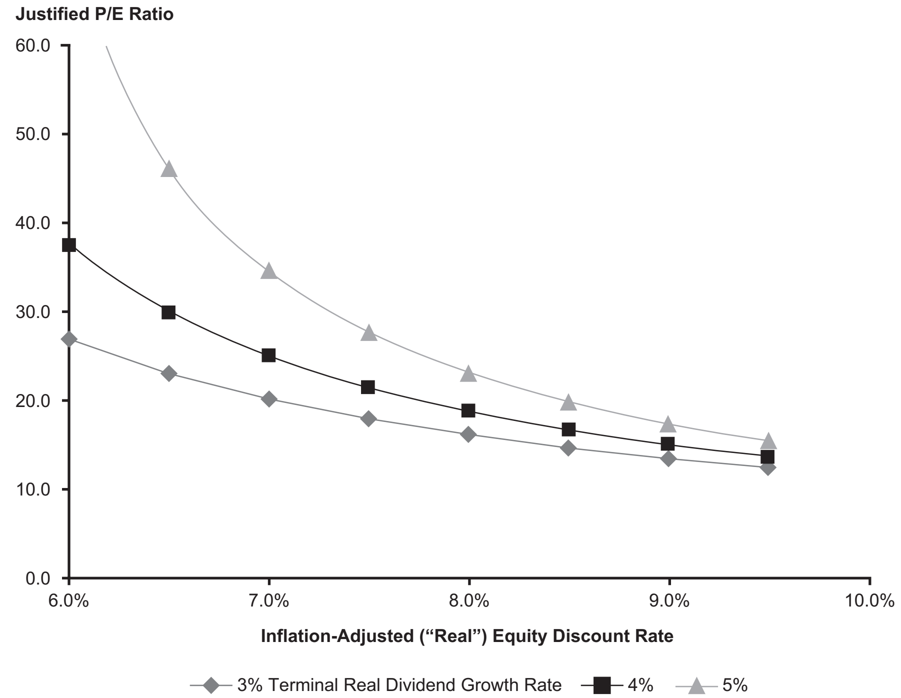
11 Analysts and practitioners may, if desired, proceed directly to forecasting a justified market index *level* based on the H-Model, the current dividend rate, and the growth-factor inputs. (Strictly speaking, the H-Model does not directly utilize earnings per share, although, indirectly, the trajectory of dividend growth is assumed to be supported by growth in EPS.) Under our approach, the relative differences in P/E ratios in Exhibits 3 and 4 translate directly into the relative differences between observed and justified market levels.

inflation-adjusted growth rates decline in a linear fashion over a 30-year time horizon from the 9.25 percent per year we estimate for year one.[12] The 30-year time horizon is selected both because it is a round number and because it is not unlike other historical instances where national economies experienced fundamental changes in political and economic structure, the notable examples being post-World War II European economies and Japan both in the late 19th century and after World War II.

Exhibit 3 Justified P/E Ratios for Chinese Equity Market at Mid-Year 2009

Terminal Real Growth Rate (%)	Real Equity Discount Rate							
	6.00%	6.50%	7.00%	7.50%	8.00%	8.50%	9.00%	9.50%
3	26.8	23.0	20.1	17.9	16.1	14.6	13.4	12.4
4	37.3	29.9	24.9	21.3	18.7	16.6	14.9	13.6
5	69.0	46.0	34.5	27.6	23.0	19.7	17.2	15.3

Exhibit 4 Justified Chinese Equity Market Valuation Multiples



Note: Chinese Equity Markets Justified P/Es: 30-Year Transition from 9.25% Real Dividend Growth Rate to Various Terminal Growth Rates to Perpetuity.

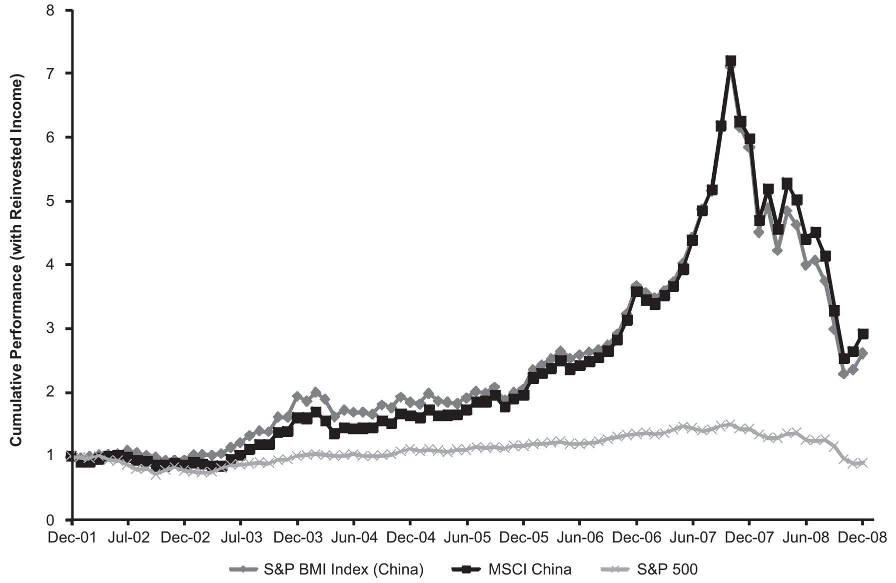
12 The geometric average growth rate during the 30-year period is around 6.7%. Also of interest, the *average* compound growth rate for the first 20 years is not far off from the ZHB 20-year 8% annual growth rate.

In Exhibit 3 and Exhibit 4, we have presented justified P/E ratios. Interpolating visually, the observed 19.1 P/E ratio on 15 July 2009, assuming a terminal 4.25 percent real dividend growth rate to perpetuity, is consistent with a real equity discount rate just under 8.0 percent.

This leads immediately to the question of what the proper discount rate should be. To answer this we would like to know a little bit about both the volatility of Chinese equity prices and how such return/volatility prospects compare with other world equity markets.

In Exhibit 5, we present cumulative return data for both the S&P China BMI Index and the S&P 500. The data series commence in 2001, the point at which the China BMI Index data are first available, and a point by which mainland Chinese equity markets had become seasoned and widely accessible to non-Chinese investors.[13]

Exhibit 5 Cumulative Performance Comparison for Chinese and US Stock Markets (in RMB for China, USD for S&P 500)



Source: Standard & Poor's, Morgan Stanley.

Over the 7-year period, the cumulative return of Chinese equities far outpaced those of the United States. However, Chinese markets also experienced more of a bubble and subsequent collapse, resulting in a much higher volatility of returns. For China, the annualized return over this time frame approached an inflation-adjusted 13 percent. This outpaced the return one would have expected even from coupling a 2 to 2.5 percent dividend yield with an aggressive core dividend real growth rate of 9 percent. The cumulative return benefited from a positive shift in P/E ratios between the beginning and end of the measurement period. In fact, the P/E expansion and subsequent contraction appear to be the chief cause of the 2007 Chinese bubble and the following collapse.

By comparison, the equity markets in the United States were much less volatile, although the cumulative return, particularly net of inflation, was negative. Assuming a 6 percent inflation-adjusted discount rate for the US equity market (in line with

13 To check the reasonableness of the data, we computed the cumulative total returns for the Morgan Stanley Capital International (MSCI) China Index and found them in good accord with the S&P China BMI Index.

long-term historical realized returns), the cumulative nominal return—incorporating the 2.6 percent average inflation rate—would have resulted in an ending cumulative return index value of around 1.80 in December 2008 (compared with the 2.61 achieved by the S&P China BMI Index). The explanation for the poor performance by US equities was largely attributable to a contraction from the high normalized P/E valuation relationships prevailing in 2001 to the below-average levels at the end of the period. Additional evidence is presented in Exhibit 6.

Exhibit 6 Return and Volatility Data

31 December 2001 – 31 December 2008	S&P China BMI (%)	MSCI China (%)	S&P 500 (%)
Annualized nominal total return[1]	14.7	16.6	–1.5
Annualized standard deviation of total returns[2]	29.4	29.4	14.3
Annualized inflation rate[3]	3.7	3.7	2.6

Notes:
[1] In RMB for Chinese composites, USD for S&P 500.
[2] Based on monthly observations.
[3] Data through 2007, reflects changes in GDP deflator.
Sources: Standard & Poor's, Morgan Stanley, Bloomberg, World Bank.

In establishing a reasonable discount rate to apply to our cash flow forecasts, we should take into account the higher volatility of Chinese markets.

The higher observed volatility also has arisen from behavioral shifts in P/E relationships that are more pronounced than those usually seen in US and European markets. Such valuation-induced price volatility is not unusual for any market that is new by historical standards and has been experienced in the past by US, European, and Japanese markets in their own long paths to economic maturity.

The effect of higher volatility on required returns might be somewhat mitigated to the extent 1) market returns are less than perfectly correlated with other international equity markets and 2) cross-border investing and divesting of equities is freely achievable by investors both inside and outside of China.

On balance, the foregoing factors suggest that the required real equity discount rate be higher in China than, for example, in the United States. This naturally leads us to investigate both what the realized real equity discount rate has been in the United States and what real equity discount rates are predicted based on alternative theoretical models. Historical studies have been undertaken[14] which indicate that a prospective inflation-adjusted equity discount rate[15] in the area of 5–7 percent is reasonable. Work based on macroeconomic, corporate finance, and financial market equilibrium[16] suggests that a slightly higher range might be possible in equilibrium. Purely theoretical models have had mixed results. The utility function models,[17] for example, find that real discount rates in the 5–7 percent area are well above what can

14 Ibbotson and Sinquefield (1989), Siegel (1992), Arnott and Bernstein (2002), Siegel (2005).
15 Geometrically compounded.
16 Ibbotson, Diermeier, and Siegel (1984), Stimes (2008).
17 Mehra and Prescott (1985), Mehra (2003).

be justified by underlying market volatility. On the other hand, a theoretical approach[18] simply based on prospective wealth accumulation under different volatility assumptions is consistent with the results actually realized in the US historical record.

In the volatile economic and market conditions at the time this reading is being written (2009), the higher end of the discount rate estimates seem to be in order for the United States. If we place these at 6–7 percent, the additional relative risk considerations for the Chinese market suggest that a required discount rate for that market might be in the range of 7.5–8.5 percent. This is a necessarily judgmental adjustment but should 1) reflect an analyst's view of differential riskiness (in the context of a well-diversified international portfolio) and 2) reflect congruence with historical realized return differentials between markets that were then seasoned and those that were then developing.

Referring back to Exhibit 4 and integrating our view of the real (i.e., inflation-adjusted) equity discount rate with the sustainable dividend growth rates obtained from our macroeconomic framework, we can conclude that the currently observed 19.1 forward P/E ratio for Chinese equities is not unreasonable. As a further check, we note that a 19.1 P/E would be somewhat on the high side for seasoned US and European stock markets. However, reflecting much higher growth prospects over the next several decades, a much higher Chinese P/E ratio would be warranted.

Inherent in our analysis of equity composites is the difficulty of specifying precise price or P/E ratios at which a "buy" or "sell" recommendation is to be made. However, the strength of this kind of relative value approach is that, in a diversified portfolio context, investors can usually make reasonable decisions—at the margin—whether it is then appropriate to raise or lower market exposures relative to the investable universe in the aggregate. The price/value relationships prevailing at the date of the analysis were such that those investing on a fundamental basis should have a weighting in Chinese equities close to their baseline or normal-strategy allocation.

EXAMPLE 2

Equity Market Valuation Using Dividend Discount Models

1. The S&P China BMI Index on 30 September 2009 is 358. Forecasted 12-month earnings per share for the composite are 18.00 RMB, and the current annual dividend rate for the composite is 7.90 RMB. Assuming an 8.0 percent inflation-adjusted equity discount rate, a 30-year decline in dividend growth rates from an initial growth rate of 8.25 percent, and a terminal sustainable growth rate to perpetuity of 4.25 percent, compute the composite index price level implied by the H-Model (Equation 4). Next compute the justified P/E ratio implied by such price level.
2. Assuming the same annualized dividend rate of 7.90 RMB, using the Gordon growth model compute the discount rate required to reproduce the prevailing index level of 358 under different growth assumptions, specifically assuming an 8 percent real growth rate of dividends to perpetuity, rather than a gradually slackening rate of growth as in Question 1. Evaluate the result.
3. Assuming the same information in Question 1, what would be the appropriate composite index price level and justified P/E ratio, if the period at which the 4.25 percent growth rate to perpetuity is reached A) at year 20, and B) at year 40?

18 Arnott (2004).

Solution to 1:

The H-Model states that:

$$V_0 = \frac{D_0}{r - g_L}\left[(1 + g_L) + \frac{N}{2}(g_s - g_L)\right]$$

Inserting the information given, we get

$$V_0 = \frac{7.90}{0.08 - 0.0425}\left[(1 + 0.0425) + \frac{30}{2}(0.0825 - 0.0425)\right] = 346.02$$

Dividing this result by forecasted earnings of 18.00 produces a justified P/E ratio of 19.2.

Solution to 2:

The standard Gordon growth model states that:

$$V_0 = \frac{D_0(1 + g)}{r - g}$$

which can be rearranged as

$$r = \frac{D_0(1 + g)}{V_0} + g$$

Substituting in the given values, we obtain:

$$r = \frac{7.90(1 + 0.08)}{358} + 0.08 = 0.1038 \approx 10.4\%$$

This result, which assumes no slackening in core growth rates, produces an implied discount rate that appears unusually large relative to the prospects of other world equity markets. Given the ability of international portfolio reallocation, even on a constrained basis, capital market equilibrium does not seem consistent with a real equity discount to perpetuity rate almost twice that of mature equity markets. The implication is that Chinese market participants are pricing the index at a lower discount rate, consistent with other worldwide investment opportunities, but also with a more restrained long-term growth outlook relative to those growth rates expected over the next few years.

Solution to 3A:

Assuming an interim period of 20 years, Equation 4 produces:

$$V_0 = \frac{7.90}{0.08 - 0.0425}\left[(1 + 0.0425) + \frac{20}{2}(0.0825 - 0.0425)\right] = 303.89$$

and a P/E = 303.89 ÷ 18 = 16.9

Solution to 3B:

Assuming an interim period of 40 years, Equation 4 produces:

$$V_0 = \frac{7.90}{0.08 - 0.0425}\left[(1 + 0.0425) + \frac{40}{2}(0.0825 - 0.0425)\right] = 388.15$$

and a justified P/E = 388.15 ÷ 18 = 21.6

The possible criticisms of our approach should not be overlooked. From a practical perspective, there may be severe problems with the accuracy of data inputs. It is difficult enough to obtain macroeconomic data in developed countries with long-established methods and facilities. In developing markets or in economies experiencing profound governmental and structural change, such as the Eastern Bloc after the fall of the

Berlin Wall, the problems of obtaining accurate and, more importantly, historically consistent, data are multiplied. The same fluidity in political and demographic fundamentals also calls into question whether companies' growth rates will track GDP growth rates. In certain instances, there can be long departures between growth rates, meaning that for long periods of time the share of corporate profits may be rising or declining relative to GDP.

The analysis in this reading has also focused on inflation-adjusted income, cash flow, and discount rates. In a global economy with reasonably robust currency exchange markets and where monetary growth is targeted to keep inflation at manageable levels, this is probably appropriate. However, hyperinflation, currency instability, and other trade disequilibria have occurred far too frequently from a historical perspective to be overlooked. In the presence of such factors, the confidence of our model's approach could be diminished.

EXAMPLE 3

Applying Valuation Methodology to a Developed Economy

In the following, assume that all growth and discount rates are stated in real terms.

1 Assume the Eurozone inflation-adjusted average growth in capital stock is 3.0 percent per annum into perpetuity. Long-term labor force growth is expected to remain stable at 0.0 percent, while TFP growth is projected to average 1.0 percent per annum over time. If the output elasticity of capital is 0.4 and the output elasticity of labor is 0.6, calculate the implied growth rate of Eurozone GDP.

2 The Dow Jones Euro Stoxx 50 Index is comprised of mature, large capital common equities domiciled in the Euro currency zone. At 30 September 2009 the index level stood at 2450. Forecasted 12-month dividends per share for the composite (net of withholding tax) are €125.00. Because of the mature nature of the economy and the particular market composite, you project that growth in both inflation-adjusted earnings and dividends will equal that of GDP. Using the Gordon constant dividend growth rate model solved for the discount rate, estimate the implied inflation-adjusted discount rate to perpetuity.

3 A Applying the Gordon growth model to value the DJ Euro Stoxx 50 Index, you assume that the appropriate discount rate to perpetuity should be 6.0 percent. If this assumption is correct, what is the fair value of the DJ Euro Stoxx 50 Index?

B As of the end of September 2009, the DJ Euro Stoxx 50 Index was trading almost 30 percent below its high of twelve months earlier. What is the likely major cause for the price decline?

Solution to 1:

In the context of Equation 1 from the text, total growth in GDP is:

$$\frac{\Delta Y}{Y} \approx \frac{\Delta A}{A} + \alpha\frac{\Delta K}{K} + (1-\alpha)\frac{\Delta L}{L}$$

Substituting the information given, the GDP growth rate is 2.2 percent, computed as follows:

$$\frac{\Delta Y}{Y} \approx 1.0\% + 0.4 \times 3.0\% + 0.6 \times 0.0\% = 2.2\%$$

Solution to 2:

The Gordon growth model can be rearranged as

$$r = \frac{D_1}{P_0} + g$$

Substituting in the given values for dividends, the index level, and our forecast for dividend growth, we obtain:

$$r = \frac{125}{2450} + 0.022 = 0.051 + 0.022 = 0.073 = 7.3\%$$

Solution to 3A:

The constant growth model gives us the following estimate of the fair value of the DJ Euro Stoxx 50:

$$V_0 = \frac{D_1}{r - g} = \frac{125}{0.06 - 0.022} = 3289$$

This estimate is more than 34 percent above the level observed at 30 September 2009.

Solution to 3B:

Given the mature nature of the underlying economic region and the companies in the composite, it is unlikely that the estimate of long-term, real dividend growth changed much, if at all. If the actual dividends paid also did not change much, the most likely major cause of the price decline is an increase in the discount rate over the period.

TOP-DOWN AND BOTTOM-UP FORECASTING

3

When it comes to predicting equity returns, analytical approaches can be divided into two major categories: top-down and bottom-up. In top-down forecasting, analysts use *macroeconomic* projections to produce return expectations for large stock market composites, such as the S&P 500, the Nikkei 225, or the FTSE 100. These can then be further refined into return expectations for various market sectors and industry groups within the composites. At the final stage, such information can, if desired, be distilled into projected returns for individual securities.

By contrast, bottom-up forecasting begins with the *microeconomic* outlook for the fundamentals of individual companies. An analyst can use this information to develop predicted investment returns for each security. If desired, the forecasts for individual security returns can be aggregated into expected returns for industry groupings, market sectors, and for the equity market as a whole.

Exhibit 7 sets forth the manner in which top-down and bottom-up approaches are typically implemented. Top-down can be characterized as moving from the general to the specific, while bottom-up forecasting moves from the specific to the general. Depending on the investment strategy and portfolio context, one of the types of forecasting may be more suitable. In other instances, both types of forecasting may be useful. In those cases where both top-down and bottom-up are used, the additional work involved may provide valuable insights.

Exhibit 7 Comparison of Top-Down and Bottom-Up Analyses

Top-Down Analysis

- Market analysis: Examine valuations in different equity markets to identify those with superior expected returns.
  - Compare relative value measures for each equity market to their historical values to identify those markets where equities are relatively cheap or expensive.
  - Examine the trends in relative value measures for each equity market to identify market momentum.
  - Compare the expected returns for those equity markets expected to provide superior performance to the expected returns for other asset classes, such as bonds, real estate, and commodities.
- Industry analysis: Evaluate domestic and global economic cycles to determine those industries expected to be top performers in the best-performing equity markets.
  - Compare relative growth rates and expected profit margins across industries.
  - Identify those industries that will be favorably impacted by expected trends in interest rates, exchange rates, and inflation.
- Company analysis: Identify the best stocks in those industries that are expected to be top-performers in the best-performing equity markets.

Bottom-Up Analysis

- Company analysis: Identify a rationale for why certain stocks should be expected to outperform, without regard to the prevailing macroeconomic conditions.
  - Identify reasons why a company's products, technology, or services should be expected to be successful.
  - Evaluate the company's management, history, business model, and growth prospects.
  - Use discounted cash flow models to determine expected returns for individual securities.
- Industry analysis: Aggregate expected returns for stocks within an industry to identify the industries that are expected to be the best performers.
- Market analysis: Aggregate expected industry returns to identify the expected returns for every equity market.

3.1 Portfolio Suitability of Each Forecasting Type

In theory and practice, it is not necessary for either top-down or bottom-up forecasting to be carried to the final step shown for each method in Exhibit 7. For example, if a portfolio focuses primarily on tactical asset allocation among different market composites (and/or different industry groups within such composites), a top-down forecast may not need to focus all the way down to the relative merits of individual securities.

Likewise, there are instances where either the investment strategy or specific portfolio constraints dictate a focus primarily on individual security returns. In such instances, the unique factors pertaining to particular securities may render the need to study industry and market composite unnecessary. In such cases, the bottom-up method stops well short of the top. The partial application of each method is developed in the following examples.

EXAMPLE 4

Growth Model Questions

Explain whether top-down or bottom-up forecasting is more appropriate for each of the different investors.

1 The MegaCosmos Mutual Fund has a stated goal of investing in the stock market composites of developed country economies in North America, Western Europe, and Japan. Its return target is expressed in euros. The Fund may or may not hedge individual country currency returns depending on its outlook for foreign exchange rates. Furthermore, the Fund attempts to track individual country stock market composites while minimizing tracking error via the use of index baskets wherever possible.

2 EMF Advisers is a boutique firm that manages a dedicated portfolio of electric, gas, and water utility companies domiciled in the United States. The portfolio EMF oversees is, in turn, a small part of the American Pipefitters Union Pension Plan.

3 Bocage International is a hedge fund that actively bets on the relative attractiveness of stocks, interest rates, currencies, and commodities. Its investment in equities is limited to futures and options on exchange-traded equity indexes.

4 Alpha Bet Partnership is an investment vehicle featuring a US long/short overlay. Specifically, the partnership may keep short positions in US equities in an amount not to exceed 30 percent of the net value of the partnership. All short positions must be invested in US equities to maintain an overall beta of 1.0. The partnership hopes for the stocks it owns to outperform the stocks it has sold short in order to generate a respectable alpha. The partnership specifies that with the objective of minimizing tracking error, every stock sold short must be matched by a stock bought in the same industry.

Solution to 1:

MegaCosmos' ability to carry out its strategy will depend on its ability to forecast economic factors at a very "macro" scale. It would employ a top-down approach involving an examination of the economic strength of different international economies, different fiscal and tax policies among the governments involved, and international trade patterns and currency flows. MegaCosmos' desire to track underlying national markets quite closely means that its holdings will not diverge materially from the particular market composites selected. Individual security selection will not be much of a focus, thereby minimizing the need for continuing the top-down analysis as low as individual market sectors, industry groups, or securities.

Solution to 2:

EMF Advisers' ability to carry out its strategy will depend on its ability to select among different securities within the very specific niche to which it has been assigned by the pension plan sponsor. As a result, bottom-up forecasting is most appropriate and probably no higher than the industry level in any great detail. The plan sponsor, however, will need to be concerned with top-down forecasting to determine the appropriate allocation to EMF's strategy.

Solution to 3:

Bocage's situation is very similar to that of MegaCosmos in Question 1, and top-down forecasting is thus appropriate. In Bocage's case, exposure to individual stocks is not permitted so the analysis need not be carried down to the level of industry groups or market sectors.

Solution to 4:

There are two parts to the answer for Alpha Bet. Because the underlying beta is targeted at 1.0, this portion of the strategy can be considered passive and very little or no top-down forecasting is required. In contrast, the remaining portion of the strategy, the long/short overlay, involves pure security selection on a "matched" or "paired trades" basis. Within each long/short combination, aggregate factors (global, market, industry) cancel each other out, because both long and short candidates must be matched to the same country, market, and industry. Only specific factors affecting each of the two paired companies matter. Therefore, a bottom-up forecast is necessary—and one that likely does not need to go above the level of individual securities.

3.2 Using Both Forecasting Types

When engaged in fundamental securities analysis, it can be wise to use both top-down and bottom-up forecasting. However, when we use both approaches, we often find ourselves in the situation of the person with two clocks, each displaying a different time. They may both be wrong, but they cannot both be right!

It is frequently the case that top-down and bottom-up forecasts provide significantly different results. In such instances, the analyst should investigate the underlying data, assumptions, and forecast methods before employing them as a basis for investment decisions. After all, if forecasts cannot be consistent with each other, at least one of them cannot be consistent with underlying reality.

Because we are fallible human beings, most forecasting discrepancies arise from our own limited knowledge, errors, and incomplete assumptions. Reconciling top-down and bottom-up forecasts is therefore a discipline that can help prevent us from taking inappropriate investment actions. In other words, most of the time, the aggregate market consensus will tend to be more accurate than the individual forecasts that comprise the consensus. The reconciling and revision process is therefore useful in helping us better understand the market consensus.

However, in rare and significant instances, we will find that carefully retracing the steps reveals a gap between the two forecast types that gives rise to significant market opportunities. In such instances, the process of reconciling the two types of forecasts creates instances where we differ significantly and correctly from the consensus.

In the early 2000s, top-down forecasts provided much more subdued outlooks compared to bottom-up projections for corporate profits, both in the aggregate and for particular industries. In the tech area, both consumer and capital spending on computer equipment were below the projected sales growth that companies, and the analysts following them, were individually expecting. After all, individual companies were optimistic about their own prospects. However, in the aggregate, many of their technologies and products competed with each other. Thus, the success of some companies meant the failure of others, and this natural, competitive offsetting tendency was correctly reflected in aggregate, top-down sales growth for the industry.

Aggregating the bottom-up forecasts of individual companies, however, produced a wildly inflated forecast of both sales quantities and average prices and profit margins. Thus, the high-tech "bubble" originated from the mistaken principle that *all* the companies could be above average. For those who recognized the inconsistency of the top-down and bottom-up forecasts, and the accuracy of the top-down forecast, much of the carnage of the bubble collapse was avoided.

The collapse of equity markets in developed countries was a case where bottom-up forecasts proved their superiority over top-down approaches. In both the United States and the United Kingdom, some banking and real estate analysts perceived the excesses in residential real estate in a microeconomic sense. Those who pushed their analyses to a macro conclusion realized that certain large financial institutions were imperiled, particularly the highly levered Fannie Mae and Freddie Mac in the United States. Those who understood the pressure on these and other large financial institutions correctly foresaw that then-prevailing worldwide forecasts of economic activity and equity market returns were dramatically overstated.

3.3 Top-Down and Bottom-Up Forecasting of Market Earnings per Share

Two different methods are employed when estimating earnings for a market index, such as the S&P 500 Index. The first is to add up the individual estimates of the companies in the index. This is referred to as the bottom-up earnings estimate. The top-down estimate relies on forecasts for various macroeconomic variables and a model that fits these forecasts to past trends in aggregate earnings.

EXAMPLE 5

Comparing and Evaluating Top-Down and Bottom-Up Forecasts

Standard & Poor's July 2009 top-down and bottom-up forecasts for operating earnings per share appear in Exhibit 8. Note that the bottom-up forecasts are more optimistic than the top-down forecasts.

Exhibit 8 Standard & Poor's Forecasts: July 2009

Quarter Ending	Operating Earnings per Share (Estimates Are Bottom-Up)	Operating Earnings per Share (Estimates Are Top-Down)	Difference
31 Dec 2010	$20.39	$12.50	$7.89
30 Sep 2010	19.11	11.42	7.69
30 Jun 2010	18.00	11.18	6.82
31 Mar 2010	16.59	10.86	5.73
31 Dec 2009	16.25	11.72	4.53
30 Sep 2009	15.05	11.68	3.37
30 Jun 2009	14.06	11.05	3.01

The bottom-up projection starts from a June 2009 level of earnings that is some 27 percent above the top-down estimate. Furthermore, the annualized growth of estimated earnings over the subsequent 18 months is 28.1 percent for the bottom-up forecast and a much lower 8.6 percent for the top-down projection.

There are several possible reasons for the forecast discrepancies. First, the bottom-up estimates may be influenced by managers believing that their own company's earnings prospects are better than those for the economy as a whole.[19] This is simple human nature as confirmed by survey results, which consistently report that 85 percent of all drivers think they are better than average.

Alternatively, the bottom-up estimates may be correctly detecting signs of a cyclical economic and profit upturn. Most top-down models are of the econometric type and rely on historical relationships to be the basis for assumptions about the future. Thus, top-down models can be slow in detecting cyclical turns. This would be particularly true if the current statistical relationships between economic variables deviate significantly from their historic norms.

In short, the data indicate that we need to investigate both forecasts in greater detail. Without further analysis, we might be unable to distinguish whether the S&P 500 composite is overvalued, undervalued, or somewhere in between due to the disparity in the two earnings forecasts.

EXAMPLE 6

Earnings Forecast Revisions

The information in Exhibit 9 was collected from the Standard & Poor's website. Note that actual Q2 2008 EPS for the S&P 500 was $17.02. The percentages for the S&P 500 represent how much Q2 2009 EPS were expected to change from the Q2 2008 amount of $17.02 on a particular forecast date. For example, the estimate for Q2 2009 EPS on 30 September 2008 was that the year-over-year change would be an increase of 49.47 percent:

$$\$17.02(1 + 0.4947) \approx \$25.44$$

On 30 June 2009, the estimate for Q2 2009 EPS was that the year-over-year change would be a decrease of 17.38 percent:

$$\$17.02(1 - 0.1738) \approx \$14.06$$

Similarly, the percentages for the various industry sectors noted in Exhibit 9 reflect how much Q2 2009 EPS were expected to change from the Q2 2008 amount on a particular forecast date. Exhibit 9 shows that earnings forecast revisions can be significant.

Exhibit 9 Revisions to Bottom-Up Estimates of Operating Earnings per Share Standard & Poor's Forecasts: July 2009

Q2 Estimates	9/30/2008	12/31/2008	3/30/2009	6/30/2009	7/7/2009
S&P 500	49.47%	17.04%	−12.82%	−17.38%	−17.42%
Consumer Discretionary	134.52%	72.02%	−40.30%	36.61%	36.61%
Consumer Staples	14.52%	12.88%	6.83%	2.96%	4.12%
Energy	16.91%	−17.06%	−43.35%	−65.11%	−65.11%
Financials	691.43%	450.48%	289.93%	297.10%	293.35%

19 For company EPS estimates, Darrough and Russell (2002) showed that bottom-up forecasts are systematically more optimistic than top-down forecasts. The authors contend this occurs because analysts rely heavily on management's assessment of future profitability and such estimates often are overly optimistic.

Exhibit 9 (Continued)

Q2 Estimates	9/30/2008	12/31/2008	3/30/2009	6/30/2009	7/7/2009
Health Care	13.56%	10.41%	4.51%	3.06%	3.06%
Industrials	1.17%	−14.07%	−34.53%	−42.89%	−42.89%
Information Technology	24.12%	−5.04%	−28.02%	−26.31%	−26.31%
Materials	9.30%	−26.85%	−51.13%	−69.52%	−69.52%
Telecommunication Services	26.22%	6.67%	−9.77%	−5.64%	−5.44%
Utilities	14.61%	8.99%	−1.10%	−4.80%	−4.80%
S&P 500 EPS	$25.44	$19.92	$14.84	$14.06	$14.06

EXAMPLE 7

Bottom-Up and Top-Down Market EPS Forecasts

What considerations might encourage a market analyst to rely more on a top-down or bottom-up forecast of S&P 500 operating earnings?

Solution:

Bottom-up forecasts are based on consensus earnings estimates from equity research analysts covering the S&P 500 stocks. Top-down estimates are often based on econometric methods rather than fundamental analysis of the companies comprising the index.

Analysts frequently wait for information from the companies they follow to change their forecasts. Thus, bottom-up estimates may be more optimistic than top-down heading into a recession, and more pessimistic than top-down coming out. If the belief exists that companies are reacting slowly to changes in economic conditions, then a market analyst may prefer a top-down forecast.

However, top-down earnings forecasting models also have limitations. Most such models rely on the extrapolation of past trends in economic data. As a result, the impact of a significant contemporaneous change in a key economic variable or variables on the stock market may not be accurately predicted by the model. If the belief exists that the economy is on the brink of a significant change, then a market analyst may prefer the bottom-up forecast.

RELATIVE VALUE MODELS

4

Relative value investing is consistent with the popular trading maxim that investors should buy what is cheap and sell what is expensive. The relative value models presented in this section can be used to support the tactical asset allocation decision. They can help to identify times when investors would be well served switching from bonds to stocks, or vice-versa. As an investor, it is important to focus on the markets in a comparative fashion.

4.1 Earnings-Based Models

In its 22 July 1997 Humphrey-Hawkins report to Congress, the United States Federal Reserve compared 1982–1997 10-year Treasury note yields to the earnings yield of the S&P 500 and showed a very close correlation between the two. The **Fed model**, so named by Edward Yardeni of (at the time) Prudential Securities, was based primarily on the results of this report. However, use of the term "Fed model" is somewhat misleading as the model has never been formally adopted by the Federal Reserve as a policy-making tool.

The Fed model is a theory of equity valuation that hypothesizes that the yield on long-term US Treasury securities (usually defined as the 10-year T-note yield) should be equal to the S&P 500 earnings yield (usually defined as forward operating earnings divided by the index level) in equilibrium. Differences in these yields identify an overpriced or underpriced equity market. The model predicts:

- US stocks are undervalued if the forward earnings yield on the S&P 500 is greater than the yield on US Treasury bonds.
- US stocks are overvalued if the forward earnings yield on the S&P 500 is less than the yield on US Treasury bonds.

For example, if the S&P 500 forward earnings yield is 5 percent and the 10-year T-note yield is 4.5 percent, stocks would be considered undervalued according to the Fed model.

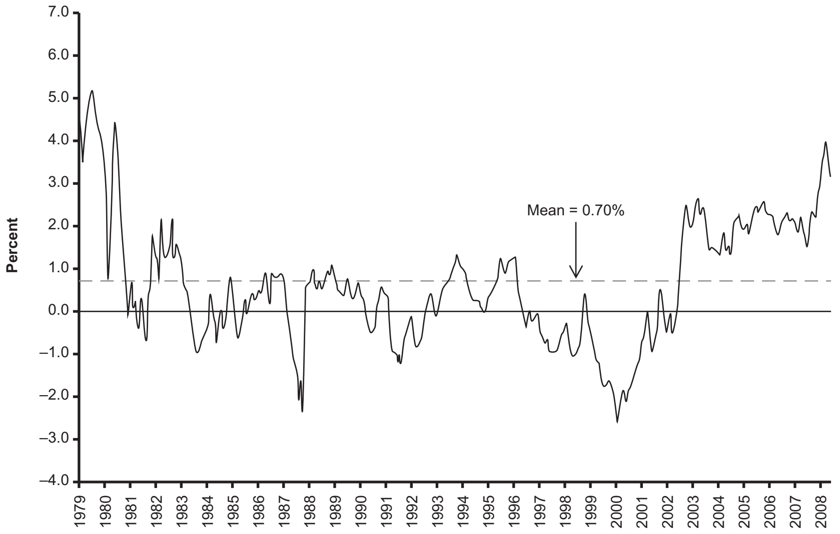
EXAMPLE 8

Fed Model with US Data

The difference between the S&P 500 earnings yield (based on forward operating earnings estimates) and the 10-year T-note yield for the time period January 1979 through December 2008 is presented in Exhibit 10.

The Fed model predicts that investors will be indifferent between investing in equities and investing in government bonds when the difference between the two yields is zero. Note that the average difference between the two yields was 0.70 percent during this time period. The positive difference between the two yields was at its greatest in December 2008. Thus, the model predicted that equities were significantly undervalued at that time, following the stock market sell-off during the second half of 2008. Similarly, the largest negative differences occurred prior to the collapse of the stock market bubbles in October 1987 and early 2000.

Exhibit 10 The Fed Model: Difference between the S&P 500 Forward Earnings Yield and Yield on 10-Year T-Note Monthly Data: January 1979–December 2008



Source: www.yardeni.com.

The key criticism of the Fed model is that it ignores the equity risk premium. (Informally, the **equity risk premium** is the compensation demanded by investors for the greater risk of investing in equities compared to investing in default-risk-free debt.) The validity of this criticism is apparent if one understands the assumptions necessary to derive the Fed model from the Gordon growth model. Equation 5 presents the Gordon growth model where V_0 is intrinsic value, D_1 is the dividend per share to be received one-year from today, r is the required return, and the constant annual dividend growth rate is g:

$$V_0 = \frac{D_1}{r - g} \tag{5}$$

Assuming markets correctly set price, P_0, equal to intrinsic value, then $P_0 = V_0$. The expected dividend, D_1, can be determined as the payout ratio, p, times expected earnings, E_1. Sustainable growth, g, can be estimated as return on equity, ROE, times the earnings retention rate, $(1 - p)$.

Substituting $D_1 = pE_1$ and $g = \text{ROE}(1 - p)$ into Equation 5 and noting that $P_0 = V_0$ we are able to derive Equation 6. Equation 6 provides a Gordon growth model estimate for the forward earnings yield, E_1/P_0:

$$\frac{E_1}{P_0} = \frac{r - \text{ROE}(1 - p)}{p} \tag{6}$$

The Fed model hypothesizes that the earnings yield, E_1/P_0, and the yield on Treasury bonds, y_T, are equal in equilibrium. One way to produce this equilibrium using Equation 6 is to assume that the required return, r, and the return on equity, ROE, are equal to the Treasury bond yield, y_T. Making these substitutions in Equation 6 shows this result:

$$\frac{E_1}{P_0} = \frac{r - \text{ROE}(1-p)}{p} = \frac{y_T - y_T(1-p)}{p} = y_T \tag{7}$$

Thus, implicit in the Fed model equilibrium are the assumptions that the required return, r, and the accounting rate of return on equity, ROE, for risky equity securities are equal to the Treasury bond yield, y_T. Historical evidence and financial theory resoundingly reject the notion that either assumption is true. For example, the long-run average return on US equities has exceeded the long-run average return on T-bonds by a significant amount.[20] Because of this, many analysts consider the Fed model flawed.

Two additional criticisms of the Fed model are that it ignores inflation and earnings growth opportunities. Asness (2003) criticized the Fed model because it compares an arguably real variable, the earnings yield, to a nominal variable, the T-bond yield. According to this argument, the earnings yield is real because it is a ratio of current period prices.[21] The T-bond yield is nominal because it is reflective of the expected rate of inflation as first noted by Fisher (1930). In the presence of inflation, investors should compare the earnings yield with a real interest rate. Asness provides evidence that the Fed model has often been a poor predictor of future equity returns.

Another criticism of the Fed model is that it ignores any earnings growth opportunities available to equity holders beyond those forecasted for the next year (as reflected by expected earnings, E_1). In the United States, long-term compound average earnings growth has been 3–4 percent nominal and 1–2 percent real.[22] Thus, the model ignores a significant portion of total equity return.[23]

In spite of the several criticisms, the Fed model still can provide some useful insights. It does suggest that equities become more attractive as an asset class when interest rates decline. This is consistent with the predictions of any discounted cash flow model and is supported by market evidence. In practice, the model typically makes use of expected earnings (a future cash flow) as an input, which is again consistent with traditional discounted cash flow analysis.

Some analysts find a comparison of the earnings yield and Treasury bond yield to be most useful when the relationship is towards the extremes of its typical range. For example, some analysts compare the current difference between the earnings yield and the Treasury bond yield with the historical average difference. Stocks are viewed as more attractive as an investment when the current period difference significantly exceeds the historical average difference.

---

20 Bodie, Kane, and Marcus (2007) show that the geometric average annual return on US large capitalization stocks was 10.23 percent over the 1926 to 2003 time period. The geometric average annual return on long-term Treasury bonds was 5.10 percent during this same time period.

21 Wilcox (2007) and Palkar and Wilcox (2009) note that accounting and debt adjustments must be made to GAAP-based reported earnings before they can be considered real.

22 Based on a dataset maintained by Professor Shiller (www.econ.yale.edu/~shiller/data.htm), the compound (geometric mean) annual earnings growth rate was 3.52% from 1872–2008. In the more recent shorter term, the growth rate has been higher: from 1990–2008, the rate was 3.55%, but stopping at 2007 (i.e., from 1990–2007), reflecting an unusually long-period of sustained growth, it was 6.89%. The long-run rate is probably most appropriate here.

23 The required return on equity for a no-growth company that pays all of its earnings out as dividends is the earnings yield (based on a constant EPS).

EXAMPLE 9

Fed Model with UK Data

The Fed model can be applied to the valuation of non-US equity markets. In early 2009, the Fed model produced a very bullish prediction for British stocks. The forecasted earnings yield on the FTSE 100 was 10.1 percent and the yield on 10-year UK government debt was 3.6 percent. The difference between these yields was much greater than the long-term average, according to Citigroup data, of 4.5 percent.

Analysts should always question the inputs to any valuation model. Reasonable questions for these results include "Can the government bond yield be expected to rise?" and "Can the forecast for earnings be expected to decline?" Most would likely agree that the latter question was of greater concern in early 2009.

EXAMPLE 10

Fed Model Questions

1 Assume the S&P 500 forward earnings yield is 5 percent and the 10-year T-note yield is 4.6 percent. Are stocks overvalued or undervalued according to the Fed model?

2 Why might the earnings yield be considered a poor measure for the true worth of equities?

Solution to 1:

According to the Fed model, stocks are undervalued because the forward S&P 500 earnings yield exceeds the 10-year T-note yield. However, recall from Example 8 that the average difference between the S&P 500 earnings yield and the 10-year T-note yield for the time period January 1979–December 2008 was 0.70 percent. In this question, the difference between the two yields is 0.40 percent. Analysts who compare the difference in yields to this average difference would contend that equities are overvalued.

Solution to 2:

The forward earnings yield measure used in the Fed model to assess the worth of equities fails to accurately capture the long-term growth opportunities available to equity investors. Although studies show that the dividend yield has been the major determinant of long-term equity returns, the impact of earnings growth has been significant and arguably should not be ignored.

The **Yardeni model** addresses some of the criticisms of the Fed model. In creating the model, Yardeni (2002) assumed investors valued earnings rather than dividends. With the assumption that markets set price equal to intrinsic value, $P_0 = V_0$, a constant growth valuation model that values earnings is presented in Equation 8. E_1 is an estimate of next year's earnings, r is the required return, and g is the earnings growth rate. Equation 8 shows that, given the assumptions of the model, the earnings yield, E_1/P_0, is equal to the required return, r, minus the growth rate, g.

$$P_0 = \frac{E_1}{r-g} \Rightarrow \frac{E_1}{P_0} = r - g \tag{8}$$

As a data input for the required return, r, Yardeni used the Moody's A-rated corporate bond yield, y_B, which allowed for risk to be incorporated into the model. The risk premium captured by the model, however, is largely a default risk premium (the credit spread between the A-rated bond, y_B, and the yield on a Treasury bond, y_T), not the unobservable equity risk premium. Thus, while an improvement over the Fed model, the Yardeni model still does not fully capture the risk of equities.

As an input for the growth rate, g, Yardeni used the consensus five-year earnings growth forecast for the S&P 500 from Thomson Financial, LTEG. Note that g is truly a perpetual or sustainable growth rate and that a five-year forecast for growth may not be sustainable.

The Yardeni model introduces an additional variable, the coefficient d. It represents a weighting factor measuring the importance the market assigns to the earnings projections. Yardeni (2000) found that the historical values for d averaged about 0.10.[24] However, depending on market conditions, d can vary considerably from its historical average. Equation 9 presents the Yardeni model stated as the justified (forward) earnings yield on equities.

$$\frac{E_1}{P_0} = y_B - d \times \text{LTEG} \tag{9}$$

A justified forward earnings yield that is below, equal to, or greater than the forward earnings yield value implied by current equity market index values (using consensus forward earnings estimates, for example) would indicate that equities are undervalued, fairly valued, or overvalued in the marketplace. A valuation judgment can also be made by using Equation 9 solved for P_0, which gives the Yardeni model expression for the fair value of the equity market: $E_1/(y_B - d \times \text{LTEG})$. The judgment would be that the equity market is undervalued, fairly valued, or overvalued if the fair value estimate is above, equal to, or below the current equity market level. Example 11 shows such an analysis.

EXAMPLE 11

The Yardeni Model (1)

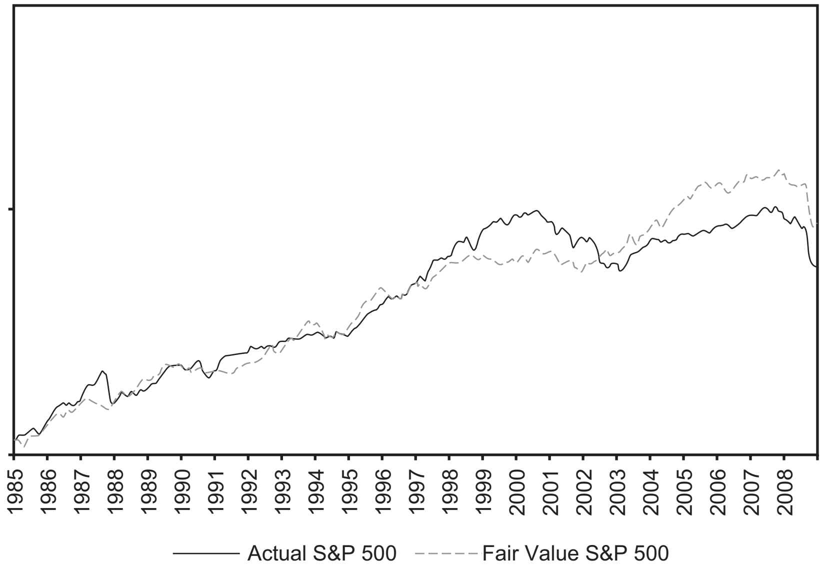
Exhibit 11 presents in logarithmic ("log") scale the actual S&P 500 Index and a fair value estimate of the S&P 500 using the Yardeni model assuming d = 0.10. The time period is January 1985 to December 2008.

As Exhibit 12 shows, the Yardeni model predicted the S&P 500 was undervalued by 39.25 percent in December 2008. The Yardeni model also did a good job predicting the overvaluation and subsequent pullbacks of October 1987 and the early 2000s. However, the model signaled the equity market was significantly undervalued in 2007 even though US and other world equity markets collapsed dramatically in the wake of a major financial crisis.

---

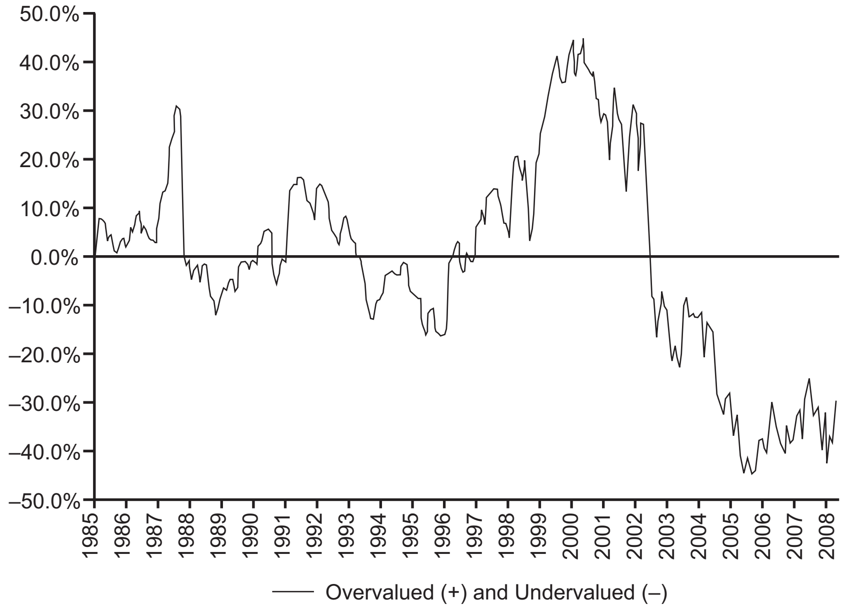
24 Note that rearranging the terms in Equation 9 so that they produce a formula for d results in $d = \dfrac{y_B - \dfrac{E_1}{P_0}}{\text{LTEG}}$. Thus, historical values for d can be estimated from market data.

Exhibit 11 S&P 500 Index and Fair Value Estimate Using Yardeni Model with $d = 0.10$ (Log Scale) Monthly Data: January 1985–December 2008



Source: www.yardeni.com.

Exhibit 12 Overvaluation (+) and Undervaluation (–) of S&P 500 Index versus Fair Value Estimate Using Yardeni Model with $d = 0.10$ Monthly Data: January 1985–December 2008



Source: www.yardeni.com.

EXAMPLE 12

The Yardeni Model (2)

1 Assume the Moody's A-rated corporate bond yield is 6.49 percent and the forecast for long-term earnings growth is 11.95 percent. Determine the Yardeni model estimate of the fair value earnings yield assuming d = 0.05 and d = 0.10. Are equities overvalued or undervalued if the S&P 500 earnings yield is 5.5 percent?

2 Assume the Moody's A-rated corporate bond yield is 6.32 percent and the forecast for long-term earnings growth is 11.5 percent. Determine the Yardeni model estimate of the fair value price–earnings (P/E) ratio assuming d = 0.10. When would equities be undervalued? When would equities be overvalued?

3 A Indicate the directional relationship predicted by the Yardeni model between changes in y_B, LTEG, and d and changes in the earnings yield.

B Indicate the directional relationship predicted by the Yardeni model between changes in y_B, LTEG, and d and changes in the P/E ratio.

Solution to 1:

For d = 0.05: 0.0649 − 0.05(0.1195) ≈ 0.0589 ≈ 5.89% is the Yardeni model estimate. Equities are overvalued as 5.5% < 5.89%.

For d = 0.10: 0.0649 − 0.10(0.1195) ≈ 0.0530 ≈ 5.30% is the Yardeni model estimate. Equities are undervalued as 5.5% > 5.30%.

Solution to 2:

P/E is the reciprocal of the earnings yield. The Yardeni estimate of the fair value P/E ratio would be 1 ÷ [0.0632 − 0.10(0.115)] or approximately 19.3. Stocks would be undervalued if the actual P/E ratio for the S&P 500 is less than 19.3. Stocks would be overvalued if the actual P/E ratio for the S&P 500 is greater than 19.3.

Solution to 3A:

Increases in y_B and decreases in d and LTEG produce higher fair value estimates of the earnings yield.

Solution to 3B:

Decreases in y_B and increases in d and LTEG produce higher fair value estimates of the P/E ratio.

Examples 11 and 12 were taken from US equity markets. To date, nearly all analysis using the Yardeni model has been limited to the US equity market. If adequate data are available, especially for forecasted earnings growth, the Yardeni model could be applied in any equity market.

Campbell and Shiller's (1998, 2005) 10-year Moving Average Price/Earnings, more commonly known as the **Cyclically Adjusted P/E Ratio (CAPE)**, has become a popular measure of market valuation. The authors defined the numerator of CAPE as the real S&P 500 price index and the denominator as the moving average of the preceding 10 years of real reported earnings. "Real" denotes that the stock index and earnings are adjusted for inflation using the Consumer Price Index (CPI). The purpose of the 10-year moving average of real reported earnings is to control for business cycle effects on earnings and is based on recommendations from the seminal Graham and Dodd (1934) text.

EXAMPLE 13

Determining CAPE: A Historical Exercise

For the purpose of illustrating the calculation of CAPE one can use data from any period. Exhibit 13 is a historical exercise showing the calculation of CAPE for US equities as of 1881. The real stock price index and real earnings are priced in 2009 US dollars and are determined using the January 2009 CPI value of 211.143. Note that:

$$\text{Real stock price index}_t = \text{Nominal stock price index}_t \times \text{CPI}_{2009} \div \text{CPI}_t$$
$$\text{Real earnings}_t = \text{Nominal earnings}_t \times \text{CPI}_{2009} \div \text{CPI}_{t+1}$$

$$\text{Real stock price index}_{1871} = 4.44 \times 211.143 \div 12.464061 = 75.21424358$$
$$\text{Real earnings}_{1871} = 0.4 \times 211.143 \div 12.654392 = 6.674141278$$

The CAPE in 1881 of 18.21479737 is the real stock price index in 1881 of 138.7532563 divided by average real earnings from 1871 to 1880 of 7.617611851.

$$\text{Average real earnings}_{1871\rightarrow 1880} = (6.674141278 + \ldots + 10.9836988) \div 10 = 7.617611851$$

$$\text{CAPE}_{1881} = 138.7532563 \div 7.617611851 = 18.21479737$$

Exhibit 13 Determining CAPE in 1881[25]

Year	Stock Price Index (January)	Earnings Accruing to Index	Consumer Price Index	Real Stock Price Index	Real Earnings	CAPE
1871	4.44	0.40	12.464061	75.21424358	6.674141278	
1872	4.86	0.43	12.654392	81.09081653	7.016448545	
1873	5.11	0.46	12.939807	83.38151643	7.852421105	
1874	4.66	0.46	12.368896	79.54843989	8.436439183	
1875	4.54	0.36	11.5126510	83.26398672	7.007878524	
1876	4.46	0.28	10.8465750	86.81982838	5.403166224	
1877	3.55	0.30	10.9417400	68.50442891	6.863396587	
1878	3.25	0.31	9.2290893	74.35346302	7.907328567	
1879	3.58	0.38	8.2776793	91.31689120	8.031199688	

(continued)

25 The number of decimal places shown reflects the precision given in the Shiller dataset referenced.

Exhibit 13 (Continued)

Year	Stock Price Index (January)	Earnings Accruing to Index	Consumer Price Index	Real Stock Price Index	Real Earnings	CAPE
1880	5.11	0.49	9.9903306	107.99850110	10.983698800	
1881	6.19	0.44	9.4194198	138.75325630		18.21479737

Source: www.econ.yale.edu/~shiller/data.htm.

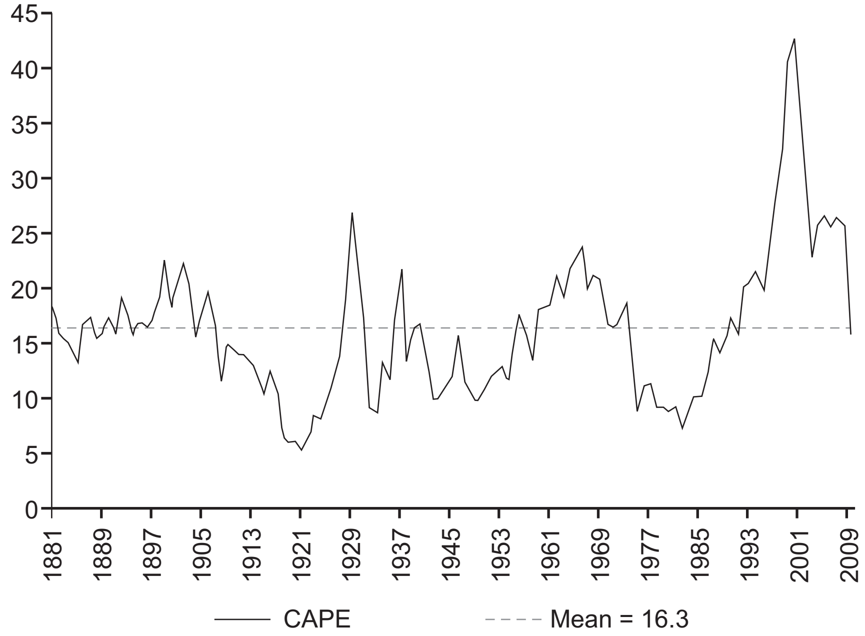
Real earnings in 1880 exceeded the 10-year moving average by a considerable amount, and this is a typical result. This discrepancy undoubtedly reflects the real growth in corporate earnings. If some smoothing for business cycle effects is necessary, the case can be made that it would be better to compute a moving average of the P/E ratio (current period price divided by current period earnings).

Many analysts believe that CAPE should be considered a mean-reverting series. Exhibit 14 presents CAPE from January 1881 to January 2009. The mean value of CAPE for this time period was 16.3 and the January 2009 CAPE was 15.8, suggesting the US equity market was slightly undervalued at that time. The highest value for CAPE was 42.5 in 2000, and the lowest value for CAPE was 5.3 in 1921.

Campbell and Shiller (1998, 2005) made the case that the US equity market was extremely overvalued in the late 1990s and provided evidence that future 10-year real price growth was negatively related to CAPE. Exhibit 15 updates the Campbell-Shiller results through 2009. Each plotted data point represents an annual observation for real price growth for the next 10 years and CAPE for that same year. A trend line is plotted and shows the ordinary least squares regression relationship between 10-year real price growth and CAPE.

The table in the upper right-hand corner of Exhibit 15 shows the predicted 10-year real price growth given some value for the explanatory variable CAPE. The regression results predict that 10-year real price growth will be 17.9 percent given the January 2009 CAPE of 15.8. The R-squared for the regression is 0.1488, which indicates that CAPE explains only 14.88 percent of the variation in 10-year real price growth. Furthermore, the traditional regression statistics for this regression are unreliable because of the serial correlation induced by the overlapping time periods used to compute returns.

Exhibit 14 CAPE Annual Data: 1881–2009

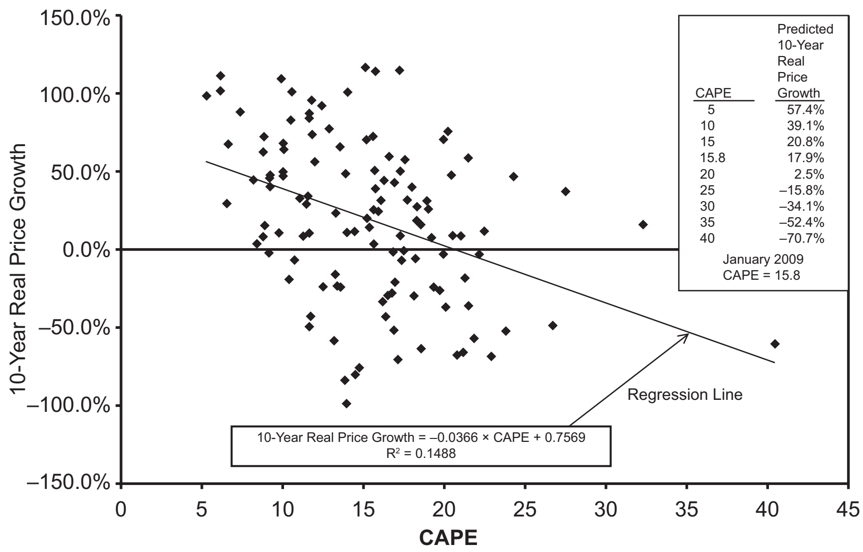


Source: www.econ.yale.edu/~shiller/data.htm.

With a 10-year projected real price growth of 17.9 percent, the annualized growth rate is less than 1.7 percent per year. Given a dividend yield of about 2.9 percent and assuming share repurchases effectively add 1 percent to the annual real cash flow to shareholders, the inflation-adjusted expected return would be approximately 5.6 percent, which is below the 6–7 percent compounded average inflation-adjusted return over long periods in the United States (Siegel 2005). Thus, in contrast to the Fed model and the Yardeni model, the Campbell and Shiller model implies below-average prospective returns.

The conflicting signals between and among various valuation models may provide valuable insights, especially if they cause us to rethink how the parameter estimates and the numerical inputs give rise to the different results. For example, assume that changes in accounting rules lead to significant differences in how earnings are reported over time. Thus, the CAPE at some given point in time might not be comparable with other time periods. A high or low CAPE relative to the long-term average at present could be due to differences in prior accounting rules, thereby resulting in stocks actually being more undervalued or overvalued than they currently appear.

Exhibit 15 CAPE and Predicted 10-Year Real Price Growth



Source: www.econ.yale.edu/~shiller/data.htm.

EXAMPLE 14

CAPE Questions

1 What adjustments are made to earnings in determining CAPE?

2 Assume CAPE reached an all-time high of 42.5 in 2000. Use the regression results in Exhibit 15 to determine predicted real price growth for the time period 2000–2009.

Solution to 1:

Following Graham and Dodd, Campbell and Shiller averaged earnings over a 10-year time period. Their goal was to normalize earnings by providing an estimate of what earnings would be under mid-cyclical conditions. The implicit assumption is that the typical business cycle lasts 10 years.

Campbell and Shiller also control for inflation by adjusting past earnings to current period dollars using the Consumer Price Index.

Solution to 2:

10-year Real price growth = $-0.0366 \times 42.5 + 0.7569 = -0.7986 = -79.86\%$.

4.2 Asset-Based Models

Tobin's *q* ratio, pioneered in Brainard and Tobin (1968) and Tobin (1969), is an asset-based valuation measure. Tobin's q has been used for several purposes, including decision-making concerning physical capital investment and equity market valuation. The first application is the simplest: At the company level, Tobin's q is calculated as the market value of a company (i.e., the market value of its debt and equity) divided by the replacement cost of its assets. According to economic theory, Tobin's q is approximately equal to 1 in equilibrium. If it is greater than 1 for a company, the marketplace values the company's assets at more than their replacement costs, so additional capital investment should be profitable for the company's suppliers of financing. By contrast, a Tobin's q below 1 indicates that further capital investment is unprofitable.

Tobin's q has also been calculated at an overall market level. In that case, the denominator involves an estimate of the replacement cost of aggregate corporate assets, and the numerator involves estimates of aggregate equity and debt market values. Some analysts have used a market-level Tobin's q to judge whether an equity market is misvalued. This application involves a comparison of the current value of market-level Tobin's q with its presumed equilibrium value of 1 or with its historical mean value. Assuming that Tobin's q will revert to the comparison value, a Tobin's q below, at, or above the comparison values is interpreted as the market being undervalued, fairly valued, or overvalued. Strong economic arguments exist that both Tobin's q and equity q, discussed later, should be mean-reverting series.

The calculation of Tobin's q often poses difficulties. At the company level, it is usually possible to get a fairly accurate estimate of market value (the numerator of Tobin's q) by summing the values of the securities a company has issued, such as its stocks and bonds. It is much more difficult to obtain an accurate estimate of replacement costs of the company's assets (the denominator of Tobin's q). Liquid markets for many assets (e.g., many kinds of industrial equipment) do not exist. Moreover, such items as human capital, trade secrets, copyrights and patents, and brand equity are intangible assets that are often difficult to value. Typically, researchers who try to construct Tobin's q ignore the replacement cost of intangible assets in their calculations.

Smithers and Wright (2000) created an **equity *q*** that is the ratio of a company's equity market capitalization divided by net worth measured at replacement cost. Their measure differs from the price-to-book value ratio because net worth is based on replacement cost rather than the historic or book value of equity. Based on a market-level equity q, Smithers and Wright made the case that the US equity market was extremely overvalued in the late 1990s. The principles of that application parallel those given for Tobin's q.

To date, much of the market-level analysis using Tobin's q or equity q has been conducted in the US equity market, but analysis based on European and Asian equity markets is increasingly available.

EXAMPLE 15

Market-Level Analysis of Tobin's *q* and Equity *q*

Data from which Tobin's q and equity q can be calculated are published in the *Flow of Funds Accounts of the United States-Z.1*, published quarterly by the Federal Reserve.[26] This data source is available from 1952 onwards. Below are data for Nonfarm Nonfinancial Corporate Business for the fourth quarter of 2008 (billions of US dollars). Based on this data, determine Tobin's q and equity q.

Assets at Market Value or Replacement Cost	Liabilities	Market Value of Equities Outstanding
28,277.33	12,887.51	9,554.05

Source: www.federalreserve.gov/releases/z1/.

26 Specifically, the data currently appear in Table B.102 of that publication.

Solution:

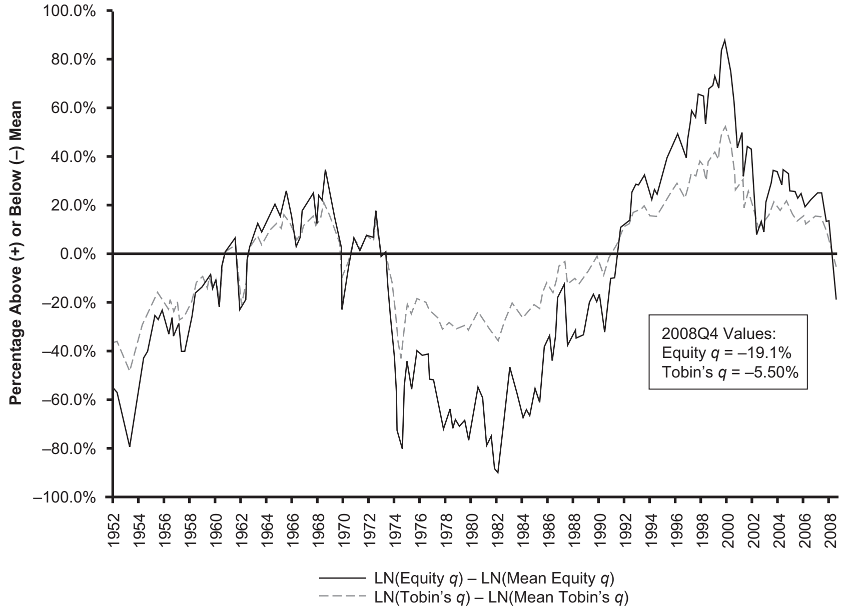
Tobin's q = (12,887.51 + 9,554.05) ÷ 28,277.33 = 0.793623726

Equity q = 9,554.05 ÷ (28,277.33 – 12,887.51) = 0.620803232

Using this data, the long-term average for Tobin's q and equity q are both significantly below 1.0. Smithers and Wright suggest this is due to the true economic rate of depreciation being underestimated, which leads to the replacement cost of assets being overstated. Such overstatement means that the denominator in both formulations of the q ratio is too high and that the correctly measured ratios should be much higher.

Exhibit 16 presents quarterly Tobin's q data and quarterly equity q data for the US equity market over the time period 1952 to 2008. Last quarter 2008 values for these two variables relative to their respective means suggest that the US equity market was slightly undervalued at that time. However, both series had declined to significantly lower levels relative to their means in the early 1950s and early 1980s.

Exhibit 16 Equity q and Tobin's q Quarterly Data: (1952 Q1 to 2008 Q4)



Source: www.federalreserve.gov/releases/z1/.

EXAMPLE 16

Tobin's q and Equity q

1 Why should Tobin's q be expected to mean revert?

2 How does equity q differ from the price-to-book ratio?

Solution to 1:

If Tobin's q is greater than 1.0, then the market is valuing a company at more than it costs to replace its assets. Either security prices must fall or the company should continue to invest in new assets until the ratio returns to its equilibrium. If Tobin's q is below 1.0, then security prices are undervalued because new businesses cannot be created as cheaply as they can be bought in the open market. Either security prices must rise or the company should sell some of its assets until the ratio returns to its equilibrium.

Solution to 2:

Book value in the price-to-book ratio reflects the value of equity that is reported on the company's balance sheet. The denominator of equity q reflects the difference between the replacement cost of assets and the market value of liabilities. Most financial reporting standards require the use of acquisition cost as a measure of asset value. Thus, the book value of assets is typically less than their replacement cost, and this is particularly true during periods of rising prices.

A summary of the relative value models appears in Exhibit 17.

Exhibit 17 Summary of Relative Value Models

Model	Predictions of the Model	Strengths	Limitations
Fed model	The equity market is undervalued if its earnings yield exceeds the yield on government securities.	■ Easy to understand and apply. ■ Consistent with discounted cash flow models that show an inverse relationship between value and the discount rate.	■ Ignores the equity risk premium. ■ Compares a real variable to a nominal variable. ■ Ignores earnings growth.
Yardeni model	Equities are overvalued if the fair value estimate of the earnings yield provided by the model exceeds the actual earnings yield for the market index.	■ Improves on the Fed model by including the yield on risky debt and a measure of expected earnings growth as determinants of value.	■ Risk premium captured by the model is largely a default risk premium that does not accurately measure equity risk. ■ The forecast for earnings growth may not be accurate or sustainable. ■ The estimate of fair value assumes the discount factor investors apply to the earnings forecast remains constant over time.

(continued)

Exhibit 17 (Continued)

Model	Predictions of the Model	Strengths	Limitations
CAPE	Future equity returns will be higher when CAPE is low.	■ Controls for inflation and business cycle effects by using a 10-year moving average of real earnings. ■ Historical data supports an inverse relationship between CAPE and future equity returns.	■ Changes in the accounting methods used to determine reported earnings may lead to comparison problems. ■ Current period or other measures of earnings may provide a better estimate for equity prices than the 10-year moving average of real earnings. ■ Evidence suggests that both low and high levels of CAPE) can persist for extended periods of time.
Tobin's *q* and equity *q*	Future equity returns will be higher when Tobin's *q* and equity *q* are low.	■ Both measures rely on a comparison of security values to asset replacement costs (minus the debt market value, in the case of equity *q*); economic theory suggests this relationship is mean-reverting. ■ Historical data supports an inverse relationship between both measures and future equity returns.	■ It is difficult to obtain an accurate measure of replacement cost for many assets because liquid markets for these assets do not exist and intangible assets are often difficult to value. ■ Evidence suggests that both low and high levels of Tobin's *q* and equity *q* can persist for extended periods of time.

EXAMPLE 17

Questions Regarding the Relative Value Models

1. Which of the models ignore the current level of market interest rates as determinants of equity market value?
2. Under what conditions might the Fed model and Yardeni model provide a different assessment of the value of the equity market?
3. Which of the models use some measure of earnings as an input? How might this lead to comparison issues?

Solution to 1:

In assessing equity market value, CAPE, Tobin's *q*, and equity *q* are typically compared to their long-term averages and not to market interest rates. While the Yardeni model compares the fair value earnings yield predicted by the model to the actual earnings yield, the A-rated corporate bond yield is an input to the model.

Solution to 2:

The Fed model compares the earnings yield to the Treasury bond yield. The Yardeni model uses the A-rated corporate bond yield and the consensus 5-year earnings growth forecast to determine a fair value earnings yield. One scenario in which the two models might differ in their predictions would be if the default risk premium on the A-rated corporate bond was currently high, the Treasury bond yield was currently low, and earnings were forecasted to grow at a slow rate. Given these assumptions, the Fed model might indicate that equities are undervalued while the Yardeni model indicates equities are overvalued.

Solution to 3:

The Fed model, Yardeni model, and CAPE all use some measure of earnings as a determinant of value. Time series comparisons will be problematic if the accounting methods used to determine earnings change over time.

SUMMARY

In this reading, we have investigated several ways in which economic theory can be applied to the valuation of equity markets. Among the major points are the following:

- The growth accounting equation allows one to decompose real GDP growth, $\Delta Y/Y$, into components that can be attributed to the observable factors: the growth of the capital stock, $\Delta K/K$, the output elasticity of capital, α, the growth in the labor force, $\Delta L/L$, the output elasticity of labor, $1 - \alpha$, and a residual factor—often called the Solow residual—that is the portion of growth left unaccounted for by increases in the standard factors of production, $\Delta A/A$.

$$\frac{\Delta Y}{Y} \approx \frac{\Delta A}{A} + \alpha\frac{\Delta K}{K} + \left(1 - \alpha\right)\frac{\Delta L}{L}$$

- The existence of TFP growth, $\Delta A/A$, means that total output can grow at a faster rate than would be predicted simply from growth in accumulated capital stock and the labor force. TFP is typically linked to innovation and technical progress. However, changes in work organization, government regulation, and the literacy and skills of the workforce, as well as many other factors, also affect TFP.
- The inputs for the H-model include the initial growth rate, g_S, a period of years, N, where the dividend growth rate declines in a linear fashion, and a long-term dividend growth rate, g_L, that is expected to prevail to perpetuity beginning at the end of period N. With an initial annualized dividend rate D_0 and a discount rate to perpetuity of r, the formula for intrinsic value, V_0, according to the H-model is:

$$V_0 = \frac{D_0}{r - g_L}\left[\left(1 + g_L\right) + \frac{N}{2}\left(g_S - g_L\right)\right]$$

- In top-down forecasting, analysts use macroeconomic forecasts to develop market forecasts and then make industry and security forecasts consistent with the market forecasts. In bottom-up forecasting, individual company forecasts are aggregated to industry forecasts, which in turn are aggregated to produce a market forecast.
- Bottom-up forecasts tend to be more optimistic than top-down forecasts. Top-down models can be slow in detecting cyclical turns if the current statistical relationships between economic variables deviate significantly from their historic norms.

- The Fed model is a theory of equity valuation that hypothesizes that the yield on long-term US Treasury securities (usually defined as the 10-year T-note yield) should be equal to the S&P 500 earnings yield (usually defined as forward operating earnings divided by the index level) in equilibrium.
- A common criticism of the Fed model equilibrium is that it fails to incorporate the equity risk premium. The earnings yield can also be a poor measure of the true value of equities if significant growth opportunities exist. Some authors have also argued that the Fed model comparison is flawed because the earnings yield is real while the Treasury yield is nominal.
- The Yardeni model addresses some of the criticisms of the Fed model. As inputs, Yardeni used the Moody's A-rated corporate bond yield, y_B, the consensus five-year earnings growth forecast for the S&P 500 from Thomson Financial, LTEG, and the coefficient d, which represents a weighting factor measuring the importance the market assigns to the earnings projections. Yardeni found that the historical values for d averaged about 0.10. The formula for the Yardeni model is:

$$\frac{E_1}{P_0} = y_B - d \times \text{LTEG}$$

- Limitations of the Yardeni model include that the risk premium captured by the model is largely a default risk premium and not the future equity risk premium, which is unobservable. Also, the consensus five-year earnings growth forecast for the S&P 500 from Thomson Financial may not be sustainable, and evidence suggests that the weighting factor varies significantly over time.
- Campbell and Shiller's CAPE has become a popular measure of market valuation. The numerator of CAPE is the real S&P 500 and the denominator is the moving average of the preceding 10 years of real reported earnings. "Real" denotes that the stock index and earnings are adjusted for inflation using the Consumer Price Index (CPI). The purpose of the 10-year moving average of real reported earnings is to control for business cycle effects on earnings and is based on recommendations from the seminal work of Graham and Dodd.
- Tobin's q is calculated at the individual company level as the market value of a company divided by the replacement cost of its assets. Smithers and Wright created an equity q that is the ratio of a company's market capitalization divided by net worth measured at replacement cost. Market-level measures may be computed for Tobin's q and equity q by a process of aggregation; these market-level measures may be used to form a valuation judgment about an equity market. Assuming that Tobin's q will revert to the comparison value, a Tobin's q below, at, or above the comparison value is interpreted as the market being under-, fairly, or overvalued. Strong economic arguments exist that both Tobin's q and equity q should be mean-reverting series.
- In practice, estimating replacement cost can be problematic due to the lack of liquid markets for many assets. Moreover, such items as human capital, trade secrets, copyrights and patents, and brand equity are intangible assets that are difficult to value.

REFERENCES

Arnott, Robert D. 2004. "The Meaning of a Slender Risk Premium." *Financial Analysts Journal*, vol. 60, no. 2:6–8.

Arnott, Robert D., and Peter L. Bernstein. 2002. "What Risk Premium is 'Normal'?" *Financial Analysts Journal*, vol. 58, no. 2:64–85.

Asness, Clifford S. 2003. "Fight the Fed Model." *Journal of Portfolio Management*, vol. 30, no. 1:11–24.

Bodie, Zvi, Alex Kane, and Alan Marcus. 2007. *Essentials of Investments*, 6th Edition. New York: McGraw-Hill Irwin.

Brainard, William C., and James Tobin. 1968. "Pitfalls in Financial Model Building." *American Economic Review*, vol. 58, no. 2:99–122.

Campbell, John Y., and Robert J. Shiller. 1998. "Valuation Ratios and the Long-Run Stock Market Outlook." *Journal of Portfolio Management*, vol. 24, no. 2:11–26.

Campbell, John Y., and Robert J. Shiller. 2005. "Valuation Ratios and the Long-Run Stock Market Outlook: An Update." *Behavioral Finance II*. Edited by R. Thaler. New York: Russell Sage Foundation.

Cobb, C.W., and Paul H. Douglas. 1928. "A Theory of Production." *American Economic Review*, vol. 18, no. Supplement:139–165.

Darrough, Masako N., and Thomas Russell. 2002. "A Positive Model of Earnings Forecasts: Top Down versus Bottom Up." *Journal of Business*, vol. 75, no. 1:127–152.

Fisher, Irving. 1930. *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*. New York: Macmillan Company.

Fuller, Russell J., and Chi-Cheng Hsia. 1984. "A Simplified Common Stock Valuation Model." *Financial Analysts Journal*, vol. 40, no. 5:49–56.

Graham, Benjamin, and David L. Dodd. 1934. *Security Analysis*. New York: McGraw-Hill.

Ibbotson, Roger G., and Rex A. Sinquefield. 1989. *Stocks, Bonds, Bills, and Inflation: Historical Returns (1926–1987)*. Chicago: Dow-Jones Irwin.

Ibbotson, Roger G., Jeffrey J. Diermeier, and Laurence B. Siegel. 1984. "The Demand for Capital Market Returns: A New Equilibrium Theory." *Financial Analysts Journal*, vol. 40, no. 1:22–33.

Mehra, Rajnish. 2003. "The Equity Premium: Why Is It a Puzzle?" *Financial Analysts Journal*, vol. 59, no. 1:54–69.

Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Risk Premium: A Puzzle." *Journal of Monetary Economics*, vol. 15, no. 2:145–162.

Palkar, Darshana, and Stephen E. Wilcox. 2009. "Adjusted Earnings Yields and Real Rates of Return." *Financial Analysts Journal*, vol. 65, no. 5:66–79.

Preston, R. 2009. "China's Foreign Reserves Top $2tn." *BBC News*, 15 July 2009.

Siegel, Jeremy J. 1992. "The Equity Premium: Stock and Bond Returns Since 1802." *Financial Analysts Journal*, vol. 48, no. 1:28–38.

Siegel, Jeremy J. 2005. "Perspectives on the Equity Risk Premium." *Financial Analysts Journal*, vol. 61, no. 6:61–73.

Smithers, Andrew, and Stephen Wright. 2000. *Valuing Wall Street*. New York: McGraw-Hill.

Solow, Robert. 1957. "Technical Change and the Aggregate Production Function." *The Review of Economics and Statistics*, vol. 39:312–320.

Stimes, Peter C. 2008. *Equity Valuation, Risk, and Investment: A Practitioner's Roadmap*. Hoboken, NJ: John Wiley & Sons.

Tobin, James. 1969. "A General Equilibrium Approach to Monetary Theory." *Journal of Money, Credit and Banking*, vol. 1, no. 1:15–29.

Wilcox, Stephen E. 2007. "The Adjusted Earnings Yield." *Financial Analysts Journal*, vol. 63, no. 5:54–68.

Xin, Z., and S. Rabinovitch. 2009. "China Central Bank Taps on Brakes as Money, Supply Surges." *Reuters*, 15 July 2009.

Yardeni, Edward E. 2000. "How to Value Earnings Growth." *Topical Study #49* (Deutsche Banc Alex Brown).

Yardeni, Edward E. 2002. "Stock Valuation Models," Prudential Financial Research, *Topical Study #56* (August 8).

Zheng, Jinghai, Arne Bigsten, and Anagang Hu. 2006. "Can China's Growth be Sustained? A Productivity Perspective." *Working Papers in Economics* 236, Göteborg University, Department of Economics.

Zheng, Jinghai, Angan Hu, and Arne Bigsten. 2009. "Potential Output in a Rapidly Developing Economy: The Case of China and a Comparison with the United States and the European Union." *Federal Reserve Bank of St. Louis Review*, vol. 91, no. 4:317–348.

PRACTICE PROBLEMS

1 Elizabeth Villeneuve is a senior economist at Proplus Financial Economics Consulting (Proplus). She is responsible for the valuation of equity markets in developing countries and is reviewing the preliminary report on Emerge Country prepared by one of her analysts, Danielle DeLaroche. Emerge Country is now experiencing stronger economic growth than most developed countries.

DeLaroche summarized in Exhibit 1 some of the assumptions made in the report. In modeling the growth in the country's real output, she used the Cobb-Douglas production function under the assumption of constant returns to scale and, in valuing the equity market, she used the standard Gordon growth model with constant dividend growth rate.

Exhibit 1 Assumptions about the Equity Market Index of Emerge Country

Annual dividend per share in 2014	450 CU*
Forecasted earnings per share in 2015	750 CU*
Forecasted annual growth in TFP	1.5%
Expected real growth rate of dividends to perpetuity	5.5%
Required real discount rate to perpetuity	7.5%

*CU = currency unit of Emerge Country

A Based on the information in Exhibit 1, calculate the equity index price level of Emerge Country implied by the Gordon growth model, as of 31 December 2014.

Villeneuve is familiar with the Gordon growth model but not the H-model.

B Identify *two* variables that are needed in the H-model but not needed in the Gordon growth model.

As an illustration of a relative value approach that can be used to support tactical asset allocation, DeLaroche estimated that the forward operating earnings yield of the equity market index in Emerge Country is 6 percent and that the medium-term government bond yield is 7 percent. She then applied the Fed model to the situation in Emerge Country.

C Based on the Fed model, determine whether the equity market is undervalued or overvalued and identify three criticisms of the Fed model.

2 Don Murray, an economist, is president of the investment committee of a large US pension plan. He is reviewing the plan's recent investment returns and finds that non-US equity returns were much higher than US equity returns. Before making any changes to the plan's asset allocation, he asks to meet with Susan McLean, CFA, who is responsible for the equity portion of the pension plan assets. Murray wants to discuss with McLean the current valuation levels of various equity markets.

© 2011 CFA Institute. All rights reserved.

Murray develops his own growth projections for the US and for a hypothetical country (Hyp Country) that has a well-developed economy but whose population is aging. These projections are shown in Exhibit 2. In addition, Murray projects that output elasticity of capital equals 0.3 and 0.5 for the United States and Hyp Country, respectively.

Exhibit 2 Growth Projections (2015–2034)

Country	Growth in Total Factor Productivity (%)	Growth in Capital Stock (%)	Growth in Labor Input (%)
United States	0.6	3.5	0.4
Hyp Country	1.0	3.3	0.1

A Based on the information in Exhibit 2, calculate the projected GDP growth for the United States for the period 2015–2034. Use the Cobb-Douglas production function and assume constant returns to scale.

Murray identifies two possible actions that the government of Hyp Country could take, and he wants to know how these actions would affect projected GDP growth for Hyp Country.

Action 1: Lower the retirement age from 65 to 63, gradually over the next four-year period.

Action 2: Reduce subsidies to higher education over the next five years.

B For each of the government actions identified by Murray in Exhibit 2, indicate which growth factor is *most* affected. Justify your answers.

Murray is surprised that the bottom-up forecasts produced by McLean for the United States in the last five years were consistently more optimistic than her top-down forecasts. As a result, he expresses doubt about the validity of either approach.

C State *one* justification for using both top-down and bottom-up models even when these models produce different forecasts, and state *one* justification for using the bottom-up approach by itself.

Murray suggests replacing earnings-based models with asset-based models in valuing equity markets. In response, McLean recommends using Tobin's q ratio and equity q ratio, although both are subject to estimation errors when applied to valuing a particular company.

D Identify *two* problems that McLean may have in estimating the Tobin's q ratio and the equity q ratio for the pension plan assets that she manages.

The following information relates to Questions 3–10

Claudia Atkinson, CFA, is chief economist of an investment management firm. In analyzing equity markets, the firm has always used a bottom-up approach, but now Atkinson is in the process of implementing a top-down approach. She is discussing this topic with her assistant, Nicholas Ryan.

At Atkinson's request, Ryan has prepared a memo comparing the top-down approach and the bottom-up approach. Ryan presents three conclusions:

Conclusion 1 The top-down approach is less optimistic when the economy is heading into a recession than the bottom-up approach.

Conclusion 2 The top-down approach is more often based on consensus earnings estimates from equity analysts than the bottom-up approach.

Conclusion 3 The top-down approach is often more accurate in predicting the effect on the stock market of a contemporaneous change in a key economic variable than is the bottom-up approach.

Atkinson explains to Ryan how the Cobb-Douglas function can be used to model GDP growth under assumptions of constant returns to scale. For illustrative purposes, she uses the data shown in Exhibit 1.

Exhibit 1 Hypothetical Data for a Developing Country

Time Period	Growth in Total Factor Productivity (%)	Output Elasticity of Capital	Growth in Capital Stock (%)	Growth in Labor Input (%)
1970–1989	2.5	0.4	4.8	3.0
1990–2014	2.8	0.4	4.4	4.6

Atkinson wants to use the data shown in Exhibit 1 as an input for estimating justified P/E ratios. Ryan expresses some criticisms about the use of historical data:

- "In a context of hyperinflation, the approach may not be appropriate."
- "Corporate growth rates may not match GDP growth for long periods."
- "Government-implemented measures may not be taken into account in any of the growth factors."

Atkinson intends to use relative value models in order to support the firm's asset allocation recommendation. The earnings-based approach that she studies is the Fed model. She asks Ryan to write a summary of the advantages of that model. Ryan's report makes the following assertions about the Fed model:

- "The model can be used for non-US equity markets."
- "The model captures the net present value of growth investment opportunities available to investors."
- "The model is most informative when the excess of the earnings yield over the Treasury bond yield is close to the historical average."

Atkinson thinks that the Yardeni model might address some of the criticisms of the Fed model and bring certain improvements. She will use that model as an alternate approach.

Because different results from various equity market valuation models may provide relevant information, Atkinson will present a third earnings-based approach, namely the CAPE model. Ryan identifies many positive features in that model, including the following:

- "The model controls for inflation."
- "The model is independent of changes in accounting rules."
- "The model controls for business cycle effects on earnings."

When evaluating the equity market in developed markets, Atkinson uses the following asset-based models: Tobin's *q* ratio and equity *q* ratio. She calculates the equity *q* ratio of the corporate sector for a particular developed market based on the data shown in Exhibit 2:

Exhibit 2 Corporate Sector for Fourth Quarter of 2014 (Billions of US Dollars)

Assets at market value or replacement cost	27.3
Assets at book value	23.4
Liabilities	13.3
Equities at market value	9.0

Atkinson notes that the Tobin's *q* ratio derived from Exhibit 2 is less than 1. She asks Ryan what conclusion could be drawn from such a low ratio if it had been obtained for a specific company.

3 Which conclusion presented by Ryan about the top-down approach and the bottom-up approach is *most likely* correct?

A Conclusion 1.

B Conclusion 2.

C Conclusion 3.

4 Based on Exhibit 1, which of the components of economic growth has contributed *most* to GDP growth during the 1970–1989 time period?

A Labor input.

B Capital stock.

C Total factor productivity.

5 Which of the following criticisms expressed by Ryan about the use of historical data is the *least* valid?

A In a context of hyperinflation, the approach may not be appropriate.

B Corporate growth rates may not match GDP growth for long periods.

C Government-implemented measures may not be taken into account in any of the growth factors.

6 Which of the following advantages listed by Ryan with respect to the earnings-based approach studied by Atkinson is *most likely* correct? The model:

A can be used for non-US equity markets.

B captures the net present value of growth investment opportunities available to investors.

C is most informative when the excess of the earnings yield over the Treasury bond yield is close to the historical average.

7 The *most likely* improvement from using the Yardeni model instead of the Fed model is that the Yardeni model captures:

A a pure equity risk premium.

B a pure default risk premium.

C the effect of long-term earnings growth on equity market values.

8 Which of the following features of the CAPE model as stated by Ryan is *least likely* to be correct? The model:

A controls for inflation.

B is independent of changes in accounting rules.

C controls for business cycle effects on earnings.

9 Based on the data shown in Exhibit 2, the equity *q* ratio is closest to:

A 0.6429.

B 0.8168.

C 0.8911.

10 The *best* conclusion that Ryan can provide to Atkinson regarding the calculated value for Tobin's *q* ratio is that, based on comparing it to an equilibrium value of 1:

A the replacement cost of the company's assets is understated.

B the company appears to be overvalued in the marketplace.

C the company appears to be undervalued in the marketplace.

The following information relates to Questions 11–16

Egon Carmichael, CFA, is a senior analyst at Supranational Investment Management (Supranational), a firm specializing in global investment analysis. He is meeting with Nicolas Schmidt, a potential client representing a life insurance company, to discuss a report prepared by Supranational on the US equity market. The report contains valuations of the US equity market based on two approaches: the justified P/E model and the Fed model.

When Carmichael informs Schmidt that Supranational applies the neoclassical approach to growth accounting, Schmidt makes the following statements about what he considers to be some limitations of that approach:

Statement 1	The growth in total factor productivity is not directly observable.
Statement 2	The growth factors must be stated in nominal (i.e., not inflation-adjusted) terms.
Statement 3	The total output may not grow at a rate faster than predicted by the growth in capital stock and in labor force.

For use in estimating the justified P/E based on the Gordon constant growth model, Carmichael develops the assumptions displayed in Exhibit 1.

Exhibit 1 Justified P/E Ratio for the US Equity Market: Assumptions	
Required real rate of return	5.0%
Inflation-adjusted dividend growth rate	2.5%

Using the assumptions in Exhibit 1, Carmichael's estimate of the justified P/E ratio for the US equity market is 13.2. Schmidt asks Carmichael, "All else equal, what would cause the justified P/E for the US equity market to fall?"

Supranational's report concludes that the US equity market is currently undervalued, based on the Fed model. Schmidt asks Carmichael, "Which of the following scenarios would result in the Fed model most likely indicating that the US equity market is overvalued?"

Scenario 1	The S&P 500 forward earnings yield is 4.5 percent, and the 10-year T-note yield is 4.75 percent.
Scenario 2	The S&P 500 forward earnings yield is 4.5 percent, the 10-year T-note yield is 4.0 percent, and the average difference between the S&P 500 forward earnings yield and the 10-year T-note over the last 20 years has been 0.25 percent.
Scenario 3	The long-term inflation rate is expected to be 2 percent, and the long-term average earnings growth is expected to be 1 percent real.

Schmidt points out that the Fed model has been the subject of criticism and recommends that Carmichael use the Yardeni model to value the US equity market. Before employing the Yardeni model, Carmichael asks Schmidt to identify criticisms of the Fed model that are addressed by the Yardeni model.

Finally, Carmichael presents a third earnings-based approach, the CAPE model, and describes many positive features of that model.

Schmidt mentions that the international life insurance company that he represents might be interested in the equity forecasts produced by Supranational. He says that his company's objective is to accumulate sufficient assets to fulfill the firm's obligations under its long-term insurance and annuity contracts. For competitive reasons, the company wants to quickly detect significant cyclical turns in equity markets and to minimize tracking errors with respect to the equity index. Schmidt asks Carmichael to identify the forecasting approach that is most appropriate.

11 Which of the statements expressed by Schmidt about the neoclassical approach to growth accounting is correct?

A Statement 1.

B Statement 2.

C Statement 3.

12 Carmichael's *most appropriate* response to Schmidt's question about the justified P/E ratio is:

A lower volatility of the US equity market.

B higher inflation-adjusted dividend growth rate.

C higher correlation of US equity market with international equity markets.

13 Carmichael's *most appropriate* response to Schmidt's question about the Fed model is:

A Scenario 1.

B Scenario 2.

C Scenario 3.

14 In response to Carmichael's question about which criticisms of the Fed model are addressed by the Yardeni model, Schmidt's *most appropriate* response is that the Yardeni model does take account of the criticism that the Fed model:

A assumes that investors value earnings rather than dividends.

B ignores long-term earnings growth opportunities available to shareholders.

C assumes that the required rate of return on equity equals the Treasury bill rate.

15 Which of the following features is *least* applicable to the third earnings-based approach presented by Carmichael? The model:

A controls for inflation.

B is independent of changes in accounting rules.

C controls for business cycle effects on earnings.

16 Carmichael's *best* answer to Schmidt's question about a recommended forecasting approach is to use:

A a top-down approach.

B a bottom-up approach.

C both top-down and bottom-up approaches.

SOLUTIONS

1 A Using the Gordon growth model we have:

$$V_0 = \frac{D_0(1+g)}{r-g}$$

Here:

$D_0 = 450$
$g = 5.5\%$
$r = 7.5\%$

so that:

$$V_0 = \frac{450(1+0.055)}{0.075-0.055} = 23{,}738$$

B One variable present in the H-model but not the Gordon growth model is the initial growth rate of the dividend, and a second is the number of years during which the dividend growth rate declines from its initial value to the long-term sustainable growth rate. In the Gordon growth model, a single dividend growth rate is assumed.

C The Fed model predicts that stocks are overvalued if the forward earnings yield on the equity index (here 6 percent) is less than the yield on Treasury bonds (here 7 percent). Therefore, in Emerge Country, stocks are overvalued.

The Fed model has three important limitations. It:

- ignores the equity risk premium;
- ignores earnings growth; and
- compares a real variable to a nominal variable.

2 A Using Equation 3, we have:

Percentage growth in GDP = Growth in total factor productivity

+ (Output of elasticity of capital) × (Growth in capital stock)

+ (1 − Output of elasticity of capital) × (Growth in labor input)

= 0.6% + (0.3 × 3.5%) + (0.7 × 0.4%)

= 1.93%

B The growth factors most affected for each government action are as follows:

Action 1: Lowering the retirement age will reduce the growth in labor participation and therefore the growth in labor input until a new steady-state labor force participation rate is attained. Subsequent growth in labor input should then track underlying population growth.

Action 2: Lowering the subsidies to higher education will most likely reduce future technical innovation and therefore reduce growth in total factor productivity. The effect may be slow at the beginning, but will increase gradually for a period that will extend well beyond the five years of reduction in subsidies.

C Top-down and bottom-up forecasts frequently differ from each other. In these cases, the reconciling and revision process of the forecasts can:

- help the analyst better understand the market consensus, and
- reveal a gap that gives rise to significant market opportunities.

In spite of the optimistic bias observed by Murray, the bottom-up approach a) may provide the opportunity to identify attractively priced securities irrespective of the attractiveness of the sectors and b) may be a better fit for the investors who focus on a market niche.

D The denominators of the Tobin's q ratio and of the equity q ratio include the replacement cost of company assets. It is difficult to obtain those replacement costs for two reasons: a) there may be no liquid markets for the assets, and b) intangible assets are often difficult to value.

3 A is correct. It has been shown that bottom-up forecasts are often more optimistic than top-down forecasts (not the other way around). This may be because analysts rely on management's assessment of future profitability.

4 C is correct. The contributions from total factor productivity are 2.5 percent and 2.8 percent, respectively, for the periods 1970–1989 and 1990–2014. The corresponding contributions from labor input are 1.8% (= 0.6 × 3.0%) and 2.76% (= 0.6 × 4.6%), and the corresponding contributions from capital stock are 1.92% (= 0.4 × 4.8%) and 1.76% (= 0.4 × 4.4%).

5 C is correct. Government-implemented measures are among the inputs that have an impact on the economy and therefore on the historical data that the analyst uses. It rests upon the analyst to establish whether these measures will continue in the future when he projects economic growth.

6 A is correct. The Fed model, although developed in the United States, can be applied to the valuation of non-US equity markets.

7 C is correct. The Yardeni model incorporates the effect on equity market value of long-term earnings growth.

8 B is correct. The CAPE model is dependent of changes in accounting rules because it averages earnings over 10 years. Therefore, the feature is not applicable to the model.

9 A is correct. The equity q ratio is equal to

Market value of equities/(Replacement cost of assets – Liabilities)

= 9.0/(27.3 – 13.3)

= 0.6429

10 C is correct. A Tobin's q value of less than 1, when 1 is used as a comparison point, indicates that the company is undervalued in the marketplace because it indicates an opportunity to buy assets at a price below their replacement cost.

11 A is correct. It is true that the growth in total productivity is not directly observable. That growth is obtained by using the following data: growth in real output, growth in capital stock, and growth in labor input.

12 C is correct. A higher correlation of the US equity market with international equity markets would increase the risk of the US equity market and thus increase the required return (r) on that market. Increasing r would reduce the justified P/E ratio:

$$\frac{P_0}{E_1} = \frac{D_0(1+g)/E_1}{r-g}$$

13 A is correct. It would be true that the Fed model predicts that US stocks are overvalued if the forward earnings yield on the S&P 500 (4.5 percent in this scenario 1) is less than the yield on US Treasury bonds (4.75 percent in this scenario 1).

14 B is correct. A criticism of the Fed model that the Yardeni model does address is that the Fed model does not take account of long-term earnings growth. The Yardeni model includes a long-term earnings growth variable.

15 B is correct. The CAPE model is dependent of changes in accounting rules because it averages earnings over 10 years. Therefore, the feature is not applicable to the model.

16 C is correct. On the one hand, because the insurance company wants to minimize tracking errors with respect to the equity indexes, Carmichael should recommend the top-down approach because the forecast does not need to focus on individual security selection. On the other hand, because the insurance company wants to detect quickly any significant turn in equity markets, Carmichael should recommend the bottom-up approach because the bottom-up approach can be effective in anticipating cyclical turning points.

PORTFOLIO MANAGEMENT
STUDY SESSION

9

Asset Allocation and Related Decisions in Portfolio Management (1)

Often considered the most important activity in the investment process, the strategic asset allocation decision takes place after the formation of capital market expectations. The portfolio's long-term asset class, or factor, exposures and the best means to achieve these exposures are determined only after considering the investor's unique financial situation and objectives, risk–return tradeoffs, and other key inputs.

This study session provides a conceptual framework for understanding asset allocation considerations and key implementation approaches. Consideration of an investor's overall financial context using an economic balance sheet to incorporate all relevant investor assets and liabilities is presented. Three major approaches to asset allocation are described: asset only, liability relative, and goals based. Concepts underlying active and passive implementation and strategic rebalancing are also introduced. The session then circles back to explain and illustrate the three approaches in greater depth.

READING ASSIGNMENTS

Reading 18	Introduction to Asset Allocation by William W. Jennings, PhD, CFA, and Eugene L. Podkaminer, CFA
Reading 19	Principles of Asset Allocation by Jean L.P. Brunel, CFA, Thomas M. Idzorek, CFA, and John M. Mulvey, PhD

© 2018 CFA Institute. All rights reserved.

READING

18

Introduction to Asset Allocation

by William W. Jennings, PhD, CFA, and Eugene L. Podkaminer, CFA

William W. Jennings, PhD, CFA, is at the US Air Force Academy (USA). Eugene L. Podkaminer, CFA, is at Callan (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** prepare an economic balance sheet for a client and interpret its implications for asset allocation;
☐	**b.** compare the investment objectives of asset-only, liability-relative, and goals-based asset allocation approaches;
☐	**c.** contrast concepts of risk relevant to asset-only, liability-relative, and goals-based asset allocation approaches;
☐	**d.** explain how asset classes are used to represent exposures to systematic risk and discuss criteria for asset class specification;
☐	**e.** explain the use of risk factors in asset allocation and their relation to traditional asset class–based approaches;
☐	**f.** select and justify an asset allocation based on an investor's objectives and constraints;
☐	**g.** describe the use of the global market portfolio as a baseline portfolio in asset allocation;
☐	**h.** discuss strategic implementation choices in asset allocation, including passive/active choices and vehicles for implementing passive and active mandates;
☐	**i.** discuss strategic considerations in rebalancing asset allocations.

INTRODUCTION

1

Asset owners are concerned with accumulating and maintaining the wealth needed to meet their needs and aspirations. In that endeavor, investment portfolios—including individuals' portfolios and institutional funds—play important roles. Asset allocation

© 2018 CFA Institute. All rights reserved.

is a strategic—and often a first or early—decision in portfolio construction. Because it holds that position, it is widely accepted as important and meriting careful attention. Among the questions addressed in this reading are the following:

- How broad a picture should an adviser have of an asset owner's assets and liabilities in recommending an asset allocation?
- How can an asset owner's objectives and sensitivities to risk be represented in asset allocation?
- What are the broad approaches available in developing an asset allocation recommendation, and when might one approach be more or less appropriate than another?
- What are the top-level decisions that need to be made in implementing a chosen asset allocation?
- How may asset allocations be rebalanced as asset prices change?

The strategic asset allocation decision determines return levels[1] in which allocations are invested, irrespective of the degree of active management. Because of its strategic importance, the investment committee, at the highest level of the governance hierarchy, typically retains approval of the strategic asset allocation decision. Often a proposal is developed only after a formal asset allocation study that incorporates obligations, objectives, and constraints; simulates possible investment outcomes over an agreed-on investment horizon; and evaluates the risk and return characteristics of the possible allocation strategies.

In providing an overview of asset allocation, this reading's focus is the alignment of asset allocation with the asset owner's investment objectives, constraints, and overall financial condition. This is the first reading in several sequences of readings that address, respectively, asset allocation and portfolio management of equities, fixed income, and alternative investments. Asset allocation is also linked to other facets of portfolio management, including risk management and behavioral finance. As coverage of asset allocation progresses in the sequence of readings, various connections to these topics, covered in detail in other areas of the curriculum, will be made.[2]

In the asset allocation sequence, the role of this reading is the "big picture." It also offers definitions that will provide a coordinated treatment of many later topics in portfolio management. The second reading—Asset Allocation in Practice—provides the basic "how" of developing an asset allocation, and the third reading explores various common, real-world complexities in developing an asset allocation.

This reading is organized as follows: Section 2 explains the modern interest in considering asset allocation from a comprehensive perspective of the asset owner's entire financial condition. Section 3 distinguishes three broad approaches to asset allocation and explains how they differ in investment objective and concepts of risk. In Section 4, these three approaches are discussed at a high level in relation to three cases. Section 5 provides a top-level orientation to how a chosen asset allocation may be implemented, providing a set of definitions that underlie subsequent readings. As a result of movements in market prices, allocations to assets typically drift away from target allocations. Section 6 discusses rebalancing, and Section 7 provides a summary of the reading.

---

1 See Ibbotson and Kaplan (2000, p. 30) and Xiong, Ibbotson, and Chen (2010). The conclusion for the aggregate follows from the premise that active management is a zero-sum game overall (Sharpe 1991).

2 Among these readings, see Blanchett, Cordell, Finke, and Idzorek (2016) concerning human capital and longevity and other risks and Pompian (2011a and 2011b) and Pompian, McLean, and Byrne (2011) concerning behavioral finance.

2 THE ECONOMIC BALANCE SHEET AND ASSET ALLOCATION

An accounting balance sheet reflects a point-in-time snapshot of an organization's financial condition and shows the assets, liabilities, and owners' equity recognized by accountants. An **economic balance sheet** includes conventional assets and liabilities (called "financial assets" and "financial liabilities" in this reading) as well as additional assets and liabilities—known as **extended portfolio assets and liabilities**—that are relevant in making asset allocation decisions but do not appear on conventional balance sheets.

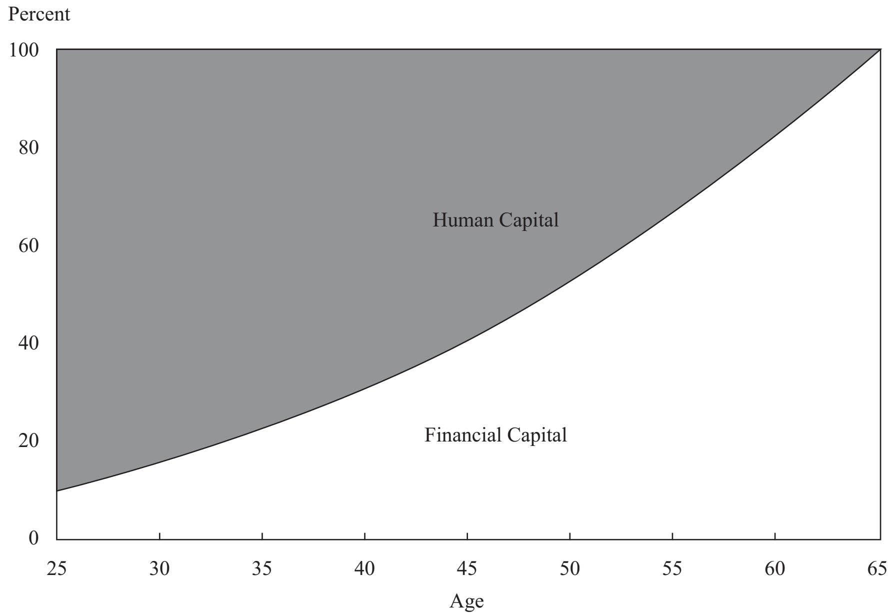
For individual investors, extended portfolio assets include human capital (the present value of future earnings), the present value of pension income, and the present value of expected inheritances. Likewise, the present value of future consumption is an extended portfolio liability.

For an institutional investor, extended portfolio assets might include underground mineral resources or the present value of future intellectual property royalties. Extended portfolio liabilities might include the present value of prospective payouts for foundations, whereas grants payable would appear as conventional liabilities.

Theory and, increasingly, practice suggest that asset allocation should consider the full range of assets and liabilities—both the financial portfolio and extended portfolio assets and liabilities—to arrive at an appropriate asset allocation choice. For example, an asset allocation process that considers the extended balance sheet, including the sensitivity of an individual investor's earnings to equity market risk (and that of the industry in which the individual is working), may result in a more appropriate allocation to equities than one that does not.

Life-cycle balanced funds (also known as target date funds) are examples of investments that seek to coordinate asset allocation with human capital. A 2040 life-cycle balanced fund that seeks to provide a retirement investment vehicle appropriate for many individuals retiring in 2040. Exhibit 1 illustrates a typical path for the composition of an individual's economic balance sheet from age 25 through age 65.

Exhibit 1 Human Capital (HC) and Financial Capital (FC) relative to Total Wealth



At age 25, with most of the individual's working life ahead of him, human capital dominates the economic balance sheet. As the individual progresses through life, the present value of human capital declines as human capital is transformed into earnings. Earnings saved and invested build financial capital balances. By a retirement age of 65, the conversion of human capital to earnings and financial capital is assumed to be complete.

Life-cycle balanced funds reflect these extended portfolio assets. Research indicates that, on average, human capital is roughly 30% equity-like and 70% bond-like, with significant variation among industries.[3] Making the simplifying assumption that investors have approximately constant risk tolerance through life, their asset allocation for total overall wealth (including human capital and financial capital) should be, in theory, constant over time. In this case, the asset allocation chosen for financial capital should reflect an increasing allocation to bonds as human capital declines to age 65, holding all else constant. Exhibit 2 shows the glide path for the equity/bond allocation chosen by one US mutual fund family. The increasing allocation to bonds is consistent with the view that human capital has preponderant bond-like characteristics.

Exhibit 2 Glide Path of Target Date Investment Funds in One Family

Assumed Age	Equity Allocation	Bond Allocation
25	85%	15%
35	82	18
45	77	23

3 See Blanchett and Straehl (2015) and Blanchett and Straehl (2017).

Exhibit 2 (Continued)

Assumed Age	Equity Allocation	Bond Allocation
55	63	37
65	49	51%

Note: Allocations as of 31 December 2009.
Source: Based on data in Idzorek, Stempien, and Voris (2013).

Although estimating human capital is quite complex, including human capital and other extended portfolio assets and economic liabilities in asset allocation decisions is good practice.[4]

EXAMPLE 1

The Economic Balance Sheet of Auldberg University Endowment

- *Name*: Auldberg University Endowment (AUE)
- *Narrative*: AUE was established in 1852 in Caflandia and largely serves the tiny province of Auldberg. AUE supports about one-sixth of Auldberg University's CAF$60 million operating budget; real estate income and provincial subsidies provide the remainder and have been relatively stable. The endowment has historically had a portfolio limited to domestic equities, bonds, and real estate holdings; that policy is under current review. Auldberg University itself (not the endowment) has a CAF$350 million investment in domestic commercial real estate assets, including office buildings and industrial parks, much of it near the campus. AUE employs a well-qualified staff with substantial diverse experience in equities, fixed income, and real estate.
- *Assets*: Endowment assets include CAF$100 million in domestic equities, CAF$60 million in domestic government debt, and CAF$40 million in Class B office real estate. The present value of expected future contributions (from real estate and provincial subsidies) is estimated to be CAF$400 million.
- *Liabilities*: These include CAF$10 million in short-term borrowings and CAF$35 million in mortgage debt related to real estate investments. Although it has no specific legal requirement, AUE has a policy to distribute to the university 5% of 36-month moving average net assets.

4 Human capital is non-tradable, cannot be hedged, is subject to unspecified future taxes, and is a function of an individual's mortality. Human capital is technically defined as the net present value of an investor's future expected labor income weighted by the probability of surviving to each future age (see Ibbotson, Milevsky, Chen, and Zhu 2007). Thus, the present value of future earnings and pensions should be valued with mortality-weighted probabilities of receiving future cash flows, not the present value over life expectancy. There is meaningful extra value from the low-odds event of extreme longevity, which has an important portfolio implication in that individual investors can outlive their financial portfolios but not lifetime annuity payments.

In effect, the endowment supports $10 million of Auldberg University's annual operating budget. The present value of expected future support is CAF$450 million.

1 Prepare an economic balance sheet for AUE.
2 Describe elements in Auldberg University's investments that might affect AUE's asset allocation choices.

Solution to 1:

The economic balance sheet for the endowment (given in the following table) does not include the real estate owned by Auldberg University. The economic net worth is found as a plug item (600 – 10 – 35 – 450 = 105).

AUE Economic Balance Sheet (in CAF$ millions) 31 December 20x6

Assets		Liabilities and Net Worth	
Financial assets		*Financial liabilities*	
Domestic equities	100	Short-term borrowing	10
Domestic fixed income	60	Mortgage debt	35
Class B office real estate	40		
Extended Assets		*Extended Liabilities*	
Present value of expected future contributions to AUE	400	Present value of expected future support	450
		Net Worth	
		Economic net worth (Economic assets – Economic liabilities)	105
Total	600		600

Solution to 2:

AUE's Class B real estate investments' value and income are likely to be stressed during the same economic circumstances as the university's own real estate investments. In such periods, the university may look to the endowment for increased operating support and AUE may not be well positioned to meet that need. Thus, the AUE's real estate investment is actually less diversifying than it may appear and the allocation to it may need to be re-examined. Similar considerations apply to AUE's holdings in equities in relation to Auldberg University's.

3 APPROACHES TO ASSET ALLOCATION

We can identify three broad approaches to asset allocation: (1) **asset-only**, (2) **liability-relative**, and (3) **goals-based**. These are decision-making frameworks that take account of or emphasize different aspects of the investment problem.

Asset-only approaches to asset allocation focus solely on the asset side of the investor's balance sheet. Liabilities are not explicitly modeled. Mean–variance optimization (MVO) is the most familiar and deeply studied asset-only approach. MVO considers only the expected returns, risks, and correlations of the asset classes in the opportunity set. In contrast, liability-relative and goals-based approaches explicitly account for the liabilities side of the economic balance sheet, dedicating assets to meet, respectively, legal liabilities and quasi-liabilities (other needs that are not strictly liabilities but are treated as such) or goals.

Liability-relative approaches to asset allocation choose an asset allocation in relation to the objective of funding liabilities. The phrase "funding of liabilities" means to provide for the money to pay liabilities when they come due. An example is surplus optimization: mean–variance optimization applied to surplus (defined as the value of the investor's assets minus the present value of the investor's liabilities). In modeling, liabilities might be represented by a short position in a bond or series of bonds matched to the present value and duration of the liabilities. Another approach involves constructing a liability-hedging portfolio focused on funding liabilities and, for any remaining balance of assets, a risky-asset portfolio (so called because it is risky or riskier in relation to liabilities—often also called a "return-seeking portfolio" because it explicitly seeks return above and beyond the liability benchmark). **Liability-driven investing** (LDI) is an investment industry term that generally encompasses asset allocation that is focused on funding an investor's liabilities. Related fixed-income techniques are covered in the fixed-income sequence under liability-based mandates.

All approaches to asset allocation can be said to address goals. In investment practice and literature, however, the term "goals based" has come be widely associated with a particular type of approach to asset allocation and investing.

Goals-based approaches to asset allocation, as discussed here, are used primarily for individuals and families, involve specifying asset allocations for sub-portfolios, each of which is aligned to specified goals ranging from supporting lifestyle needs to aspirational. Each goal is associated with regular, irregular, or bulleted cash flows; a distinct time horizon; and a risk tolerance level expressed as a required probability of achieving the goal.[5] For example, a middle-aged individual might specify a goal of maintaining his current lifestyle and require a high level of confidence that this goal will be attained. That same individual might express a goal of leaving a bequest to his alma mater. This would be a very long-term goal and might have a low required probability. Each goal is assigned to its own sub-portfolio, and an asset allocation strategy specific to that sub-portfolio is derived. The sum of all sub-portfolio asset allocations results in an overall strategic asset allocation for the total portfolio. **Goals-based investing** (GBI) is an investment industry term that encompasses the asset allocation focused on addressing an investor's goals.

Institutions and Goals-Based Asset Allocation

Asset segmentation as practiced by some life insurers has some similarities to goals-based investing. Asset segmentation involves notionally or actually segmenting general account assets into sub-portfolios associated with specific lines of business or blocks of liabilities. On one hand, such an approach may be distinguished from goals-based asset allocation for individual investors in being motivated by competitive concerns (to facilitate offering competitive crediting rates on groups of contracts) rather than behavioral ones. On the other hand, Fraser and Jennings (2006) described a behaviorally motivated goals-based approach to asset allocation for foundations and endowments.

5 See Shefrin and Statman (2000) and Brunel (2015).

Following their approach, components of an overall appropriate mean–variance optimal portfolio are allocated to time-based sub-portfolios such that uncomfortably novel or risky positions for the entity's governing body are made acceptable by being placed in longer-term sub-portfolios.

Although any asset allocation approach that considers the liabilities side of the economic balance sheet might be termed "liability relative," there are several important distinctions between liabilities for an institutional investor and goals for an individual investor. These distinctions have meaningful implications for asset allocation:[6]

- Liabilities of institutional investors are legal obligations or debts, whereas goals, such as meeting lifestyle or aspirational objectives, are not. Failing to meet them does not trigger similar consequences.
- Whereas institutional liabilities, such as life insurer obligations or pension benefit obligations, are uniform in nature (all of a single type), an individual's goals may be many and varied.
- Liabilities of institutional investors of a given type (e.g., the pension benefits owed to retirees) are often numerous and so, through averaging, may often be forecast with confidence. In contrast, individual goals are not subject to the law of large numbers and averaging. Contrast an estimate of expected death benefits payable for a group of life insurance policies against an individual's uncertainty about the resources needed in retirement: For a 65-year-old individual, the number of remaining years of life is very uncertain, but insurers can estimate the average for a group of 65-year-olds with some precision.

Liability-Relative and Goals-Based Approaches to Investing

Various perspectives exist concerning the relationship between liability-relative and goals-based approaches to investing. Professor Lionel Martellini summarizes one perspective in the following three statements:[7]

1. Goals-based investing is related to a new paradigm that advocates more granular and investor-centric investment solutions.
2. This new investment solutions paradigm translates into goals-based investing (GBI) approaches in individual money management, in which investors' problems can be summarized in terms of their goals, and it translates into liability-driven investing (LDI) approaches in institutional money management, where the investors' liability is treated as a proxy for their goal.
3. GBI and LDI are therefore related, but each of these approaches has its own specific characteristics. For example, GBI implies the capacity to help individual investors identify a hierarchical list of goals, with a distinction between different types of goals (affordable versus non affordable, essential versus aspirational, etc.) for which no exact counterpart exists in institutional money management.

6 See Rudd and Siegel (2013), which recognizes goals-based planning as a distinct approach. This discussion draws on Brunel (2015).
7 Communication of 3 June 2016, used with permission.

3.1 Relevant Objectives

All three of the asset allocation approaches listed here seek to make optimal use of the amount of risk that the asset owner is comfortable bearing to achieve stated investment objectives, although they generally define risk differently. Exhibit 3 summarizes typical objectives.

Exhibit 3 Asset Allocation Approaches: Investment Objective

Asset Allocation Approach	Relation to Economic Balance Sheet	Typical Objective	Typical Uses and Asset Owner Types
Asset only	Does not explicitly model liabilities or goals	Maximize Sharpe ratio for acceptable level of volatility	Liabilities or goals not defined and/or simplicity is important ■ Some foundations, endowments ■ Sovereign wealth funds ■ Individual investors
Liability relative	Models legal and quasi-liabilities	Fund liabilities and invest excess assets for growth	Penalty for not meeting liabilities high ■ Banks ■ Defined benefit pensions ■ Insurers
Goals based	Models goals	Achieve goals with specified required probabilities of success	Individual investors

In a mean–variance asset-only approach, the objective is to maximize expected portfolio return per unit of portfolio volatility over some time horizon, consistent with the investor's tolerance for risk and consistent with any constraints stated in the IPS. A portfolio's Sharpe ratio is a characteristic metric for evaluating portfolios in an asset-only mean–variance approach.

The basic objective of a liability-relative asset allocation approach is to ensure payment of liabilities when they are due.

A goals-based approach is similar to a liability-relative approach in that it also seeks to ensure that there are sufficient assets to meet the desired payouts. In goals-based approaches, however, goals are generally associated with individual sub-portfolios, and an asset allocation is designed for each sub-portfolio that reflects the time horizon and required probability of success such that the sum of the sub-portfolios addresses the totality of goals satisfactorily.

3.2 Relevant Risk Concepts

Asset-only approaches focus on asset class risk and effective combinations of asset classes. The baseline asset-only approach, mean–variance optimization, uses volatility (standard deviation) of portfolio return as a primary measure of risk, which is a function of component asset class volatilities and the correlations of asset class returns. A mean–variance asset allocation can also incorporate other risk sensitivities, including risk relative to benchmarks and downside risk. Risk relative to benchmarks is usually measured by tracking risk (tracking error). Downside risk can be represented in various ways, including semi-variance, peak-to-trough maximum drawdown, and measures that focus on the extreme (tail) segment of the downside, such as value at risk.

Mean–variance results, although often the starting point for understanding portfolio risk, are regularly augmented by Monte Carlo simulation. By providing information about how an asset allocation performs when one or more variables are changed—for example, to values representing conditions of financial market stress—simulation helps complete the picture of risk, including downside and tail risk. Insights from simulation can then be incorporated as refinements to the asset allocation.

Liability-relative approaches focus on the risk of having insufficient assets to pay obligations when due, which is a kind of shortfall risk. Other risk concerns include the volatility of contributions needed to fund liabilities. Risk in a liability-relative context is generally underpinned by the differences between asset and liability characteristics (e.g., their relative size, their interest rate sensitivity, their sensitivity to inflation).

Goals-based approaches are concerned with the risk of failing to achieve goals.[8] The risk limits can be quantified as the maximum acceptable probability of not achieving a goal.[9] The plural in "liabilities" and "goals" underscores that these risks are generally related to multiple future points in time. Overall portfolio risk is thus the weighted sum of the risks associated with each goal.

Generally, a given statistical risk measure may be relevant in any of the three approaches. For example, standard deviation can be used to assess overall portfolio volatility in asset-only approaches, and it may be used to measure surplus volatility (the volatility of the difference between the values of assets and liabilities) or the volatility of the funded ratio (the ratio of the values of assets and liabilities) in liability-relative asset allocation.

3.3 Modeling Asset Class Risk

Asset classes are one of the most widely used investment concepts but are often interpreted in distinct ways. Greer (1997) defines an asset class as "a set of assets that bear some fundamental economic similarities to each other, and that have characteristics that make them distinct from other assets that are not part of that class." He specifies three "super classes" of assets:

- *Capital assets*. An ongoing source of something of value (such as interest or dividends); capital assets can be valued by net present value.
- *Consumable/transformable assets*. Assets, such as commodities, that can be consumed or transformed, as part of the production process, into something else of economic value, but which do not yield an ongoing stream of value.
- *Store of value assets*. Neither income generating nor valuable as a consumable or an economic input; examples include currencies and art, whose economic value is realized through sale or exchange.

EXAMPLE 2

Asset Classes (1)

Classify the following investments based on Greer's (1997) framework, or explain how they *do not* fit in the framework:

1 Precious metals

8 See Das, Markowitz, Scheid, and Statman (2010), who call goals "mental accounts."
9 See Brunel (2015).

2 Petroleum
3 Hedge funds
4 Timberland
5 Inflation-linked fixed-income securities
6 Volatility

Solutions:

1 Precious metals are a store of value asset except in certain industrial applications (e.g., palladium and platinum in the manufacture of catalytic converters).
2 Petroleum is a consumable/transformable asset; it can be consumed to generate power or provide fuel for transport.
3 Hedge funds do not fit into Greer's (1997) super class framework; a hedge fund strategy invests in underlying asset classes.
4 Timberland is a capital asset or consumable/transformable asset. It is a capital asset in the sense that timber can be harvested and replanted cyclically to generate a stream of cash flows; it is a consumable asset in that timber can be used to produce building materials/ packaging or paper.
5 Inflation-linked fixed-income securities is a capital asset because cash flows can be determined based on the characteristics of the security.
6 Volatility does not fit; it is a measurable investment characteristic. Because equity volatility is the underlying for various derivative contracts and an investable risk premium may be associated with it, it is mentioned by some as an asset.

Greer (1997) approaches the classification of asset classes in an abstract or generic sense. The next question is how to specify asset classes to support the purposes of strategic asset allocation.[10] For example, if a manager lumps together very different investments, such as distressed credit and Treasury securities, into an asset class called "fixed income," asset allocation becomes less effective in diversifying and controlling risk. Furthermore, the investor needs a logical framework for distinguishing an asset class from an investment strategy. The following are five criteria that will help in effectively *specifying asset classes for the purpose of asset allocation*:[11]

1 *Assets within an asset class should be relatively homogeneous.* Assets within an asset class should have similar attributes. In the example just given, defining equities to include both real estate and common stock would result in a non-homogeneous asset class.
2 *Asset classes should be mutually exclusive.* Overlapping asset classes will reduce the effectiveness of strategic asset allocation in controlling risk and could introduce problems in developing asset class return expectations. For example, if one asset class for a US investor is domestic common equities, then world equities ex-US is more appropriate as another asset class rather than global equities, which include US equities.

10 See Kritzman (1999).
11 As opposed to criteria for asset class definition in an absolute sense.

3 *Asset classes should be diversifying.* For risk control purposes, an included asset class should not have extremely high expected correlations with other asset classes or with a linear combination of other asset classes. Otherwise, the included asset class will be effectively redundant in a portfolio because it will duplicate risk exposures already present. In general, a pairwise correlation above 0.95 is undesirable (given a sufficient number of observations to have confidence in the correlation estimate).

4 *The asset classes as a group should make up a preponderance of world investable wealth.* From the perspective of portfolio theory, selecting an asset allocation from a group of asset classes satisfying this criterion should tend to increase expected return for a given level of risk. Furthermore, the inclusion of more markets expands the opportunities for applying active investment strategies, assuming the decision to invest actively has been made. However, such factors as regulatory restrictions on investments and government-imposed limitations on investment by foreigners may limit the asset classes an investor can invest in.

5 *Asset classes selected for investment should have the capacity to absorb a meaningful proportion of an investor's portfolio.* Liquidity and transaction costs are both significant considerations. If liquidity and expected transaction costs for an investment of a size meaningful for an investor are unfavorable, an asset class may not be practically suitable for investment.

Note that Criteria 1 through 3 strictly focus on assets themselves, while Criterion 5, and to some extent Criterion 4, involve potential investor-specific considerations.

Asset Classes Should Be Diversifying

Pairwise asset class correlations are often useful information and are readily obtained. However, in evaluating an investment's value as a diversifier at the portfolio level, it is important to consider an asset in relation to all other assets as a group rather than in a one-by-one (pairwise) fashion. It is possible to reach limited or incorrect conclusions by solely considering pairwise correlations. To give an example, denote the returns to three assets by X, Y, and Z, respectively. Suppose that $Z = aX + bY$; a and b are constants, not both equal to zero. Asset Z is an exact weighted combination of X and Y and so has no value as a diversifier added to a portfolio consisting of assets X and Y. Yet, if the correlation between X and Y is -0.5, it can be shown that Z has a correlation of just 0.5 with X as well as with Y.

Examining return series' correlations during times of financial market stress can provide practically valuable insight into potential diversification benefits beyond typical correlations that average all market conditions.

In current professional practice, the listing of asset classes often includes the following:

- *Global public equity*—composed of developed, emerging, and sometimes frontier markets and large-, mid-, and small-cap asset classes; sometimes treated as several sub-asset classes (e.g., domestic and non-domestic).
- *Global private equity*—includes venture capital, growth capital, and leveraged buyouts (investment in special situations and distressed securities often occurs within private equity structures too).

- *Global fixed income*—composed of developed and emerging market debt and further divided into sovereign, investment-grade, and high-yield sub-asset classes, and sometimes inflation-linked bonds (unless included in real assets; see the following bullet). Cash and short-duration securities can be included here.
- *Real assets*—includes assets that provide sensitivity to inflation, such as private real estate equity, private infrastructure, and commodities. Sometimes, global inflation-linked bonds are included as a real asset rather than fixed income because of their sensitivity to inflation.

Emerging Market Equities and Fixed Income

Investment practice distinguishes between developed and emerging market equities and fixed income within global equities. The distinction is based on practical differences in investment characteristics, which can be related to typical market differences including the following:

- diversification potential, which is related to the degree to which investment factors driving market returns in developed and emerging markets are not identical (a topic known as "market integration");
- perceived level of informational efficiency; and
- corporate governance, regulation, taxation, and currency convertibility.

As of mid-2016, emerging markets represent approximately 10% of world equity value based on MSCI indices.[12] In fixed income, investment opportunities have expanded as governments and corporations domiciled in emerging markets have increasingly issued debt in their own currency. Markets in local currency inflation-indexed emerging market sovereign debt have become more common.[13]

"Asset classes" are, by definition, groupings of assets. Investment vehicles, such as hedge funds, that apply strategies to asset classes and/or individual investments with the objective of earning a return to investment skill or providing attractive risk characteristics may be treated as a category called "strategies" or "diversifying strategies." When that is the case, this category is assigned a percentage allocation of assets, similar to a true asset class. Economically, asset classes contrast with "strategies" by offering, in general, an inherent, non-skill-based *ex ante* expected return premium.[14]

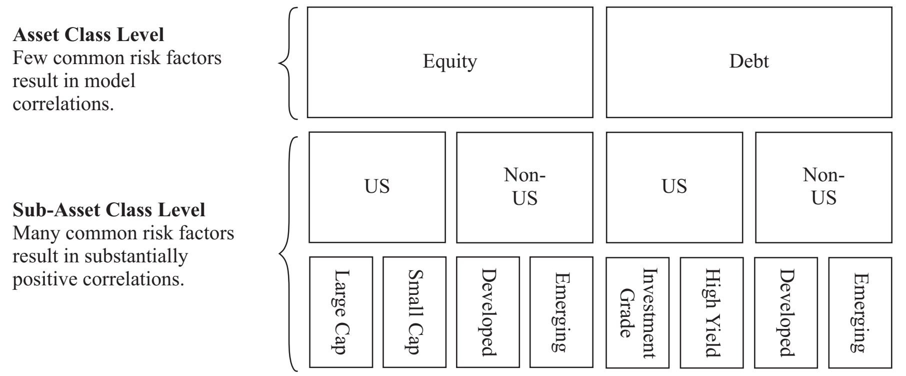
Effective portfolio optimization and construction may be hindered by excessive asset class granularity. Consider Exhibit 4.

12 MSCI uses three broad definitions to sort countries into developed, emerging, and frontier: 1) economic development, 2) size and liquidity requirements, and 3) market accessibility criteria (see the MSCI Market Classification Framework at www.msci.com/market-classification).

13 For a discussion of their potential benefits, see Burger, Warnock, and Warnock (2012), Perry (2011), and Swinkels (2012). Kozhemiakin (2011) discusses how emerging market bonds can facilitate broader representation than an equity-only portfolio because some countries (e.g., Argentina) have small equity markets but larger bond markets.

14 See Idzorek and Kowara (2013), p.20.

Exhibit 4 Examples of Asset Classes and Sub-Asset Classes



As more and more sub-asset classes are defined, they become less distinctive. In particular, the sources of risk for more broadly defined asset classes are generally better distinguished than those for narrowly defined subgroups. For example, the overlap in the sources of risk of US large-cap equity and US small-cap equity would be greater than the overlap between US and non-US equity. Using broadly defined asset classes with fewer risk source overlaps in optimization is consistent with achieving a diversified portfolio. Additionally, historical data for broadly defined asset classes may be more readily available or more reliable. The question of how much to allocate to equity versus fixed income versus other assets is far more important in strategic asset allocation than *precisely* how much to allocate to the various sub-classes of equity and fixed income. However, when the investor moves from the strategic asset allocation phase to policy implementation, sub-asset class choices become relevant.

EXAMPLE 3

Asset Classes (2)

Discuss a specification of asset classes that distinguishes between "domestic intermediate-duration fixed income" and "domestic long-duration fixed income." Contrast potential relevance in asset-only and liability-relative contexts.

Solution:

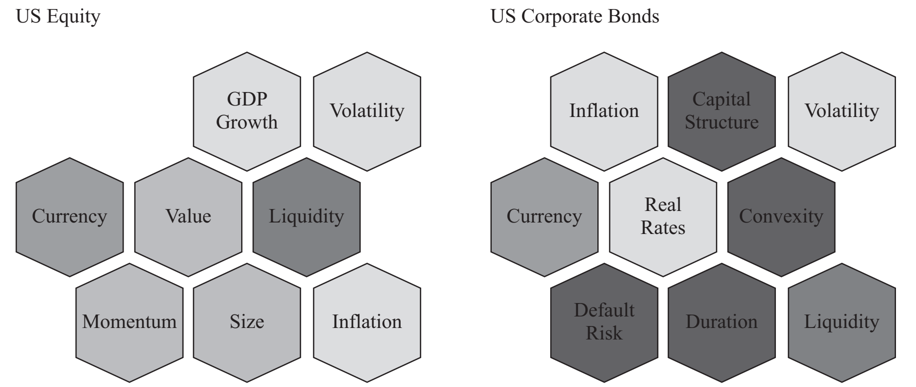
These two groups share key risk factors, such as interest rate and credit risk. For achieving diversification in asset risk—for example, in an asset-only context—asset allocation using domestic fixed income, which includes intermediate and long duration, should be effective and simple. Subsequently, allocation within domestic fixed income could address other considerations, such as interest rate views. When investing in relation to liabilities, distinctions by duration could be of first-order importance and the specification could be relevant.

Any asset allocation, by whatever means arrived at, is expressed ultimately in terms of money allocations to assets. Traditionally—and still in common practice—asset allocation uses asset classes as the unit of analysis. Thus, mean–variance optimization based on four asset classes (e.g., global public equity, global private equity, global fixed income, and real assets) would be based on expected return, return volatility, and return correlation estimates for these asset classes. (The development of such capital market assumptions is the subject of another reading.) Factor-based approaches, discussed in more detail later, do not use asset classes as the basis for portfolio construction.

Technically, the set of achievable investment outcomes cannot be enlarged simply by developing an asset allocation by a different means (for instance, using asset classes as the unit of analysis), all else being equal, such as constraints against short selling (non-negativity constraints).[15] Put another way, adopting a factor-based asset allocation approach does not, by default, lead to superior investment outcomes.

There are allocation methods that focus on assigning investments to the investor's desired exposures to specified risk factors. These methods are premised on the observation that asset classes often exhibit some overlaps in sources of risk, as illustrated in Exhibit 5.[16]

Exhibit 5 Common Factor Exposures across Asset Classes



The overlaps seen in Exhibit 5 help explain the correlation of equity and credit assets. Modeling using asset classes as the unit of analysis tends to obscure the portfolio's sensitivity to overlapping risk factors, such as inflation risk in this example. As a result, controlling risk exposures may be problematic. Multifactor risk models, which have a history of use in individual asset selection, have been brought to bear on the issue of controlling systematic risk exposures in asset allocation.

In broad terms, when using factors as the units of analysis, we begin with specifying risk factors and the desired exposure to each factor. Asset classes can be described with respect to their sensitivities to each of the factors. Factors, however, are not directly investable. On that basis, asset class portfolios that isolate exposure to the risk factor are constructed; these factor portfolios involve both long and short positions. A choice of risk exposures in factor space can be mapped back to asset class space for implementation. Uses of multifactor risk models in asset allocation have been labeled "factor-based asset allocation" in contrast to "asset class-based asset allocation," which uses asset classes directly as the unit of analysis.

Factor Representation

Although risk factors can be thought of as the basic building blocks of investments, most are not directly investable. In this context, risk factors are associated with expected return premiums. Long and short positions in assets (spread positions) may be needed

15 Stated more formally and demonstrated in Idzorek and Kowara (2013).
16 See Podkaminer (2013).

to isolate the respective risks and associated expected return premiums. Other risk factors may be accessed through derivatives. The following are a few examples of how risk factor exposures can be achieved.

- *Inflation*. Going long nominal Treasuries and short inflation-linked bonds isolates the inflation component.
- *Real interest rates*. Inflation-linked bonds provide a proxy for real interest rates.
- *US volatility*. VIX (Chicago Board Options Exchange Volatility Index) futures provide a proxy for implied volatility.
- *Credit spread*. Going long high-quality credit and short Treasuries/government bonds isolates credit exposure.
- *Duration*. Going long 10+ year Treasuries and short 1–3 year Treasuries isolates the duration exposure being targeted.

Factor Models in Asset Allocation

The interest in using factors for asset allocation stems from a number of considerations, including the following:

- The desire to shape the asset allocation based on goals and objectives that cannot be expressed by asset classes (such as matching liability characteristics in a liability-relative approach).
- An intense focus on portfolio risk in all of its various dimensions, helped along by availability of commercial factor-based risk measurement and management tools.
- The acknowledgment that many highly correlated so-called asset classes are better defined as parts of the same high-level asset class. For example, domestic and foreign equity may be better seen as sub-classes of global public equity.
- The realization that equity risk can be the dominant risk exposure even in a seemingly well-diversified portfolio.

4 STRATEGIC ASSET ALLOCATION

An asset allocation that arises in long-term investment planning is often called the "strategic asset allocation" or "policy portfolio": It is an asset allocation that is expected to be effective in achieving an asset owner's investment objectives, given his or her investment constraints and risk tolerance, as documented in the investment policy statement.

A theoretical underpinning for quantitative approaches to asset allocation is utility theory, which uses a utility function as a mathematical representation of preferences that incorporates the investor's risk aversion. According to utility theory, the optimal asset allocation is the one that is expected to provide the highest utility to the investor at the investor's investment time horizon. The optimization program, in broad terms, is

$$\underset{\text{by choice of asset class weights } w_i}{\text{Maximize}} \quad E\left[U\left(W_T\right)\right] = f\begin{pmatrix} W_0, w_i, \text{asset class return distributions,} \\ \text{degree of risk aversion} \end{pmatrix}$$

$$\text{subject to} \sum_{i=1}^{n} w_i = 1 \text{ and any other constraints on } w_i$$

The first line is the objective function, and the second line consists of constraints on asset class weights; other constraints besides those on weights can also be incorporated (for example, specified levels of bond duration or portfolio yield may be targeted). With W_0 and W_T (the values of wealth today and at time horizon T, respectively) the investor's problem is to select the asset allocation that maximizes the expected utility of ending wealth, $E[U(W_T)]$, subject to the constraints that asset class weights sum to 1 and that weights observe any limits the investor places on them. Beginning wealth, asset class weights, and asset class returns imply a distribution of values for ending wealth, and the utility function assigns a value to each of them; by weighting these values by their probability of occurrence, an expected utility for the asset allocation is determined.

An expected utility framework underlies many, but not all, quantitative approaches to asset allocation. A widely used group in asset allocation consists of power utility functions,[17] which exhibit the analytically convenient characteristic that risk aversion does not depend on the level of wealth. Power utility can be approximated by mean–variance utility, which underlies mean–variance optimization.

Optimal Choice in the Simplest Case

The simplest asset allocation decision problem involves one risky asset and one risk-free asset. Let λ, μ, r_f, and σ^2 represent, respectively, the investor's degree of risk aversion, the risk asset's expected return, the risk-free interest rate, and the variance of return. With mean–variance utility, the optimal allocation to the risky asset, w^*, can be shown to equal

$$w^* = \frac{1}{\lambda}\left(\frac{\mu - r_f}{\sigma^2}\right)$$

The allocation to the risky asset is inversely proportional to the investor's risk aversion and directly proportional to the risk asset's expected return per unit of risk (represented by return variance).[18]

Selection of a strategic asset allocation generally involves the following steps:[19]

1 Determine and quantify the investor's objectives. What is the pool of assets meant for (e.g., paying future benefit payments, contributing to a university's budget, securing ample assets for retirement)? What is the investor trying to achieve? What liabilities or needs or goals need to be recognized (explicitly or implicitly)? How should objectives be modeled?

2 Determine the investor's risk tolerance and how risk should be expressed and measured. What is the investor's overall tolerance for risk and specific risk sensitivities? How should these be quantified in the process of developing an appropriate asset allocation (risk measures, factor models)?

3 Determine the investment horizon(s). What are the appropriate planning horizons to use for asset allocation; that is, over what horizon(s) should the objectives and risk tolerance be evaluated?

---

17 Power utility has the form $U = \frac{w_T^{1-\lambda}}{1-\lambda}$, where $\lambda > 0$ is the parameter of risk aversion (if $\lambda \to 0$, the investor is risk neutral).

18 See Ang (2014), Chapter 4, for further analysis.

19 Arjan Berkelaar, CFA, contributed to this formulation of steps.

4 Determine other constraints and the requirements they impose on asset allocation choices. What is the tax status of the investor? Should assets be managed with consideration given to ESG issues? Are there any legal and regulatory factors that need to be considered? Are any political sensitivities relevant? Are there any other constraints that the investor has imposed in the IPS and other communications?

5 Determine the approach to asset allocation that is most suitable for the investor.

6 Specify asset classes, and develop a set of capital market expectations for the specified asset classes.

7 Develop a range of potential asset allocation choices for consideration. These choices are often developed through optimization exercises. Specifics depend on the approach taken to asset allocation.

8 Test the robustness of the potential choices. This testing often involves conducting simulations to evaluate potential results in relation to investment objectives and risk tolerance over appropriate planning horizon(s) for the different asset allocations developed in Step 7. The sensitivity of the outcomes to changes in capital market expectations is also tested.

9 Iterate back to Step 7 until an appropriate and agreed-on asset allocation is constructed.

Subsequent readings on asset allocation in practice will address the "how." The following sections give an indication of thematic considerations. We use investors with specific characteristics to illustrate the several approaches distinguished: sovereign wealth fund for asset-only allocation; a frozen corporate DB plan for liability-relative allocation; and an ultra-high-net-worth family for goals-based allocation. In practice, any type of investor could approach asset allocation with varying degrees of focus on modeling and integrating liabilities-side balance sheet considerations. How these cases are analyzed in this reading should not be viewed as specifying normative limits of application for various asset allocation approaches. For example, a liability-relative perspective has wide potential relevance for institutional investors because it has the potential to incorporate all information on the economic balance sheet. Investment advisers to high-net-worth investors may choose to use any of the approaches.

4.1 Asset Only

Asset-only allocation is based on the principle of selecting portfolios that make efficient use of asset risk. The focus here is mean–variance optimization, the mainstay among such approaches. Given a set of asset classes and assumptions concerning their expected returns, volatilities, and correlations, this approach traces out an efficient frontier that consists of portfolios that are expected to offer the greatest return at each level of portfolio return volatility. The Sharpe ratio is a key descriptor of an asset allocation: If a portfolio is efficient, it has the highest Sharpe ratio among portfolios with the same volatility of return.

An example of an investor that might use an asset-only approach is the (hypothetical) Government Petroleum Fund of Caflandia (GPFC) introduced next.

Investor Case Facts: GPFC, A Sovereign Wealth Fund

- *Name*: Government Petroleum Fund of Caflandia (GPFC)

- *Narrative*: The emerging country of Caflandia has established a sovereign wealth fund to capture revenue from its abundant petroleum reserves. The government's goal in setting up the fund is to promote a fair sharing of the benefits between current and future generations (intergenerational equity) from the export of the country's petroleum resources. Caflandia's equity market represents 0.50% of global equity market capitalization. Economists estimate that distributions in the interest of intergenerational equity may need to begin in 20 years. Future distribution policy is undetermined.
- *Tax status*: Non-taxable.
- *Financial assets and financial liabilities*: Financial assets are CAF$40 billion at market value, making GPFC among the largest investors in Caflandia. GPFC has no borrowings.
- *Extended assets and liabilities*: Cash inflows from petroleum exports are assumed to grow at inflation + 1% for the next 15 years and may change depending on reserves and global commodity demand. The present value of expected future income from state-owned reserves is estimated to be CAF$60 billion. Future spending needs are positively correlated with consumer inflation and population growth. In Exhibit 6, the amount for the present value (PV) of future spending, which GPFC has not yet determined, is merely a placeholder to balance assets and liabilities; as a result, no equity is shown.

Exhibit 6 GPFC Economic Balance Sheet (in CAF$ billions) 31 December 20x6

Assets		Liabilities and Net Worth	
Financial assets		*Financial liabilities*	
Investments (includes cash, equities, fixed income, and other investments)	40		
Extended assets		*Extended liabilities*	
PV of expected future income	60	PV of future spending	100
		Net worth	
		Economic net worth	0
Total	100		100

For GPFC, the amount and timing of funds needed for future distributions to Caflandia citizens are, as yet, unclear. GPFC can currently focus on asset risk and its efficient use to grow assets within the limits of the fund's risk tolerance. In addition to considering expected return in relation to volatility in selecting an asset allocation, GPFC might include such considerations as the following:

- diversification across global asset classes (possibly quantified as a constraint on the proportion allocated to any given asset classes);
- correlations with the petroleum sources of income to GPFC;
- the potential positive correlation of future spending with inflation and population growth in Caflandia;

- long investment horizon (as a long-term investor, GPFC may be well positioned to earn any return premium that may be associated with the relatively illiquid asset classes); and
- return outcomes in severe financial market downturns.

Suppose GPFC quantifies its risk tolerance in traditional mean–variance terms as willingness to bear portfolio volatility of up to 17% per year. This risk tolerance is partly based on GPFC's unwillingness to allow the fund to fall below 90% funded. GPFC's current strategic asset allocation, along with several alternatives that have been developed by its staff during an asset allocation review, are shown in Exhibit 7. The category "Diversifying strategies" consists of a diversified allocation to hedge funds.

Exhibit 7 GPFC Strategic Asset Allocation Decision[20]

	Asset Allocation			
		Proposed		
	Current	A	B	C
Investment				
Equities				
Domestic	50%	40%	45%	30%
Global ex-domestic		10%	20%	25%
Bonds				
Nominal	30%	30%	20%	10%
Inflation linked				10%
Real estate	20%	10%	15%	10%
Diversifying strategies		10%		15%
Portfolio statistics				
Expected arithmetic return	8.50%	8.25%	8.88%	8.20%
Volatility (standard deviation)	15.57%	14.24%	16.63%	14.06%
Sharpe ratio	0.353	0.369	0.353	0.370
One-year 5% VaR	−17.11%	−15.18%	−18.48%	−14.93%

Notes: The government bond rate is 3%. The acceptable level of volatility is ≤ 17% per year. The value at risk (VaR) is stated as a percent of the initial portfolio value over one year (e.g., −16% means a decline of 16%).

GPFC decides it is willing to tolerate a 5% chance of losing 22% or more of portfolio value in a given year. This risk is evaluated by examining the one-year 5% VaR of potential asset allocations.

20 The assumed expected returns and return volatilities are (given in that order in parentheses and expressed as decimals, rather than percentages): domestic equities (0.11, 0.25), non-domestic equities (0.09, 0.18), nominal bonds (0.05, 0.10), inflation-linked bonds (0.035, 0.06), real estate (0.075, 0.16), and diversifying strategies (0.07, 0.09). A correlation matrix with hypothetical values and a hypothetical relationship between the allocations and VaR also lies behind the exhibit. Because the purpose here is to illustrate concepts rather than mechanics, inputs are not discussed although they are very important in asset allocation.

Let us examine GPFC's decision. The current asset allocation and the alternatives developed by staff all satisfy the GPFC's tolerance for volatility and VaR limit. The staff's alternatives appear to represent incremental, rather than large-scale, changes from the current strategic asset allocation. We do not know whether capital market assumptions have changed since the current strategic asset allocation was approved.

Mix A, compared with the current asset allocation, diversifies the equity allocation to include non-domestic (global ex-domestic) equities and spreads the current allocation to real estate over real estate and diversifying strategies. Given GPFC's long investment horizon and absence of liquidity needs, an allocation to diversifying strategies at 10% should not present liquidity concerns. Because diversifying strategies are more liquid than private real estate, the overall liquidity profile of the fund improves. It is important to note that given the illiquid nature of real estate, it could take considerable time to reallocate from real estate to diversifying strategies. Mix A has a lower volatility (by 133 bps) than the current allocation and slightly lower tail risk (the 5% VaR for Mix A is −15%, whereas the 5% VaR for the current asset mix is −17%). Mix A's Sharpe ratio is slightly higher. On the basis of the facts given, Mix A appears to be an incremental improvement on the current asset allocation.

Compared with Mix A and the current asset allocation, Mix B increases the allocation to equities by 15 percentage points and pulls back from the allocation to bonds and, in relation to Mix A, diversifying strategies. Although Mix B has a higher expected return and its VaR is within GPFC's tolerance of 22%, Mix B's lower Sharpe ratio indicates that it makes inefficient use of its additional risk. Mix B does not appear to deserve additional consideration.

Compared with the current asset allocation and Mix A, Mix C's total allocation to equities, at 55%, is higher and the mix is more diversified considering the allocation of 25% non-domestic equities. Mix C's allocation to fixed income is 20% compared with 30% for Mix A and the current asset mix. The remaining fixed-income allocation has been diversified with an exposure to both nominal and inflation-linked bonds. The diversifying strategies allocation is funded by a combination of the reduced weights to fixed income and real estate. The following observations may be made:

- Mix C's increase in equity exposure (compared with the equity exposure of Mix A and the current mix) has merit because more equity-like choices in the asset allocation could be expected to give GPFC more exposure to such a factor as a GDP growth factor (see Exhibit 6); population growth is one driver of GDP.
- Within fixed income, Mix C's allocation to inflation-linked bonds could be expected to hedge the inflation risk inherent in future distributions.
- Mix C has the lowest volatility and the lowest VaR among the asset allocations, although the differences compared with Mix A are very small. Mix C's Sharpe ratio is comparable to (insignificantly higher than) Mix A's.

Based on the facts given, Mix A and Mix C appear to be improvements over the current mix. Mix C may have the edge over Mix A based on the discussion. As a further step in the evaluation process, GPFC may examine the robustness of the forecasted results by changing the capital market assumptions and simulating shocks to such variables as inflation. The discussion of Mix C shows that there are means for potential liability concerns (the probable sensitivity of spending to inflation and population growth) to enter decision making even from a mean–variance optimization perspective.

EXAMPLE 4

Asset-Only Asset Allocation

1. Describe how the Sharpe ratio, considered in isolation, would rank the asset allocation in Exhibit 7.
2. State a limitation of basing a decision only on the Sharpe ratio addressed in Question 1.
3. An assertion is heard in an investment committee discussion that because the Sharpe ratio of diversifying strategies (0.55) is higher than real estate's (0.50), any potential allocation to real estate would be better used in diversifying strategies. Describe why the argument is incomplete.

Solution to 1:

The ranking by Sharpe ratios in isolation is C (3.70), A (3.69), and current and B (both 3.53). Using only the Sharpe ratio, Mix C appears superior to the other choices, but such an approach ignores several important considerations.

Solution to 2:

The Sharpe ratio, while providing a means to rank choices on the basis of return per unit of volatility, does not capture other characteristics that are likely to be important to the asset owner, such as VaR and funded ratio. Furthermore, the Sharpe ratio by itself cannot confirm that the absolute level of portfolio risk is within the investor's specified range.

Solution to 3:

It is true that the higher the Sharpe ratio of an investment, the greater its contribution to the Sharpe ratio of the overall portfolio, *holding all other things equal*. However, that condition is not usually true. Diversification potential in a portfolio (quantified by correlations) may differ. For example, including both diversifying strategies and real estate in an allocation may ultimately decrease portfolio-level risk through favorable correlation characteristics. Also, as in the solution to Question 2, other risk considerations besides volatility may be relevant.

Financial theory suggests that investors should consider the global market-value weighted portfolio as a baseline asset allocation. This portfolio, which sums all investable assets (global stocks, bonds, real estate, and so forth) held by investors, reflects the balancing of supply and demand across world markets. In financial theory, it is the portfolio that minimizes non-diversifiable risk, which in principle is uncompensated. Because of that characteristic, theory indicates that the global market portfolio should be the available portfolio that makes the most efficient use of the risk budget.[21] Other arguments for using it as a baseline include its position as a reference point for a highly diversified portfolio and the discipline it provides in relation to mitigating any investment biases, such as home-country bias (discussed below).

At a minimum, the global market portfolio serves as a starting point for discussion and ensures that the investor articulates a clear justification for moving away from global capitalization market weights. The global market portfolio is expressed in two phases. The first phase allocates assets in proportion to the global portfolio of stocks, bonds, and real assets. The second phase disaggregates each of these broad asset

21 According to the two-fund separation theorem, all investors optimally hold a combination of a risk-free asset and an optimal portfolio of all risky assets. This optimal portfolio is the global market value portfolio.

classes into regional, country, and security weights using capitalization weights. The second phase is typically used within a global equity portfolio where an asset owner will examine the global capitalization market weights and either accept them or alter them. Common tilts (biases) include overweighting the home-country market, value, size (small cap), and emerging markets. For many investors, allocations to foreign fixed income have been adopted more slowly than allocations to foreign equity. Most investors have at least some amount in non-home-country equity.

Home-Country Bias

A given for GPFC was that Caflandia's equity markets represent only 0.50% of the value of world equity markets. However, in all asset allocations in Exhibit 7, the share of domestic equity ranged from 50% for the current asset allocation to 30% for Mix C. The favouring of domestic over non-domestic investment relative to global market value weights is called **home-country bias** and is very common. Even relatively small economies feature pension plans, endowments, and other funds, which are disproportionately tilted toward the equity and fixed-income offerings in the domestic market. The same tendency is true for very large markets, such as the United States and the eurozone. By biasing toward the home market, asset owners may not be optimally aligning regional weights with the global market portfolio and are implicitly implementing a market view. Investment explanations for the bias, such as offsetting liabilities that are denominated in the home currency, may be relevant in some cases, however.

For reference, the MSCI All Country World Portfolio (ACWI), a proxy for the public equities portion of the global equity market portfolio, contains the following capitalization weights as of 31 December 2015:

- Developed Europe and the Middle East: 22.8%
- Developed Pacific: 11.7%
- North America: 55.9%
- Emerging markets: 9.6%

Investing in a global market portfolio faces several implementation hurdles. First, estimating the size of each asset class on a global basis is an imprecise exercise given the uneven availability of information on non-publicly traded assets. Second, the practicality of investing proportionately in residential real estate, much of which is held in individual homeowners' hands, has been questioned. Third, private commercial real estate and global private equity assets are not easily carved into pieces of a size that is accessible to most investors. Practically, proxies for the global market portfolio are often based only on traded assets, such as portfolios of exchange-traded funds (ETFs). Furthermore, some investors have implemented alternative weighting schemes, such as GDP weight or equal weight. However, it is a useful discipline to articulate a justification for any deviation from the capitalization-weighted global market portfolio.

4.2 Liability Relative

To illustrate the liability-relative approach, we take the defined benefit (DB) pension plan of (hypothetical) GPLE Corporation, with case facts given below.

A Frozen DB Plan, GPLE Corporation Pension

- *Name*: GPLE Corporation Pension
- *Narrative*: GPLE is a machine tool manufacturer with a market value of $2 billion. GPLE is the sponsor of a $1.25 billion legacy DB plan, which is now frozen (i.e., no new plan participants and no new benefits accruing for existing plan participants). GPLE Pension has a funded ratio (the ratio of pension assets to liabilities) of 1.15. Thus, the plan is slightly overfunded. Responsibility for the plan's management rests with the firm's treasury department (which also has responsibility for GPLE Corporation treasury operations).
- *Tax status*: Non-taxable.
- *Financial assets and financial liabilities*: Assets amount to $1.25 billion at market values. Given a funded ratio of 1.15, that amount implies that liabilities are valued at about $1.087 billion. Projected distributions to pension beneficiaries have a present value of $1.087 billion at market value.

GPLE does not reflect any extended assets or liabilities; thus, economic net worth is identical to traditional accounting net worth.

Exhibit 8 GPLE Pension Economic Balance Sheet (in US$ billions) 31 December 20x6

Assets		Liabilities and Net Worth	
Financial assets		*Financial liabilities*	
Pension assets	1.250	PV of pension liability	1.087
		Net worth	
		Economic net worth	0.163
Total	1.250		1.250

GPLE, the plan sponsor, receives two asset allocation recommendations. Recommendation A does not explicitly consider GPLE's pension's liabilities but is instead based on an asset-only perspective: the mean–variance efficient frontier given a set of capital market assumptions. A second recommendation, "Recommendation B," does explicitly consider liabilities, incorporating a liability-hedging portfolio based on an analysis of GPLE pension liabilities and a return-seeking portfolio.

In evaluating asset allocation choices, consider the pensioners' and the plan sponsor's interests. Pensioners want to receive the stream of promised benefits with as little risk, or chance of interruption, as possible. Risk increases as the funded ratio declines. When the funded ratio is 1.0, pension assets just cover pension liabilities with no safety buffer. When the funded ratio is less than 1.0, the plan sponsor generally needs to make up the deficit in pension assets by contributions to the plan. For example, with a 10-year investment time horizon and a choice between two asset allocations, the allocation with the lower expected present value of cumulative contributions to Year 10 would generally be preferred by the sponsor, all else being equal. In practice, all else is usually not equal. For example, the alternative with the lower *expected* present value of contributions may involve more risk to the level of contributions in adverse market conditions. For example, the 5% of *worst outcomes* for the present value of cumulative contributions may be more severe for the lower expected contribution alternative. Thus, possible asset allocations generally involve risk trade-offs.[22] Now consider the recommendations.

Recommendation A, based on asset-only analysis, involves a 65% allocation to global equities and a 35% allocation to global fixed income. Assume that this asset allocation is mean–variance efficient and has the highest Sharpe ratio among portfolios that meet the pension's assumed tolerance for asset return volatility. Capital market assumptions indicate that equities have a significantly higher expected return and volatility than fixed income.

Recommendation B, based on a liability-relative approach to asset allocation, involves an allocation of \$1.125 billion to a fixed-income portfolio that is very closely matched in interest rate sensitivity to the present value of plan liabilities (and to any other liability factor risk exposures)—the liability hedging portfolio—and a \$0.125 allocation to equities (the return-seeking portfolio). This is a proportional allocation of 10% to equities and 90% to fixed income. The equities allocation is believed to provide potential for increasing the size of the buffer between pension assets and liabilities with negligible risk to funded status. Recommendation B lies below the asset-only efficient frontier with a considerably lower expected return vis-à-vis Recommendation A.

What are the arguments for and against each of these recommendations? Recommendation A is expected, given capital market assumptions, to increase the size of the buffer between pension assets and liabilities. But the sponsor does not benefit from increases in the buffer if the current buffer is adequate.[23] However, with a 0.65 × \$1.25 billion = \$0.8125 allocation to equities and a current buffer of assets of \$1.25 billion – \$1.087 billion = \$0.163 billion, a decline of that amount or more in equity values (a 20% decline) would put the plan into underfunded status (assuming no commensurate changes in the liability). Thus, Recommendation A creates contribution risk for the plan sponsor without a potential upside clearly benefiting either the sponsor or beneficiaries.

For Recommendation B, because the risk characteristics of the \$1.125 billion fixed-income portfolio are closely matched with those of the \$1.087 billion of pension liabilities with a buffer, the plan sponsor should not face any meaningful risk of needing to make further contributions to the pension. Pensioners expect the plan to be fully funded on an ongoing basis without any reliance on the sponsor's ability to make additional contributions. This is an excellent outcome for both. The pension liabilities are covered (defeased). Consider next Recommendation A.

The example is highly stylized—the case facts were developed to make points cleanly—but does point to the potential value of managing risk in asset allocation explicitly in relation to liabilities. A typical use of fixed-income assets in liability-relative asset allocation should be noted: Liability-relative approaches to asset allocation tend to give fixed income a larger role than asset-only approaches in such cases as the one examined here because interest rates are a major financial market driver of both liability and bond values. Thus, bonds can be important in hedging liabilities, but equities can be relevant for liability hedging too. With richer case facts, as when liabilities accrue with inflation (not the case in the frozen DB example), equities may have a long-term role in matching the characteristics of liabilities. In underfunded plans, the potential upside of equities would often have greater value for the plan sponsor than in the fully funded case examined.

22 Collie and Gannon (2009) explore the contribution risk trade-off considered here in more detail.

23 Real-world complexities, such as DB plan termination to capture a positive surplus or pension risk transfer (annuitization), are beyond the scope of this reading; generally, there are restrictions and penalties involved in such actions, and the point made here is valid.

Liability Glide Paths

If GPLE were underfunded, it might consider establishing a liability glide path. A **liability glide path** is a technique in which the plan sponsor specifies in advance the desired proportion of liability-hedging assets and return-seeking assets and the duration of the liability hedge as funded status changes and contributions are made. The technique is particularly relevant to underfunded pensions. The idea reflects the fact that the optimal asset allocation in general is sensitive to changes in the funded status of the plan. The objective is to increase the funded status by reducing surplus risk over time. Although a higher contribution rate may be necessary to align assets with liabilities, the volatility of contributions should decrease, providing more certainty for cash flow planning purposes and decreasing risk to plan participants. Eventually, GPLE would hope to achieve and maintain a sufficiently high funded ratio so that there would be minimal risk of requiring additional contributions or transferring pension risk to an annuity provider.

The importance of such characteristics as interest rate sensitivity (duration), inflation, and credit risk in constructing a liability-hedging asset portfolio suggests the relevance of risk-factor modeling in liability-relative approaches. A risk factor approach can be extended to the return-seeking portfolio in order to minimize unintentional overlap among common factors across both portfolios—for example, credit. Exploring these topics is outside the scope of the current reading.

The next section addresses an approach to asset allocation related to liability relative in its focus on funding needs.

4.3 Goals Based

We use the hypothetical Lee family to present some thematic elements of a goals-based approach.

Investor Case Facts: The Lee Family

- *Name*: Ivy and Charles Lee
- *Narrative*: Ivy is a 54-year-old life sciences entrepreneur. Charles is 55 years old and employed as an orthopedic surgeon. They have two unmarried children aged 25 (Deborah) and 18 (David). Deborah has a daughter with physical limitations.
- *Financial assets and financial liabilities*: Portfolio of $25 million with $1 million in margin debt as well as residential real estate of $3 million with $1 million in mortgage debt.
- *Other assets and liabilities*:
  - Pre-retirement earnings are expected to total $16 million in present value terms (human capital).
  - David will soon begin studying at a four-year private university; the present value of the expected parental contribution is $250,000.
  - The Lees desire to give a gift to a local art museum in five years. In present value terms, the gift is valued at $750,000.
  - The Lees want to establish a trust for their granddaughter with a present value of $3 million to be funded at the death of Charles.
  - The present value of future consumption expenditures is estimated at $20 million.

Exhibit 9 Lee Family Economic Balance Sheet (in US$ millions) 31 December 20x6

Assets		**Liabilities and Net Worth**	
Financial assets		*Financial liabilities*	
Investment portfolio	25	Margin debt	1
Real estate	3	Mortgage	1
Extended assets		*Extended liabilities*	
Human capital	16	David's education	0.25
		Museum gift	0.75
		Special needs trust	3
		PV of future consumption	20
		Net worth	
		Economic net worth (economic assets less economic liabilities)	18
Total	44		44

The financial liabilities shown are legal liabilities. The extended liabilities include funding needs that the Lees want to meet. The balance sheet includes an estimate of the present value of future consumption, which is sometimes called the "consumption liability." The amount shown reflects expected values over their life expectancy given their ages. If they live longer, consumption needs will exceed the $20 million in the case facts and erode the $18 million in equity. If their life span is shorter, $18 million plus whatever they do not consume of the $20 million in PV of future consumption becomes part of their estate. Note that for the Lees, the value of assets exceeds the value of liabilities, resulting in a positive economic net worth (a positive difference between economic assets and economic liabilities); this is analogous to a positive owners' equity on a company's financial balance sheet.

From Exhibit 9, we can identify four goals totaling $24 million in present value terms: a lifestyle goal (assessed as a need for $20 million in present value terms), an education goal ($0.25 million), a charitable goal ($0.75 million), and the special needs trust ($3 million).

The present value of expected future earnings (human capital) at $16 million is less than the lifestyle present value of $20 million, which means that some part of the investment portfolio must fund the Lees' standard of living. It is important to note that although the Lee family has $18 million of economic net worth, most of this comes from the $16 million extended asset of human capital. Specific investment portfolio assets have not yet been dedicated to specific goals.

Goals-based asset allocation builds on several insights from behavioral finance. The approach's characteristic use of sub-portfolios is grounded in the behavioral finance insight that investors tend to ignore money's fungibility[24] and assign specific dollars to specific uses—a phenomenon known as mental accounting. Goals-based asset allocation, as described here, systemizes the fruitful use of mental accounts. This approach may help investors embrace more-optimal portfolios (as defined in an asset-only or

24 "Fungibility" is the property of an asset that a quantity of it may be replaced by another equal quantity in the satisfaction of an obligation. Thus, any 5,000 Japanese yen note can be used to pay a yen obligation of that amount, and the notes can be said to be fungible.

asset–liability framework) by adding higher risk assets—that, without context, might frighten the investor—to longer-term, aspirational sub-portfolios while adopting a more conservative allocation for sub-portfolios that address lifestyle preservation.

In Exhibit 10, the Lees' lifestyle goal is split into three components: a component called "lifestyle—minimum" intended to provide protection for the Lees' lifestyle in a disaster scenario, a component called "lifestyle—baseline" to address needs outside of worst cases, and a component called "lifestyle—aspirational" that reflects a desire for a chance at a markedly higher lifestyle. These sum to the present value of future consumption shown in the preceding Exhibit 9. Exhibit 10 describes these qualitatively; a numerical characterization could be very relevant for some advisers, however. By eliciting information on the Lees' perception of the goals' importance, the investment adviser might calibrate the required probabilities of achieving the goals quantitatively. For example, the three lifestyle goals might have 99%, 90%, and 50% assigned probabilities of success, respectively.

Exhibit 10 Lee Family: Required Probability of Meeting Goals and Goal Time Horizons

Goal	Required Probability of Achieving	Time Horizon
Lifestyle—minimum	Extremely high	Short to distant
Lifestyle—baseline	Very high	Short to distant
Lifestyle—aspirational	Moderate	Distant
Education	Very high	Short
Trust	High	Long
Charitable	Moderate	Short

Because the Lees might delay or forego making a gift to the museum if it would affect the trust goal, the trust goal is more urgent for the Lees. Also note that although parts of the Lees' lifestyle goals run the full time horizon spectrum from short to distant, they also have significant current earnings and human capital (which transforms into earnings as time passes). This fact puts the investment portfolio's role in funding the lifestyle goal further into the future.

Goals-based approaches generally set the strategic asset allocation in a bottom-up fashion. The Lees' lifestyle goal might be addressed with three sub-portfolios, with the longest horizon sub-portfolio being less liquid and accepting more risk than the others. Although for the GPLE pension, no risk distinction was made among different parts of the pension liability vis-à-vis asset allocation, such distinctions are made in goals-based asset allocation.

What about the Lees' other goals? Separate sub-portfolios could be assigned to the special needs and charitable goals with asset allocations that reflect the associated time horizons and required probabilities of not attaining these goals. A later reading on asset allocation in practice addresses implementation processes in detail.

Types of Goals

As goals-based asset allocation has advanced, various classification systems for goals have been proposed. Two of those classification systems are as follows.

Brunel (2012):

- *Personal goals*—to meet current lifestyle requirements and unanticipated financial needs
- *Dynastic goals*—to meet descendants' needs
- *Philanthropic goals*

Chhabra (2005):

- *Personal risk bucket*—to provide protection from a dramatic decrease in lifestyle (i.e., safe-haven investments)
- *Market risk bucket*—to ensure the current lifestyle can be maintained (allocations for average risk-adjusted market returns)
- *Aspirational risk bucket*—to increase wealth substantially (greater than average risk is accepted)

EXAMPLE 5

Goals-Based Asset Allocation

The Lees are presented with the following optimized asset allocations:

Asset Allocation	Cash	Global Bonds	Global Equities	Diversifying Strategies
A	40%	50%	10%	0%
B	10%	30%	45%	15%

Assume that a portfolio of 70% global equities and 30% bonds reflects an appropriate balance of expected return and risk for the Lees with respect to a 10-year time horizon for most moderately important goals. Based on the information given:

1 What goal(s) may be addressed by Allocation A?

2 What goal(s) may be addressed by Allocation B?

Because of her industry connections in the life sciences, Ivy Lee is given the opportunity to be an early-stage venture capital investor in what she assesses is a very promising technology.

3 What insights does goals-based asset allocation offer on this opportunity?

Solution to 1:

Allocation A stresses liquidity and stability. It may be appropriate to meet short-term lifestyle and education goals.

Solution to 2:

Allocation B has a greater growth emphasis, although it is somewhat conservative in relation to a 70/30 equity/bond baseline. It may be appropriate for funding the trust because of the goal's long time horizon and the Lees' desire for a high probability of achieving it.

Solution to 3:

Early-stage venture capital investments are both risky and illiquid; therefore, they belong in the longer-term and more risk-tolerant sub-portfolios. Ivy's decision about how much money she can commit should relate to how much excess capital remains after addressing goals that have a higher priority associated with them. Note that economic balance sheet thinking would stress that the life sciences opportunity is not particularly diversifying to her human capital.

Discount Rates and Longevity Risk

Although calculation of assets needed for sub-portfolios is outside the scope of this reading, certain themes can be indicated. Consider a retiree with a life expectancy of 20 years. The retiree has two goals:

- To maintain his current lifestyle upon retirement. This goal has a high required probability of achievement that is evaluated at 95%.
- To gift $1 million to a university in five years. This is viewed as a "desire" rather than a "need" and has a required probability evaluated at 75%.

Suppose that the investor's adviser specifies sub-portfolios as follows:

- for the first decade of lifestyle spending, a 3% expected return;
- for the second decade of lifestyle spending, a 4.6% expected return; and
- for the planned gift to the university, a 5.4% expected return.

Based on an estimate of annual consumption needs and the amount of the gift and given expected returns for the assigned sub-portfolios, the assets to be assigned to each sub-portfolio could be calculated by discounting amounts back to the present using their expected returns. However, this approach does not reflect the asset owner's required probability of achieving a goal. The higher the probability requirement for a future cash need, the greater the amount of assets needed in relation to it. Because of the inverse relation between present value and the discount rate, to reflect a 95% required probability, for example, the discount rates could be set at a lower level so that more assets are assigned to the sub-portfolio, increasing the probability of achieving the goal to the required level of 95% level.

Another consideration in determining the amount needed for future consumption is longevity risk. Life expectancies are median (50th percentile) outcomes. The retiree may outlive his life expectancy. To address longevity risk, the calculation of the present value of liabilities might use a longer life expectancy, such as a 35-year life expectancy instead of his actuarial 20-year expectation. Another approach is to transfer the risk to an insurer by purchasing an annuity that begins in 20 years and makes payments to the retiree for as long as he lives. Longevity risk and this kind of deferred annuity (sometimes called a "longevity annuity") are discussed in another curriculum reading on risk management.[25]

There are some drawbacks to the goals-based approach to asset allocation. One is that the sub-portfolios add complexity. Another is that goals may be ambiguous or may change over time. Goals-based approaches to asset allocation raise the question

25 See Blanchett et al. (2016) for the management of longevity risk. Milevsky (2016) is a further reference.

of how sub-portfolios coordinate to constitute an efficient whole. The subject will be taken up in a later reading, but the general finding is that the amount of sub-optimality is small.[26]

IMPLEMENTATION CHOICES

5

Having established the strategic asset allocation policy, the asset owner must address additional strategic considerations before moving to implementation. One of these is the passive/active choice.

There are two dimensions of passive/active choices. One dimension relates to the management of the strategic asset allocation itself—for example, whether to deviate from it tactically or not. The second dimension relates to passive and active implementation choices in investing the allocation to a given asset class. The first dimension is covered in Section 5.1. The second dimension is the subject of Section 5.2.

In an advisory role, asset managers have an unequivocal responsibility to make implementation and asset selection choices that are initially, and on an ongoing basis, suitable for the client.[27]

5.1 Passive/Active Management of Asset Class Weights

Tactical asset allocation (TAA) involves deliberate short-term deviations from the strategic asset allocation. Whereas the strategic asset allocation incorporates an investor's long-term, equilibrium market expectations, tactical asset allocation involves short-term tilts away from the strategic asset mix that reflect short-term views—for example, to exploit perceived deviations from equilibrium.

Tactical asset allocation is active management at the asset class level because it involves intentional deviations from the strategic asset mix to exploit perceived opportunities in capital markets to improve the portfolio's risk–return trade-off. TAA mandates are often specified to keep deviations from the strategic asset allocation within rebalancing ranges or within risk budgets. Tactical asset allocation decisions might be responsive to price momentum, perceived asset class valuation, or the particular stage of the business cycle. A strategy incorporating deviations from the strategic asset allocation that are motivated by longer-term valuation signals or economic views is sometimes distinguished as **dynamic asset allocation** (DAA).

Tactical asset allocation may be limited to tactical changes in domestic stock–bond or stock–bond–cash allocations or may be a more comprehensive multi-asset approach, as in a global tactical asset allocation (GTAA) model. Tactical asset allocation inherently involves market timing as it involves buying and selling in anticipation of short-term changes in market direction; however, TAA usually involves smaller allocation tilts than an invested-or-not-invested market timing strategy.

Tactical asset allocation is a source of risk when calibrated against the strategic asset mix. An informed approach to tactical asset allocation recognizes the trade-off of any potential outperformance against this tracking error. Key barriers to successful tactical asset allocation are monitoring and trading costs. For some investors, higher short-term capital gains taxes will prove a significant obstacle because taxes are an additional trading cost. A program of tactical asset allocation must be evaluated through a cost–benefit lens. The relevant cost comparisons include the expected costs of simply following a rebalancing policy (without deliberate tactical deviations).

26 This is addressed technically in Das et al. (2010). See also Brunel (2015).
27 See Standard III (C) in the Standards of Practice Handbook (CFA Institute 2014).

5.2 Passive/Active Management of Allocations to Asset Classes

In addition to active and passive decisions about the asset class mix, there are active and passive decisions about how to implement the individual allocations within asset classes. An allocation can be managed passively or actively or incorporate both active and passive sub-allocations. For investors who delegate asset management to external firms, these decisions would come under the heading of manager structure,[28] which includes decisions about how capital and active risk are allocated to points on the passive/active spectrum and to individual external managers selected to manage the investor's assets.[29] (Risk budgeting is described more fully in Section 5.3.)

With a **passive management** approach, portfolio composition does not react to changes in the investor's capital market expectations or to information on or insights into individual investments. (The word *passive* means *not reacting.*) For example, a portfolio constructed to track the returns of an index of European equities might add or drop a holding in response to a change in the index composition but not in response to changes in the manager's expectations concerning the security's investment value; the market's expectations reflected in market values and index weights are taken as is. Indexing is a common passive approach to investing. (Another example would be buying and holding a fixed portfolio of bonds to maturity.)

In contrast, a portfolio manager for an active management strategy will respond to changing capital market expectations or to investment insights resulting in changes to portfolio composition. The objective of active management is to achieve, after expenses, positive excess risk-adjusted returns relative to a passive benchmark.

The range of implementation choices can be practically viewed as falling along a passive/active spectrum because some strategies use both passive and active elements. In financial theory, the pure model of a passive approach is indexing to a broad market-cap-weighted index of risky assets—in particular, the global market portfolio discussed in Section 4.1. This portfolio sums all investments in index components and is macro-consistent in the sense that all investors could hold it, and it is furthermore self-rebalancing to the extent it is based on market-value-weighted indices. A buy-and-hold investment as a proxy for the global market portfolio would represent a theoretical endpoint on the passive/active spectrum. However, consider an investor who indexes an equity allocation to a broad-based value equity style index. The investment could be said to reflect an active decision in tilting an allocation toward value but be passive in implementation because it involves indexing. An even more active approach would be investing the equity allocation with managers who have a value investing approach and attempt to enhance returns through security selection. Those managers would show positive tracking risk relative to the value index in general. Unconstrained active investment would be one that is "go anywhere" or not managed with consideration of any traditional asset class benchmark (i.e., "benchmark agnostic"). The degree of active management has traditionally been quantified by tracking risk and, from a different perspective, by active share.

Indexing is generally the lowest-cost approach to investing. Indexing involves some level of transaction costs because, as securities move in and out of the index, the portfolio holdings must adjust to remain in alignment with the index. Although indexing to a market-cap-weighted index is self-rebalancing, tracking an index based on other weighting schemes requires ongoing transactions to ensure the portfolio remains in alignment with index weights. An example is tracking an equally weighted index: As changes in market prices affect the relative weights of securities in the portfolio over

---

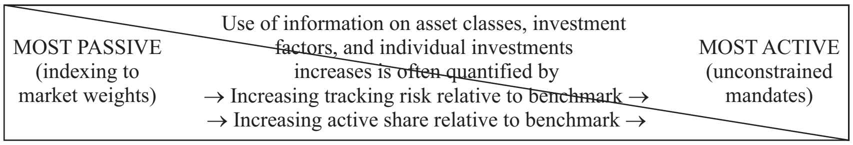
28 Manager structure is defined by the number of managers, types of managers, as well as which managers are selected.

29 See, for example, Waring, Whitney, Pirone, and Castille (2000).

time, the portfolio will need to be rebalanced to restore equal weights. Portfolios tracking fixed-income indices also incur ongoing transaction costs as holdings mature, default, or are called away by their issuers.

Exhibit 11 diagrams the passive/active choice as a continuum rather than binary (0 or 1) characteristic. Tracking risk and active share are widely known quantitative measures of the degree of active management that capture different aspects of it. Each measure is shown as tending to increase from left to right on the spectrum; however, they do not increase (or decrease) in lockstep with each other, in general.

Exhibit 11 Passive/Active Spectrum



Asset class allocations may be managed with different approaches on the spectrum. For example, developed market equities might be implemented purely passively, whereas emerging market bonds might be invested with an unconstrained, index-agnostic approach.

Factors that influence asset owners' decisions on where to invest on the passive/active spectrum include the following:

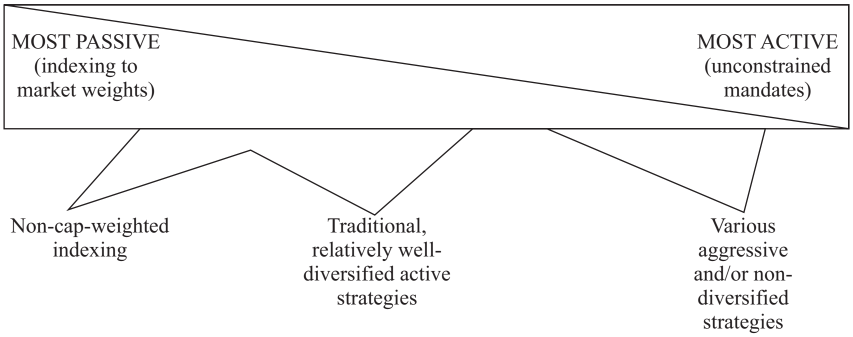
- *Available investments.* For example, the availability of an investable and representative index as the basis for indexing.
- *Scalability of active strategies being considered.* The prospective value added by an active strategy may begin to decline at some level of invested assets. In addition, participation in it may not be available below some asset level, a consideration for small investors.
- *The feasibility of investing passively while incorporating client-specific constraints.* For example, an investor's particular ESG investing criteria may not align with existing index products.
- *Beliefs concerning market informational efficiency.* A strong belief in market efficiency for the asset class(es) under consideration would orient the investor away from active management.
- *The trade-off of expected incremental benefits relative to incremental costs and risks of active choices.* Costs of active management include investment management costs, trading costs, and turnover-induced taxes; such costs would have to be judged relative to the lower costs of index alternatives, which vary by asset class.
- *Tax status.* Holding other variables constant, taxable investors would tend to have higher hurdles to profitable active management than tax-exempt investors.[30] For taxable investors who want to hold both passive and active investments, active investments would be held, in general, in available tax-advantaged accounts.

The curriculum readings on equity, fixed-income, and alternative investments will explore many strategies and the nature of any active decisions involved. Investors do need to understand the nature of the active decisions involved in implementing

30 See Jeffrey and Arnott (1993).

their strategic asset allocations and their appropriateness given the factors described. Exhibit 12 shows qualitatively (rather than precisely) some choices that investors may consider for equity and fixed-income allocations. In the exhibit, non-cap-weighted indexing includes such approaches as equal weighting and quantitative rules-based indexing approaches (discussed further in the equity readings).[31]

Exhibit 12 Placement on the Passive/Active Spectrum: Examples of Possible Choices



EXAMPLE 6

Implementation Choices (1)

1. Describe two kinds of passive/active choices faced by investors related to asset allocation.
2. An equity index is described as "a rules-based, transparent index designed to provide investors with an efficient way to gain exposure to large-cap and small-cap stocks with low total return variability." Compared with the market-cap weighting of the parent index (with the same component securities), the weights in the low-volatility index are proportional to the inverse of return volatility, so that the highest-volatility security receives the lowest weight. Describe the active and passive aspects of a decision to invest an allocation to equities in ETFs tracking such indices.
3. Describe how investing in a GDP-weighted global bond index involves both active and passive choices.

Solution to 1:

One choice relates to whether to allow active deviations from the strategic asset allocation. Tactical asset allocation and dynamic asset allocation are examples of active management of asset allocations. A second set of choices relates to where to invest allocations to asset classes along the passive/active spectrum.

Solution to 2:

The active element is the decision, relative to the parent index, to overweight securities with low volatility and underweight securities with high volatility. This management of risk is distinct from reducing portfolio volatility by combining a

31 Podkaminer (2015) provides a survey.

market-cap-weighted index with a risk-free asset proxy because it implies a belief in some risk–return advantage to favoring low-volatility equities on an individual security basis. The passive element is a transparent rules-based implementation of the weighting scheme based on inverse volatilities.

Solution to 3:

The passive choice is represented by the overall selection of the universe of global bonds; however, the active choice is represented by the weighting scheme, which is to use GDP rather than capital market weights. This is a tilt toward the real economy and away from fixed-income market values.

EXAMPLE 7

Implementation Choices (2)

Describe characteristic(s) of each of the following investors that are likely to influence the decision to invest passively or actively. See Exhibit 3 for Questions 1–3 and Example 1 for Question 4.

1 Caflandia sovereign wealth fund
2 GPLE corporate pension
3 The Lee family
4 Auldberg University Endowment

Solution:

1 For a large investor like the Caflandia sovereign wealth fund (CAF$40 billion), the scalability of active strategies that it may wish to employ may be a consideration. If only a small percentage of portfolio assets can be invested effectively in an active strategy, for example, the potential value added for the overall portfolio may not justify the inherent costs and management time. Although the equities and fixed-income allocations could be invested using passive approaches, investments in the diversifying strategies category are commonly active.
2 The executives responsible for the GPLE corporate pension also have other, non-investment responsibilities. This is a factor favoring a more passive approach; however, choosing an outsourced chief investment officer or delegated fiduciary consultant to manage active manager selection could facilitate greater use of active investment.
3 The fact that the Lees are taxable investors is a factor generally in favor of passive management for assets not held in tax-advantaged accounts. Active management involves turnover, which gives rise to taxes.
4 According to the vignette in Example 1, the Auldberg University Endowment has substantial staff resources in equities, fixed income, and real estate. This fact suggests that passive/active decisions are relatively unconstrained by internal resources. By itself, it does not favor passive or active, but it is a factor that allows active choices to be given full consideration.

5.3 Risk Budgeting Perspectives in Asset Allocation and Implementation

Risk budgeting addresses the questions of which types of risks to take and how much of each to take. Risk budgeting provides another view of asset allocation—through a risk lens. Depending on the focus, the risk may be quantified in various ways. For example, a concern for volatility can be quantified as variance or standard deviation of returns, and a concern for tail risk can be quantified as VaR or drawdown. Risk budgets (budgets for risk taking) can be stated in absolute or in relative terms and in money or percent terms. For example, it is possible to state an overall risk budget for a portfolio in terms of volatility of returns, which would be an example of an absolute risk budget stated in percent terms (for example, 20% for portfolio return volatility). Risk budgeting is a tool that may be useful in a variety of contexts and asset allocation approaches.

Some investors may approach asset allocation with an exclusive focus on risk. A risk budgeting approach to asset allocation has been defined as an approach in which the investor specifies how risk (quantified by some measure, such as volatility) is to be distributed across assets in the portfolio, without consideration of the assets' expected returns.[32] An example is aiming for equal expected risk contributions to overall portfolio volatility from all included asset classes as an approach to diversification, which is a risk parity (or equal risk contribution) approach. A subsequent reading in asset allocation addresses this in greater detail.

More directly related to the choice of passive/active implementation are active risk budgets and active risk budgeting. **Active risk budgeting** addresses the question of how much benchmark-relative risk an investor is willing to take in seeking to outperform a benchmark. This approach is risk budgeting stated in benchmark-relative terms. In parallel to the two dimensions of the passive/active decision outlined previously are two levels of active risk budgeting, which can be distinguished as follows:

- At the level of the overall asset allocation, active risk can be defined relative to the strategic asset allocation benchmark. This benchmark may be the strategic asset allocation weights applied to specified (often, broad-based market-cap-weighted) indices.
- At the level of individual asset classes, active risk can be defined relative to the asset class benchmark.

Active risk budgeting at the level of overall asset allocation would be relevant to tactical asset allocation. Active risk budgeting at the level of each asset class is relevant to how the allocation to those asset classes is invested. For example, it can take the form of expected-alpha versus tracking-error optimization in a manner similar to classic mean–variance optimization. If investment factor risks are the investor's focus, risk budgeting can be adapted to have a focus on allocating factor risk exposures instead. Later readings revisit risk budgeting in investing in further detail.

6 REBALANCING: STRATEGIC CONSIDERATIONS

Rebalancing is the discipline of adjusting portfolio weights to more closely align with the strategic asset allocation. Rebalancing is a key part of the monitoring and feedback step of the portfolio construction, monitoring, and revision process. An investor's rebalancing policy is generally documented in the IPS.

32 See Roncalli (2013).

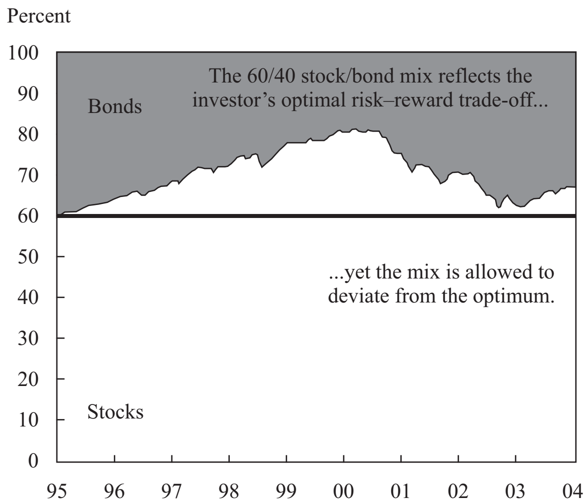
Even in the absence of changing investor circumstances, a revised economic outlook, or tactical asset allocation views, normal changes in asset prices cause the portfolio asset mix to deviate from target weights. Industry practice defines “rebalancing” as portfolio adjustments triggered by such price changes. Other portfolio adjustments, even systematic ones, are not rebalancing.

Ordinary price changes cause the assets with a high forecast return to grow faster than the portfolio as a whole. Because high-return assets are typically also higher risk, in the absence of rebalancing, overall portfolio risk rises. The mix of risks within the portfolio becomes more concentrated as well. Systematic rebalancing maintains the original strategic risk exposures. The discipline of rebalancing serves to control portfolio risks that have become different from what the investor originally intended.

Consider the example from the internet bubble (1995–2001) in Exhibit 13. The example assumes a 60/40 stock/bond portfolio, in which stocks are represented by the large-cap US growth stocks that characterized the internet bubble. In Panel B, the left-hand scale and upper two lines show month-by-month total portfolio *values* with and without monthly rebalancing (“wealth rebalanced” and “wealth unrebalanced,” respectively). The right-hand scale and lower two lines show month-by-month portfolio *risk* as represented by the 5th percentile drawdown (in a VaR model) with and without monthly rebalancing (“risk rebalanced” and “risk unrebalanced,” respectively).

Exhibit 13 Rebalancing

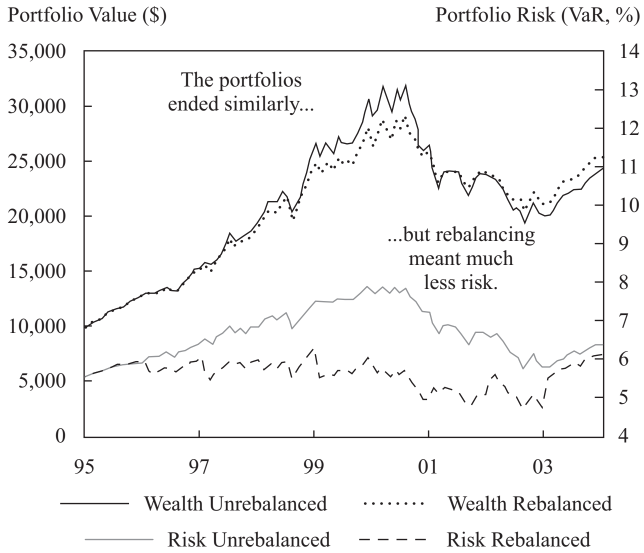
Panel A. Asset Mix



(continued)

Exhibit 13 (Continued)

Panel B. Portfolio Value and Risk



Note: The data are a 60/40 mix of the S&P 500 Growth Index and the Barclays Capital Aggregate Bond Index.

Panel A shows that, without rebalancing, the asset mix deviates dramatically from the target. Panel B shows that although the portfolios' values ended similarly (the upper two lines), disciplined rebalancing meant more-stable risks (illustrated by the lower two lines).

This risk perspective is important. Taken to the extreme, *never rebalancing* allows the high-return (and presumably higher-risk) assets to grow and dominate the portfolio. Portfolio risk rises and concentrates. Taken even further, such a philosophy of never rebalancing may suggest it would have been simpler to have invested only in the highest-expected-return asset class back when the asset mix decision was made. Not rebalancing could negate an intended level of diversification.

Because rebalancing is countercyclical, it is fundamentally a contrarian investment approach.[33] Behavioral finance tells us that such contrarianism will be uncomfortable; no one likes to sell the most recently best-performing part of the portfolio to buy the worst. Thus, rebalancing is a *discipline* of adjusting the portfolio to better align with the strategic asset allocation in both connotations of discipline—the sense of a typical practice and the sense of a strengthening regime.

6.1 A Framework for Rebalancing

The actual mechanics of rebalancing are more complex than they first appear. A number of questions arise: How often should the portfolio be rebalanced? What levels of imbalance are worth tolerating? Should the portfolio be rebalanced to the edge of the policy range or to some other point? These non-trivial questions represent the key strategic decisions in rebalancing.

33 A quantitative interpretation of rebalancing, given by Ang (2014), is that the return to rebalancing is selling out of the money puts and calls.

The simplest approach to rebalancing is **calendar rebalancing**, which involves rebalancing a portfolio to target weights on a periodic basis—for example, monthly, quarterly, semiannually, or annually. The choice of rebalancing frequency may be linked to the schedule of portfolio reviews. Although simple, rebalancing points are arbitrary and have other disadvantages.

Percent-range rebalancing permits tighter control of the asset mix compared with calendar rebalancing. Percent-range approach involves setting rebalancing thresholds or trigger points, stated as a percentage of the portfolio's value, around target values. For example, if the target allocation to an asset class is 50% of portfolio value, **trigger points** at 45% and 55% of portfolio value define a 10 percentage point **rebalancing range** (or corridor) for the value of that asset class. The rebalancing range creates a no-trade region. The portfolio is rebalanced when an asset class's weight first passes through one of its trigger points. Focusing on percent-range rebalancing, the following questions are relevant:

- How frequently is the portfolio valued?
- What size deviation triggers rebalancing?
- Is the deviation from the target allocation fully or partially corrected?

How frequently is the portfolio valued? The percent-range discipline requires monitoring portfolio values for breaches of a trigger point at an agreed-on frequency; the more frequent the monitoring, the greater the precision in implementation. Such monitoring may be scheduled daily, weekly, monthly, quarterly, or annually. A number of considerations—including governance resources and asset custodian resources—can affect valuation frequency. For many investors, monthly or quarterly evaluation efficiently balances the costs and benefits of rebalancing.

What size deviation triggers rebalancing? Trigger points take into account such factors as traditional practice, transaction costs, asset class volatility, volatility of the balance of the portfolio, correlation of the asset class with the balance of the portfolio, and risk tolerance.[34]

Before the rise of modern multi-asset portfolios, the stock/bond split broadly characterized the asset allocation and a traditional $\pm x$% rebalancing band was common. These fixed ranges would apply no matter the size or volatility of the allocation target. For example, both a 40% domestic equity allocation and a 15% real asset allocation might have ±5% rebalancing ranges. Alternatively, proportional bands reflect the size of the target weight. For example, a 60% target asset class might have a ±6% band, whereas a 5% allocation would have a ±0.5% band. Proportional bands might also be set to reflect the relative volatility of the asset classes. A final approach is the use of cost–benefit analysis to set ranges.

Is the deviation from the target allocation fully or partially corrected? Once the portfolio is evaluated and an unacceptably large deviation found, the investor must determine rebalancing trade size, as well as the timeline for implementing the rebalancing. In practice, three main approaches are used: rebalance back to target weights, rebalance to range edge, or rebalance halfway between the range-edge trigger point and the target weight.

34 See Masters (2003) for details on these factors apart from traditional factors.

6.2 Strategic Considerations in Rebalancing

The four-part rebalancing framework just described highlights important questions to address in setting rebalancing policy. Strategic considerations generally include the following, all else being equal:

- Higher transaction costs for an asset class imply wider rebalancing ranges.
- More risk-averse investors will have tighter rebalancing ranges.
- Less correlated assets also have tighter rebalancing ranges.
- Beliefs in momentum favor wider rebalancing ranges, whereas mean reversion encourages tighter ranges.
- Illiquid investments complicate rebalancing.
- Derivatives create the possibility of synthetic rebalancing.
- Taxes, which are a cost, discourage rebalancing and encourage asymmetric and wider rebalancing ranges.

Asset class volatility is also a consideration in the size of rebalancing ranges.[35]

A cost–benefit approach to rebalancing sets ranges, taking transaction costs, risk aversion, asset class risks, and asset class correlations into consideration. For example, an asset that is more highly correlated with the rest of the portfolio than another would merit a wider rebalancing range, all else equal, because it would be closer to being a substitute for the balance of the portfolio; thus, larger deviations would have less impact on portfolio risk.

EXAMPLE 8

Different Rebalancing Ranges

The table shows a simple four-asset strategic mix along with rebalancing ranges created under different approaches.

Explain why the international equity range is wider than the domestic equity range under proportional ranges using the cost–benefit approach.[36]

Asset Class	Strategic Target	Fixed Width Ranges	Proportional Ranges (±1,000 bps)	Cost–Benefit Ranges
Domestic equity	40%	35%–45%	36%–44%	35%–45%
International equity	25%	20%–30%	22½%–27½%	19%–31%
Emerging markets	15%	10%–20%	13½%–16½%	12%–18%
Fixed income	20%	15%–25%	18%–22%	19%–21%

Solution:

Higher transaction costs for international equity compared with domestic equity could explain the wider range for international equity compared with domestic equity under the cost–benefit approach. Another potential explanation relates to the possibility that international equity has a higher correlation with the balance of the portfolio (i.e., the portfolio excluding international equity)

35 The effect of an asset's volatility on its appropriate rebalancing range involves a trade-off between transaction costs and risk control. This complex issue is discussed further in the curriculum reading "Asset Allocation in Practice."

36 The cost-benefit approach calculations are based on Masters (2003).

than does domestic equity (i.e., with the portfolio excluding domestic equity). If that is the case then, all else being equal, a wider band would be justified for international equity.

Investors' perspectives on capital markets can affect their approach to rebalancing. A belief in momentum and trend following, for example, encourages wider rebalancing ranges. In contrast, a belief in mean reversion encourages stricter adherence to rebalancing, including tighter ranges.

Illiquid assets complicate rebalancing. Relatively illiquid investments, such as hedge funds, private equity, or direct real estate, cannot be readily traded without substantial trading costs and/or delays. Accordingly, illiquid investments are commonly assigned wide rebalancing ranges. However, rebalancing of an illiquid asset may be affected indirectly when a highly correlated liquid asset can be traded or when exposure can be adjusted by means of positions in derivatives. For example, public equity could be reduced to offset an overweight in private equity. Rebalancing by means of highly correlated liquid assets and derivatives, however, involves some imprecision and basis risk.

This insight about liquidity is an instance where thinking ahead about rebalancing can affect the strategic asset allocation. It is one reason that allocations to illiquid assets are often smaller than if trading were possible.

Factor-based asset allocation, liability-relative investing, and goals-based investing, each a valid approach to asset allocation, can give rise to different rebalancing considerations. Factor exposures and liability hedges require monitoring (and rebalancing) the factors weights and surplus duration in addition to asset class weights. Goals-based investing in private wealth management may require both asset class rebalancing and moving funds between different goal sub-portfolios.

Tax considerations also complicate rebalancing. Rebalancing typically realizes capital gains and losses, which are taxable events in many jurisdictions. For private wealth managers, any rebalancing benefit must be compared with the tax cost. Taxes, as a cost, are much larger than other transaction costs, which often leads to wider rebalancing ranges in taxable portfolios than in tax-exempt portfolios. Because loss harvesting generates tax savings and realizing gains triggers taxes, rebalancing ranges in taxable accounts may also be asymmetric. (For example, a 25% target asset class might have an allowable range of 24%–28%, which is −1% to +3%.)

Modern cost–benefit approaches to rebalancing suggest considering derivatives as a rebalancing tool. Derivatives can often be used to rebalance synthetically at much lower transaction costs than the costs of using the underlying stocks and bonds. Using a derivatives overlay also avoids disrupting the underlying separate accounts in a multi-manager implementation of the strategic asset allocation. Tax considerations are also relevant; it may be more cost effective to reduce an exposure using a derivatives overlay than to sell the underlying asset and incur the capital gains tax liability. Lastly, trading a few derivatives may be quicker and easier than hundreds of underlying securities. Of course, using derivatives may require a higher level of risk oversight, but then risk control is the main rationale for rebalancing.

Estimates of the benefits of rebalancing vary. Many portfolios are statistically indistinguishable from each other, suggesting that much rebalancing is unnecessary. In contrast, Willenbrock (2011) demonstrates that even zero-return assets can, in theory, generate positive returns through rebalancing, which is a demonstrable (and surprising) benefit. Whatever the return estimate for the value added from rebalancing, the key takeaway is that rebalancing is chiefly about risk control, not return enhancement.

SUMMARY

This reading has introduced the subject of asset allocation. Among the points made are the following:

- The economic balance sheet includes non-financial assets and liabilities that can be relevant for choosing the best asset allocation for an investor's financial portfolio.
- The investment objectives of asset-only asset allocation approaches focus on the asset side of the economic balance sheet; approaches with a liability-relative orientation focus on funding liabilities; and goals-based approaches focus on achieving financial goals.
- The risk concepts relevant to asset-only asset allocation approaches focus on asset risk; those of liability-relative asset allocation focus on risk in relation to paying liabilities; and a goals-based approach focuses on the probabilities of not achieving financial goals.
- Asset classes are the traditional units of analysis in asset allocation and reflect systematic risks with varying degrees of overlap.
- Assets within an asset class should be relatively homogeneous; asset classes should be mutually exclusive; asset classes should be diversifying; asset classes as a group should make up a preponderance of the world's investable wealth; asset classes selected for investment should have the capacity to absorb a meaningful proportion of an investor's portfolio.
- Risk factors are associated with non-diversifiable (i.e., systematic) risk and are associated with an expected return premium. The price of an asset and/or asset class may reflect more than one risk factor, and complicated spread positions may be necessary to identify and isolate particular risk factors. Their use as units of analysis in asset allocation is driven by considerations of controlling systematic risk exposures.
- The global market portfolio represents a highly diversified asset allocation that can serve as a baseline asset allocation in an asset-only approach.
- There are two dimensions of passive/active choices. One dimension relates to the management of the strategic asset allocation itself—for example, whether to deviate from it tactically or not. The second dimension relates to passive and active implementation choices in investing the allocation to a given asset class. Tactical and dynamic asset allocation relate to the first dimension; active and passive choices for implementing allocations to asset classes relate to the second dimension.
- Risk budgeting addresses the question of which types of risks to take and how much of each to take. Active risk budgeting addresses the question of how much benchmark-relative risk an investor is willing to take. At the level of the overall asset allocation, active risk can be defined relative to the strategic asset allocation benchmark. At the level of individual asset classes, active risk can be defined relative to the benchmark proxy.
- Rebalancing is the discipline of adjusting portfolio weights to more closely align with the strategic asset allocation. Rebalancing approaches include calendar-based and range-based rebalancing. Calendar-based rebalancing rebalances the portfolio to target weights on a periodic basis. Range-based rebalancing sets

rebalancing thresholds or trigger points around target weights. The ranges may be fixed width, percentage based, or volatility based. Range-based rebalancing permits tighter control of the asset mix compared with calendar rebalancing.

- Strategic considerations in rebalancing include transaction costs, risk aversion, correlations among asset classes, volatility, and beliefs concerning momentum, taxation, and asset class liquidity.

REFERENCES

Ang, Andrew. 2014. *Asset Management: A Systematic Approach to Factor Investing*. New York: Oxford University Press.

Blanchett, David M., and Philip U. Straehl. 2015. "No Portfolio is an Island." *Financial Analysts Journal*, vol. 71, no. 3 (May/June): 15–33.

Blanchett, David M., and Philip U. Straehl. 2017. "Portfolio Implications of Job-Specific Human Capital Risk." *Journal of Asset Management*, vol. 18, no. 1:1–15.

Blanchett, David M., David M. Cordell, Michael S. Finke, and Thomas Idzorek. 2016. "Risk Management for Individuals." CFA Institute.

Brunel, Jean L.P. 2012. "Goals-Based Wealth Management in Practice." *Conference Proceedings Quarterly*, vol. 29, no. 1 (March): 57–65.

Brunel, Jean L.P. 2015. *Goals-Based Wealth Management: An Integrated and Practical Approach to Changing the Structure of Wealth Advisory Practices*. Hoboken, NJ: John Wiley.

Burger, John D., Francis E. Warnock, and Veronica Cacdac Warnock. 2012. "Emerging Local Currency Bond Markets." *Financial Analysts Journal*, vol. 68, no. 4 (July/August): 73–93.

Chhabra, Ashvin. 2005. "Beyond Markowitz: A Comprehensive Wealth Allocation Framework for Individual Investors." *Journal of Wealth Management*, vol. 7, no. 4 (Spring): 8–34.

Collie, Bob, and James A. Gannon. 2009. "Liability-Responsive Asset Allocation: A Dynamic Approach to Pension Plan Management." Russell Investments Insights (April).

Das, Sanjiv, Harry Markowitz, Jonathan Scheid, and Meir Statman. 2010. "Portfolio Optimization with Mental Accounts." *Journal of Financial and Quantitative Analysis*, vol. 45, no. 2 (April): 311–334.

Fraser, Steve P., and William W. Jennings. 2006. "Behavioral Asset Allocation for Foundations and Endowments." *Journal of Wealth Management*, vol. 9, no. 3 (Winter): 38–50.

Greer, Robert J. 1997. "What is an Asset Class, Anyway?" *Journal of Portfolio Management*, vol. 23, no. 2 (Winter): 86–91.

Ibbotson, Roger G., and Paul D. Kaplan. 2000. "Does Asset Allocation Policy Explain 40, 90, or 100 Percent of Performance?" *Financial Analysts Journal*, vol. 56, no. 1 (January/February): 26–33.

Ibbotson, Roger G., Moshe A. Milevsky, Peng Chen, and Kevin X. Zhu. 2007. *Lifetime Financial Advice: Human Capital, Asset Allocation, and Insurance*. Charlottesville, VA: CFA Institute Research Foundation.

Idzorek, Thomas M., and Maciej Kowara. 2013. "Factor-Based Asset Allocation versus Asset-Class-Based Asset Allocation." *Financial Analysts Journal*, vol. 69, no. 3 (May/June): 19–29.

Idzorek, Thomas, Jeremy Stempien, and Nathan Voris. 2013. "Bait and Switch: Glide Path Instability." *Journal of Investing*, vol. 22, no. 1 (Spring): 74–82.

Jeffrey, Robert H., and Robert D. Arnott. 1993. "Is Your Alpha Big Enough to Cover Its Taxes?" *Journal of Portfolio Management*, vol. 19, no. 3 (Spring): 15–25.

Kozhemiakin, Alexander. 2011. "Emerging Markets Local Currency Debt: Capitalizing on Improved Sovereign Fundamentals." BNY Mellon Asset Management.

Kritzman, Mark. 1999. "Toward Defining an Asset Class." *Journal of Alternative Investments*, vol. 2, no. 1 (Summer): 79–82.

Masters, Seth J. 2003. "Rebalancing." *Journal of Portfolio Management*, vol. 29, no. 3 (Spring): 52–57.

Milevsky, Moshe A. 2016. "It's Time to Retire Ruin (Probabilities)." *Financial Analysts Journal*, vol. 72, no. 2 (March/April): 8–12.

Perry, William. 2011. "The Case For Emerging Market Corporates." *Journal of Indexes*, vol. 14, no. 5 (September/October): 10–17.

Podkaminer, Eugene. 2013. "Risk Factors as Building Blocks for Portfolio Diversification: The Chemistry of Asset Allocation." CFA Institute Investment Risk and Performance papers (January).

Podkaminer, Eugene. 2015. "The Education of Beta: Can Alternative Indexes Make Your Portfolio Smarter?" *Journal of Investing*, vol. 24, no. 2 (Summer): 7–34.

Pompian, Michael M. 2011a. "The Behavioral Biases of Individuals." CFA Institute.

Pompian, Michael M. 2011b. "The Behavioral Finance Perspective." CFA Institute.

Pompian, Michael, Colin McLean, and Alistair Byrne. 2011. "Behavioral Finance and Investment Processes." CFA Institute.

Roncalli, Thierry. 2013. *Introduction to Risk Parity and Budgeting*. New York: CRC Press.

Rudd, Andrew, and Laurence B. Siegel. 2013. "Using an Economic Balance Sheet for Financial Planning." *Journal of Wealth Management*, vol. 16, no. 2 (Fall): 15–23.

Sharpe, William F. 1991. "The Arithmetic of Active Management." *Financial Analysts Journal*, vol. 47, no. 1:7–9.

Shefrin, H., and M. Statman. 2000. "Behavioral Portfolio Theory." *Journal of Financial and Quantitative Analysis*, vol. 35, no. 2 (June): 127–151.

CFA Institute.*Standards of Practice Handbook*, 11th edition. 2014. CFA Institute.

Swinkels, Laurens. 2012. “Emerging Market Inflation-Linked Bonds.” *Financial Analysts Journal*, vol. 68, no. 5 (September/October): 38–56.

Waring, M. Barton, Duane Whitney, John Pirone, and Charles Castille. 2000. “Optimizing Manager Structure and Budgeting Manager Risk.” *Journal of Portfolio Management*, vol. 26, no. 3 (Spring): 90–104.

Willenbrock, Scott. 2011. “Diversification Return, Portfolio Rebalancing, and the Commodity Return Puzzle.” *Financial Analysts Journal*, vol. 67, no. 4 (July/August): 42–49.

Xiong, James X., Roger G. Ibbotson, Thomas M. Idzorek, and Peng Chen. 2010. “The Equal Importance of Asset Allocation and Active Management.” *Financial Analysts Journal*, vol. 66, no. 2 (March/April): 22–30.

PRACTICE PROBLEMS

The following information relates to Questions 1–8

Meg and Cramer Law, a married couple aged 42 and 44, respectively, are meeting with their new investment adviser, Daniel Raye. The Laws have worked their entire careers at Whorton Solutions (WS), a multinational technology company. The Laws have two teenage children who will soon begin college.

Raye reviews the Laws' current financial position. The Laws have an investment portfolio consisting of $800,000 in equities and $450,000 in fixed-income instruments. Raye notes that 80% of the equity portfolio consists of shares of WS. The Laws also own real estate valued at $400,000, with $225,000 in mortgage debt. Raye estimates the Laws' pre-retirement earnings from WS have a total present value of $1,025,000. He estimates the Laws' future expected consumption expenditures have a total present value of $750,000.

The Laws express a very strong desire to fund their children's college education expenses, which have an estimated present value of $275,000. The Laws also plan to fund an endowment at their alma mater in 20 years, which has an estimated present value of $500,000. The Laws tell Raye they want a high probability of success funding the endowment. Raye uses this information to prepare an economic balance sheet for the Laws.

In reviewing a financial plan written by the Laws' previous adviser, Raye notices the following asset class specifications.

Equity:	US equities
Debt:	Global investment-grade corporate bonds and real estate
Derivatives:	Primarily large-capitalization foreign equities

The previous adviser's report notes the asset class returns on equity and derivatives are highly correlated. The report also notes the asset class returns on debt have a low correlation with equity and derivative returns.

Raye is concerned that the asset allocation approach followed by the Laws' previous financial adviser resulted in an overlap in risk factors among asset classes for the portfolio. Raye plans to address this by examining the portfolio's sensitivity to various risk factors, such as inflation, liquidity, and volatility, to determine the desired exposure to each factor.

Raye concludes that a portfolio of 75% global equities and 25% bonds reflects an appropriate balance of expected return and risk for the Laws with respect to a 20-year time horizon for most moderately important goals. Raye recommends the Laws follow a goals-based approach to asset allocation and offers three possible portfolios for the Laws to consider. Selected data on the three portfolios are presented in Exhibit 1.

© 2017 CFA Institute. All rights reserved.

Exhibit 1 Proposed Portfolio Allocations for the Law Family

	Cash	Fixed Income	Global Equities	Diversifying Strategies*
Portfolio 1	35%	55%	10%	0%
Portfolio 2	10%	15%	65%	10%
Portfolio 3	10%	30%	40%	20%

* Diversifying strategies consists of hedge funds

Raye uses a cost–benefit approach to rebalancing and recommends that global equities have a wider rebalancing range than the other asset classes.

1 Using the economic balance sheet approach, the Laws' economic net worth is *closest* to:

A $925,000.

B $1,425,000.

C $1,675,000.

2 Using an economic balance sheet, which of the Laws' current financial assets is *most* concerning from an asset allocation perspective?

A Equities

B Real estate

C Fixed income

3 Raye believes the previous adviser's specification for debt is incorrect given that, for purposes of asset allocation, asset classes should be:

A diversifying.

B mutually exclusive.

C relatively homogeneous.

4 Raye believes the previous adviser's asset class specifications for equity and derivatives are inappropriate given that, for purposes of asset allocation, asset classes should be:

A diversifying.

B mutually exclusive.

C relatively homogeneous.

5 To address his concern regarding the previous adviser's asset allocation approach, Raye should assess the Laws' portfolio using:

A a homogeneous and mutually exclusive asset class–based risk analysis.

B a multifactor risk model to control systematic risk factors in asset allocation.

C an asset class–based asset allocation approach to construct a diversified portfolio.

6 Based on Exhibit 1, which portfolio *best* meets the Laws' education goal for their children?

A Portfolio 1

B Portfolio 2

C Portfolio 3

7 Based on Exhibit 1, which portfolio *best* meets the Laws' goal to fund an endowment for their alma mater?

A Portfolio 1

B Portfolio 2

C Portfolio 3

8 Raye's approach to rebalancing global equities is consistent with:

A the Laws' being risk averse.

B global equities' having higher transaction costs than other asset classes.

C global equities' having lower correlations with other asset classes.

SOLUTIONS

1 A is correct. The Laws' economic net worth is closest to $925,000. An economic balance sheet includes conventional financial assets and liabilities, as well as extended portfolio assets and liabilities that are relevant in making asset allocation decisions. The economic balance sheet for the Law family is shown in the following exhibit.

Assets		Liabilities and Net Worth	
Financial Assets		*Financial Liabilities*	
Fixed income	450,000	Mortgage debt	225,000
Real estate	400,000		
Equity	800,000		
Extended Assets		*Extended Liabilities*	
Human capital	1,025,000	Children's education	275,000
		Endowment funding	500,000
		Present value of consumption	750,000
Total Economic Assets	2,675,000	*Total Economic Liabilities*	1,750,000
		Economic Net Worth	925,000

Economic net worth is equal to total economic assets minus total economic liabilities ($2,675,000 – $1,750,000 = $925,000).

2 A is correct. The Laws' equity portfolio is heavily concentrated in WS stock (80% of the equity portfolio), and both Laws work at WS. Should WS encounter difficult economic circumstances, the investment value of WS stock and the Laws' human capital are both likely to be adversely affected. Thus, their investment in WS should be reviewed and their equity portfolio diversified further.

3 C is correct. In order to effectively specify asset classes for the purpose of asset allocation, assets within an asset class should be relatively homogeneous and have similar attributes. The previous adviser's specification of the debt asset class includes global investment-grade corporate bonds and real estate. This definition results in a non-homogeneous asset class.

4 A is correct. For risk control purposes, an asset class should be diversifying and should not have extremely high expected correlations with other classes. Because the returns to the equity and the derivatives asset classes are noted as being highly correlated, inclusion of both asset classes will result in duplication of risk exposures. Including both asset classes is not diversifying to the asset allocation.

5 B is correct. Raye believes the Laws' previous financial adviser followed an asset allocation approach that resulted in an overlap in risk factors among asset classes. A multifactor risk model approach can be used to address potential risk factor overlaps. Risk factor approaches to asset allocation focus on assigning investments to the investor's desired exposures to specified risk factors. These methods are premised on the observation that asset classes often exhibit some overlaps in sources of risk.

6 A is correct. Portfolio 1 best meets the Laws' education goal for their children. The estimated present value of the Laws' expected education expense is \$275,000. Given that the children will be starting college soon, and the Laws have a very strong desire to achieve this goal, Portfolio 1, which stresses liquidity and stability, is most appropriate to meet the Laws' short-term education goal.

7 B is correct. Portfolio 2 best meets the Laws' goal to fund an endowment for their alma mater in 20 years. In present value terms, the gift is valued at \$500,000, with the Laws desiring a high probability of achieving this goal. Although slightly more conservative than the 75/25 global equity/bond mix, Portfolio 2 has a greater growth emphasis compared with Portfolios 1 and 3. Therefore, Portfolio 2 is best for funding the endowment at their alma mater given the goal's long-term horizon and the Laws' desire for a high probability of achieving it.

8 B is correct. Using the cost–benefit approach, higher transaction costs for an asset class imply wider rebalancing ranges. Raye's recommendation for a wider rebalancing range for global equities is consistent with the presence of higher transaction costs for global equities.

READING

19

Principles of Asset Allocation

by Jean L.P. Brunel, CFA, Thomas M. Idzorek, CFA, and John M. Mulvey, PhD

Jean L.P. Brunel, CFA, is at Brunel Associates LLC (USA). Thomas M. Idzorek, CFA, is at Morningstar (USA). John M. Mulvey, PhD, is at the Bendheim Center for Finance at Princeton University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe and critique the use of mean–variance optimization in asset allocation;
☐	**b.** recommend and justify an asset allocation using mean–variance optimization;
☐	**c.** interpret and critique an asset allocation in relation to an investor's economic balance sheet;
☐	**d.** discuss asset class liquidity considerations in asset allocation;
☐	**e.** explain absolute and relative risk budgets and their use in determining and implementing an asset allocation;
☐	**f.** describe how client needs and preferences regarding investment risks can be incorporated into asset allocation;
☐	**g.** discuss the use of Monte Carlo simulation and scenario analysis to evaluate the robustness of an asset allocation;
☐	**h.** describe the use of investment factors in constructing and analyzing an asset allocation;
☐	**i.** recommend and justify an asset allocation based on the global market portfolio;
☐	**j.** describe and evaluate characteristics of liabilities that are relevant to asset allocation;
☐	**k.** discuss approaches to liability-relative asset allocation;
☐	**l.** recommend and justify a liability-relative asset allocation;

(continued)

© 2017 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	The candidate should be able to:
☐	**m.** recommend and justify an asset allocation using a goals-based approach;
☐	**n.** describe and critique heuristic and other approaches to asset allocation;
☐	**o.** discuss factors affecting rebalancing policy.

1 INTRODUCTION

Determining a strategic asset allocation is arguably the most important aspect of the investment process. This reading builds on the "Introduction to Asset Allocation" reading and focuses on several of the primary frameworks for developing an asset allocation, including asset-only mean–variance optimization, various liability-relative asset allocation techniques, and goals-based investing. Additionally, it touches on various other asset allocation techniques used by practitioners, as well as important related topics, such as rebalancing.

The process of creating a diversified, multi-asset class portfolio typically involves two separate steps. The first step is the asset allocation decision, which can refer to both the process and the result of determining long-term (strategic) exposures to the available asset classes (or risk factors) that make up the investor's opportunity set. Asset allocation is the first and primary step in translating the client's circumstances, objectives, and constraints into an appropriate portfolio (or, for some approaches, multiple portfolios) for achieving the client's goals within the client's tolerance for risk. The second step in creating a diversified, multi-asset-class portfolio involves implementation decisions that determine the specific investments (individual securities, pooled investment vehicles, and separate accounts) that will be used to implement the targeted allocations.

Although it is possible to carry out the asset allocation process and the implementation process simultaneously, in practice, these two steps are often separated for two reasons. First, the frameworks for simultaneously determining an asset allocation and its implementation are often complex. Second, in practice, many investors prefer to revisit their strategic asset allocation policy somewhat infrequently (e.g., annually or less frequently) in a dedicated asset allocation study, while most of these same investors prefer to revisit/monitor implementation vehicles (actual investments) far more frequently (e.g., monthly or quarterly).

Section 2 covers the traditional mean–variance optimization (MVO) approach to asset allocation. We apply this approach in what is referred to as an "asset-only" setting, in which the goal is to create the most efficient mixes of asset classes in the absence of any liabilities. We highlight key criticisms of mean–variance optimization and methods used to address them. This section also covers risk budgeting in relation to asset allocation, factor-based asset allocation, and asset allocation with illiquid assets. The observation that almost all portfolios exist to help pay for what can be characterized as a "liability" leads to the next subject.

Section 3 introduces liability-relative asset allocation—including a straightforward extension of mean–variance optimization known as surplus optimization. Surplus optimization is an economic balance sheet approach extended to the liability side of the balance sheet that finds the most efficient asset class mixes in the presence of liabilities. Liability-relative optimization is simultaneously concerned with the return of the assets, the change in value of the liabilities, and how assets and liabilities interact to determine the overall value or health of the total portfolio.

Section 4 covers an increasingly popular approach to asset allocation called goals-based asset allocation. Conceptually, goals-based approaches are similar to liability-relative asset allocation in viewing risk in relation to specific needs or objectives associated with different time horizons and degrees of urgency.

Section 5 introduces some informal (heuristic) ways that asset allocations have been determined and other approaches to asset allocation that emphasize specific objectives.

Section 6 addresses the factors affecting choices that are made in developing specific policies relating to rebalancing to the strategic asset allocation. Factors discussed include transaction costs, correlations, volatility, and risk aversion.[1]

Section 7 summarizes important points and concludes the reading.

DEVELOPING ASSET-ONLY ASSET ALLOCATIONS

2

In this section, we discuss several of the primary techniques and considerations involved in developing strategic asset allocations, leaving the issue of considering the liabilities to Section 3 and the issue of tailoring the strategic asset allocation to meet specific goals to Section 4.

We start by introducing mean–variance optimization, beginning with unconstrained optimization, prior to moving on to the more common mean–variance optimization problem in which the weights, in addition to summing to 1, are constrained to be positive (no shorting allowed). We present a detailed example, along with several variations, highlighting some of the important considerations in this approach. We also identify several criticisms of mean–variance optimization and the major ways these criticisms have been addressed in practice.

2.1 Mean–Variance Optimization: Overview

Mean–variance optimization (MVO), as introduced by Markowitz (1952, 1959), is perhaps the most common approach used in practice to develop and set asset allocation policy. Widely used on its own, MVO is also often the basis for more sophisticated approaches that overcome some of the limitations or weaknesses of MVO.

Markowitz recognized that whenever the returns of two assets are not perfectly correlated, the assets can be combined to form a portfolio whose risk (as measured by standard deviation or variance) is less than the weighted-average risk of the assets themselves. An additional and equally important observation is that as one adds assets to the portfolio, one should focus not on the individual risk characteristics of the additional assets but rather on those assets' effect on the risk characteristics of the entire portfolio. Mean–variance optimization provides us with a framework for determining how much to allocate to each asset in order to maximize the *expected* return of the portfolio for an *expected* level of risk. In this sense, mean–variance optimization is a risk-budgeting tool that helps investors to spend their risk budget—the amount of risk they are willing to assume—wisely. We emphasize the word "expected" because the inputs to mean–variance optimization are necessarily forward-looking estimates, and the resulting portfolios reflect the quality of the inputs.

1 In this reading, "volatility" is often used synonymously with "standard deviation."

Mean–variance optimization requires three sets of inputs: returns, risks (standard deviations), and pair-wise correlations for the assets in the opportunity set. The objective function is often expressed as follows:

$$U_m = E(R_m) - 0.005\lambda\sigma_m^2 \tag{1}$$

where

U_m = the investor's utility for asset mix (allocation) m
R_m = the return for asset mix m
λ = the investor's risk aversion coefficient
σ_m^2 = the expected variance of return for asset mix m

The risk aversion coefficient (λ) characterizes the investor's risk–return trade-off; in this context, it is the rate at which an investor will forgo expected return for less variance. The value of 0.005 in Equation 1 is based on the assumption that $E(R_m)$ and σ_m are expressed as percentages rather than as decimals. If those quantities were expressed as decimals, the 0.005 would change to 0.5. For example, if $E(R_m)$ = 0.10, λ = 2, and σ = 0.20 (variance is 0.04), then U_m is 0.06, or 6% [= 0.10 – 0.5(2)(0.04)]. In this case, U_m can be interpreted as a certainty-equivalent return—that is, the utility value of the risky return offered by the asset mix, stated in terms of the risk-free return that the investor would value equally. In Equation 1, 0.005 merely scales the second term appropriately.

In words, the objective function says that the value of an asset mix for an investor is equal to the expected return of the asset mix minus a penalty that is equal to one-half of the expected variance of the asset mix scaled by the investor's risk aversion coefficient. Optimization involves selecting the asset mix with the highest such value (certainty equivalent). Smaller risk aversion coefficients result in relatively small penalties for risk, leading to aggressive asset mixes. Conversely, larger risk aversion coefficients result in relatively large penalties for risk, leading to conservative asset mixes. A value of λ = 0 corresponds to a risk-neutral investor because it implies indifference to volatility. Most investors' risk aversion is consistent with λ between 1 and 10.[2] Empirically, λ = 4 can be taken to represent a moderately risk-averse investor, although the specific value is sensitive to the opportunity set in question and to market volatility.

In the absence of constraints, there is a closed-form solution that calculates, for a given set of inputs, the single set of weights (allocation) to the assets in the opportunity set that maximizes the investor's utility. Typically, this single set of weights is relatively extreme, with very large long and short positions in each asset class. Except in the special case in which the expected returns are derived using the reverse-optimization process of Sharpe (1974), the expected-utility-maximizing weights will not add up to 100%. We elaborate on reverse optimization in Section 2.4.1.

In most real-world applications, asset allocation weights must add up to 100%, reflecting a fully invested, non-leveraged portfolio. From an optimization perspective, when seeking the asset allocation weights that maximize the investor's utility, one must constrain the asset allocation weights to sum to 1 (100%). This constraint that weights sum to 100% is referred to as the "budget constraint" or "unity constraint." The inclusion of this constraint, or any other constraint, moves us from a problem that has a closed-form solution to a problem that must be solved numerically using optimization techniques.

In contrast to the single solution (single set of weights) that is often associated with unconstrained optimization (one could create an efficient frontier using unconstrained weights, but it is seldom done in practice), Markowitz's mean–variance optimization paradigm is most often identified with an efficient frontier that plots all

2 See Ang (2014, p. 44).

potential efficient asset mixes subject to some common constraints. In addition to a typical budget constraint that the weights must sum to 1 (100% in percentage terms), the next most common constraint allows only positive weights or allocations (i.e., no negative or short positions).

Efficient asset mixes are combinations of the assets in the opportunity set that maximize expected return per unit of expected risk or, alternatively (and equivalently), minimize expected risk for a given level of expected return. To find all possible efficient mixes that collectively form the efficient frontier, *conceptually* the optimizer iterates through all the possible values of the risk aversion coefficient (λ) and for each value finds the combination of assets that maximizes expected utility. We have used the word *conceptually* because there are different techniques for carrying out the optimization that may vary slightly from our description, even though the solution (efficient frontier and efficient mixes) is the same. The efficient mix at the far left of the frontier with the lowest risk is referred to as the global minimum variance portfolio, while the portfolio at the far right of the frontier is the maximum expected return portfolio. In the absence of constraints beyond the budget and non-negativity constraints, the maximum expected return portfolio consists of a 100% allocation to the single asset with the highest expected return (which is not necessarily the asset with the highest level of risk).

Risk Aversion

Unfortunately, it is extremely difficult to precisely estimate a given investor's risk aversion coefficient (λ). Best practices suggest that when estimating risk aversion (or, conversely, risk tolerance), one should examine both the investor's *preference* for risk (willingness to take risk) and the investor's *capacity* for taking risk. Risk preference is a subjective measure and typically focuses on how an investor feels about and potentially reacts to the ups and downs of portfolio value. The level of return an investor hopes to earn can influence the investor's willingness to take risk, but investors must be realistic when setting such objectives. Risk capacity is an objective measure of the investor's ability to tolerate portfolio losses and the potential decrease in future consumption associated with those losses.[3] The psychometric literature has developed validated questionnaires, such as that of Grable and Joo (2004), to approximately locate an investor's risk preference, although this result then needs to be blended with risk capacity to determine risk tolerance. For individuals, risk capacity is affected by factors such as net worth, income, the size of an emergency fund in relation to consumption needs, and the rate at which the individual saves out of gross income, according to the practice of financial planners noted in Grable (2008).

With this guidance in mind, we move forward with a relatively global opportunity set, in this case defined from the point of view of an investor from the United Kingdom with an approximate 10-year time horizon. The analysis is carried out in British pounds (GBP), and none of the currency exposure is hedged. Exhibit 1 identifies 12 asset classes within the universe of available investments and a set of plausible forward-looking capital market assumptions: expected returns, standard deviations,

3 *Risk preference* and *risk capacity* are sometimes referred to as the willingness and the ability to take risk, respectively.

and correlations. The reading on capital market expectations covers how such inputs may be developed.[4] In the exhibit, three significant digits at most are shown, but the subsequent analysis is based on full precision.

Time Horizon

Mean–variance optimization is a "single-period" framework in which the single period could be a week, a month, a year, or some other time period. When working in a "strategic" setting, many practitioners typically find it most intuitive to work with annual capital market assumptions, even though the investment time horizon could be considerably longer (e.g., 10 years). If the strategic asset allocation will not be re-evaluated within a long time frame, capital market assumptions should reflect the average annual distributions of returns expected over the entire investment time horizon. In most cases, investors revisit the strategic asset allocation decision more frequently, such as annually or every three years, rerunning the analysis and making adjustments to the asset allocation; thus, the annual capital market assumption often reflects the expectations associated with the evaluation horizon (e.g., one year or three years).

Exhibit 1 Hypothetical UK-Based Investor's Opportunity Set with Expected Returns, Standard Deviations, and Correlations

Panel A: Expected Returns and Standard Deviations

Asset Class	Expected Return (%)	Standard Deviation (%)
UK large cap	6.6	14.8
UK mid cap	6.9	16.7
UK small cap	7.1	19.6
US equities	7.8	15.7
Europe ex UK equities	8.6	19.6
Asia Pacific ex Japan equities	8.5	20.9
Japan equities	6.4	15.2
Emerging market equities	9.0	23.0
Global REITs	9.0	22.5
Global ex UK bonds	4.0	10.4
UK bonds	2.9	6.1
Cash	2.5	0.7

4 The standard deviations and correlations in Exhibit 1 are based on historical numbers, while expected returns come from reverse optimization (described later).

Exhibit 1 (Continued)

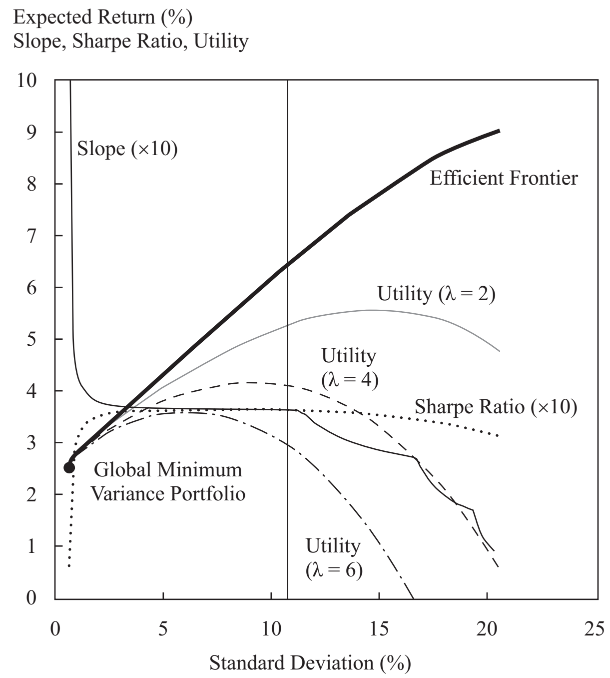
Panel B: Correlations

	UK Large Cap	UK Mid Cap	UK Small Cap	US Equities	Europe ex UK Equities	Asia Pacific ex Japan Equities	Japan Equities	Emerging Market Equities	Global REITs	Global ex UK Bonds	UK Bonds	Cash
UK large cap	1.00	0.86	0.79	0.76	0.88	0.82	0.55	0.78	0.64	–0.12	–0.12	–0.06
UK mid cap	0.86	1.00	0.95	0.76	0.84	0.75	0.51	0.74	0.67	–0.16	–0.10	–0.17
UK small cap	0.79	0.95	1.00	0.67	0.79	0.70	0.49	0.71	0.61	–0.22	–0.15	–0.17
US equities	0.76	0.76	0.67	1.00	0.81	0.72	0.62	0.69	0.77	0.14	0.00	–0.12
Europe ex UK equities	0.88	0.84	0.79	0.81	1.00	0.82	0.60	0.80	0.72	0.04	–0.04	–0.03
Asia Pacific ex Japan equities	0.82	0.75	0.70	0.72	0.82	1.00	0.54	0.94	0.67	0.00	–0.02	0.02
Japan equities	0.55	0.51	0.49	0.62	0.60	0.54	1.00	0.56	0.52	0.18	0.07	–0.01
Emerging market equities	0.78	0.74	0.71	0.69	0.80	0.94	0.56	1.00	0.62	–0.02	–0.03	0.04
Global REITs	0.64	0.67	0.61	0.77	0.72	0.67	0.52	0.62	1.00	0.16	0.18	–0.15
Global ex UK bonds	–0.12	–0.16	–0.22	0.14	0.04	0.00	0.18	–0.02	0.16	1.00	0.62	0.24
UK bonds	–0.12	–0.10	–0.15	0.00	–0.04	–0.02	0.07	–0.03	0.18	0.62	1.00	0.07
Cash	–0.06	–0.17	–0.17	–0.12	–0.03	0.02	–0.01	0.04	–0.15	0.24	0.07	1.00

The classification of asset classes in the universe of available investments may vary according to local practices. For example, in the United States and some other larger markets, it is common to classify equities by market capitalization, whereas the practice of classifying equities by valuation ("growth" versus "value") is less common outside of the United States. Similarly, with regard to fixed income, some asset allocators may classify bonds based on various attributes—nominal versus inflation linked, corporate versus government issued, investment grade versus non-investment grade (high yield)—and/or by maturity/duration (short, intermediate, and long). By means of the non-negativity constraint and using a reverse-optimization procedure (to be explained later) based on asset class market values to generate expected return estimates, we control the typically high sensitivity of the composition of efficient portfolios to expected return estimates (discussed further in Section 2.4). Without such precautions, we would often find that efficient portfolios are highly concentrated in a subset of the available asset classes.

Running this set of capital market assumptions through a mean–variance optimizer with the traditional non-negativity and unity constraints produces the efficient frontier depicted in Exhibit 2. We have augmented this efficient frontier with some non-traditional information that will assist with the understanding of some key concepts related to the efficient frontier. A risk-free return of 2.5% is used in calculating the reserve-optimized expected returns as well as the Sharpe ratios in Exhibit 2.

Exhibit 2 Efficient Frontier—Base Case



The slope of the efficient frontier is greatest at the far left of the efficient frontier, at the point representing the global minimum variance portfolio. Slope represents the rate at which expected return increases per increase in risk. As one moves to the right, in the direction of increasing risk, the slope decreases; it is lowest at the point representing the maximum return portfolio. Thus, as one moves from left to right along the efficient frontier, the investor takes on larger and larger amounts of risk for smaller and smaller increases in expected return. The "kinks" in the line representing the slope (times 10) of the efficient frontier correspond to portfolios (known as corner portfolios) in which an asset either enters or leaves the efficient mix.

For most investors, at the far left of the efficient frontier, the increases in expected return associated with small increases in expected risk represent a desirable trade-off. The risk aversion coefficient identifies the specific point on the efficient frontier at which the investor refuses to take on additional risk because he or she feels the associated increase in expected return is not high enough to compensate for the increase in risk. Of course, each investor makes this trade-off differently.

For this particular efficient frontier, the three expected utility curves plot the solution to Equation 1 for three different risk aversion coefficients: 2.0, 4.0, and 6.0, respectively.[5] For a given risk aversion coefficient, the appropriate efficient mix from the efficient frontier is simply the mix in which expected utility is highest (i.e.,

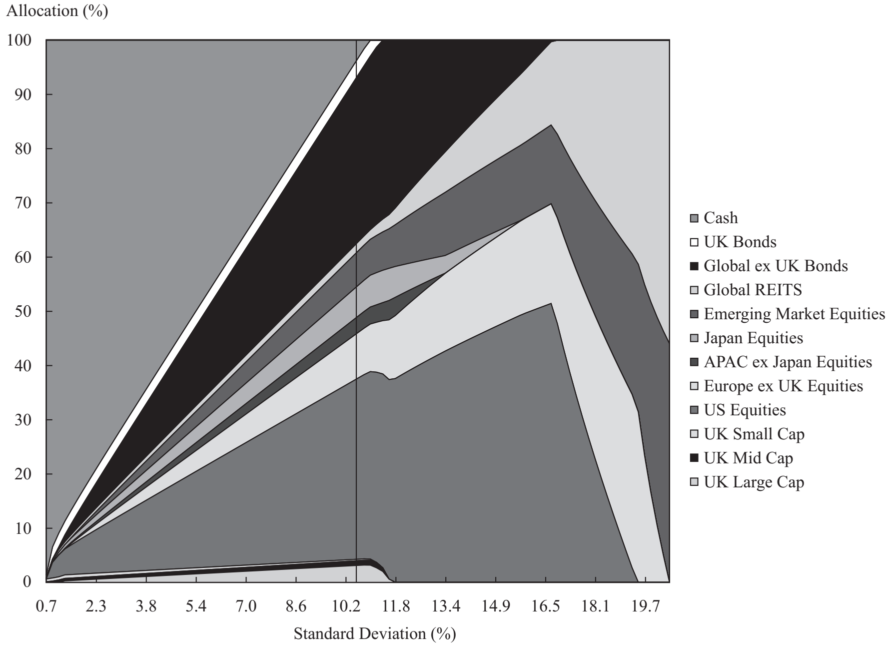
5 Numbers have been rounded to increase readability.

maximized). As illustrated in Exhibit 2, a lower risk aversion coefficient leads to a riskier (higher) point on the efficient frontier, while a higher risk aversion coefficient leads to a more conservative (lower) point on the efficient frontier.

The vertical line (at volatility of 10.88%) identifies the asset mix with the highest Sharpe ratio; it intersects the Sharpe ratio line at a value of 3.7 (an unscaled value of 0.37). This portfolio is also represented by the intersection of the slope line and the Sharpe ratio line.

Exhibit 3 is an efficient frontier asset allocation area graph. Each vertical cross section identifies the asset allocation at a point along the efficient frontier; thus, the vertical cross section at the far left, with nearly 100% cash, is the asset allocation of the minimum variance portfolio, and the vertical cross section at the far right, with 45% in emerging markets and 55% in global REITs, is the optimal asset allocation for a standard deviation of 20.5%, the highest level of portfolio volatility shown. In this example, cash is treated as a risky asset; although its return volatility is very low, because it is less than perfectly correlated with the other asset classes, mixing it with small amounts of other asset classes reduces risk further. The vertical line identifies the asset mix with the highest Sharpe ratio and corresponds to the similar line shown on the original efficient frontier graph (Exhibit 2). The asset allocation mixes are well diversified for most of the first half of the efficient frontier, and in fact, for a large portion of the efficient frontier, all 12 asset classes in our opportunity set receive a positive allocation.[6]

Exhibit 3 Efficient Frontier Asset Allocation Area Graph—Base Case



6 Studying Exhibit 3 closely, one notices distinct regime shifts where the rate at which allocations are made to asset classes changes so that a line segment with a different slope begins. These regime shifts occur at what are called *corner portfolios*. The efficient mixes between two adjacent corner portfolios are simply linear combinations of those portfolios. The efficient frontier asset allocation area graph helps to clarify this result. More formally, corner portfolios are points on the efficient frontier at which an asset class either enters or leaves the efficient mix or a constraint either becomes binding or is no longer binding.

The investment characteristics of potential asset mixes based on mean–variance theory are often further investigated by means of Monte Carlo simulation, as discussed in Section 2.2. Several observations from theory and practice are relevant to narrowing the choices.

Equation 1 indicates that the basic approach to asset allocation involves estimating the investor's risk aversion parameter and then finding the efficient mix that maximizes expected utility. When the risk aversion coefficient has not been estimated, the investor may be able to identify the maximum tolerable level of portfolio return volatility. If that level is 10% per annum, for example, only the part of the efficient frontier associated with volatility less than or equal to 10% is relevant. This approach is justifiable because for a given efficient frontier, every value of the risk aversion coefficient can be associated with a value of volatility that identifies the best point on the efficient frontier for the investor; the investor may also have experience with thinking in terms of volatility. In addition, when the investor has a numerical return objective, he or she can further narrow the range of potential efficient mixes by identifying the efficient portfolios expected to meet that return objective. For example, if the return objective is 5%, one can select the asset allocation with a 5% expected return.

Example 1 illustrates the use of Equation 1 and shows the adaptability of MVO by introducing the choice problem in the context of an investor who also has a shortfall risk concern.

EXAMPLE 1

Mean–Variance-Efficient Portfolio Choice 1

An investment adviser is counseling Aimée Goddard, a client who recently inherited €1,200,000 and who has above-average risk tolerance ($\lambda = 2$). Because Goddard is young and one of her goals is to fund a comfortable retirement, she wants to earn returns that will outpace inflation in the long term. Goddard expects to liquidate €60,000 of the inherited portfolio in 12 months to fund the down payment on a house. She states that it is important for her to be able to take out the €60,000 without invading the initial capital of €1,200,000. Exhibit 4 shows three alternative strategic asset allocations.

Exhibit 4 Strategic Asset Allocation Choices for Goddard

	Investor's Forecasts	
Asset Allocation	**Expected Return**	**Standard Deviation of Return**
A	10.00%	20%
B	7.00	10
C	5.25	5

1 Based only on Goddard's risk-adjusted expected returns for the asset allocations, which asset allocation would she prefer?

2 Recommend and justify a strategic asset allocation for Goddard.

Note: In addressing 2, calculate the minimum return, R_L, that needs to be achieved to meet the investor's objective not to invade capital, using the expression ratio $[E(R_P) - R_L]/\sigma_P$, which reflects the probability of exceeding the minimum given a normal return distribution assumption in a safety-first approach.[7]

Solution to 1:

Using Equation 1,

$$\begin{aligned} U_m &= E(R_m) - 0.005\lambda\sigma_m^2 \\ &= E(R_m) - 0.005(2)\sigma_m^2 \\ &= E(R_m) - 0.01\sigma_m^2 \end{aligned}$$

So Goddard's utility for Asset Allocations A, B, and C are as follows:

$$\begin{aligned} U_A &= E(R_A) - 0.01\,\sigma_A^2 \\ &= 10.0\% - 0.01(20\%)^2 \\ &= 10.0\% - 4.0\% \\ &= 6.0\% \\ U_B &= E(R_B) - 0.01\,\sigma_B^2 \\ &= 7.0\% - 0.01(10\%)^2 \\ &= 7.0\% - 1.0\% \\ &= 6.0\% \\ U_C &= E(R_C) - 0.01\,\sigma_C^2 \\ &= 5.25\% - 0.01(5\%)^2 \\ &= 5.25\% - 0.25\% \\ &= 5.0\% \end{aligned}$$

Goddard would be indifferent between A and B based only on their common perceived certainty-equivalent return of 6%.

Solution to 2:

Because €60,000/€1,200,000 is 5.0%, for any return less than 5.0%, Goddard will need to invade principal when she liquidates €60,000. So 5% is a threshold return level.

To decide which of the three allocations is best for Goddard, we calculate the ratio $[E(R_P) - R_L]/\sigma_P$:

Allocation A	(10% – 5%)/20% = 0.25
Allocation B	(7% – 5%)/10% = 0.20
Allocation C	(5.25% – 5%)/5% = 0.05

Both Allocations A and B have the same expected utility, but Allocation A has a higher probability of meeting the threshold 5% return than Allocation B. Therefore, A would be the recommended strategic asset allocation.

There are several different approaches to determining an allocation to cash and cash equivalents, such as government bills. Exhibit 1 included cash among the assets for which we conducted an optimization to trace out an efficient frontier. The return to cash over a short time horizon is essentially certain in nominal terms. One approach

7 See the Level I CFA Program reading "Common Probability Distributions" for coverage of Roy's safety-first criterion.

to asset allocation separates out cash and cash equivalents as a (nominally) risk-free asset and calculates an efficient frontier of risky assets. Alternatively, a ray from the risk-free rate (a point on the return axis) tangent to the risky-asset efficient frontier (with cash excluded) then defines a linear efficient frontier. The efficient frontier then consists of combinations of the risk-free asset with the tangency portfolio (which has the highest Sharpe ratio among portfolios on the risky-asset efficient frontier).

A number of standard finance models (including Tobin two-fund separation) adopt this treatment of cash. According to two-fund separation, if investors can borrow or lend at the risk-free rate, they will choose the tangency portfolio for the risky-asset holdings and borrow at the risk-free rate to leverage the position in that portfolio to achieve a higher expected return, or they will split money between the tangency portfolio and the risk-free asset to reach a position with lower risk and lower expected return than that represented by the tangency portfolio. Since over horizons that are longer than the maturity of a money market instrument, the return earned would not be known, another approach that is well established in practice and reflected in Exhibit 1 is to include cash in the optimization. The amount of cash indicated by an optimization may be adjusted in light of short-term liquidity needs; for example, some financial advisers advocate that individuals hold an amount of cash equivalent to six months of expenses. All of these approaches are reasonable alternatives in practice.

Although we will treat cash as a risky asset in the following discussions, in Example 2, we stop to show the application of the alternative approach based on distinguishing a risk-free asset.

EXAMPLE 2

A Strategic Asset Allocation Based on Distinguishing a Nominal Risk-Free Asset

The Caflandia Foundation for the Fine Arts (CFFA) is a hypothetical charitable organization established to provide funding to Caflandia museums for their art acquisition programs.

CFFA's overall investment objective is to maintain its portfolio's real purchasing power after distributions. CFFA targets a 4% annual distribution of assets. CFFA has the following current specific investment policies.

Return objective

CFFA's assets shall be invested with the objective of earning an average nominal 6.5% annual return. This level reflects a spending rate of 4%, an expected inflation rate of 2%, and a 40 bp cost of earning investment returns. The calculation is $(1.04)(1.02)(1.004) - 1 = 0.065$, or 6.5%.

Risk considerations

CFFA's assets shall be invested to minimize the level of standard deviation of return subject to satisfying the expected return objective.

The investment office of CFFA distinguishes a nominally risk-free asset. As of the date of the optimization, the risk-free rate is determined to be 2.2%.

Exhibit 5 gives key outputs from a mean–variance optimization in which asset class weights are constrained to be non-negative.

Exhibit 5 Corner Portfolios Defining the Risky-Asset Efficient Frontier

Portfolio Number	Expected Nominal Returns	Standard Deviation	Sharpe Ratio
1	9.50%	18.00%	0.406
2	8.90	15.98	0.419
3	8.61	15.20	0.422
4	7.24	11.65	0.433
5	5.61	7.89	0.432
6	5.49	7.65	0.430
7	3.61	5.39	0.262

The portfolios shown are corner portfolios (see footnote 6), which as a group define the risky-asset efficient frontier in the sense that any portfolio on the frontier is a combination of the two corner portfolios that bracket it in terms of expected return.

Based only on the facts given, determine the most appropriate strategic asset allocation for CFFA given its stated investment policies.

Solution:

An 85%/15% combination of Portfolio 4 and the risk-free asset is the most appropriate asset allocation. This combination has the required 6.5% expected return with the minimum level of risk. Stated another way, this combination defines the efficient portfolio at a 6.5% level of expected return based on the linear efficient frontier created by the introduction of a risk-free asset.

Note that Portfolio 4 has the highest Sharpe ratio and is the tangency portfolio. With an expected return of 7.24%, it can be combined with the risk-free asset, with a return of 2.2%, to achieve an expected return of 6.5%:

$$6.50 = 7.24w + 2.2(1 - w)$$
$$w = 0.853$$

Placing about 85% of assets in Portfolio 4 and 15% in the risk-free asset achieves an efficient portfolio with expected return of 6.4 with a volatility of 0.853(11.65) = 9.94%. (The risk-free asset has no return volatility by assumption and, also by assumption, zero correlation with any risky portfolio return.) This portfolio lies on a linear efficient frontier formed by a ray from the risk-free rate to the tangency portfolio and can be shown to have the same Sharpe ratio as the tangency portfolio, 0.433. The combination of Portfolio 4 with Portfolio 5 to achieve a 6.5% expected return would have a lower Sharpe ratio and would not lie on the efficient frontier.

Asset allocation decisions have traditionally been made considering only the investor's investment portfolio (and financial liabilities) and not the total picture that includes human capital and other non-traded assets (and liabilities), which are missing in a traditional balance sheet. Taking such extended assets and liabilities into account can lead to improved asset allocation decisions, however.

Depending on the nature of an individual's career, human capital can provide relatively stable cash flows similar to bond payments. At the other extreme, the cash flows from human capital can be much more volatile and uncertain, reflecting a lumpy, commission-based pay structure or perhaps a career in a seasonal business. For many individuals working in stable job markets, the cash flows associated with their human capital are somewhat like those of an inflation-linked bond, relatively consistent and tending to increase with inflation. If human capital is a relatively large component of the individual's total economic worth, accounting for this type of hidden asset in an asset allocation setting is extremely important and would presumably increase the individual's capacity to take on risk.

Let us look at a hypothetical example. Emma Beel is a 45-year-old tenured university professor in London. Capital market assumptions are as before (see Exhibit 1). Beel has GBP 1,500,000 in liquid financial assets, largely due to a best-selling book. Her employment as a tenured university professor is viewed as very secure and produces cash flows that resemble those of a very large, inflation-adjusted, long-duration bond portfolio. The net present value of her human capital is estimated at GBP 500,000. Beel inherited her grandmother's home on the edge of the city, valued at GBP 750,000. The results of a risk tolerance questionnaire that considers both risk preference and risk capacity suggest that Beel should have an asset allocation involving moderate risk. Furthermore, given our earlier assumption that the collective market risk aversion coefficient is 4.0, we assume that the risk aversion coefficient of a moderately risk-averse investor is approximately 4.0, from a total wealth perspective.

To account for Beel's human capital and residential real estate, these two asset classes were modeled and added to the optimization. Beel's human capital of GBP 500,000 was modeled as 70% UK long-duration inflation-linked bonds, 15% UK corporate bonds, and 15% UK equities.[8] Residential real estate was modeled based on a de-smoothed residential property index for London. (We will leave the complexities of modeling liabilities to Section 3.) Beel's assets include those shown in Exhibit 6.

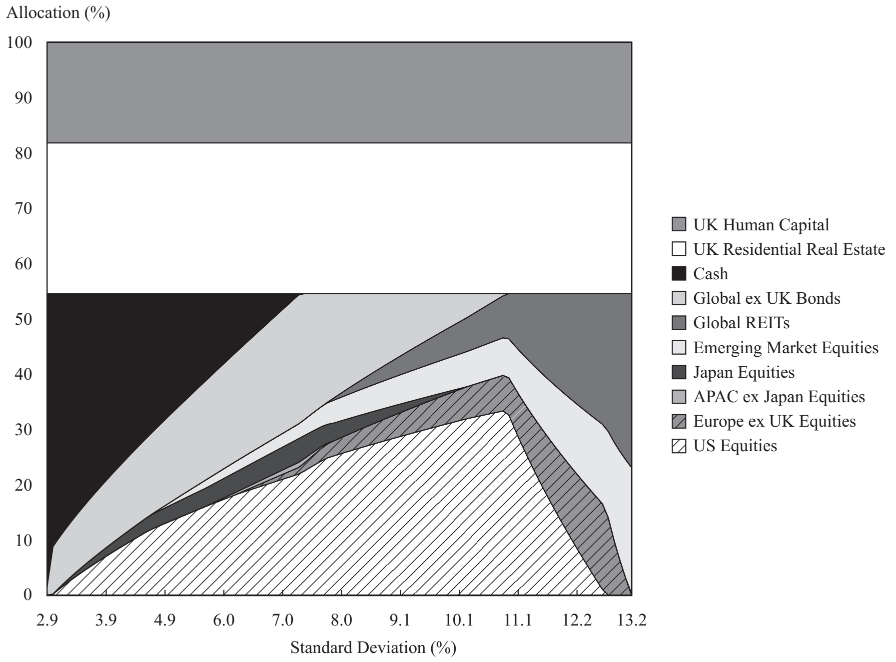
Exhibit 6 Emma Beel's Assets

Asset	Value (GBP)	Percentage
Liquid financial assets	1,500,000	54.55
UK residential real estate	750,000	27.27
Human capital	500,000	18.18
	2,750,000	100

Beel's UK residential real estate (representing the London house) and human capital were added to the optimization opportunity set. Additionally, working under the assumption that Beel's house and human capital are non-tradable assets, the optimizer was forced to allocate 27.27% or more to UK residential real estate and 18.18% to human capital and then determined the optimal asset allocation based on a risk aversion coefficient of 4. Beel's expected utility is maximized by an efficient asset allocation with volatility of approximately 8.2%. Exhibit 7 displays the resulting asset allocation area graph.

8 These weights were used to create the return composite representing Beel's human capital that was used in the asset allocation optimization.

Exhibit 7 Efficient Frontier Asset Allocation Area Graph—Balance Sheet Approach



Looking past the constrained allocations to human capital and UK residential real estate, the remaining allocations associated with Beel's liquid financial assets do not include UK equities, UK fixed income, or global REITs. Each of these three asset classes is relatively highly correlated with either UK residential real estate or UK human capital.[9]

2.2 Monte Carlo Simulation

Monte Carlo simulation complements MVO by addressing the limitations of MVO as a single-period framework. Additionally, in the case in which the investor's risk tolerance is either unknown or in need of further validation, Monte Carlo simulation can help paint a realistic picture of potential future outcomes, including the likelihood of meeting various goals, the distribution of the portfolio's expected value through time, and potential maximum drawdowns. Simulation also provides a tool for investigating the effects of trading/rebalancing costs and taxes and the interaction of evolving financial markets with asset allocation. It is important to note that not all Monte Carlo simulation tools are the same: They vary significantly in their ability to model non-normal multivariate returns, serial and cross-correlations, tax rates, distribution requirements, an evolving asset allocation schedule (target-date glide path), non-traditional investments (e.g., annuities), and human capital (based on age, geography, education, and/or occupation).

Using Monte Carlo simulation, an investment adviser can effectively grapple with a range of practical issues that are difficult or impossible to formulate analytically. Consider rebalancing to a strategic asset allocation for a taxable investor. We can readily calculate the impact of taxes during a single time period. Also, in a single-period

9 For additional information on applying a total balance sheet approach, see, for example, Blanchett and Straehl (2015) or Rudd and Siegel (2013).

setting, as assumed by MVO, rebalancing is irrelevant. In the multi-period world of most investment problems, however, the portfolio will predictably be rebalanced, triggering the realization of capital gains and losses. Given a specific rebalancing rule, different strategic asset allocations will result in different patterns of tax payments (and different transaction costs too). Formulating the multi-period problem mathematically would be a daunting challenge. We could more easily incorporate the interaction between rebalancing and taxes in a Monte Carlo simulation.

We will examine a simple multi-period problem to illustrate the use of Monte Carlo simulation, evaluating the range of outcomes for wealth that may result from a strategic asset allocation (and not incorporating taxes).

The value of wealth at the terminal point of an investor's time horizon is a possible criterion for choosing among asset allocations. Future wealth incorporates the interaction of risk and return. The need for Monte Carlo simulation in evaluating an asset allocation depends on whether there are cash flows into or out of the portfolio over time. For a given asset allocation with no cash flows, the sequence of returns is irrelevant; ending wealth will be path independent (unaffected by the sequence or path of returns through time). With cash flows, the sequence is also irrelevant if simulated returns are independent, identically distributed random variables. We could find expected terminal wealth and percentiles of terminal wealth analytically.[10] Investors save/deposit money in and spend money out of their portfolios; thus, in the more typical case, terminal wealth is path dependent (the sequence of returns matters) because of the interaction of cash flows and returns. When terminal wealth is path dependent, an analytical approach is not feasible but Monte Carlo simulation is. Example 3 applies Monte Carlo simulation to evaluate the strategic asset allocation of an investor who regularly withdraws from the portfolio.

EXAMPLE 3

Monte Carlo Simulation for a Retirement Portfolio with a Proposed Asset Allocation

Malala Ali, a resident of the hypothetical country of Caflandia, has sought the advice of an investment adviser concerning her retirement portfolio. At the end of 2017, she is 65 years old and holds a portfolio valued at CAF$1 million. Ali would like to withdraw CAF$40,000 a year to supplement the corporate pension she has begun to receive. Given her health and family history, Ali believes she should plan for a retirement lasting 25 years. She is also concerned about passing along a portion of her portfolio to the families of her three children; she hopes that at least the portfolio's current real value can go to them. Consulting with her adviser, Ali has expressed this desire quantitatively: She wants the median value of her bequest to her children to be no less than her portfolio's current value of CAF$1 million in real terms. The median is the 50th percentile outcome. The asset allocation of her retirement portfolio is currently 50/50 Caflandia equities/Caflandia intermediate-term government bonds. Ali and her adviser have decided on the following set of capital market expectations (Exhibit 8):

10 Making a plausible statistical assumption, such as a lognormal distribution, for ending wealth.

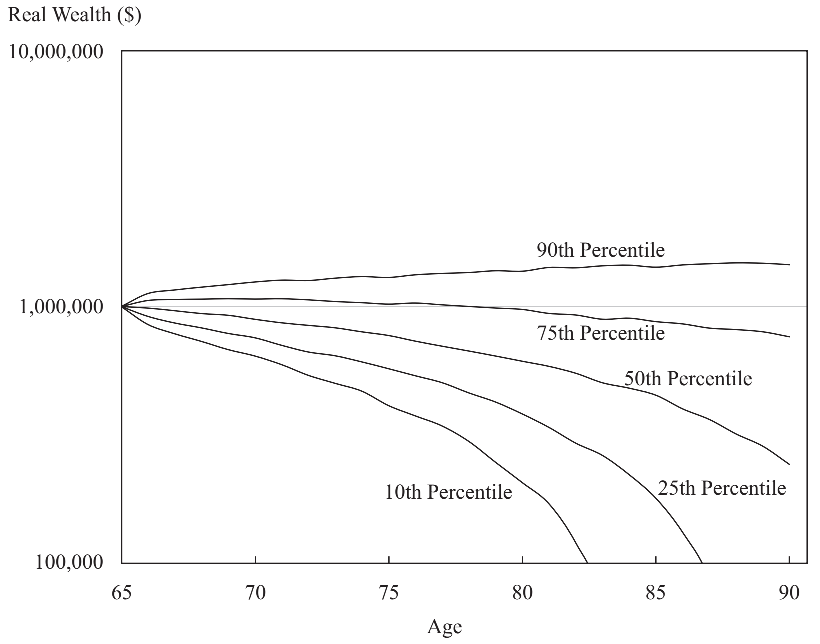
Exhibit 8 Caflandia Capital Market Expectations

	Investor's Forecasts	
Asset Class	**Expected Return**	**Standard Deviation of Return**
Caflandia equities	9.4%	20.4%
Caflandia bonds	5.6%	4.1%
Inflation	2.6%	

The predicted correlation between returns of Caflandia equities and Caflandia intermediate-term government bonds is 0.15.

With the current asset allocation, the expected nominal return on Ali's retirement portfolio is 7.5% with a standard deviation of 11%. Exhibit 9 gives the results of the Monte Carlo simulation.[11] In Exhibit 9, the lowest curve represents, at various ages, levels of real wealth at or below which the 10% of worst real wealth outcomes lie (i.e., the 10th percentile for real wealth); curves above that represent, respectively, 25th, 50th, 75th, and 90th percentiles for real wealth.

Exhibit 9 Monte Carlo Simulation of Ending Real Wealth with Annual Cash Outflows



11 Note that the y-axis in this exhibit is specified using a logarithmic scale. The quantity CAF$1 million is the same distance from CAF$100,000 as CAF$10 million is from CAF$1 million because CAF$1 million is 10 times CAF$100,000, just as CAF$10 million is 10 times CAF$1 million. CAF$100,000 is 10^5, and CAF $1 million is 10^6. In Exhibit 9, a distance halfway between the CAF$100,000 and CAF$1 million hatch marks is $10^{5.5}$ = CAF$316,228.

Based on the information given, address the following:

1 Justify the presentation of ending wealth in terms of real rather than nominal wealth in Exhibit 9.
2 Is the current asset allocation expected to satisfy Ali's investment objectives?

Solution to 1:

Ali wants the median real value of her bequest to her children to be "no less than her portfolio's current value of CAF$1 million." We need to state future amounts in terms of today's values (i.e., in real dollars) to assess the purchasing power of those amounts relative to CAF$1 million today. Exhibit 9 thus gives the results of the Monte Carlo simulation in real dollar terms. The median real wealth at age 90 is clearly well below the target ending wealth of real CAF$1 million.

Solution to 2:

From Exhibit 9, we see that the median terminal (at age 90) value of the retirement portfolio in real dollars is less than the stated bequest goal of CAF$1 million. Therefore, the most likely bequest is less than the amount Ali has said she wants. The current asset allocation is not expected to satisfy all her investment objectives. Although one potential lever would be to invest more aggressively, given Ali's age and risk tolerance, this approach seems imprudent. An adviser may need to counsel that the desired size of the bequest may be unrealistic given Ali's desired income to support her expenditures. Ali will likely need to make a relatively tough choice between her living standard (spending less) and her desire to leave a CAF$1 million bequest in real terms. A third alternative would be to delay retirement, which may or may not be feasible.

2.3 Criticisms of Mean–Variance Optimization

With this initial understanding of mean–variance optimization, we can now elaborate on some of the most common criticisms of it. The following criticisms and the ways they have been addressed motivate the balance of the coverage of MVO:

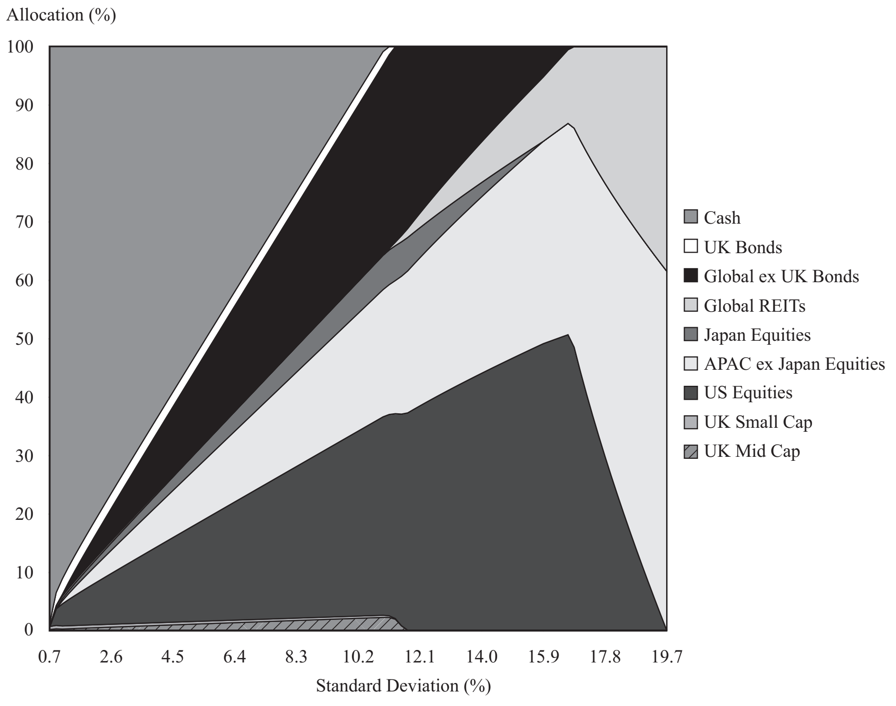
1 The outputs (asset allocations) are highly sensitive to small changes in the inputs.
2 The asset allocations tend to be highly concentrated in a subset of the available asset classes.
3 Many investors are concerned about more than the mean and variance of returns, the focus of MVO.
4 Although the asset allocations may appear diversified across assets, the sources of risk may not be diversified.
5 Most portfolios exist to pay for a liability or consumption series, and MVO allocations are not directly connected to what influences the value of the liability or the consumption series.
6 MVO is a single-period framework that does not take account of trading/rebalancing costs and taxes.

In the rest of Section 2, we look at various approaches to addressing criticisms 1 and 2, giving some attention also to criticisms 3 and 4. Sections 3 and 4 present approaches to addressing criticism 5. "Asset Allocation with Real World Constraints" addresses some aspects of criticism 6.

It is important to understand that the first criticism above is not unique to MVO. Any optimization model that uses forward-looking quantities as inputs faces similar consequences of treating input values as capable of being determined with certainty. Sensitivity to errors in inputs is a problem that cannot be fully solved because it is inherent in the structure of optimization models that use as inputs forecasts of uncertain quantities.

To illustrate the importance of the quality of inputs, the sensitivity of asset weights in efficient portfolios to small changes in inputs, and the propensity of mean–variance optimization to allocate to a relatively small subset of the available asset classes, we made changes to the expected return of two asset classes in our base-case UK-centric opportunity set in Exhibit 1. We increased the expected return of Asia Pacific ex Japan equities from 8.5% to 9.0% and decreased the expected return of Europe ex UK equities from 8.6% to 8.1% (both changes are approximately 50 bps). We left all of the other inputs unchanged and reran the optimization. The efficient frontier as depicted in mean–variance space appears virtually unchanged (not shown); however, the efficient asset mixes of this new efficient frontier are dramatically different. Exhibit 10 displays the efficient frontier asset allocation area graph based on the slightly changed capital market assumptions. Notice the dramatic difference between Exhibit 10 and Exhibit 3. The small change in return assumptions has driven UK large cap, Europe ex-UK equities, and emerging market equities out of the efficient mixes, and the efficient mixes are now highly concentrated in a smaller subset of the available asset classes. Given that the expected returns of UK large cap and emerging market equities were unchanged, their disappearance from the efficient frontier is not intuitive.

Exhibit 10 Efficient Frontier Asset Allocation Area Graph—Changed Expected Returns



To aid with the comparison of Exhibit 10 with Exhibit 3, we identified three specific efficient asset allocation mixes and compared the version based on the ad hoc modification of expected returns to that of the base case. This comparison is shown in Exhibit 11.

Exhibit 11 Comparison of Select Efficient Asset Allocations—Ad Hoc Return Modification Allocations vs. Base-Case Allocations

	Modified 25/75	Base Case 25/75	Difference	Modified 50/50	Base Case 50/50	Difference	Modified 75/25	Base Case 75/25	Difference
UK large cap	0.0%	1.2%	–1.2%	0.0%	2.5%	–2.5%	0.0%	0.0%	0.0%
UK mid cap	0.8%	0.6%	0.3%	1.7%	0.8%	0.9%	0.0%	0.0%	0.0%
UK small cap	0.5%	0.5%	–0.1%	0.4%	0.4%	0.0%	0.0%	0.0%	0.0%
US equities	13.7%	13.8%	–0.1%	26.6%	26.8%	–0.2%	40.1%	40.5%	–0.4%
Europe ex UK equities	0.0%	2.7%	–2.7%	0.0%	6.5%	–6.5%	0.0%	13.2%	–13.2%
Asia Pacific ex Japan equities	7.5%	1.0%	6.5%	16.6%	2.3%	14.2%	26.8%	1.5%	25.3%
Japan equities	2.2%	2.3%	–0.1%	4.5%	4.5%	0.0%	4.4%	4.3%	0.1%
Emerging market equities	0.0%	2.0%	–2.0%	0.0%	4.9%	–4.9%	0.0%	10.0%	–10.0%
Global REITs	0.3%	0.9%	–0.6%	0.2%	1.4%	–1.3%	3.8%	5.6%	–1.8%
Global ex UK bonds	10.9%	10.6%	0.3%	24.7%	23.9%	0.7%	25.0%	25.0%	0.0%
UK bonds	2.5%	2.7%	–0.2%	2.4%	3.0%	–0.6%	0.0%	0.0%	0.0%
Cash	61.6%	61.7%	–0.1%	22.9%	23.1%	–0.1%	0.0%	0.0%	0.0%
Subtotal equities	25.0%	25.0%		50.0%	50.0%		75.0%	75.0%	
Subtotal fixed income	75.0%	75.0%		50.0%	50.0%		25.0%	25.0%	

2.4 Addressing the Criticisms of Mean–Variance Optimization

In this section, we explore several methods for overcoming some of the potential short-comings of mean–variance optimization. Techniques that address the first two criticisms mostly take three approaches: improving the quality of inputs, constraining the optimization, and treating the efficient frontier as a statistical construct. These approaches are treated in the following three subsections.

In MVO, the composition of efficient portfolios is typically more sensitive to expected return estimates than it is to estimates of volatilities and correlations. Furthermore, expected returns are generally more difficult to estimate accurately than are volatilities and correlations. Thus, in addressing the first criticism of MVO—that outputs are highly sensitive to small changes in inputs—the reading will focus on expected return inputs. However, volatility and correlation inputs are also sources of potential error.

2.4.1 *Reverse Optimization*

Reverse optimization is a powerful tool that helps explain the implied returns associated with any portfolio. It can be used to estimate expected returns for use in a forward-looking optimization. MVO solves for optimal asset weights based on expected returns, covariances, and a risk aversion coefficient. Based on predetermined inputs, an optimizer solves for the optimal asset allocation weights. As the name implies, *reverse* optimization works in the opposite direction. Reverse optimization takes as its inputs a set of asset allocation weights *that are assumed to be optimal* and, with the additional inputs of covariances and the risk aversion coefficient, solves for expected returns. These reverse-optimized returns are sometimes referred to as implied or imputed returns.

When using reverse optimization to estimate a set of expected returns for use in a forward-looking optimization, the most common set of starting weights is the observed market-capitalization value of the assets or asset classes that form the opportunity set. The market capitalization of a given asset or asset classes should reflect the collective information of market participants. In representing the world market portfolio, the use of non-overlapping asset classes representing the majority of the world's investable assets is most consistent with theory.

Some practitioners will find the link between reverse optimization and CAPM equilibrium elegant, while others will see it as a shortcoming. For those who truly object to the use of market-capitalization weights in estimating inputs, the mechanics of reverse optimization can work with any set of starting weights—such as those of an existing policy portfolio, the average asset allocation policy of a peer group, or a fundamental weighting scheme. For those with more minor objections, we will shortly introduce the Black–Litterman model, which allows the expression of alternative forecasts or views.

In order to apply reverse optimization, one must create a working version of the all-inclusive market portfolio based on the constituents of the opportunity set. The market size or capitalization for most of the traditional stock and bond asset classes can be easily inferred from the various indexes that are used as asset class proxies. Many broad market-capitalization-weighted indexes report that they comprise over 95% of the securities, by market capitalization, of the asset classes they are attempting to represent. Exhibit 12 lists approximate values and weights for the 12 asset classes in our opportunity set, uses the weights associated with the asset classes to form a working version of the global market portfolio, and then uses the beta of each asset relative to our working version of the global market portfolio to infer what expected returns would be if all assets were priced by the CAPM according to their market beta. We assume a risk-free rate of 2.5% and a global market risk premium of 4%. Note that expected returns are rounded to one decimal place from the more precise values shown later (in Exhibit 13); expected returns cannot in every case be exactly reproduced based on Exhibit 12 alone because of the approximations mentioned. Also, notice in the final row of Exhibit 12 that the weighted average return and beta of the assets are 6.5% and 1, respectively.

Exhibit 12 Reverse-Optimization Example (Market Capitalization in £ billions)

Asset Class	Mkt Cap	Weight	Return $E[R_i]$		Risk-Free Rate r_f		Beta $\beta_{i,mkt}$	Market Risk Premium
UK large cap	£1,354.06	3.2%	6.62%	=	2.5%	+	1.03	(4%)
UK mid cap	£369.61	0.9%	6.92%	=	2.5%	+	1.11	(4%)
UK small cap	£108.24	0.3%	7.07%	=	2.5%	+	1.14	(4%)
US equities	£14,411.66	34.4%	7.84%	=	2.5%	+	1.33	(4%)
Europe ex UK equities	£3,640.48	8.7%	8.63%	=	2.5%	+	1.53	(4%)
Asia Pacific ex Japan equities	£1,304.81	3.1%	8.51%	=	2.5%	+	1.50	(4%)
Japan equities	£2,747.63	6.6%	6.43%	=	2.5%	+	0.98	(4%)
Emerging market equities	£2,448.60	5.9%	8.94%	=	2.5%	+	1.61	(4%)
Global REITs	£732.65	1.8%	9.04%	=	2.5%	+	1.64	(4%)

(continued)

Exhibit 12 (Continued)

Asset Class	Mkt Cap	Weight	Return $E[R_i]$		Risk-Free Rate r_f		Beta $\beta_{i,mkt}$	Market Risk Premium
Global ex UK bonds	£13,318.58	31.8%	4.05%	=	2.5%	+	0.39	(4%)
UK bonds	£1,320.71	3.2%	2.95%	=	2.5%	+	0.112	(4%)
Cash	£83.00	0.2%	2.50%	=	2.5%	+	0.00	(4%)
	£41,840.04	100.0%	6.50%				1	

Notes: For the Mkt Cap and Weight columns, the final row is the simple sum. For the Return and Beta columns, the final row is the weighted average.

Looking back at our original asset allocation area graph (Exhibit 3), the reason for the well-behaved and well-diversified asset allocation mixes is now clear. By using reverse optimization, we are consistently relating assets' expected returns to their systematic risk. If there isn't a consistent relationship between the expected return and systematic risk, the optimizer will see this inconsistency as an opportunity and seek to take advantage of the more attractive attributes. This effect was clearly visible in our second asset allocation area graph after we altered the expected returns of Asia Pacific ex Japan equities and Europe ex UK equities.

As alluded to earlier, some practitioners find that the reverse-optimization process leads to a nice starting point, but they often have alternative forecasts or views regarding the expected return of one or more of the asset classes that differ from the returns implied by reverse optimization based on market-capitalization weights. One example of having views that differ from the reverse-optimized returns has already been illustrated, when we altered the returns of Asia Pacific ex Japan equities and Europe ex UK equities by approximately 50 bps. Unfortunately, due to the sensitivity of mean–variance optimization to small changes in inputs, directly altering the expected returns caused relatively extreme and unintuitive changes in the resulting asset allocations. If one has strong views on expected returns that differ from the reverse-optimized returns, an alternative or additional approach is needed; the next section presents one alternative.

2.4.2 *Black–Litterman Model*

A complementary addition to reverse optimization is the Black–Litterman model, created by Fischer Black and Robert Litterman (see Black and Litterman 1990, 1991, 1992). Although the Black–Litterman model is often characterized as an asset allocation model, it is really a model for deriving a set of expected returns that can be used in an unconstrained or constrained optimization setting. The Black–Litterman model starts with excess returns (in excess of the risk-free rate) produced from reverse optimization and then provides a technique for altering reverse-optimized expected returns in such a way that they reflect an investor's own distinctive views yet still behave well in an optimizer.

The Black–Litterman model has helped make the mean–variance optimization framework more useful. It enables investors to combine their unique forecasts of expected returns with reverse-optimized returns in an elegant manner. When coupled with a mean–variance or related framework, the resulting Black–Litterman expected returns often lead to well-diversified asset allocations by improving the consistency between each asset class's expected return and its contribution to systematic risk. These asset allocations are grounded in economic reality—via the market capitalization of the assets typically used in the reverse-optimization process—but still reflect the information contained in the investor's unique forecasts (or views) of expected return.

The mathematical details of the Black–Litterman model are beyond the scope of this reading, but many practitioners have access to asset allocation software that includes the Black–Litterman model.[12] To assist with an intuitive understanding of the model and to show the model's ability to blend new information (views) with reverse-optimized returns, we present an example based on the earlier views regarding the expected returns of Asia Pacific ex Japan equities and Europe ex UK equities. The Black–Litterman model has two methods for accepting views: one in which an absolute return forecast is associated with a given asset class and one in which the return differential of an asset (or group of assets) is expressed relative to another asset (or group of assets). Using the relative view format of the Black–Litterman model, we expressed the view that we believe Asia Pacific ex Japan equities will outperform Europe ex UK equities by 100 bps. We placed this view into the Black–Litterman model, which blends reverse-optimized returns with such views to create a new, mixed estimate.

Exhibit 13 compares the Black–Litterman model returns to the original reverse-optimized returns (as in Exhibit 12 but showing returns to the second decimal place based on calculations with full precision). The model accounts for the correlations of the assets with each other, and as one might expect, all of the returns change slightly (the change in return on cash was extremely small).

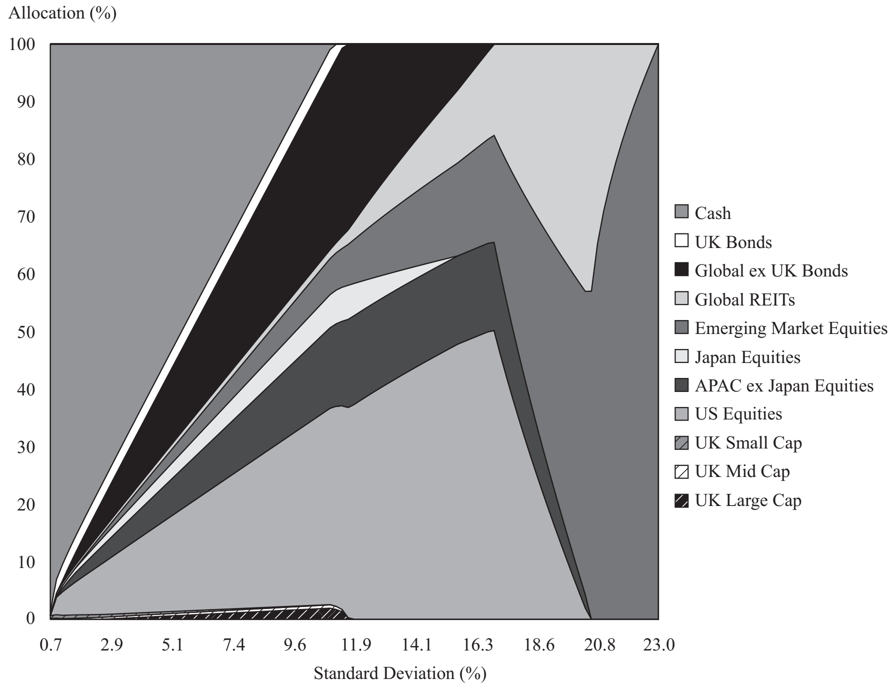
Exhibit 13 Comparison of Black–Litterman and Reverse-Optimized Returns

Asset Class	Reverse-Optimized Returns	Black–Litterman Returns	Difference
UK large cap	6.62%	6.60%	–0.02%
UK mid cap	6.92	6.87	–0.05
UK small cap	7.08	7.03	–0.05
US equities	7.81	7.76	–0.05
Europe ex UK equities	8.62	8.44	–0.18
Asia Pacific ex Japan equities	8.53	8.90	0.37
Japan equities	6.39	6.37	–0.02
Emerging market equities	8.96	9.30	0.33
Global REITs	9.02	9.00	–0.01
Global ex UK bonds	4.03	4.00	–0.03
UK bonds	2.94	2.95	0.01
Cash	2.50	2.50	0.00

Next, we created another efficient frontier asset allocation area graph based on these new returns from the Black–Litterman model, as shown in Exhibit 14. The allocations look relatively similar to those depicted in Exhibit 3. However, if you compare the allocations to Asia Pacific ex Japan equities and Europe ex UK equities to their allocations in the original efficient frontier asset allocation graph, you will notice that allocations to Asia Pacific ex Japan equities have increased across the frontier and allocations to Europe ex UK equities have decreased across the frontier with very little impact on the other asset allocations.

12 For those interested in the mathematical details of the Black–Litterman model, see Idzorek (2007); a pre-publication version is available here: http://corporate.morningstar.com/ib/documents/MethodologyDocuments/IBBAssociates/BlackLitterman.pdf).

Exhibit 14 Efficient Frontier Asset Allocation Area Graph, Black–Litterman Returns



As before, to aid in the comparison of Exhibit 14 (Black–Litterman allocations) with Exhibit 3 (the base-case allocations), we identified three specific mixes in Exhibit 14 and compared those efficient asset allocation mixes based on the expected returns from the Black–Litterman model to those of the base case. The results are shown in Exhibit 15.

Exhibit 15 Comparison of Select Efficient Asset Allocations, Black–Litterman Allocations vs. Base-Case Allocations

	Modified 25/75	Base Case 25/75	Difference	Modified 50/50	Base Case 50/50	Difference	Modified 75/25	Base Case 75/25	Difference
UK large cap	0.4%	1.2%	–0.8%	1.4%	2.5%	–1.1%	0.0%	0.0%	0.0%
UK mid cap	0.4	0.6	–0.2	0.5	0.8	–0.3	0.0	0.0	0.0
UK small cap	0.4	0.5	–0.1	0.2	0.4	–0.2	0.0	0.0	0.0
US equities	13.8	13.8	0.0	26.8	26.8	0.0	40.0	40.5	–0.5
Europe ex UK equities	0.0	2.7	–2.7	0.0	6.5	–6.5	0.0	13.2	–13.2
Asia Pacific ex Japan equities	5.2	1.0	4.2	10.8	2.3	8.5	15.4	1.5	14.0
Japan equities	2.2	2.3	0.0	4.5	4.5	0.0	4.2	4.3	–0.1
Emerging market equities	1.8	2.0	–0.1	4.6	4.9	–0.2	9.8	10.0	–0.1
Global REITs	0.8	0.9	–0.1	1.3	1.4	–0.2	5.5	5.6	–0.1
Global ex UK bonds	10.3	10.6	–0.2	23.6	23.9	–0.3	25.0	25.0	0.0

Exhibit 15 (Continued)

	Modified 25/75	Base Case 25/75	Difference	Modified 50/50	Base Case 50/50	Difference	Modified 75/25	Base Case 75/25	Difference
UK bonds	3.1	2.7	0.3	3.5	3.0	0.5	0.0	0.0	0.0
Cash	61.6	61.7	–0.1	22.9	23.1	–0.1	0.0	0.0	0.0
Subtotal equities	25.0%	25.0%		50.0%	50.0%		75.0%	75.0%	
Subtotal fixed income	75.0%	75.0%		50.0%	50.0%		25.0%	25.0%	

2.4.3 *Adding Constraints beyond the Budget Constraints*

When running an optimization, in addition to the typical budget constraint and the non-negativity constraint, one can impose additional constraints. There are two primary reasons practitioners typically apply additional constraints: (1) to incorporate real-world constraints into the optimization problem and (2) to help overcome some of the potential shortcomings of mean–variance optimization elaborated above (input quality, input sensitivity, and highly concentrated allocations).

Most commercial optimizers accommodate a wide range of constraints. Typical constraints include the following:

1 Specify a set allocation to a specific asset—for example, 30% to real estate or 45% to human capital. This kind of constraint is typically used when one wants to include a non-tradable asset in the asset allocation decision and optimize around the non-tradable asset.

2 Specify an asset allocation range for an asset—for example, the emerging market allocation must be between 5% and 20%. This specification could be used to accommodate a constraint created by an investment policy, or it might reflect the user's desire to control the output of the optimization.

3 Specify an upper limit, due to liquidity considerations, on an alternative asset class, such as private equity or hedge funds.

4 Specify the relative allocation of two or more assets—for example, the allocation to emerging market equities must be less than the allocation to developed equities.

5 In a liability-relative (or surplus) optimization setting, one can constrain the optimizer to hold one or more assets representing the systematic characteristics of the liability short. (We elaborate on this scenario in Section 3.)

In general, good constraints are those that model the actual circumstances/context in which one is attempting to set asset allocation policy. In contrast, constraints that are simply intended to control the output of a mean–variance optimization should be used cautiously. A perceived need to add constraints to control the MVO output would suggest a need to revisit one's inputs. If a very large number of constraints are imposed, one is no longer optimizing but rather specifying an asset allocation through a series of binding constraints.

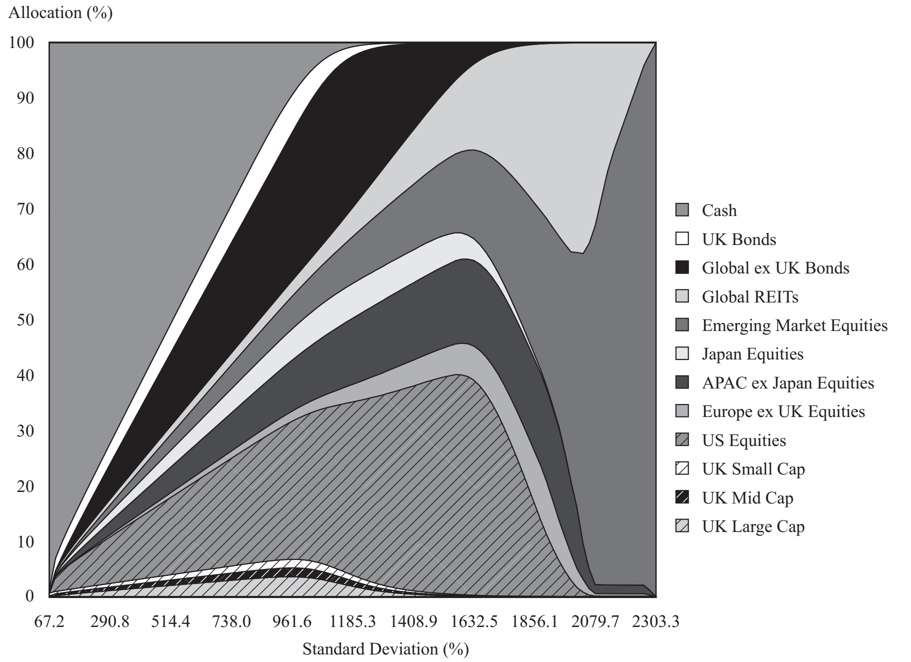
2.4.4 *Resampled Mean–Variance Optimization*

Another technique used by asset allocators is called resampled mean–variance optimization (or sometimes "resampling" for short).[13] Resampled mean–variance optimization combines Markowitz's mean–variance optimization framework with Monte Carlo simulation and, all else equal, leads to more-diversified asset allocations. In contrast to reverse optimization, the Black–Litterman model, and constraints, resampled mean–variance optimization is an attempt to build a better optimizer that recognizes that forward-looking inputs are inherently subject to error.

Resampling uses Monte Carlo simulation to estimate a large number of potential capital market assumptions for mean–variance optimization and, eventually, for the resampled frontier. Conceptually, resampling is a large-scale sensitivity analysis in which hundreds or perhaps thousands of variations on baseline capital market assumptions lead to an equal number of mean–variance optimization frontiers based on the Monte Carlo–generated capital market assumptions. These intermediate frontiers are referred to as simulated frontiers. The resulting asset allocations, or portfolio weights, from these simulated frontiers are saved and averaged (using a variety of methods). To draw the resampled frontier, the averaged asset allocations are coupled with the starting capital market assumptions.

To illustrate how resampling can be used with other techniques, we conducted a resampled mean–variance optimization using the Black–Litterman returns from Exhibit 10, above. Exhibit 16 provides the asset allocation area graph from this optimization. Notice that the resulting asset allocations are smoother than in any of the previous asset allocation area graphs. Additionally, relative to Exhibit 15, based on the same inputs, the smallest allocations have increased in size while the largest allocations have decreased somewhat.

Exhibit 16 Efficient Frontier Asset Allocation Area Graph, Black–Litterman Returns with Resampling



13 The current embodiments of resampling grew out of the work of Jobson and Korkie (1980, 1981); Jorion (1992); DiBartolomeo (1993); and Michaud (1998).

The asset allocations from resampling as depicted in Exhibit 16 are appealing. Criticisms include the following: (1) Some frontiers have concave "bumps" where expected return decreases as expected risk increases; (2) the "riskier" asset allocations are over-diversified; (3) the asset allocations inherit the estimation errors in the original inputs; and (4) the approach lacks a foundation in theory.[14]

2.4.5 *Other Non-Normal Optimization Approaches*

From our list of shortcomings/criticisms of mean–variance optimization, the third is that investor preferences may go beyond the first two moments (mean and variance) of a portfolio's return distribution. The third and fourth moments are, respectively, skewness and kurtosis. Skewness measures the degree to which return distributions are asymmetrical, and kurtosis measures the thickness of the distributions' tails (i.e., how frequently extreme events occur). A normal distribution is fully explained by the first two moments because the skewness and (excess) kurtosis of the normal distribution are both zero.

Returning to the discussion of Equation 1, the mean–variance optimization program involves maximizing expected utility, which is equal to expected return minus a penalty for risk, where risk is measured as variance (standard deviation). Unfortunately, variance or standard deviation is an incomplete measure of risk when returns are not normally distributed. By studying historical return distributions for the major asset classes and comparing those historical distributions to normal distributions, one will quickly see that, historically, asset class returns are not normally distributed. In fact, empirically extreme returns seem to occur approximately 10 times more often than the normal distribution would suggest. Coupling this finding with the asymmetrical risk preferences observed in investors—whereby the pain of a loss is approximately twice as significant as the joy from an equivalent gain (according to Prospect theory)—has led to more complex utility functions and optimizers that expressly account for non-normal returns and asymmetric risk preference.[15] A number of variations of these more sophisticated optimization techniques have been put forth, making them challenging to cover. In general, most of them consider the non-normal return distribution characteristics and use a more sophisticated definition of risk, such as conditional value-at-risk. We view these as important advancements in the toolkit available to practitioners.

Exhibit 17 summarizes selected extensions of quantitative asset allocation approaches outside the sphere of traditional mean–variance optimization.

Exhibit 17 Selected Non-Mean–Variance Developments

Key Non-Normal Frameworks	Research/Recommended Reading
Mean–semivariance optimization	Markowitz (1959)
Mean–conditional value-at-risk optimization	Goldberg, Hayes, and Mahmoud (2013) Rockafellar and Uryasev (2000) Xiong and Idzorek (2011)
Mean–variance-skewness optimization	Briec, Kerstens, and Jokung (2007) Harvey, Liechty, Liechty, and Müller (2010)
Mean–variance-skewness-kurtosis optimization	Athayde and Flôres (2003) Beardsley, Field, and Xiao (2012)

14 For more details, see Scherer (2002).
15 For more on prospect theory, see Kahneman and Tversky (1979) and Tversky and Kahneman (1992).

Long-Term versus Short-Term Inputs

Strategic asset allocation is often described as "long term," while tactical asset allocation involves short-term movements away from the strategic asset allocation. In this context, "long term" is often defined as 10 or perhaps 20 or more years, yet in practice, very few asset allocators revisit their strategic asset allocation this infrequently. Many asset allocators update their strategic asset allocation annually, which makes it a bit more challenging to distinguish between strategic and tactical asset allocations. This frequent revisiting of the asset allocation policy brings up important questions about the time horizon associated with the inputs. In general, long-term (10-plus-year) capital market assumptions that ignore current market conditions, such as valuation levels, the business cycle, and interest rates, are often thought of as *unconditional* inputs. Unconditional inputs focus on the average capital market assumptions over the 10-plus-year time horizon. In contrast, shorter-term capital market assumptions that explicitly attempt to incorporate current market conditions (i.e., that are "conditioned" on them) are conditional inputs. For example, a practitioner who believes that the market is overvalued and that as a result we are entering a period of low returns, high volatility, and high correlations might prefer to use conditional inputs that reflect these beliefs.[16]

EXAMPLE 4

Problems in Mean–Variance Optimization

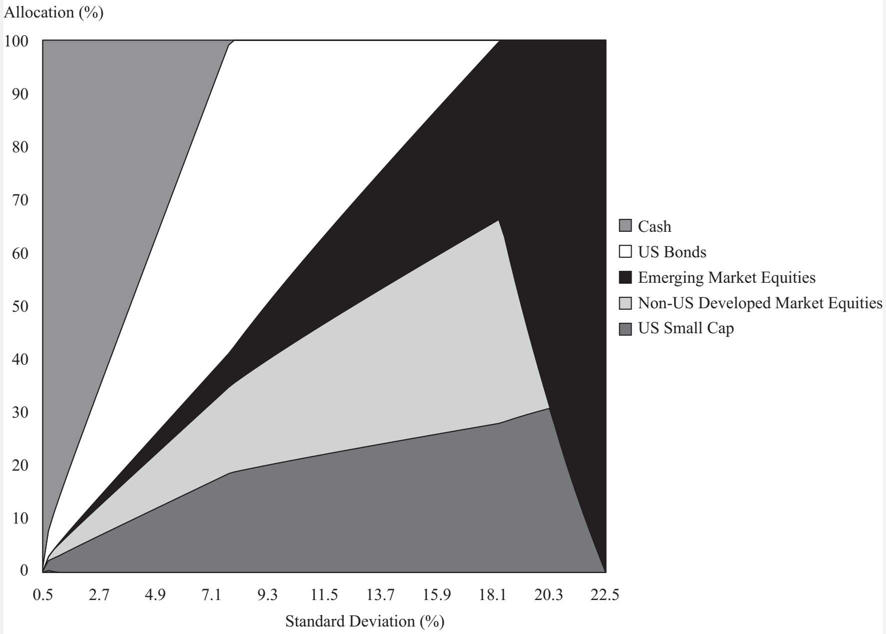
In a presentation to US-based investment clients on asset allocation, the results of two asset allocation exercises are shown, as presented in Exhibit 18.

Exhibit 18 Asset Allocation Choices

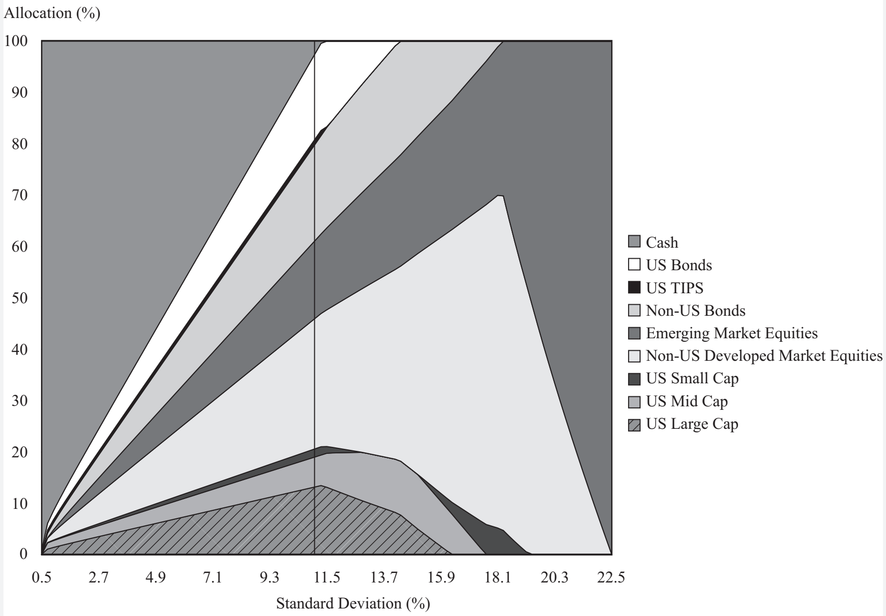
Panel A: Area Graph 1

16 Relatedly, Chow, Jacquier, Kritzman, and Lowry (1999) showed a procedure for blending the optimal portfolios for periods of normal and high return volatility. The approach accounts for the tendency of asset returns to be more highly correlated during times of high volatility.

Exhibit 18 (Continued)



Panel B: Area Graph 2



1 Based on Panel A, address the following:

- **A** Based on mean–variance analysis, what is the asset allocation that would most likely be selected by a risk-neutral investor?
- **B** Based only on the information that can be inferred from Panel A, discuss the investment characteristics of non-US developed market equity (NUSD) in efficient portfolios.
- **C** Critique the efficient asset mixes represented in Panel A.

2 Compare the asset allocations shown in Panel A with the corresponding asset allocations shown in Panel B. (Include a comparison of the panels at the level of risk indicated by the line in Panel B.)

3 **A** Identify three techniques that the asset allocations in Panel B might have incorporated to improve the characteristics relative to those of Panel A.

B Discuss how the techniques described in your answer to 3A address the high input sensitivity of MVO.

Solution to 1A:

For a risk-neutral investor, the optimal asset allocation is 100% invested in emerging market equities. For a risk-neutral investor ($\lambda = 0$), expected utility is simply equal to expected return. The efficient asset allocation that maximizes expected return is the one with the highest level of volatility, as indicated on the x-axis. Panel A shows that that asset allocation consists entirely of emerging market equities.

Solution to 1B:

The weights of NUSD as the efficient frontier moves from its minimum to its maximum risk point suggest NUSD's investment characteristics. This asset class is neither the lowest-volatility asset (which can be inferred to be cash) nor the highest-volatility asset (which is emerging market equity). At the point of the peak of NUSD, when the weight in NUSD is about to begin its decline in higher-risk efficient portfolios, US bonds drop out of the efficient frontier. Further, NUSD leaves the efficient frontier portfolio at a point at which US small cap reaches its highest weight. These observations suggest that NUSD provided diversification benefits in portfolios including US bonds—a relatively low correlation with US bonds can be inferred—that are lost at this point on the efficient frontier. Beyond a volatility level of 20.3%, representing a corner portfolio, NUSD drops out of the efficient frontier.

Solution to 1C:

Of the nine asset classes in the investor's defined opportunity set, five at most are represented by portfolios on the efficient frontier. Thus, a criticism of the efficient frontier associated with Panel A is that the efficient portfolios are highly concentrated in a subset of the available asset classes, which likely reflects the input sensitivity of MVO.

Solution to 2:

The efficient asset mixes in Panels A and B cover a similar risk range: The risk levels of the two minimum-variance portfolios are similar, and the risk levels of the two maximum-return portfolios are similar. Over most of the range of volatility, however, the efficient frontier associated with Panel B is better diversified. For example, at the line in Panel B, representing a moderate level of volatility likely relevant to many investors, the efficient portfolio contains nine asset classes rather than four, as in Panel A. At that point, for example, the allocation to fixed income is spread over US bonds, non-US bonds, and US TIPS in Panel B, as opposed to just US bonds in Panel A.

Solution to 3A:

To achieve the better-diversified efficient frontier shown in Panel B, several methods might have been used, including reverse optimization, the Black–Litterman model, and constrained asset class weights.

Solution to 3B:

Reverse optimization and the Black–Litterman model address the issue of MVO's sensitivity to small differences in expected return estimates by anchoring expected returns to those implied by the asset class weights of a proxy for the global market portfolio. The Black–Litterman framework provides a disciplined way to tilt the expected return inputs in the direction of the investor's own views. These approaches address the problem by improving the balance between risk and return that is implicit in the inputs.

A very direct approach to the problem can be taken by placing constraints on weights in the optimization to force an asset class to appear in a constrained efficient frontier within some desired range of values. For example, non-US bonds did not appear in any efficient portfolio in Panel A. The investor could specify that the weight on non-US bonds be strictly positive. Another approach would be to place a maximum on the weight in US bonds to make the optimizer spread the fixed-income allocation over other fixed-income assets besides US bonds.

2.5 Allocating to Less Liquid Asset Classes

Large institutional investors have the ability to invest in less liquid asset classes, such as direct real estate, infrastructure, and private equity. These less liquid asset classes represent unique challenges to many of the common asset allocation techniques, such as mean–variance optimization.

For traditional, highly liquid asset classes, such as publicly listed equities and bonds, almost all of the major index providers have indexes that do an outstanding job of representing the performance characteristics of the asset class (and its various sub–asset classes). For example, over any reasonably long time period, the risk and return characteristics of a given asset class are nearly identical across the major global equity indexes and the correlations between the returns of the indexes are close to 1. Additionally, in most cases, there are passive, low-cost investment vehicles that allow investors to capture the performance of the asset class with very little tracking error.

Cash, the Risk-Free Asset, and Liquidity Needs

The so called "risk-free asset" has a special and somewhat tricky spot in the world of finance. Asset allocators typically use indexes for either 30-day or 90-day government bills to represent the characteristics associated with holding cash, which they may or may not treat as the risk-free asset. The volatility associated with these total return indexes is extremely low, but it isn't zero. An alternative to using a cash index as a proxy for the risk-free asset is to use a government bond with a duration/maturity that matches the time horizon of the investor. Some asset allocators like to include cash or another asset that could be considered a risk-free asset in the optimization and to allow the optimizer to determine how to mix it with the other asset classes included in the optimization. Other asset allocators prefer to exclude the risk-free asset from the optimization and allow real-world needs, such as liquidity needs, to determine how much to allocate to cash-like assets.

Illiquid assets may offer an expected return premium as compensation for illiquidity as well as diversification benefits. Determining an appropriate allocation to these assets is associated with various challenges, however. Common illiquid asset classes cannot be readily diversified to eliminate idiosyncratic risk, so representing an overall asset class performance is problematic. Furthermore, for less liquid asset classes, such as direct real estate, infrastructure, and private equity, there are, in general, far fewer indexes that attempt to represent aggregate performance. If one were to compare the performance characteristics of multiple indexes representing one of these less liquid asset classes, there would be noticeable risk and return differences, suggesting that it is difficult to accurately measure the risk and return characteristics of these asset classes. Also, due to the illiquid nature of the constituents that make up these asset classes, it is widely believed that the indexes don't accurately reflect their true volatility. In contrast to the more traditional, highly liquid asset classes, there are no low-cost passive investment vehicles that would allow investors to closely track the aggregate performance of these less liquid asset classes.

Thus, the problem is twofold: (1) Due to the lack of accurate indexes, it is more challenging to make capital market assumptions for these less liquid asset classes, and (2) even if there were accurate indexes, there are no low-cost passive investment vehicles to track them.

Compounding the asset allocator's dilemma is the fact that the risk and return characteristics associated with actual investment vehicles, such as direct real estate funds, infrastructure funds, and private equity funds, are typically significantly different from the characteristics of the asset classes themselves. For example, the private equity "asset class" should represent the risk and return characteristics of owning all private equity, just as the MSCI All Country World Index represents the risk and return characteristics of owning all public equity. Purchasing the exchange-traded fund (ETF) that tracks the MSCI All Country World Index completely diversifies public company-specific risk. This scenario is in direct contrast to the typical private equity fund, in which the risk and return characteristics are often dominated by company-specific (idiosyncratic) risk.

In addressing asset allocation involving less liquid asset classes, practical options include the following:

1. Exclude less liquid asset classes (direct real estate, infrastructure, and private equity) from the asset allocation decision and then consider real estate funds, infrastructure funds, and private equity funds as potential implementation vehicles when fulfilling the target strategic asset allocation.
2. Include less liquid asset classes in the asset allocation decision and attempt to model the inputs to represent the *specific risk* characteristics associated with the likely *implementation vehicles*.
3. Include less liquid asset classes in the asset allocation decision and attempt to model the inputs to represent the *highly diversified* characteristics associated with the *true asset classes*.

Related to this last option, some practitioners use listed real estate indexes, listed infrastructure, and public equity indexes that are deemed to have characteristics similar to their private equity counterparts to help estimate the risk of the less liquid asset classes and their correlation with the other asset classes in the opportunity set. It should be noted that the use of listed alternative indexes often violates the recommendation that asset classes be mutually exclusive—the securities in these indexes are likely also included in indexes representing other asset classes—and thus typically results in higher correlations among different asset classes, which has the negative impact of increasing input sensitivity in most optimization settings.

For investors who do not have access to direct real estate funds, infrastructure funds, and private equity funds—for example, small investors—the most common approach is to use one of the indexes based on listed equities to represent the asset class and then to implement the target allocation with a fund that invests similarly. Thus global REITs might be used to represent (approximately) global real estate.

2.6 Risk Budgeting

> [A] risk budget is simply a particular allocation of portfolio risk. An optimal risk budget is simply the allocation of risk such that the first order of conditions for portfolio optimization are satisfied. The risk budgeting process is the process of finding an optimal risk budget.
>
> Kurt Winkelmann (2003, p. 173)

As this quote from Kurt Winkelmann suggests, there are three aspects to risk budgeting:

- The risk budget identifies the total amount of risk and allocates the risk to a portfolio's constituent parts.
- An optimal risk budget allocates risk efficiently.
- The process of finding the optimal risk budget is risk budgeting.

Although its name suggests that risk budgeting is all about risk, risk budgeting is really using risk in relation to seeking return. The goal of risk budgeting is to maximize return per unit of risk—whether overall market risk in an asset allocation setting or active risk in an asset allocation implementation setting.

The ability to determine a position's marginal contribution to portfolio risk is a powerful tool that helps one to better understand the sources of risk. The marginal contribution to a type of risk is the partial derivative of the risk in question (total risk, active risk, or residual risk) with respect to the applicable type of portfolio holding (asset allocation holdings, active holdings, or residual holdings). Knowing a position's marginal contribution to risk allows one to (1) approximate the change in portfolio risk (total risk, active risk, or residual risk) due to a change in an individual holding, (2) determine which positions are optimal, and (3) create a risk budget. *Risk-budgeting tools assist in the optimal use of risk in the pursuit of return.*

Exhibit 19 contains risk-budgeting information for the Sharpe ratio–maximizing asset allocation from our original UK example. The betas are from Exhibit 12. The marginal contribution to total risk (MCTR) identifies the rate at which risk would change with a small (or marginal) change in the current weights. For asset class *i*, it is calculated as MCTR_i = (Beta of asset class *i* with respect to portfolio)(Portfolio return volatility). The absolute contribution to total risk (ACTR) for an asset class measures how much it contributes to portfolio return volatility and can be calculated as the weight of the asset class in the portfolio times its marginal contribution to total risk: $\text{ACTR}_i = (\text{Weight}_i)(\text{MCTR}_i)$. Critically, beta takes account not only of the asset's own volatility but also of the asset's correlations with other portfolio assets.

The sum of the ACTR in Exhibit 19 is approximately 10.88%, which is equal to the expected standard deviation of this asset allocation mix. Dividing each ACTR by the total risk of 10.88% gives the percentage of total risk that each position contributes. Finally, an asset allocation is optimal from a risk-budgeting perspective when the ratio of excess return (over the risk-free rate) to MCTR is the same for all assets and matches the Sharpe ratio of the tangency portfolio. So in this case, which is based on reverse-optimized returns, we have an optimal risk budget.

Exhibit 19 Risk-Budgeting Statistics

Asset Class	Weight	MCTR	ACTR	Percent Contribution to Total Standard Deviation	Ratio of Excess Return to MCTR
UK large cap	3.2%	11.19%	0.36%	3.33%	0.368
UK mid cap	0.9	12.02	0.11	0.98	0.368

(continued)

Exhibit 19 (Continued)

Asset Class	Weight	MCTR	ACTR	Percent Contribution to Total Standard Deviation	Ratio of Excess Return to MCTR
UK small cap	0.3	12.44	0.03	0.30	0.368
US equities	34.4	14.51	5.00	45.94	0.368
Europe ex UK equities	8.7	16.68	1.45	13.34	0.368
Asia Pacific ex Japan equities	3.1	16.35	0.51	4.69	0.368
Japan equities	6.6	10.69	0.70	6.46	0.368
Emerging market equities	5.9	17.51	1.02	9.42	0.368
Global REITs	1.8	17.79	0.31	2.86	0.368
Global ex UK bonds	31.8	4.21	1.34	12.33	0.368
UK bonds	3.2	1.22	0.04	0.35	0.368
Cash	0.2	0.00	0.00	0.00	0.368
	100.0		10.88	100.00	

For additional clarity, the following are the specific calculations used to derive the calculated values for UK large-cap equities (where we show some quantities with an extra decimal place in order to reproduce the values shown in the exhibit):

- Marginal contribution to risk (MCTR):

  Asset beta relative to portfolio × Portfolio standard deviation

  1.0289 × 10.876 = 11.19%

- ACTR:

  Asset weight in portfolio × MCTR

  3.2% × 11.19% = 0.36%

- Ratio of excess return to MCTR:

  (Expected return – Risk-free rate)/MCTR

  (6.62% – 2.5%)/11.19% = 0.368

EXAMPLE 5

Risk Budgeting in Asset Allocation

1. Describe the objective of risk budgeting in asset allocation.
2. Consider two asset classes, A and B. Asset class A has two times the weight of B in the portfolio. Under what condition would B have a larger ACTR than A?
3. When is an asset allocation optimal from a risk-budgeting perspective?

Solution to 1:

The objective of risk budgeting in asset allocation is to use risk efficiently in the pursuit of return. A risk budget specifies the total amount of risk and how much of that risk should be budgeted for each allocation.

Solution to 2:

Because ACTR_i = (Weight_i)(Beta with respect to portfolio)$_i$(Portfolio return volatility), the beta of B would have to be more than twice as large as the beta of A for B to contribute more to portfolio risk than A.

Solution to 3:

An asset allocation is optimal when the ratio of excess return (over the risk-free rate) to MCTR is the same for all assets.

2.7 Factor-Based Asset Allocation

Until now, we have primarily focused on the mechanics of asset allocation optimization as applied to an opportunity set consisting of traditional, non-overlapping asset classes. An alternative approach used by some practitioners is to move away from an opportunity set of *asset classes* to an opportunity set consisting of investment *factors*.

In factor-based asset allocation, the factors in question are typically similar to the fundamental (or structural) factors in widely used multi-factor investment models. Factors are typically based on observed market premiums and anomalies. In addition to the all-important market (equity) exposure, typical factors used in asset allocation include size, valuation, momentum, liquidity, duration (term), credit, and volatility. Most of these factors were identified as return drivers that help to explain returns that were not explained by the CAPM. These factors can be constructed in a number of different ways, but with the exception of the market factor, typically, the factor represents what is referred to as a zero (dollar) investment, or self-financing investment, in which the underperforming attribute is sold short to finance an offsetting long position in the better-performing attribute. For example, the size factor is the combined return from shorting large-cap stocks and going long small-cap stocks (Size factor return = Small-cap stock return – Large-cap stock return). Of course, if large-cap stocks outperform small-cap stocks, the realized size return would be negative. Constructing factors in this manner removes most market exposure from the factors (because of the short positions that offset long positions); as a result, the factors generally have low correlations with the market and with one another.

We next present an example of a factor-based asset allocation optimization. Exhibit 20 shows the list of factors, how they were specified, and their historical returns and standard deviations (in excess of the risk-free rate as proxied by the return on three-month Treasury bills). The exhibit also includes historical statistics for three-month Treasury bills.

Thus far, our optimization examples have taken place in "total return space," where the expected return of each asset has equaled the expected return of the risk-free asset plus the amount of expected return in excess of the risk-free rate. In order to stay in this familiar total return space when optimizing with risk factors, the factor return needs to include the return on the assumed collateral (in this example, cash, represented by three-month Treasury bills). This adjustment is also needed if one plans to include both risk factors and some traditional asset classes in the same optimization, so that the inputs for the risk factors and traditional asset classes are similarly specified. Alternatively, one could move in the opposite direction, subtracting the return of the three-month Treasury bills from asset class returns and then conducting the optimization in excess-return space. One way to think about a self-financing allocation to a risk factor is that in order to invest in the risk factor, one must put up an equivalent amount of collateral that is invested in cash.

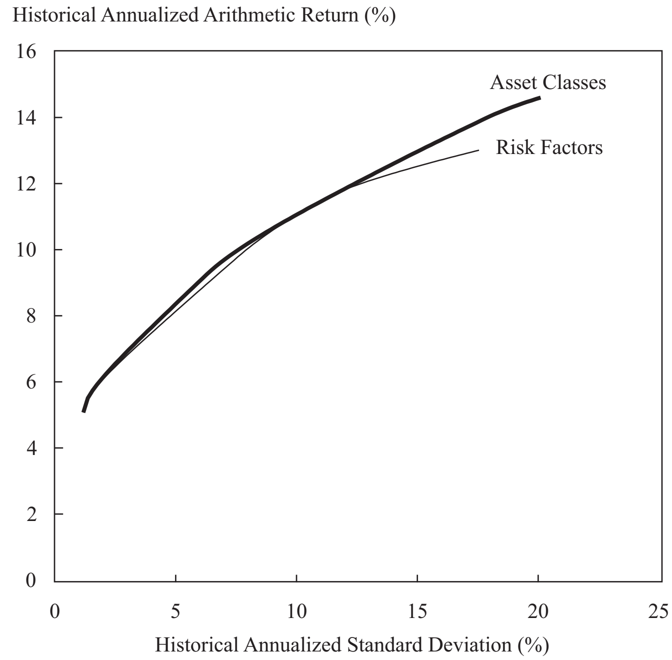
Exhibit 20 Factors/Asset Classes, Factor Definitions, and Historical Statistics (US data, January 1979 to March 2016)

Factor/Asset Class	Factor Definition	Compound Annual Factor Return	Standard Deviation	Total Return	Standard Deviation
Cash	3-Month Treasury bills			7.77%	5.66%
Market	Total market return – Cash	7.49%	16.56%	12.97	17.33
Size	Small cap – Large cap	0.41	10.15	5.56	10.65
Valuation	Value – Growth	0.68	9.20	5.84	9.76
Credit	Corporate – Treasury	0.70	3.51	5.87	3.84
Duration	Long Treasury bonds – Treasury bills	4.56	11.29	9.91	11.93
Mortgage	Mortgage-backed – Treasury bonds	0.30	3.38	5.45	3.83
Large growth	—	—	—	12.64	19.27
Large value	—	—	—	13.23	16.52
Small growth	—	—	—	12.30	25.59
Small value	—	—	—	14.54	19.84
Mortgage-backed sec.	—	—	—	8.09	6.98
Corporate bonds	—	—	—	8.52	7.52
Treasury bonds	—	—	—	7.77	5.66
Cash	—	—	—	5.13	1.23

Because of space considerations, we have not included the full correlation matrix, but it is worth noting that the average pair-wise correlation of the risk factor–based opportunity set (in excess of the risk-free rate collateral return) is 0.31, whereas that of the asset class–based opportunity set is 0.57. Given the low pair-wise correlations of the risk factors, there has been some debate among practitioners around whether it is better to optimize using asset classes or risk factors. The issue was clarified by Idzorek and Kowara (2013), who demonstrated that in a proper comparison, neither approach is inherently superior. To help illustrate risk factor optimization and to demonstrate that if the two opportunity sets are constructed with access to similar exposures, neither approach has an inherent advantage, we present two side-by-side optimizations. These optimizations are based on the data given in Exhibit 20.

Exhibit 21 contains the two efficient frontiers. As should be expected, given that the opportunity sets provide access to similar exposures, the two historical efficient frontiers are very similar. This result illustrates that when the same range of potential exposures is available in two opportunity sets, the risk and return possibilities are very similar.

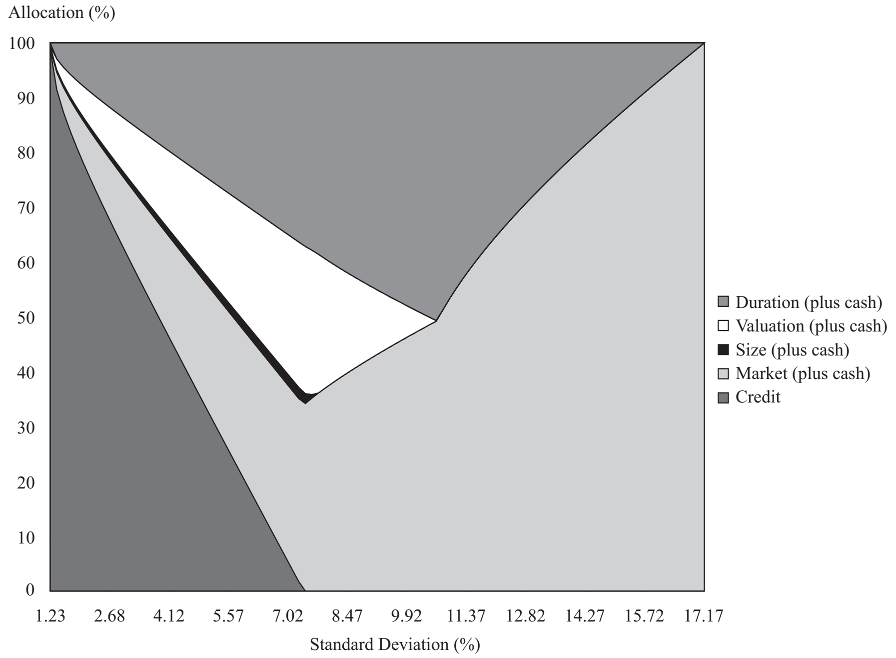
Exhibit 21 Efficient Frontiers Based on Historical Capital Market Assumptions (January 1979 to March 2016)



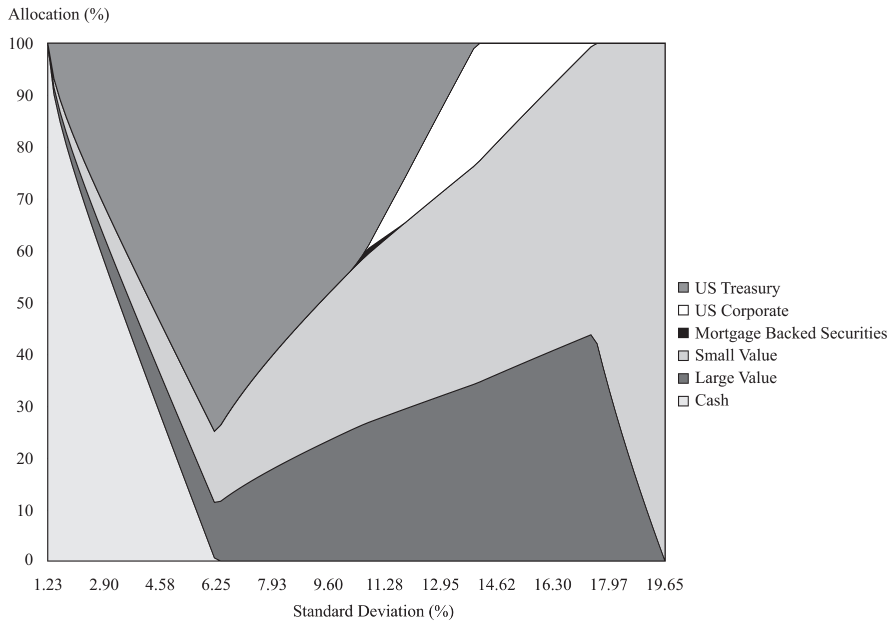
Moving to Exhibit 22, examining the two asset allocation area graphs associated with the two efficient frontiers reveals that the efficient mixes have some relatively clear similarities. For example, in Panel A (risk factors), the market, size, and valuation exposures mirror the pattern (allocations) of Panel B (asset classes) large value and small value exposures, respectively.

Exhibit 22 Asset Allocation Area Graphs—Risk Factors and Asset Classes

Panel A: Risk Factor Asset Allocation Area Graph



Panel B: Asset Class Asset Allocation Area Graph



Practitioners should choose to carry out asset allocation in the particular space—risk factors or asset classes—in which they are most equipped to make capital market assumptions. Regardless of which space a practitioner prefers, expanding one's opportunity set to include new, weakly correlated risk factors or asset classes should improve the potential risk–return trade-offs.

DEVELOPING LIABILITY-RELATIVE ASSET ALLOCATIONS

3

Liability-relative asset allocation is aimed at the general issue of rendering decisions about asset allocation in conjunction with the investor's liabilities. Liability-relative investors view assets as an inventory of capital, sometimes increased by additions, which is available to achieve goals and to pay future liabilities. What is the chance that an institution's capital is sufficient to cover future cash flow liabilities? This type of question is critical for liability-relative asset allocation because many large institutional investors—for example, banks, insurance companies, and pension plans—possess legal liabilities and operate in regulated environments in which an institution's inability to meet its liabilities with current capital has serious consequences. This concern gives rise to unique risk measures, such as the probability of meeting future cash flow requirements, and the restatement of traditional risk metrics, such as volatility, in relation to liabilities.

Liability-relative methods were developed in an institutional investor context, but these ideas have also been applied to individual investors. This section will focus on institutional investors. A later section addresses a thematically similar approach with behavioral finance roots—goals-based asset allocation.

3.1 Characterizing the Liabilities

To be soundly applied, liability-relative asset allocation requires an accurate understanding of the liabilities. A liability is a promise by one party to pay a counterparty based on a prior agreement. Liabilities may be fixed or contingent. When the amounts and timing of payments are fixed in advance by the terms of a contract, the liability is said to be fixed or non-contingent. A corporate bond with a fixed coupon rate is an example.

In many cases relevant to asset allocation, payments depend upon future, uncertain events. In such cases, the liability is a contingent liability.[17] An important example involves the liabilities of a defined benefit (DB) pension plan. The plan sponsor has a legal commitment to pay the beneficiaries of the plan during their retirement years. However, the exact dates of the payments depend on the employees' retirement dates, longevity, and cash payout rules. Insurance companies' liabilities—created by the sale of insurance policies—are also contingent liabilities: The insurance company promises to pay its policyholders a specified amount contingent on the occurrence of a predefined event.

We distinguish legal liabilities from cash payments that are expected to be made in the future and are essential to the mission of an institution but are not legal liabilities. We call these quasi-liabilities. The endowment of a university can fit this category because, in many cases, the endowment contributes a major part of the university's operating budget. The endowment assures its stakeholders that it will continue to support its essential activities through spending from the endowment capital, and failure to provide such support will often lead to changes in how the endowment is managed. Accordingly, the asset allocation decisions are made in conjunction with the university's spending rules and policies. Asset allocation is just one portion of the investment problem. Although we do not explicitly discuss them here, as suggested in Section 2, the spending needs of an individual represent another type of quasi-liability. Exhibit 23 summarizes the characteristics of liabilities that can affect asset allocation.

---

17 Note that the term "contingent liability" has a specific definition in accounting. We are using the term more broadly here.

Exhibit 23 Characteristics of Liabilities That Can Affect Asset Allocation

1 Fixed versus contingent cash flows
2 Legal versus quasi-liabilities
3 Duration and convexity of liability cash flows
4 Value of liabilities as compared with the size of the sponsoring organization
5 Factors driving future liability cash flows (inflation, economic conditions, interest rates, risk premium)
6 Timing considerations, such as longevity risk
7 Regulations affecting liability cash flow calculations

The above liability characteristics are relevant to liability-relative asset allocation in various ways. For example, they affect the choice of appropriate discount rate(s) to establish the present value of the liabilities and thus the degree to which assets are adequate in relation to those liabilities. Liability characteristics determine the composition of the liability-matching portfolio and that portfolio's basis risk with respect to the liabilities. (Basis risk in this context quantifies the degree of mismatch between the hedging portfolio and the liabilities.)

We will discuss the following case study in detail. It involves a frozen pension plan for LOWTECH, a hypothetical US company. The company has decided to close its defined benefit pension plan and switch to a defined contribution plan. The DB plan has the fixed liabilities (accumulated benefit obligations) shown in Exhibit 24.

Exhibit 24 Projected Liability Cash Flows for Company LOWTECH (US$ billions)

		PV(Liabilities)	
Beginning of Year	Cash Outflow (Liability)	4% Discount Rate	2% Discount Rate
2015	—	$2.261	$3.039
2016	$0.100	2.352	3.10
2017	0.102	2.342	3.06
2018	0.104	2.329	3.02
2019	0.106	2.314	2.97
2020	0.108	2.297	2.92
2021	0.110	2.276	2.87
2022	0.113	2.252	2.82
2023	0.115	2.225	2.76
2024	0.117	2.195	2.69
2025	0.120	2.161	2.63
2026	0.122	2.123	2.56
2027	0.124	2.081	2.49
2028	0.127	2.035	2.41
2029	0.129	1.984	2.33
2030	0.132	1.929	2.24
2031	0.135	1.869	2.15

Exhibit 24 (Continued)

Beginning of Year	Cash Outflow (Liability)	PV(Liabilities) 4% Discount Rate	2% Discount Rate
2032	0.137	1.804	2.06
2033	0.140	1.733	1.96
2034	0.143	1.657	1.86
2035	0.146	1.575	1.75
2036	0.149	1.486	1.63
2037	0.152	1.391	1.52
2038	0.155	1.289	1.39
2039	0.158	1.180	1.26
2040	0.161	1.063	1.13
2041	0.164	0.938	0.98
2042	0.167	0.805	0.84
2043	0.171	0.663	0.68
2044	0.174	0.512	0.52
2045	0.178	0.352	0.36
2046	0.181	0.181	0.181

In the Cash Outflow (Liability) column, the assumption is made that payments for a given year are made at the beginning of the year (in the exhibit, outflows have a positive sign). As of the beginning of 2015, the present value of these liabilities, given a 4% discount rate for high-quality corporate bonds (required in the United States by the Pension Protection Act of 2006, which applies to private DB pension plans), is US$2.261 billion. The current market value of the assets is assumed to equal US$2.5 billion, for a surplus of US$0.239 billion. On the other hand, if the discount rate is equal to the long-term government bond rate at 2% (required before the 2006 US legislation), the surplus becomes a deficit at –$0.539 billion. In many cases, regulations set the appropriate discount rates; these rates have an impact on the determination of surplus or deficit and thus on future contribution rules.

Like other institutions with legal liabilities, the LOWTECH company must analyze its legal future cash flows under its DB pension system and evaluate them in conjunction with the current market value of its assets on an annual basis. The following steps of the valuation exercise for a DB pension plan occur on a fixed annual date:

1 Calculate the market value of assets.
2 Project liability cash flows (via actuarial principles and rules).
3 Determine an appropriate discount rate for liability cash flows.
4 Compute the present value of liabilities, the surplus value, and the funding ratio.

Surplus = Market value (assets) – Present value (liabilities).

The surplus for the LOWTECH company is US$2.500 billion – US$2.261 billion = US$0.239 billion, given the 4% discount rate assumption.

The funding ratio is another significant measure: Funding ratio = Market value (assets)/Present value (liabilities). We say that an investor is fully funded if the investor's funding ratio equals 1 (or the surplus is 0). A state of overfunding occurs when the funding ratio is greater than 1, and a state of underfunding takes place when the funding ratio is less than 1. Based on a discount rate of 4%, the funding ratio for LOWTECH = US$2.5 billion/US$2.261 billion = 1.1057, so that the company is about 10.6% overfunded.

The surplus value and the funding ratio are highly dependent upon the discount rate assumption. For example, if the discount rate is equal to 2.0% (close to the 10-year US Treasury bond rate in early 2016), the surplus drops to –US$0.539 billion and the funding ratio equals 0.8226. The company's status changes from overfunded to underfunded. The choice of discount rate is generally set by regulations and tradition. Rate assumptions are different across industries, countries, and domains. From the standpoint of economic theory, if the liability cash flows can be hedged perfectly by a set of market-priced assets, the discount rate can be determined by reference to the discount rate for the assets. For example, if the pension plan liabilities are fixed (without any uncertainty), the discount rate should be the risk-free rate with reference to the duration of the liability cash flows—for example, a five-year zero-coupon bond yield for a liability with a (modified) duration of 5. In other cases, it can be difficult to find a fully hedged portfolio because an ongoing DB pension plan's liabilities will depend upon future economic growth and inflation, which are clearly uncertain. Even a frozen pension plan can possess uncertainty due to the changing longevity of the retirees over the long-term future.

3.2 Approaches to Liability-Relative Asset Allocation

Various approaches to liability-relative asset allocation exist. These methods are influenced by tradition, regulations, and the ability of the stakeholders to understand and extend portfolio models that come from the asset-only domain.

There are several guiding principles. The first is to gain an understanding of the make-up of the investor's liabilities and especially the factors that affect the amount and timing of the cash outflows. Given this understanding, the present value of the liabilities is calculated, along with the surplus and funding ratio. These measures are used to track the results of ongoing investment and funding policies and for other tasks. Next come the decisions regarding the asset allocation taking account of the liabilities. There are a number of ways to proceed. We will discuss three major approaches:

- *Surplus optimization*. This approach involves applying mean–variance optimization (MVO) to an efficient frontier based on the volatility of the surplus ("surplus volatility," or "surplus risk") as the measure of risk. Surplus optimization is thus an extension of MVO based on asset volatility.[18] Depending on context, surplus risk may be stated in money or percentage terms ("surplus return volatility" is then another, more precise term for this measure).
- *Hedging/return-seeking portfolios approach*. This approach involves separating assets into two groups: a hedging portfolio and a return-seeking portfolio. The reading also refers to this as the two-portfolio approach. The concept of allocating assets to two distinct portfolios can be applied for various funding ratios, but the reading distinguishes as the basic approach the case in which there is a positive surplus available to allocate to the return-seeking portfolio.
- *Integrated asset–liability approach*. For some institutional investors, such as banks and insurance companies and long–short hedge funds, asset and liability decisions can be integrated and jointly optimized.

We cover these three approaches in turn.

18 Among the papers that discuss the surplus optimization model are Leibowitz and Henriksson (1988); Mulvey (1989, 1994); Sharpe and Tint (1990); Elton and Gruber (1992).

3.2.1 *Surplus Optimization*

Surplus optimization involves adapting asset-only mean–variance optimization by substituting surplus return for asset return over any given time horizon. The quadratic optimization program involves choosing the asset allocation (mix) that maximizes expected surplus return net of a penalty for surplus return volatility at the chosen time horizon. The objective function is

$$U_m^{LR} = E\left(R_{s,m}\right) - 0.005\lambda\sigma^2\left(R_{s,m}\right) \quad (2)$$

where U_m^{LR} is the surplus objective function's expected value for a particular asset mix m; $E(R_{s,m})$ is the expected surplus return for asset mix m, with surplus return defined as (Change in asset value – Change in liability value)/(Initial asset value); and the parameter λ (lambda) indicates the investor's risk aversion. The more risk averse the investor, the greater the penalty for surplus return volatility. Note that the change in liability value (liability return) measures the time value of money for the liabilities plus any expected changes in the discount rate and future cash flows over the planning horizon.

This surplus efficient frontier approach is a straightforward extension of the asset-only portfolio model. Surplus optimization assumes that the relationship between the value of liabilities and the value of assets can be approximated through a correlation coefficient. Surplus optimization exploits natural hedges that may exist between assets and liabilities as a result of their systematic risk characteristics.

The following steps describe the surplus optimization approach:

1. Select asset categories and determine the planning horizon. One year is often chosen for the planning exercise, although funding status analysis is based on an analysis of all cash flows.
2. Estimate expected returns and volatilities for the asset categories and estimate liability returns (expanded matrix).
3. Determine any constraints on the investment mix.
4. Estimate the expanded correlation matrix (asset categories and liabilities) and the volatilities.[19]
5. Compute the surplus efficient frontier and compare it with the asset-only efficient frontier.
6. Select a recommended portfolio mix.

Exhibit 25 lists LOWTECH's asset categories and current allocation for a one-year planning horizon. The current allocation for other asset categories, such as cash, is zero. LOWTECH has been following an asset-only approach but has decided to adopt a liability-relative approach. The company is exploring several liability-relative approaches. With respect to surplus optimization, the trustees want to maintain surplus return volatility at a level that tightly controls the risk that the plan will become underfunded, and they would like to keep volatility of surplus below US$0.25 billion (10%).

19 A covariance matrix is computed by combining the correlation matrix and the volatilities.

Exhibit 25 Asset Categories and Current Allocation for LOWTECH

	Private Equity	Real Estate	Hedge Funds	Real Assets	US Equities	Non-US Equities (Developed Markets)	Non-US Equities (Emerging Markets)	US Corporate Bonds
Allocation	20.0%	12.0%	18.0%	7.0%	15.0%	12.0%	8.0%	8.0%

The second step is to estimate future expected asset and liability returns, the expected present value of liabilities, and the volatility of both assets and PV(liabilities). The capital market projections can be made in several ways—based on historical data, economic analysis, or expert judgment, for example. The plan sponsor and its advisers are responsible for employing one or a blend of these approaches. Exhibit 26 shows the plan sponsor's capital market assumptions over a three- to five-year horizon. Note the inclusion of the present value of liabilities in Exhibit 26.

Exhibit 26 LOWTECH's Capital Market Assumptions: Expected Annual Compound Returns and Volatilities

	Private Equity	Real Estate	Hedge Funds	Real Assets	US Equities	Non-US Equities (Developed Markets)	Emerging Markets	US Corporate Bonds	Cash	PV (Liabilities)
Expected returns	8.50%	7.50%	7.00%	6.00%	7.50%	7.20%	7.80%	4.90%	1.00%	4.90%
Volatilities	14.20%	9.80%	7.70%	6.10%	18.00%	19.50%	26.30%	5.60%	1.00%	5.60%

Typically, in the third step, the investor imposes constraints on the composition of the asset mix, including policy and legal limits on the amount of capital invested in individual assets or asset categories (e.g., a constraint that an allocation to equities must not exceed 50%). In our example, we simply constrain portfolio weights to be non-negative and to sum to 1.

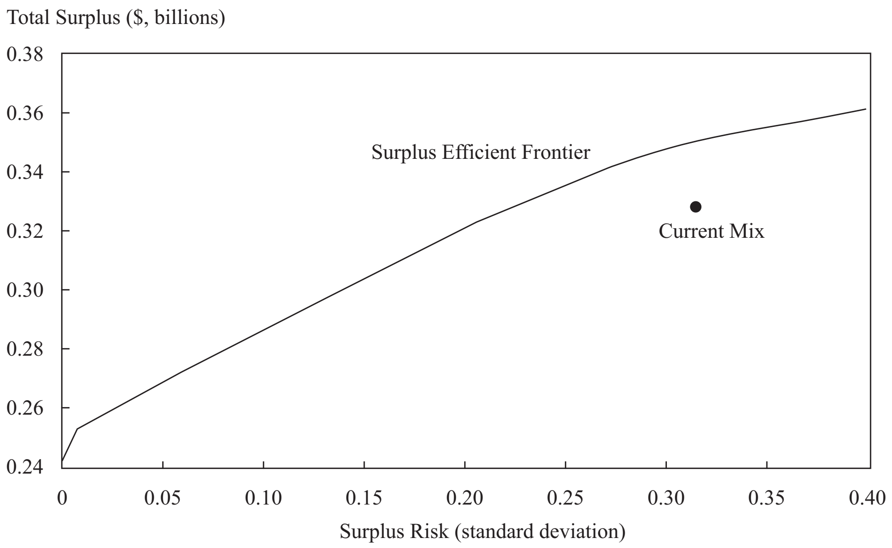
The fourth step is to estimate the correlation matrix and volatilities. We assume that the liabilities have the same expected returns and volatilities as US corporate bonds; thus, the expanded matrix has a column and a row for liabilities with values equal to the corporate bond values. For simplicity, the investor may employ historical performance. Exhibit 27 shows the correlation matrix of asset categories based on historical quarterly returns. Recall that we assume that liability returns (changes in liabilities) are driven by changes in the returns of US corporate bonds. An alternative approach is to deploy a set of underlying factors that drive the returns of the assets. Factors include changes in nominal and real interest rates, changes in economic activity (such as employment levels), and risk premiums. This type of factor investment model can be applied in an asset-only or a liability-relative asset allocation context.

Exhibit 27 Correlation Matrix of Returns

	Private Equity	Real Estate	Hedge Funds	Real Assets	US Equities	Non-US Equities (Developed Markets)	Non-US Equities (Emerging Markets)	US Corporate Bonds	Cash	PV (Liabilities)
Private equity	1	0.41	0.57	0.32	0.67	0.59	0.49	−0.27	0	−0.27
Real estate	0.41	1	0.45	0.41	0.31	0.33	0.17	−0.08	0	−0.08
Hedge funds	0.57	0.45	1	0.11	0.68	0.61	0.54	−0.23	0	−0.23
Real assets	0.32	0.41	0.11	1	0.04	0.06	−0.06	0.34	0	0.34
US equities	0.67	0.31	0.68	0.04	1	0.88	0.73	−0.38	0	−0.38
Non-US equities (developed)	0.59	0.33	0.61	0.06	0.88	1	0.81	−0.39	0	−0.39
Non-US equities (emerging)	0.49	0.17	0.54	−0.06	0.73	0.81	1	−0.44	0	−0.44
US corporate bonds	−0.27	−0.08	−0.23	0.34	−0.38	−0.39	−0.44	1	0	1
Cash	0	0	0	0	0	0	0	0	1	0
PV(liabilities)	−0.27	−0.08	−0.23	0.34	−0.38	−0.39	−0.44	1	0	1

Exhibit 28 shows a surplus efficient frontier that results from the optimization program based on the inputs from Exhibits 26 and 27. Surplus risk (i.e., volatility of surplus) in money terms (US$ billions) is on the x-axis, and expected surplus in money terms (US$ billions) is on the y-axis. By presenting the efficient frontier in money terms, we can associate the level of risk with the level of plan surplus, US$0.239 billion. Like the asset-only efficient frontier, the surplus efficient frontier has a concave shape.

Exhibit 28 Surplus Efficient Frontier



The first observation is that the current mix in Exhibit 28 lies below the surplus efficient frontier and is thus suboptimal.[20] We can attain the same expected total surplus as that of the current mix at a lower level of surplus volatility by choosing the portfolio on the efficient frontier at the current mix's level of expected total surplus. Another observation is that by uncovering the implications of asset mixes for surplus and surplus volatility, this approach allows the deliberate choice of an asset allocation in terms of the tolerable level of risk in relation to liabilities. It may be the case, for example, that neither the surplus volatility of the current mix nor that of the efficient mix with equal expected surplus is the appropriate level of surplus risk for the pension.

The surplus efficient frontier in Exhibit 28 shows efficient reward–risk combinations but does not indicate the asset class composition of the combinations. Exhibit 29 shows the asset class weights for surplus efficient portfolios.

Exhibit 29 Surplus Efficient Frontier Asset Allocation Area Graph

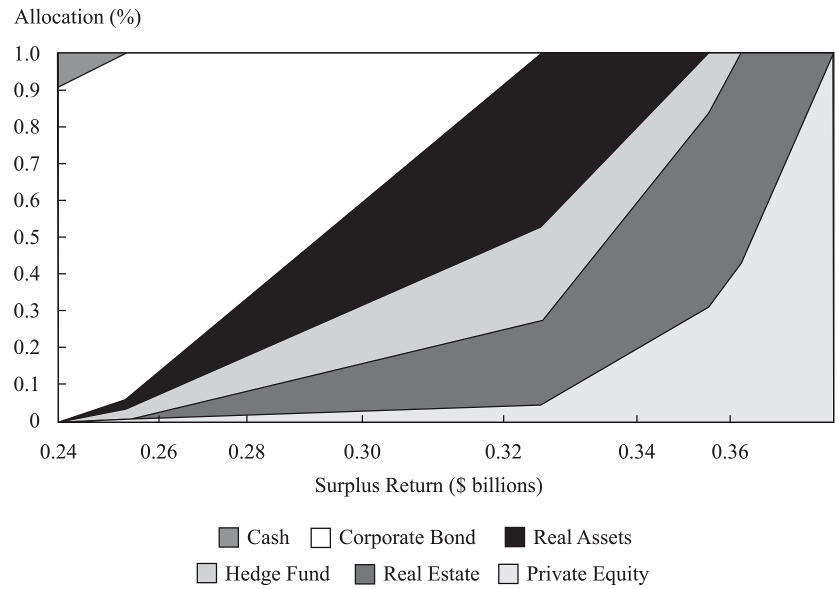
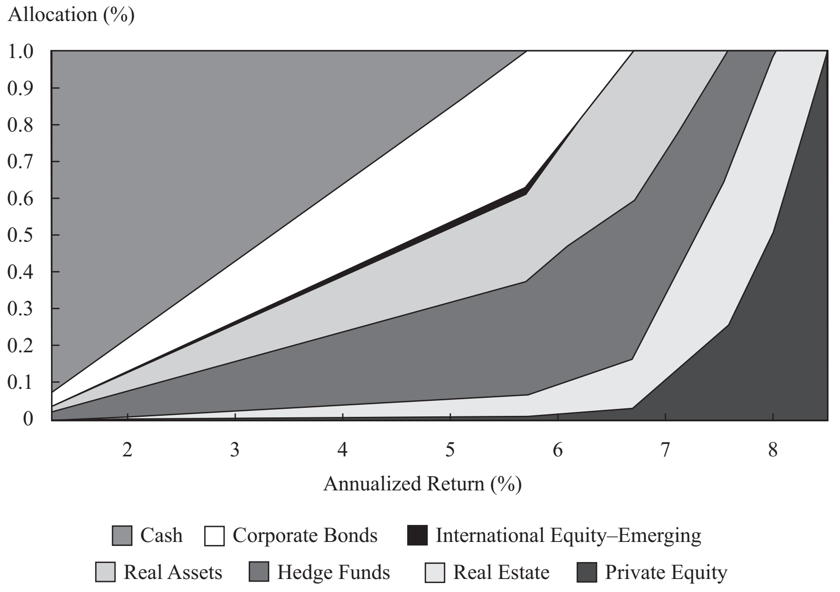


Exhibit 30, showing weights for portfolios on the usual *asset-only* efficient frontier based on the same capital market assumptions reflected in Exhibit 29, makes the point that efficient portfolios from the two perspectives are meaningfully different.[21]

20 The current mix can also be shown to lie below the asset-only mean–variance frontier.

21 In Exhibit 30, the annualized percentage returns can be equated to monetary surplus returns by multiplying by the asset value, US$2.5 billion.

Exhibit 30 Asset-Only Efficient Frontier Asset Allocation Area Graph



The asset mixes are very different on the conservative side of the two frontiers. The most conservative mix for the surplus efficient frontier (in Exhibit 29) consists mostly of the US corporate bond index (the hedging asset) because it results in the lowest volatility of surplus over the one-year horizon. Bonds are inversely correlated with changes in the present value of the frozen liability cash flows (because the liabilities indicate negative cash flows). In contrast, the most conservative mix for the asset-only efficient frontier (in Exhibit 30) consists chiefly of cash. As long as there is a hedging asset and adequate asset value, the investor can achieve a very low volatility of surplus, and for conservative investors, the asset value at the horizon will be uncertain but the surplus will be constant (or as constant as possible).

The two asset mixes (asset-only and surplus) become similar as the degree of risk aversion decreases, and they are identical for the most aggressive portfolio (private equity). Bonds disappear from the frontier about halfway between the most conservative and the most aggressive mixes, as shown in Exhibits 29 and 30.

To summarize, the current asset mix is moderately aggressive and below the surplus efficient frontier. Thus, a mean–variance improvement is possible: either higher expected surplus with the same surplus risk or lower surplus risk for the same expected surplus. The current portfolio is also poorly hedged with regard to surplus volatility; the hedging asset (long bonds in this case) has a low commitment.

The LOWTECH plan has been frozen, and the investment committee is interested in lowering the volatility of the surplus. Accordingly, it seems appropriate to choose an asset allocation toward the left-hand side of the surplus efficient frontier. For instance, a surplus efficient portfolio with about 60% bonds and the remainder in other assets (as can be approximately identified from Exhibit 29) will drop surplus volatility by about 50%.

In the end, the investment committee for the plan sponsor and its advisers and stakeholders are responsible for rendering the best decision, taking into account all of the above considerations. And as always, the recommendations of a portfolio-modeling exercise are only as good as the input data and assumptions.

Multi-Period Portfolio Models

The traditional mean–variance model assumes that the investor follows a buy-and-hold strategy over the planning horizon. Thus, the portfolio is not rebalanced at intermediate dates. A portfolio investment model requires multiple time periods if rebalancing decisions are to be directly incorporated into the model. Mulvey, Pauling, and Madey (2003) discuss the pros and cons of building and implementing multi-period portfolio models. Applicable to both asset-only and liability-relative asset allocation, multi-period portfolio models are more comprehensive than single-period models but are more complex to implement. These models are generally implemented by means of the integrated asset–liability methods discussed in Section 3.2.1

EXAMPLE 6

Surplus Optimization

1 Explain how surplus optimization solutions differ from mean–variance optimizations based on asset class risk alone.
2 What is a liability return?
3 Compare the composition of a surplus optimal portfolio at two points on the surplus efficient frontier. In particular, take one point at the lower left of the surplus frontier (surplus return = US$0.26 billion) and the other point higher on the surplus efficient frontier (surplus return = US$0.32 billion). Refer to Exhibit 29. Explain the observed relationship in terms of the use of corporate bonds as the hedging asset for the liabilities.

Solution to 1:

The surplus optimization model considers the impact of asset decisions on the (Market value of assets − Present value of liabilities) at the planning horizon.

Solution to 2:

Liability returns measure the time value of money for the liabilities plus any expected changes in the discount rate over the planning horizon.

Solution to 3:

Whereas the portfolio at the US$0.26 billion surplus return point on the efficient frontier has a substantial position in corporate bonds, the efficient mix with US$0.32 billion surplus return does not include them. The observed relationship that the allocation to corporate bonds declines with increasing surplus return can be explained by the positive correlation of bond price with the present value of liabilities. The hedging asset (corporate bonds) is employed to a greater degree at the low end of the surplus efficient frontier.

3.2.2 *Hedging/Return-Seeking Portfolio Approach*

In this approach, the liability-relative asset allocation task is divided into two parts. We distinguish as "basic" the two-portfolio approach in the case in which there is a surplus available to allocate to a return-seeking portfolio and as "variants" the approach as applied when there is not a positive surplus. In the basic case, the first part of the asset allocation task consists of hedging the liabilities through a hedging portfolio.

In the second part, the surplus (or some part of it) is allocated to a return-seeking portfolio, which can be managed independently of the hedging portfolio (for example, using mean–variance optimization or another method). An essential issue involves the composition of the hedging portfolio. In some cases, such as the LOWTECH frozen DB pension plan, the hedging portfolio is straightforward to identify. The designated cash flows can be hedged via cash flow matching, duration matching, or immunization (as explained in the fixed-income readings). This hedge will support the future cash flows with little or no risk.

In LOWTECH's application of the basic two-portfolio approach, the small surplus causes the pension plan to invest most of its capital in the hedging portfolio. The hedging portfolio can be approximated by the long-bond indexed investment as a first cut. Thus, given a 4% discount rate, US$2.261 billion is placed in long bonds. The remaining US$0.239 billion is invested in a portfolio of higher expected return assets, such as stocks, real estate, and hedge funds. This approach guarantees that the capital is adequate to pay future liabilities, as long as the hedging portfolio does not experience defaults.

Note that if the discount rate were 2% rather than 4%, the pension plan would be underfunded even if all assets were placed in a hedging portfolio. In such a case, the pension plan sponsor would either develop a strategy to increase the funding ratio so that the liabilities would be eventually paid or apply a variant of the two-portfolio approach. An underfunded plan will require higher contributions from the sponsor than a plan that is fully funded or overfunded.

The basic two-portfolio approach is most appropriate for conservative investors, such as insurance companies, and for overfunded pension plans that wish to reduce or eliminate the risk of not being able to pay future liabilities.

Several variants of the two-portfolio approach are possible. These include a partial hedge, whereby capital allocated to the hedging portfolio is reduced in order to generate higher expected returns, and dynamic versions whereby the investor increases the allotment to the hedging portfolio as the funding ratio increases. The specification of this allotment is often referred to as the liability glide path. These variants do not hedge the liabilities to the full extent possible given the assets and thus are less conservative than the basic approach discussed above. Still, there can be benefits to a partial hedge when the sponsor is able to increase contributions if the funding ratio does not increase in the future to 1 or above.

In the following discussion, we focus on determining the hedging portfolio.

Forming the Hedging Portfolio The hedging portfolio must include assets whose returns are driven by the same factor(s) that drive the returns of the liabilities. Otherwise, even if the assets and liabilities start with equal values, the assets and liabilities will likely become inconsistent over time. One example involves promises (cash outflows) that are dependent upon future inflation. The hedging portfolio in this situation would often include index-linked (inflation-linked) Treasury bonds, again cash matched to the liabilities or immunized to the degree possible.

If there is an active market for the hedging portfolio (securities) in question, the present value of future cash flows is equal to a market value of the assets contained in the hedging portfolio. In this case, the date of valuation for the assets must be the same as the date of valuation for the liabilities. Absent market values, some form of appraised value is used.

The task of forming the hedging portfolio is complicated by the discount rate assumption and by the need to identify assets that are driven by the same factors that affect the liabilities. For example, if the discount rate is set by reference to a marketable instrument, such as the long government bond index, but the liability cash flows are driven by a factor such as inflation, the hedging task may require the use of instruments beyond nominal bonds (perhaps multiple instruments, such as interest rate swaps, inflation-linked bonds, and real assets). And in many applications, the hedge cannot be fully accomplished due to the nature of the driving factors (e.g., if they are non-marketable factors, such as economic growth).

If the uncertainties in the cash flows are related to non-market factors, such as future salary increases, the discount rate will depend upon regulations and tradition. Clearly, high discount rates lead to high funding ratios and in most cases require lower contributions from the sponsoring organization (at least in the short run). Conversely, lower discount rates give rise to lower funding ratios and thereby higher contributions. In the former case, investors with high discount rates will need to generate higher asset returns to achieve their promises if the pension plan sponsor wishes to avoid future contributions. A more conservative route is to designate a lower discount rate, as is the case in much of Europe and Asia. In all cases, it is the regulator's responsibility to set the guidelines, rules, and penalties involved in determining contribution policy.

Several issues complicate the valuation of liability cash flows. In many situations, investors must satisfy their promises without being able to go to a market and purchase a security with positive cash flows equal in magnitude to the liability cash flows.

At times, uncertain liabilities can be made more certain through the law of large numbers. For example, life insurance companies promise to pay beneficiaries when a policyholder dies. The life insurance company can minimize the risk of unexpected losses by insuring large numbers of individuals. Then, valuation of liabilities will use present value of expected cash flows based on a low (or even zero) risk premium in the discount rate. The field of application of the law of large numbers can be limited. For example, averages do not eliminate longevity risk.

Limitations The basic two-portfolio approach cannot be directly applied under several circumstances. First, if the funding ratio is less than 1, the investor cannot create a fully hedging portfolio unless there is a sufficiently large positive cash flow (contribution). In this case, the sponsor might increase contributions enough to generate a positive surplus. As an alternative, there are conditional strategies that might help improve the investor's funding ratio, such as the glide path rules.[22]

A second barrier occurs when a true hedging portfolio is unavailable. An example involves losses due to weather-related causes, such as hurricanes or earthquakes. In these cases, the investor might be able to partially hedge the portfolio with instruments that share some of the same risks. The investor has "basis risk" when imperfect hedges are employed. (As an aside, the investor might be able to set up a contract with someone who, for a fee, will take on the liability risk that cannot be hedged. Insurance contracts have this defining characteristic.)

EXAMPLE 7

The Hedging/Return-Seeking Portfolios Approach

1. Compare how surplus optimization and the hedging/return-seeking portfolio approach take account of liabilities.
2. How does funding status affect the use of the basic hedging/return-seeking portfolio approach?

22 See Gannon and Collins (2009).

Solution to 1:

The surplus optimization approach links assets and the present value of liabilities through a correlation coefficient. The two-portfolio model does not require this input. Surplus optimization considers the asset allocation problem in one step; the hedging/return-seeking portfolio approach divides asset allocation into two steps.

Solution to 2:

Implementation of the basic two-portfolio approach depends on having an overfunded plan. A variant of the two-portfolio approach might be applied, however. Surplus optimization does not require an overfunded status. Both approaches address the present value of liabilities, but in different ways.

3.2.3 *Integrated Asset–Liability Approach*

The previous two approaches are most appropriate when asset allocation decisions are made after, and relatively independently of, decisions regarding the portfolio of liabilities. However, there are numerous applications of the liability-relative perspective in which the institution must render significant decisions regarding the composition of its liabilities *in conjunction with the asset allocation*. Banks, long–short hedge funds (for which short positions constitute liabilities), insurance companies, and re-insurance companies routinely fall into this situation. Within this category, the liability-relative approaches have several names, including asset–liability management (ALM) for banks and some other investors and dynamic financial analysis (DFA) for insurance companies. These approaches are often implemented in the context of multi-period models. Using the following two cases, we review the major issues.

Integrated Asset–Liability Approach for Property/Casualty Insurance Companies

A property/casualty insurance company must make asset investment decisions in conjunction with business decisions about the portfolio of insured properties, its liabilities. To that end, asset and liability decisions are frequently integrated in an enterprise risk management system. In fact, the liability portfolio is essential to the company's long-term viability. For example, a particular property/casualty (PC) insurance company might engage (accept) liabilities for catastrophic risks such as earthquakes and hurricanes. In this case, the liabilities depend upon rare events and thus are most difficult to hedge against. Specialized firms calculate insured losses for a chosen set of properties for property/casualty insurance companies, and these firms provide liability cash flows on a probabilistic (scenario) basis. In this way information is gathered about the probability of losses over the planning horizon and the estimated losses for each loss event. An important issue involves the amount of capital needed to support the indicated liabilities. This issue is addressed by evaluating the tail risks, such as the 1% Value-at-Risk or Conditional-Value-at-Risk amount. To reduce this risk, there are major advantages to forming a diversified global portfolio of liabilities and rendering asset allocation decisions in conjunction with the liability portfolio decisions. The hedging portfolio in this case is not well defined. Therefore, it is difficult to hedge liabilities for a book of catastrophic risk policies. Liabilities might be addressed via customized products or by purchasing re-insurance. The assets and liabilities are integrated so that the worst-case events can be analyzed with regard to both sides of the balance sheet.

Integrated Asset–Liability Approach for Banks

Large global banks are often required to analyze their ability to withstand stress scenarios, in accordance with the Basel III framework. These institutions must be able to show that their current capital is adequate to withstand losses in their business units, such as asset trading, in conjunction with increases in liabilities. The chief risk officer evaluates these scenarios by means of integrated asset–liability approaches. The asset and liability decisions are linked in an enterprise manner. Both the portfolio of assets and the portfolio of liabilities have major impacts on the organization's risk. Thus, decisions to take on new products or expand an existing product—thereby generating liabilities—must take into account the associated decisions on the asset side. The integrated asset–liability management system provides a mechanism for discovering the optimal mix of assets and liabilities (products). These applications often employ multi-period models via a set of projected scenarios.

Decisions about asset allocation will affect the amount of business available to a financial intermediary, such as a bank or insurance company. Similarly, decisions about the portfolio of liabilities and concentration risks will feed back to the asset allocation decisions. Accordingly, we can set up a linked portfolio model. In a similar fashion, the performance of the assets of an institution possessing quasi-liabilities, such as a university endowment, will affect the spending rules for the institution. We can reduce worst-case outcomes by adjusting spending during crash periods, for example. Portfolio models linked to liabilities can provide significant information, helping the institution make the best compromise decisions for both the assets and the liabilities under its control. The twin goals are to maximize the growth of surplus over time subject to constraints on worst-case and other risk measures relative to the institution's surplus.

3.2.4 *Comparing the Approaches*

We have introduced three approaches for addressing asset allocation decisions in the context of liability issues; Exhibit 31 summarizes their characteristics. Each of these approaches has been applied in practice. The surplus optimization approach is a straightforward extension of the traditional (asset-only) mean–variance model. Surplus optimization demonstrates the importance of the hedging asset for risk-averse investors and provides choices for investors who are less risk averse in the asset mixes located on the middle and the right-hand side of the efficient frontier. The assumptions are similar to those of the traditional Markowitz model, where the inputs are expected returns and a covariance matrix. Thus, the assets and liabilities are linked through correlation conditions. The second approach, separating assets into two buckets, has the advantage of simplicity. The basic approach is most appropriate for conservative investors, such as life insurance companies, and for overfunded/fully funded institutional investors that can fully hedge their liabilities. Another advantage of this approach is a focus on the hedging portfolio and its composition. The hedging portfolio can be constructed using a factor model and then linked to the assets via the same factors. Unfortunately, underfunded investors do not have the luxury of fully hedging their liabilities and investing the surplus in the risky portion; they must apply variants of the two-portfolio approach. The third approach, integrating the liability portfolio with the asset portfolio, is the most comprehensive of the three. It requires a formal method for selecting liabilities and for linking the asset performance with changes in the liability values. This approach can be implemented in a factor-based model, linking the assets and liabilities to the underlying driving factors. It has the potential to improve the institution's overall surplus. It does not require the linear correlation assumption and is capable of modeling transaction costs, turnover constraints, and other real-world constraints. The capital required for this approach is often determined by reference to the output of integrated asset–liability systems in banks and property/casualty insurance and re-insurance companies.

Exhibit 31 Characteristics of the Three Liability-Relative Asset Allocation Approaches

Surplus Optimization	Hedging/Return-Seeking Portfolios	Integrated Asset–Liability Portfolios
Simplicity	Simplicity	Increased complexity
Linear correlation	Linear or non-linear correlation	Linear or non-linear correlation
All levels of risk	Conservative level of risk	All levels of risk
Any funded ratio	Positive funded ratio for basic approach	Any funded ratio
Single period	Single period	Multiple periods

EXAMPLE 8

Liability-Relative Asset Allocation: Major Approaches

1 Discuss how the probability of not being able to pay future liabilities when they come due is or is not addressed by each of the major approaches to liability-relative asset allocation.

2 What are the advantages of the three approaches for investors who are more interested in protecting the surplus than growing their assets? Assume that the investor has a positive surplus.

Solution to 1:

Such issues are best addressed by means of multi-period integrated asset–liability models. Surplus optimization and the two-portfolio approach, being single-period models, have difficulty estimating the probability of meeting future obligations.

Solution to 2:

The three liability-relative approaches are appropriate for conservative investors (investors who are more interested in protecting the surplus than growing their assets). All of the three approaches force investors to understand the nature of their liabilities. This type of information can help inform the decision-making process.

3.3 Examining the Robustness of Asset Allocation Alternatives

As part of a liability-relative asset allocation study, the institutional investor can evaluate performance over selected events and “simulated” historical time periods. Each of the selected events can be interpreted as a “what if” sensitivity analysis. For example, we might wish to consider the effect of a 100 bp increase in interest rates across all maturities—that is, a parallel shift in the yield curve. This event would have a significant impact on the value of government bonds, clearly. Also, there would be a corresponding positive impact on the present discounted value of liabilities that are discounted at the government bond rate. The effect on other liability-relative asset allocation elements is less direct, and assumptions must be made. Suppose, for example, that the investor must discount at the high-quality corporate rate. In that case, we need to estimate the effect of changing government rates on corporate rates. These designated studies are part of the stress tests required by banking and other regulators.

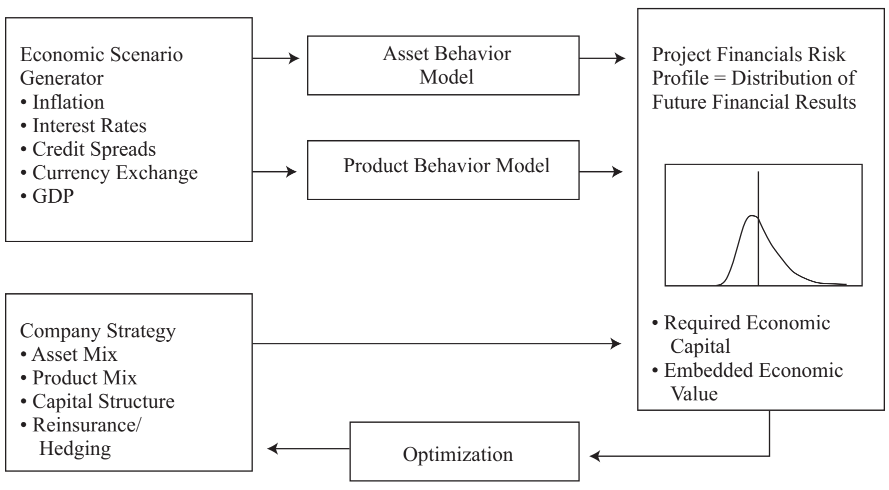
Another type of event study is the construction of scenarios based on carefully selected historical time periods. For example, we might select late 2008 as a reference point. In such a scenario, we are interested in the changes in the economic factors and the associated changes in the values of the institution's assets and liabilities. What would be the impact on our current (or projected) portfolio—assets and PV(liabilities)—if the conditions seen in late 2008 occurred again?

A more comprehensive method for examining robustness involves setting up a multi-stage simulation analysis. Here, we use scenarios to model uncertainty and replace decisions with "rules." The process begins with a set of scenarios for the underlying driving economic factors. Each scenario designates a path for the asset returns and the liability values at each stage of the planning horizon. The result is a set of probabilistic outcomes for the institutional investor's asset portfolio and the cash flows for its liabilities. In such modeling, one must take care to be consistent between asset returns and corresponding liabilities within a scenario; for example, if interest rates are a common factor driving both asset performance and the PV (liabilities), the interest rate effects should be based on the same assumptions.

Through the scenario analysis, the probability of both good and bad outcomes can be estimated. For example, we can measure the probability that an institutional investor will make a capital contribution in the future. Exhibit 32 shows the decision structure for the simulation of an insurance company over several periods, including modeling of the company's business strategy and the required capital rules.

To evaluate robustness, we can apply the simulation system with different assumptions. For instance, if we change the expected return of US equities, what is the effect on the probability of meeting the liabilities over an extended horizon, such as 10 years? This type of sensitivity analysis is routinely done in conjunction with the modeling exercise.

Exhibit 32 Simulation Analysis



3.4 Factor Modeling in Liability-Relative Approaches

A factor-based approach for liability-relative asset allocation has gained interest and credibility for several reasons. First, in many applications, the liability cash flows are dependent on multiple uncertainties. The two primary macro factors are future economic conditions and inflation. Many pension payments to beneficiaries will be based on inflation and salary changes over the employees' work span. A fully hedged portfolio cannot be constructed when the liabilities are impacted by these uncertain factors. Recall that a hedged portfolio can be constructed for a frozen plan with fixed liabilities. For ongoing pension schemes, the best that can be done is to add asset categories to the portfolio that are positively correlated with the underlying driving risk factors, such as inflation-linked bonds. A factor-based approach can be implemented with any of the three liability-relative asset allocation methods discussed above.

EXAMPLE 9

Robustness and Risk Assessment in Liability-Relative Asset Allocation

What types of sensitivity analysis can be evaluated with a multi-period ALM simulation system?

Solution:

To provide estimates of the probability of meeting future obligations and the distribution of outcomes, several types of sensitivity analysis are likely to be performed.

- For example, the expected returns could be increased or decreased to evaluate the impact on future contributions to the plan.
- Likewise, by analyzing historical events, the investor can estimate the size of losses during crash periods and make decisions about the best asset allocation to protect against these worst-case events. Multiple risk measures over time (temporal risk measures) can be readily included in a simulation system.

4 DEVELOPING GOALS-BASED ASSET ALLOCATIONS

In this section, we review the concept of goals-based asset allocation, focusing first on the rationale behind this different approach and its investment implications. We then discuss the major elements of the process, illustrating them with specific, simplified examples when necessary. We conclude with a discussion of the applicability of the approach and its major shortcomings.

A goals-based asset allocation process disaggregates the investor's portfolio into a number of sub-portfolios, each of which is designed to fund an individual goal (or "mental account") with its own time horizon and required probability of success. The literature behind the development of this approach is very rich. Initially, goals-based wealth management was specifically proposed by a small group of practitioners,[23] each of whom offered his own solution for taking into account the tendency of individuals

23 See Brunel (2003, 2005); Nevins (2004); Pompian and Longo (2004); Chhabra (2005).

to classify money into non-fungible mental accounts. Shefrin and Statman (2000) developed the concept of the behavioral portfolio, which can be related to the Maslow (1943) hierarchy of needs. Das, Markowitz, Scheid, and Statman (2010, 2011) showed that traditional and behavioral finance could be viewed as equivalent if one were prepared to change the definition of risk from volatility of returns to the probability of not achieving a goal.[24] The essential point is that optimality requires both a suitably structured portfolio that can meet the given need *and* the correct capital allocation based on an appropriate discount rate, reflecting considerations of time horizon and the required probability of success.

Individuals have needs that are different from those of institutions. The most important difference is that individuals often have multiple goals, each with its own time horizon and its own "urgency," which can be expressed as a specific required probability of success. Exhibit 33 summarizes differences in institutional and individual investor definitions of goals. An individual's goals are not necessarily mutually compatible in two senses: The investor may not be able to address them all given the financial assets available, and there may be internal contradictions among the goals. An alternative process using one set of overall investment objectives—and thus effectively ignoring or "averaging" the different time horizons and required probabilities of success of individual goals—ostensibly loses the granular nature of client goals; as a result, the inherent complexities of the investment problem are less likely to be addressed fully. An approach that breaks the problem into sub-portfolios carries a higher chance of fully addressing an investor's goals, although it may require several iterations to ensure that the investor's portfolio is internally consistent and satisfactory.

Exhibit 33 Institutional and Individual Ways of Defining Goals

	Institutions	Individuals
Goals	Single	Multiple
Time horizon	Single	Multiple
Risk measure	Volatility (return or surplus)	Probability of missing goal
Return determination	Mathematical expectations[a]	Minimum expectations
Risk determination	Top-down/bottom-up	Bottom-up
Tax status	Single, often tax-exempt	Mostly taxable

[a] "Mathematical expectations" here means the weighted expected return of portfolio components.

The characteristics of individuals' goals have three major implications for an investment process that attempts to address the characteristics directly:

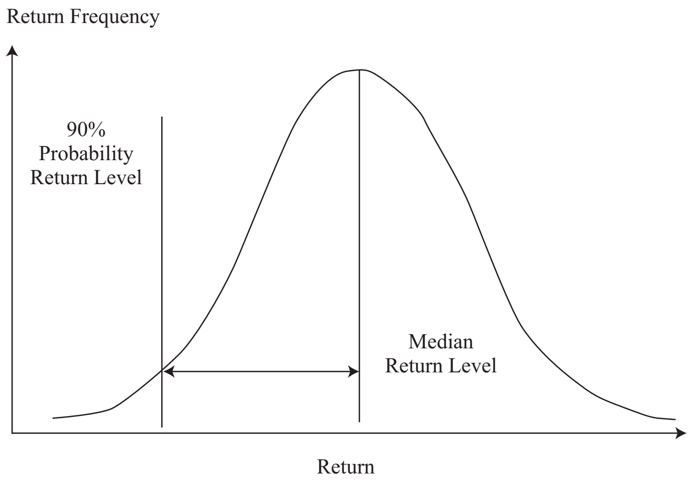
- The overall portfolio needs to be divided into sub-portfolios to permit each goal to be addressed individually.
- Both taxable and tax-exempt investments are important.
- Probability- and horizon-adjusted expectations (called "minimum expectations" in Exhibit 33) replace the typical use of mathematically expected average returns in determining the appropriate funding cost for the goal (or "discount rate" for future cash flows).

24 We apologize to these authors for grossly oversimplifying their work, but our aim is to make their insights more readily available without going into excruciating detail.

Compared with average return expectations—the median or average return anticipated for a combination of assets that is appropriate to address a goal—minimum expectations reflect a more complex concept. Minimum expectations are defined as the minimum return expected to be earned over the given time horizon with a given minimum required probability of success.

To illustrate, assume that a portfolio associated with a goal has an expected return of 7% with 10% expected volatility and the investor has indicated that the goal is to be met over the next five years with at least 90% confidence. Over the next five years, that portfolio is expected to produce returns of 35% with a volatility of 22.4%.[25] In short, this portfolio is expected to experience an average compound return of only 1.3% per year over five years with a probability of 90%; this result is quite a bit lower than the portfolio's average 7% expected return (see Exhibit 34). Thus, rather than discounting expected cash outflows by 7% to compute the dollar amount needed to defease the goal over that five-year horizon, one must use a considerably lower discount rate and by implication reserve a higher level of capital to meet that goal. Under moderate simplifying assumptions, that computation is valid whether or not return and volatility numbers are pretax or after-tax. Exhibit 34 shows, for the case of a normal distribution of returns, a return level that is expected to be exceeded 90% of the time (the 40% of the probability that lies between the vertical lines plus the 50% to the right of the median).

Exhibit 34 Probability-Weighted Return vs. Expected (= Median) Return



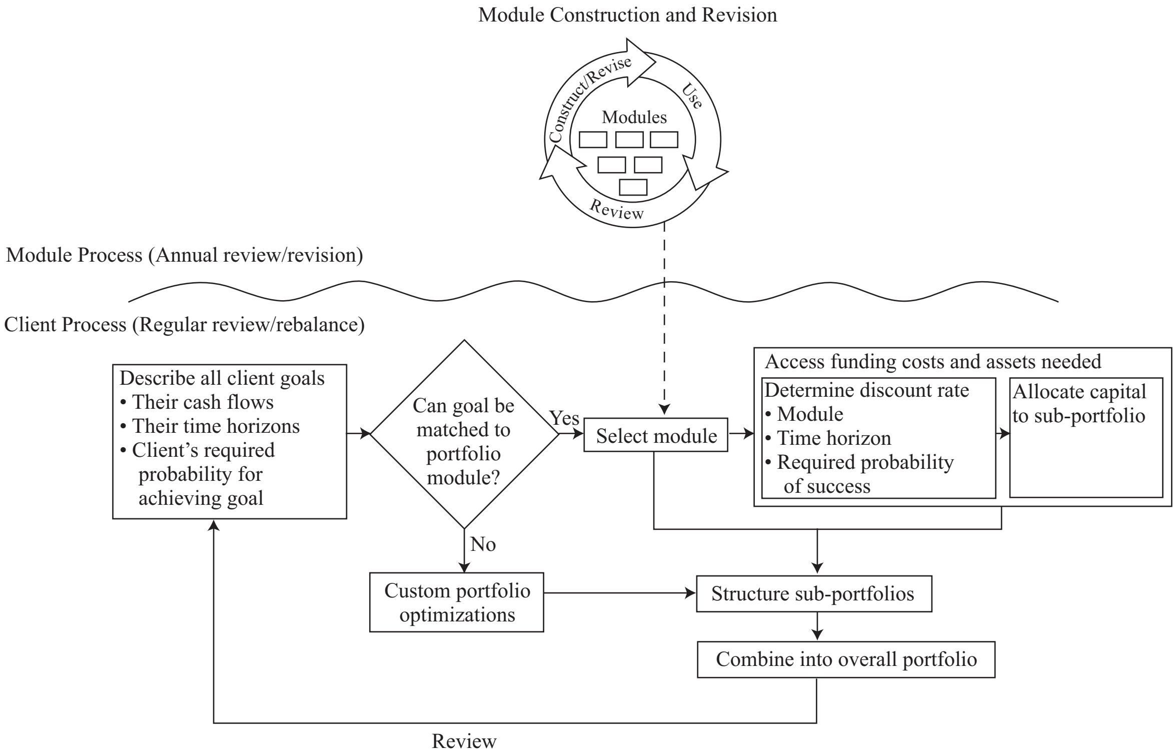
4.1 The Goals-Based Asset Allocation Process

Investment advisers taking a goals-based approach to investing client assets may implement this approach in a variety of ways. Exhibit 35 illustrates the major elements of the goals-based asset allocation process described in this reading. Ostensibly, there are two fundamental parts to this process. The first centers on the creation of portfolio modules, while the second involves identifying client goals and matching each of these goals to the appropriate sub-portfolio of a suitable asset size.

---

25 The return is the product of the annual return times the number of years, while the volatility is the product of the annual volatility times the square root of the number of years (under the assumption of independently and identically distributed returns).

Exhibit 35 A Stylized Representation of the Goals-Based Asset Allocation Process



Determining the lowest-cost funding for any given goal requires the formulation of an optimized portfolio that will be used to defease that goal optimally in the sense that risks are not taken for which the investor is not fairly compensated. Note that this process is most often generic and internal to the adviser and his or her firm. The adviser will typically not create a specific sub-portfolio for each goal of each client but rather will select, from a pre-established set, one of a few modules—or model portfolios—that best meet each goal.[26] As discussed above, adjusting the expected return on that portfolio to account for the time horizon and the required probability of success allows one to formulate the relevant discount rate which, when applied to the expected cash flows, will help determine the capital required at the outset. That capital will then be invested in the optimized portfolio asset allocation, where the balance will decline until the end of the horizon, when it runs out.[27] Note that the process is somewhat iterative because individual investors may describe a certain horizon as set when in fact they view it as "the next *x* years," with the horizon rolling by one year every year. Note also that discounting needs based on probability- and

26 See the next paragraph for a discussion of when it makes sense to create specific optimal sub-portfolios.
27 An important reason for the use of a declining-balance portfolio relates to the need for individuals and families to plan for the transfer of assets at death. In order for the income from assets to be used by an individual, these assets must be in the individual's name, or at least in a structure of which he or she is a beneficiary. Such assets would then be a part of the estate of the individual. Using a declining-balance portfolio allows the individual to receive the income—and some of the principal liquidated every year—while still ensuring that the amount of assets kept in the individual's name remains as low as appropriate given the individual's goals. An exception to this scenario would be the case of families whose income needs are so modest in relation to total assets that there is no need to provide income in planning for generational transfers or families that have such large eventual philanthropic intentions that assets kept in some beneficiaries' names are meant to be transferred to charity at death.

horizon-adjusted minimum expectations naturally means that these expectations will be exceeded under "normal circumstances." Thus, it is not unusual for the funding for a goal to seem excessive with the benefit of hindsight.

Although the great majority of advisers will likely create individual client portfolios using model portfolios—precisely, pre-optimized modules—a greater degree of customization is possible. Such customization involves creating specific sub-portfolios for each goal of each client. Indeed, it is conceivable, and mathematically possible, to create an optimal sub-portfolio for each goal. In fact, in practice, one would often proceed in this way when dealing with complex situations and with clients who have highly differentiated needs and constraints.[28] The adviser may find it impossible to use pre-optimized modules if the investment constraints imposed by the client are incompatible with those used in the creation of the module set. These might include, for instance, geographical or credit emphases—or de-emphases—that conflict with the market portfolio concept. Other restrictions might concern base currency, the use of alternative strategies, or the acceptability of illiquid investments, for example. Thus, although it is feasible for advisers to create client-specific modules, this approach can become prohibitively expensive. In short, one would likely use standardized modules for most individuals, except for those whose situation is so complex as to require a fully customized approach.

Many multi-client advisers may prefer to create a set of "goal modules" whose purpose is, collectively, to cover a full range of capital market opportunities and, individually, to represent a series of return–risk trade-offs that are sufficiently differentiated to offer adequate but not excessive choices to meet all the goals they expect their clients to express. These modules should therefore collectively appear to create a form of efficient frontier, though the frontier they depict in fact does not exist because the modules may well be based on substantially different sets of optimization constraints.

The two most significant differences from one module to the next, besides the implied return–risk trade-offs, are liquidity requirements and the eligibility of certain asset classes or strategies. Additionally, while intra–asset class allocation to individual sub–asset classes or strategies may typically be guided by the market portfolio for that asset class, one can conceive of instances where the selection of a specific sub–asset class or strategy is justified, even though the asset class per se may seem inappropriate. For instance, one might agree to hold high-yield bonds in an equity-dominated portfolio because of the equity risk factor exposure inherent in lower-credit fixed income. Conversely, the fixed-income market portfolio might be limited to investment-grade bonds and possibly the base-currency-hedged variant of non-domestic investment-grade bonds. We will return to the construction of these modules in Section 4.5.

4.2 Describing Client Goals

At this point, it is important to note that individual investors do not always consider all goals as being equal and similarly well-formulated in their own minds. Thus, while certain investors will have a well-thought-out set of goals—which may at times not be simultaneously achievable given the financial assets available—others will focus only on a few "urgent" goals and keep other requirements in the background.

Thus, a first step is to distinguish between goals for which anticipated cash flows are available—whether regularly or irregularly timed across the horizon or represented by a bullet payment at some future point—and those we call "labeled goals," for which details are considerably less precise. The term "labeled" here simply means that the individual has certain "investment features" in mind—such as minimal risk, capital

28 Note that such an approach, being more complex, is also costlier. It would therefore be more likely to be economically feasible for those advisory clients who also have the ability to pay a higher fee.

preservation, purchasing power preservation, and long-term growth—but has not articulated the actual need that stands behind each label. The individual may already have mentally allocated some portion of his or her assets, in currency or percentage terms, to one or several of these labels. For cash flow–based goals,[29] the time horizon over which the goal is to be met is usually not difficult to ascertain: It is either the period over which cash outflows are expected to be made or the point in time at which a bullet payment is expected. More complex, however, is the issue of the urgency of the goal and thus of the required minimum probability of success.

By working to preserve a human (as opposed to a technical) tone in the advisory conversations, the adviser can serve the client without forcing him or her to come up with a quantified probability of success. The adviser may start with the simple observation that there are two fundamental types of goals: those that one seeks to achieve and those whose consequences one seeks to avoid. Dividing the goals the investor seeks to achieve into "needs, wants, wishes, and dreams" provides the adviser with an initial sense of the urgency of each goal. A need typically must be met and so should command a 90%–99% probability of success, while at the other end of the spectrum, it is an unfortunate fact that we all live with unfulfilled dreams, whose required probabilities of success probably fall below 60%. A parallel—and analogous—structure can be created to deal with goals one seeks to avoid:[30] "nightmares, fears, worries, and concerns," with similar implications in terms of required probabilities of success. In short, while some discussion of probability level may well take place, it can be informed and guided by the use of commonly accepted everyday words that will ensure that the outcome is internally consistent. The adviser avoids the use of jargon, which many clients dislike, and yet is able to provide professional advice.[31]

The simplest way to bring this concept to life is to work with a basic case study. Imagine a family, the Smiths, with financial assets of US$25 million. (For the sake of simplicity, we are assuming that they do not pay taxes and that all assets are owned in a single structure.) The parents are in their mid-fifties, and the household spends about US$500,000 a year. They expect that inflation will average about 2% per year for the foreseeable future. They express four important goals and are concerned that they may not be able to meet all of them:

1 They *need* a 95% chance of being able to maintain their current expenditures over the next five years.
2 They *want* an 85% chance of being able to maintain their current expenditures over the ensuing 25 years, which they see as a reasonable estimate of their joint life expectancy.
3 They *need* a 90% chance of being able to transfer US$10 million to their children in 10 years.
4 They *wish* to have a 75% chance to be able to create a family foundation, which they wish to fund with US$10 million in 20 years.

29 Note that all cash flows do not have to be negative (i.e., outflows). One can easily imagine circumstances where certain future inflows are anticipated and yet are not seen, individually, as sufficient to meet the specified goal.

30 Although negative goals may sound surprising, they do exist and play a double role. First, when a negative goal is explicitly stated, it can be "replaced" by a specific positive goal: Avoiding the nightmare of running out of capital, for example, can be turned into the need to meet a certain expense budget. Second, negative goals serve as a useful feedback loop to check the internal consistency of the investor's goal set.

31 Note that the adviser can also identify a series of "secondary" words to help determine whether a need, for instance, means that the required probability of success should be set at 99%, 95%, or 90%. An *indispensable* need could require a 99% probability of being met, while an *urgent* need might require only a 95% probability of success, and a *serious* need a 90% probability.

EXAMPLE 10

Understanding Client Goals

1. A client describes a desire to have a reserve of €2 million for business opportunities that may develop when he retires in five years. What are the important features of this goal?
2. A 70-year-old client discusses the need to be able to maintain her lifestyle for the balance of her life and wishes to leave US$3 million to be split among her three grandchildren at her death. What are the important features of this situation?

Solution to 1:

The time horizon is five years. Words such as "desire" in describing a goal, compared with expressions indicating "need," indicate that there is room for "error" in the event that capital markets are not supportive. The portfolio required to meet the goal described as a desire will likely be able to involve a riskier profile. One would want to verify this assumption by comparing the size of that goal compared with the total financial assets available to the client.

Solution to 2:

The key takeaway is that although the two goals have the same time horizon, the two portfolios designed to defease them will have potentially significantly different risk profiles. The time horizon is approximately 20 years. The first goal relates to maintaining the client's lifestyle and must be defeased with an appropriately structured portfolio. The second goal, relating to the wish to leave some money to grandchildren, will allow more room for risk taking.

4.3 Constructing Sub-Portfolios

Having defined the needs of the investor in as much detail as possible, the next step in the process is to identify the amount of money that needs to be allocated to each goal and the asset allocation that will apply to that sum. For most advisers, the process will start with a set of sub-portfolio modules (such as those we briefly discussed in Section 4.1 and will study in more depth in Section 4.5). When using a set of pre-optimized modules, the adviser will then need to identify the module best suited to each of the specific goals of the client. That process is always driven by the client's time horizon and required probability of success, and it involves identifying the module that offers the highest possible return given the investor's risk tolerance as characterized by a given required probability of success over a given time horizon.

To illustrate, consider the set of six modules shown in Exhibit 36;[32] these modules result from an optimization process that will be explained later.[33] In the exhibit, the entries for minimum expected return are shown rounded to one decimal place; subsequent calculations for required capital are based on full precision.

32 The different ranges of required probabilities of success for various time horizons reflect the fact that the differentiation across modules can occur more or less rapidly, reflecting the different ratios of return per unit of risk.

33 Exhibit 38 presents the details of the asset allocation of these modules and the constraints underpinning their optimization.

Exhibit 36 "Highest Probability- and Horizon-Adjusted Return" Sub-Portfolio Module under Different Horizon and Probability Scenarios

	A	B	C	D	E	F
Portfolio Characteristics						
Expected return	4.3%	5.5%	6.4%	7.2%	8.0%	8.7%
Expected volatility	2.7%	4.5%	6.0%	7.5%	10.0%	12.5%

	Annualized Minimum Expectation Returns					
Time Horizon (years)				5		
Required Success						
99%	1.5%	0.9%	0.2%	–0.6%	–2.4%	–4.3%
95	2.3	2.2	2.0	1.7	0.7	–0.5
90	2.7	3.0	3.0	2.9	2.3	1.5
75	3.5	4.2	4.6	4.9	5.0	4.9
Time Horizon (years)				10		
Required Success						
99%	2.3%	2.2%	2.0%	1.7%	0.7%	–0.5%
90	3.2	3.7	4.0	4.1	4.0	3.6
75	3.7	4.6	5.1	5.6	5.9	6.0%
60	4.1	5.2	5.9	6.6	7.2	7.7
Time Horizon (years)				20		
Required Success						
95%	3.3%	3.9%	4.2%	4.4%	4.4%	4.1%
90	3.5	4.3	4.7	5.0	5.2	5.1
85	3.7	4.5	5.0	5.4	5.7	5.8
75	3.9	4.9	5.5	6.0	6.5	6.8
Time Horizon (years)				25		
Required Success						
95%	3.4%	4.1%	4.4%	4.7%	4.7%	4.6%
90	3.6	4.4	4.9	5.2	5.5	5.5
85	3.7	4.6	5.2	5.6	6.0	6.1
75	3.9	4.9	5.6	6.2	6.7	7.0

In Exhibit 36, the top section, on portfolio characteristics, presents the expected return and expected volatility of each module. Below that are four sections, one for each of four time horizons: 5, 10, 20, and 25 years. In a given section, the entries are the returns that are expected for a given required probability of achieving success. For example, at a 10-year horizon and a 90% required probability of success, Modules A, B, C, D, E, and F are expected to return, respectively, 3.2%, 3.7%, 4.0%, 4.1%, 4.0%, and 3.6%. In this case, Module D would be selected to address a goal with this time horizon and required probability of success because its 4.1% expected return is higher than those of all the other modules. Thus, Module D offers the lowest "funding cost" for the given goal. The highest expected return translates to the lowest initially required capital when the expected cash flows associated with the goal are discounted using that expected return.

EXAMPLE 11

Selecting a Module

Address the following module selection problems using Exhibit 36:

1. A client describes a desire to have a reserve of €2 million for business opportunities that may develop when he retires in five years. Assume that the word "desire" points to a wish to which the adviser will ascribe a probability of 75%.
2. A 70-year-old client with a 20-year life expectancy discusses the need to be able to maintain her lifestyle for the balance of her life and wishes to leave US$3 million to be split among her three grandchildren at her death.

Solution to 1:

The time horizon is five years. Exhibit 36 shows that Module E has the highest expected return (5.0%) over the five-year period and with the assumed 75% required probability of success.

Solution to 2:

The time horizon is 20 years. The first goal is a need, while the second is a wish. We assume a required probability of success of 95% for a need and 75% for a wish. Exhibit 36 shows that Module D provides the highest horizon- and required-probability-adjusted return (4.4%) for the first goal. Module F is better suited to the second goal because, even though the second goal has the same time horizon, it involves only a 75% required probability of success; the appropriately adjusted return is 6.8%, markedly the highest, which means the initially required capital is lower.

Returning to the Smiths, let us use that same set of modules to look at their four specific goals. The results of our analysis are presented in Exhibit 37.

1. The first goal is a need, with a five-year time horizon and a 95% required probability of success. Looking at the 95% required probability line in the five-year time horizon section of Exhibit 36, we can see that the module with the highest expected return on a time horizon- and required probability-adjusted basis is Module A and that the appropriately adjusted expected return for that module is 2.3%. Discounting a US$500,000 annual cash flow, inflated by 2% a year from Year 2 onwards, required a US$2,430,000 initial investment. This amount represents 9.7% of the total financial wealth of the Smiths.
2. The second goal is a want, with a 25-year time horizon and an 85% required probability of success. The corresponding line of the table in Exhibit 36 points to Module F and a discount rate of 6.1%. Discounting their current expenses with the same assumption over the 25 years starting in Year 6 with a 6.1% rate points to an initially required capital of US$4,978,000, representing 19.9% of the Smiths' wealth.

3 The third goal is another need, with a 10-year time horizon and a 90% required probability of success. Module D is the best module, and the US$6,671,000 required capital reflects the discounting of a US$10 million payment in 10 years at the 4.1% indicated in Exhibit 36.

4 Finally, the fourth goal is a wish with a 20-year time horizon and a 75% required probability of success. Module F is again the best module, and the discounting of a US$10 million payment 20 years from now at the 6.8% expected return from Exhibit 36 points to a required capital of US$2,679,000 today.

Note that different goals may, in fact, be optimally addressed using the same module; thus, an individual module may be used more than once in the allocation of the individual's overall financial assets. Here, Goals 2 and 4 can both be met with the riskiest of the six modules, although their time horizons differ, as do the required probabilities of success, with Goal 2 being characterized as a want and Goal 4 as a wish.

Exhibit 37 Module Selection and Dollar Allocations (US$ thousands)

	Total Financial Assets				25,000	
	Goals					Overall Asset Allocation
	1	2	3	4	Surplus	
Horizon (years)	5	25	10	20		
Required probability of success	95%	85%	90%	75%	$E(R_t)$	7.1%
Discount rate	2.3%	6.1%	4.1%	6.8%	$\sigma(R_t)$	7.6%
Module	**A**	**F**	**D**	**F**	**C**	
Required capital						
In currency	2,430	4,978	6,671	2,679	8,242	25,000
As a % of total	9.7%	19.9%	26.7%	10.7%	33.0%	100.0%

Note also that the Smiths' earlier worry, that they might not be able to meet all their goals, can be addressed easily. Our assumptions suggest that, in fact, they have excess capital representing 33% of their total financial wealth. They can either revisit their current goals and bring the timing of payments forward or raise their probability of success. The case suggests that they would rather think of additional goals but will want to give themselves some time to refine their intentions. Their adviser then suggests that a "middle of the road" module be used as a "labeled goal" for that interim period, and they call this module (Module C) "capital preservation."

4.4 The Overall Portfolio

Assuming the same six modules, with their detailed composition shown in Exhibit 38, one can then derive the overall asset allocation by aggregating the individual exposures to the various modules. In short, the overall allocation is simply the weighted average exposure to each of the asset classes or strategies within each module, with the weight being the percentage of financial assets allocated to each module. Exhibit 39 presents these computations and the overall asset allocation, which is given in bold in the right-most column. The overall portfolio's expected return and volatility are also shown. In Exhibit 38, liquidity[34] is measured as the ratio of the average number of days that might be needed to liquidate a position to the number of trading days in a year. (Note that the column B values add up to 101 because of rounding.)

Exhibit 38 Asset Allocation of Each Module

	A	B	C	D	E	F
Portfolio Characteristics						
Expected return	4.3%	5.5%	6.4%	7.2%	8.0%	8.7%
Expected volatility	2.7%	4.5%	6.0%	7.5%	10.0%	12.5%
Expected liquidity	100.0%	96.6%	90.0%	86.1%	83.6%	80.0%
Portfolio Allocations						
Cash	80%	26%	3%	1%	1%	1%
Global investment-grade bonds	20	44	45	25	0	0
Global high-yield bonds	0	5	11	25	34	4
Lower-volatility alternatives	0	9	13	0	0	0
Global developed equities	0	9	13	19	34	64
Global emerging equities	0	2	2	3	6	11
Equity-based alternatives	0	0	0	8	0	0
Illiquid global equities	0	0	5	10	15	20
Trading strategy alternatives	0	1	3	6	7	0
Global real estate	0	5	5	3	3	0
Total	100%	100%	100%	100%	100%	100%

Exhibit 39 Goals-Based Asset Allocation (US$ thousands)

	Total Financial Assets				25,000	
	Goals					Overall Asset
	1	2	3	4	Surplus	Allocation
Horizon	5	25	10	20		
Required success	95%	85%	90%	75%	$E(R_t)$	7.1%
Discount rate	2.3%	6.1%	4.1%	6.8%	$\sigma(R_t)$	7.6%

(continued)

34 Note that we need to incorporate some estimate of liquidity for all asset classes and strategies to ensure that the client's and the goals' liquidity constraints can be met.

Exhibit 39 (Continued)

Module	A	F	D	F	C	
Required capital						
In currency	2,430	4,978	6,671	2,679	8,242	25,000
As a % of total	9.7	19.9	26.7	10.7	33.0	100.0
Cash	80%	1%	1%	1%	3%	**9%**
Global investment-grade bonds	20	0	25	0	45	**24**
Global high-yield bonds	0	4	25	4	11	**12**
Lower-volatility alternatives	0	0	0	0	13	**4**
Global developed equities	0	64	19	64	13	**28**
Global emerging equities	0	11	3	11	2	**5**
Equity-based alternatives	0	0	8	0	0	**2**
Illiquid global equities	0	20	10	20	5	**10**
Trading strategy alternatives[a]	0	0	6	0	3	**3**
Global real estate	0	0	3	0	5	**2**
Total	100	100	100	100	100	**100**

[a] "Trading strategy alternatives" refers to discretionary or systematic trading strategies such as global macro and managed futures.

4.5 Revisiting the Module Process in Detail

Having explained and illustrated the client process in Exhibit 35, we now explore how modules are developed. Creating an appropriate set of optimized modules starts with the formulation of capital market assumptions. Exhibit 40 presents a possible set of forward-looking pretax capital market expectations for expected return, volatility, and liquidity[35] in Panel A and a historical 15-year correlation matrix in Panel B.[36]

35 For clients who might invest in traditional asset classes by means of vehicles such as mutual funds or ETFs, these asset classes can be treated as providing virtually instant liquidity. For clients with particularly large asset pools who might use separately managed accounts, the liquidity factor for high-yield or emerging market bonds, small-capitalization equities, and certain real assets might be adjusted downward.
36 For illiquid equities, data availability reduces the time period to seven years. The correlation matrix is based on the 15 years ending with March 2016.

Exhibit 40 Example of Capital Market Expectations for a Possible Asset Class Universe

Panel A

	Expected		
	Return	Volatility	Liquidity
Cash	4.0%	3.0%	100%
Global investment-grade bonds	5.5	6.5	100
Global high-yield bonds	7.0	10.0	100
Lower-volatility alternatives	5.5	5.0	65
Global developed equities	8.0	16.0	100
Global emerging equities	9.5	22.0	100
Equity-based alternatives	6.0	8.0	65
Illiquid global equities	11.0	30.0	0
Trading strategy alternatives	6.5	10.0	80
Global real estate	7.0	15.0	100

Panel B

		Global			Global					
	Cash	IG Bonds	HY Bonds	Lower-Volatility Alts	Developed Equities	Emerging Equities	Equity-Based Alts	Trading Strategy Alts	Illiquid Equities	Global Real Estate
Cash	**1.00**	0.00	–0.12	0.08	–0.06	–0.04	0.02	0.04	–0.26	–0.01
Global investment-grade bonds	0.00	**1.00**	0.27	0.14	0.28	0.09	0.07	0.16	0.20	0.24
Global high-yield bonds	–0.12	0.27	**1.00**	0.46	0.70	0.17	0.31	–0.08	0.35	0.28
Lower-volatility alternatives	0.08	0.14	0.46	**1.00**	0.44	0.61	0.86	0.12	0.65	0.47
Global developed equities	–0.06	0.28	0.70	0.44	**1.00**	0.17	0.32	–0.03	0.47	0.38
Global emerging equities	–0.04	0.09	0.17	0.61	0.17	**1.00**	0.72	–0.03	0.67	0.49
Equity-based alternatives	0.02	0.07	0.31	0.86	0.32	0.72	**1.00**	0.11	0.72	0.45
Trading strategy alternatives	0.04	0.16	–0.08	0.12	–0.03	–0.03	0.11	**1.00**	–0.09	0.07
Illiquid global equities	–0.26	0.20	0.35	0.65	0.47	0.67	0.72	–0.09	**1.00**	0.88
Global real estate	–0.01	0.24	0.28	0.47	0.38	0.49	0.45	0.07	0.88	**1.00**

Ostensibly, in the real world, the process ought to be associated with a set of after-tax expectations, which usually cannot be limited to broad asset classes or sub–asset classes. Indeed, the tax impact of management processes within individual asset classes or strategies (for instance, index replication, index replication with systematic tax-loss harvesting, broadly diversified portfolios, or concentrated portfolios) requires that each management process within each asset class or strategy be given its own expected return and volatility. We will dispense with that step here for the sake of simplicity, both in absolute terms and with respect to jurisdictional differences.

Exhibit 41 presents a possible set of such modules based on the capital market expectations from Exhibit 40. The optimization uses a mean–variance process and is subject to a variety of constraints that are meant to reflect both market portfolio considerations and reasonable asset class or strategy suitability given the goals that we expect to correspond to various points on the frontier. Note that the frontier is not "efficient" in the traditional sense of the term because the constraints applied to the portfolios differ from one to the next. Three elements within the set of constraints deserve special mention. The first is the need to be concerned with the liquidity of the various strategies: It would make little sense, even if it were appropriate based on other considerations, to include any material exposure to illiquid equities in a declining-balance portfolio expected to "mature" within 10 years, for instance. Any exposure thus selected would be bound to increase through time because portfolio liquidation focuses on more-liquid assets. The second relates to strategies whose return distributions are known not to be "normal." This point applies particularly to a number of alternative strategies that suffer from skew and kurtosis,[37] which a mean–variance optimization process does not take into account (see Section 2.4.4). Finally, the constraints contain a measure of drawdown control to alleviate the problems potentially associated with portfolios that, although apparently optimal, appear too risky in overly challenging market circumstances. Drawdown controls are an important element in that they help deal with the often-observed asymmetric tolerance of investors for volatility: upward volatility is much preferred to downward volatility.

Exhibit 41 Six Possible Sub-Portfolio Modules

	A	B	C	D	E	F
Portfolio Characteristics						
Expected return	4.3%	5.5%	6.4%	7.2%	8.0%	8.7%
Expected volatility	2.7	4.5	6.0	7.5	10.0	12.5
Expected liquidity	100.0	96.6	90.0	86.1	83.6	80.0
Portfolio Allocations						
Cash	80%	26%	3%	1%	1%	1%
Global investment-grade bonds	20	44	45	25	0	0
Global high-yield bonds	0	5	11	25	34	4
Lower-volatility alternatives	0	9	13	0	0	0
Global developed equities	0	9	13	19	34	64
Global emerging equities	0	2	2	3	6	11
Equity-based alternatives	0	0	0	8	0	0
Illiquid global equities	0	0	5	10	15	20
Trading strategy alternatives	0	1	3	6	7	0
Global real estate	0	5	5	3	3	0
Total	100%	100%	100%	100%	100%	100%

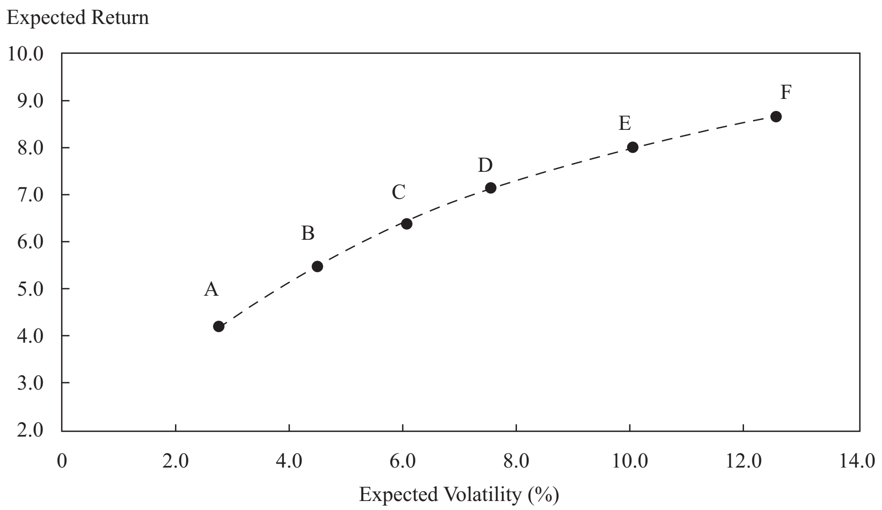
37 Kat (2003) described the challenge, and Davies, Kat, and Lu (2009) presented a solution that involves the use of mean–variance-skew-kurtosis optimization, which is typically too complex for most real-life circumstances.

Exhibit 41 (Continued)

Constraints						
Maximum volatility	3.0%	4.5%	6.0%	7.5%	10.0%	12.5%
Minimum liquidity	100.0	95.0	90.0	85.0	80.0	70.0
Maximum alternatives	0.0	10.0	20.0	30.0	30.0	30.0
Minimum cash	80.0	20.0	0.3	0.5	0.7	1.0
Maximum HY as a percent of total fixed income	0.0	10.0	20.0	50.0	100.0	100.0
Maximum equity spectrum	0.0	10.0	20.0	40.0	75.0	100.0
Maximum EM as a percent of public equities	15.0	15.0	15.0	15.0	15.0	15.0
Maximum illiquid equities	0.0	0.0	5.0	10.0	15.0	20.0
Maximum trading as a percent of equity spectrum	0.0	10.0	15.0	15.0	20.0	25.0
Maximum real estate	0.0	5.0	10.0	15.0	20.0	25.0
Escrow cash as a percent of illiquid equities	5.0	5.0	5.0	5.0	5.0	5.0
Maximum probability of return < drawdown	1.0	1.5	2.0	2.0	2.5	2.5
Drawdown horizon	3	3	3	3	3	3
Drawdown amount	0.0	–5.0	–7.5	–10.0	–15.0	–20.0

The six sub-portfolios shown in Exhibit 41 satisfy two major design goals: First, they cover a wide spectrum of the investment universe, ranging from a nearly all-cash portfolio (Portfolio A) to an all-equity alternative (Portfolio F). Second, they are sufficiently differentiated to avoid creating distinctions without real differences. These portfolios are graphed in Exhibit 42.

Exhibit 42 Sub-Portfolio Modules Cover a Full Range



Returning to an earlier point about "labeled goals," one can easily imagine "aspirations" to describe each of these modules, ranging from "immediate- to short-term lifestyle" for Module A to "aggressive growth" for Module F. Module B might be labeled "long-term lifestyle," while C and D might represent forms of capital preservation and E a form of "balanced growth."

A final point deserves special emphasis: Modules need to be revisited on a periodic basis. While equilibrium assumptions will likely not change much from one year to the next, the need to identify one's position with respect to a "normal" market cycle can lead to modest changes in forward-looking assumptions. It would indeed be foolish to keep using long-term equilibrium assumptions when it becomes clear that one is closer to a market top than to a market bottom. The question of the suitability of revisions becomes moot when using a systematic approach such as the Black–Litterman model. One may also need to review the continued suitability of constraints, not to mention (when applicable) the fact that the make-up of the market portfolio may change in terms of geography or credit distribution.

4.6 Periodically Revisiting the Overall Asset Allocation

Once set, the goals-based allocation must be regularly reviewed. Two considerations dominate:

1 Goals with an initially fixed time horizon are not necessarily one year closer to maturity after a year. Superficially, one would expect that someone who says that his or her need is to meet lifestyle expenditures over the next five years, for instance, means exactly this. Accordingly, next year, the time horizon should shift down to four years. Yet experience suggests that certain horizons are "placeholders": One year on, the time horizon remains five years. This is particularly—and understandably—relevant when the horizon reflects the anticipated death of an individual.

2 The preference for upward rather than downward volatility, combined with perceptions that goals may have higher required probabilities of success than is truly the case, leads to portfolios that typically outperform the discount rate used to compute the required initial capital. Thus, one would expect there to be some need for portfolio rebalancing when the assets allocated to certain goals appear excessive, at least in probability- and horizon-adjusted terms. This situation gives rise to important discussions with taxable clients because any form of portfolio rebalancing is inherently more complex and costly in a taxable environment than when taxes do not come into consideration.

4.7 Issues Related to Goals-Based Asset Allocation

Although goals-based asset allocation offers an elegant and mathematically sound way to deal with the circumstances of individuals, it is not a panacea. By definition, goals-based asset allocation applies best to individuals who have multiple goals, time horizons, and urgency levels. The classic example of the professional who is just starting to save for retirement and who has no other significant goal (as in the case of Aimée Goddard in Example 1) can be easily be handled with the traditional financial tools discussed in the earlier sections of this reading.[38] However, one should always be cautious to ensure that there is no "hidden" goal that should be brought out and that the apparently "single" retirement goal is not in fact an aggregation of several elements with different levels of urgency, if not also different time horizons. Single-goal circumstances may still be helped by the goals-based asset allocation process when there are sustainability or behavioral questions. In that case, one can look at the single goal as being made up of several similar goals over successive time periods with different required probabilities of success. For instance, one might apply a higher sense of urgency—and thus require a lower risk profile—to contributions made in the first few years, on the ground that adverse market circumstances might negatively affect the willingness of the client to stay with the program. In many ways, this approach can be seen as a conceptual analog to the dollar-cost-averaging investment framework.

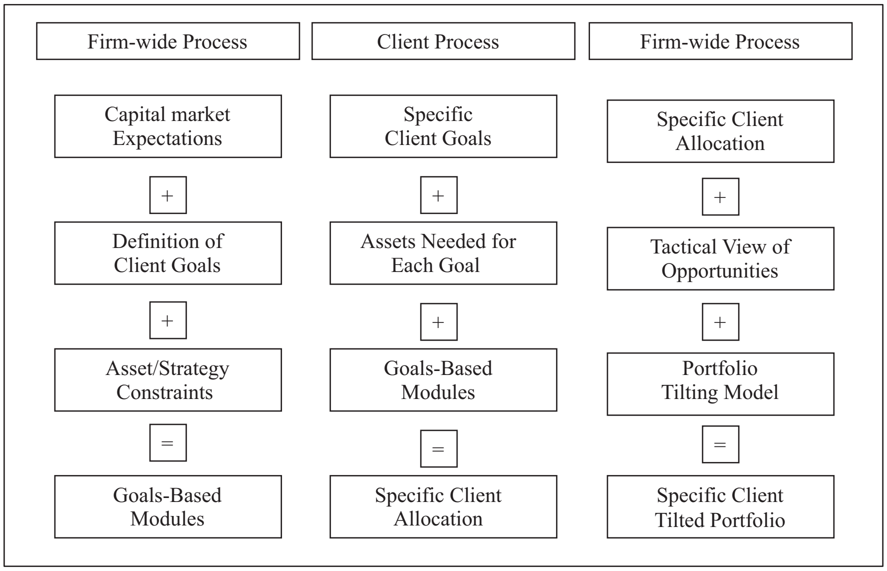
Goals-based asset allocation is ideally suited to situations involving multiple goals, time horizons, and urgency levels, whether the assets are large or more modest. In fact, in cases where "human capital" is considered, a multi-goal approach can help investors understand the various trade-offs they face. Ostensibly, the larger the assets, the more complex the nature of the investment problem, the more diverse the list of investment structures, and the more one should expect a client-focused approach to offer useful benefits. However, the ratio of cash outflows to assets under consideration is a more germane issue than the overall size of the asset pool.

Advisers using goals-based wealth management must contend with a considerably higher level of business management complexity. They will naturally expect to have a different policy for each client and potentially more than one policy per client. Thus, managing these portfolios day to day and satisfying the usual regulatory requirement that all clients be treated in an equivalent manner can appear to be a major quandary.

Typically, the solution would involve developing a systematic approach to decision making such that it remains practical for advisers to formulate truly individual policies that reflect their investment insights. Exhibit 43 offers a graphical overview of advisers' activities, divided into those that involve "firm-wide" processes, defined as areas where no real customization is warranted, and those that must remain "client focused." The result is analogous to a customized racing bicycle, whose parts are mass produced but then combined into a truly unique bike custom-designed for the individual racer.

38 However, an adviser may find it appropriate to help the individual divide the funds he or she believes are needed for retirement into several categories. For instance, there may be some incompressible lifestyle expenditure that represents a minimum required spending level, but there may also be some luxury or at least compressible spending that does not have such a high level of urgency or that applies over a different time frame (say, the early or late years). Thus, one could still describe the problem as involving multiple goals, multiple time horizons, and multiple urgency levels. Then, one could compare the costs associated with the funding of these goals and have the individual weigh potential future satisfaction against the loss of current purchasing power.

Exhibit 43 Goals-Based Wealth Management Advisory Overview



5 HEURISTICS AND OTHER APPROACHES TO ASSET ALLOCATION

In addition to the various asset allocation approaches already covered, a variety of heuristics (rules that provide a reasonable but not necessarily optimal solution) and other techniques deserve mention:

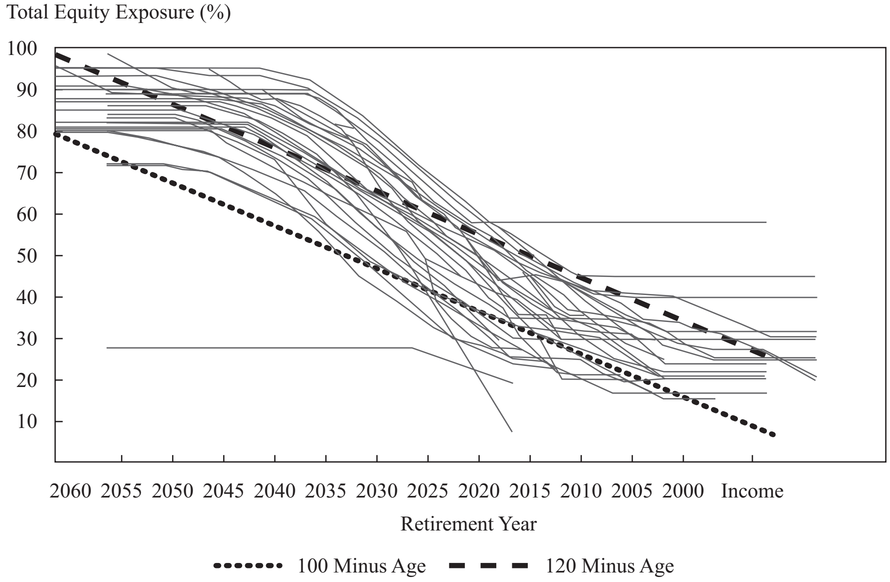
The "120 minus your age" rule. The phrase "120 minus your age" is a heuristic for inferring a hidden, age-driven risk tolerance coefficient that then leads directly to an age-based stock versus fixed income split: 120 – Age = Percentage allocated to stocks. Thus, a 25-year-old man would allocate 95% of his investment portfolio to stocks. Although we are aware of no theoretic basis for this heuristic—or its older and newer cousins, "100 minus your age" and "125 minus your age," respectively—it results in a linear decrease in equity exposure that seems to fit the general equity glide paths associated with target-date funds, including those that are based on a total balance sheet approach that includes human capital. A number of target-date funds (sometimes called life-cycle or age-based funds) and some target-date index providers report that their glide path (the age-based change in equity exposure) is based on the evolution of an individual's human capital. For example, one set of indexes[39] explicitly targets an investable proxy for the world market portfolio in which the glide path is the result of the evolving relationship of financial capital to human capital.[40]

Exhibit 44 displays the glide paths of the 60 largest target-date fund families in the United States. The retirement year (typically part of the fund's name) on the x-axis denotes the year in which the investor is expected to retire, which is almost always assumed to be the year the investor turns 65. Thus, as of 2016, the 2060 allocations correspond to a 21-year-old investor (79% equity, using the heuristic), whereas the 2005

39 Morningstar's Lifetime Allocation (target-date) indexes.
40 See Idzorek (2008).

allocation corresponds to a 76-year-old investor (24% equity, using the heuristic).[41] One dashed line represents the equity allocation based on the "100 minus your age" heuristic, while another dashed line represents the "120 minus your age" heuristic. The heuristic lines lack some of the nuances of the various glide path lines, but it would appear that an age-based heuristic leads to asset allocations that are broadly similar to those used by target-date funds.

Exhibit 44 Target-Date Funds and Age Heuristics (as of January 2016)



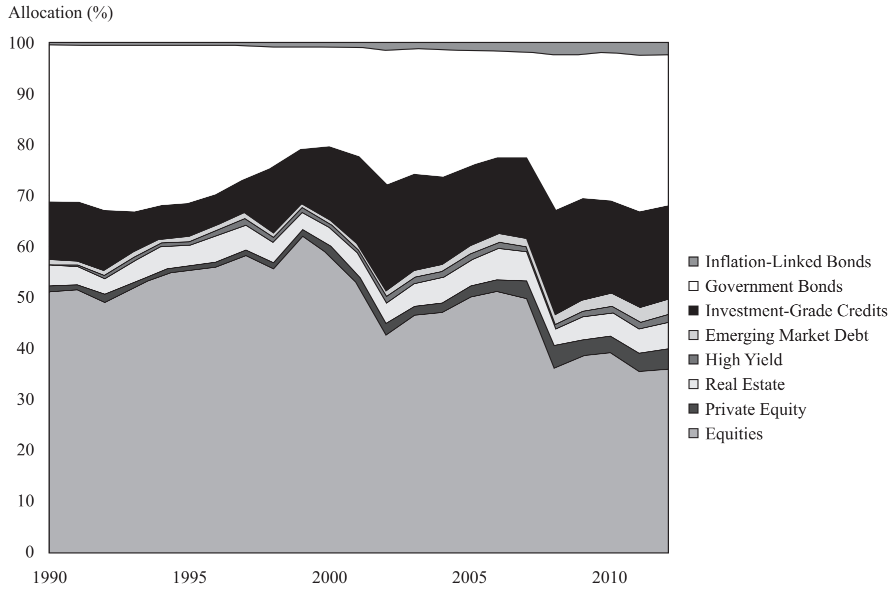
The 60/40 stock/bond heuristic. Some investors choose to skip the various optimization techniques and simply adopt an asset allocation consisting of 60% equities and 40% fixed income.

The equity allocation is viewed as supplying a long-term growth foundation, and the fixed-income allocation as supplying risk reduction benefits. If the stock and bond allocations are themselves diversified, an overall diversified portfolio should result.

There is some evidence that the global financial asset market portfolio is close to this prototypical 60/40 split. Exhibit 45 displays the estimated market value of eight major components of the market portfolio from 1990 to 2012. In approximately 7 of the 23 years, equities, private equity, and real estate account for slightly more than 60%, while for the rest of the time, the combined percentage is slightly less.

41 Many target-date funds continue to offer a "2005" vintage that would have been marketed/sold to people retiring in 2005.

Exhibit 45 Global Market Portfolio, 1990 to 2012



Source: Doeswijk, Lam, and Swinkels (2014).

The endowment model. An approach to asset allocation that emphasizes large allocations to non-traditional investments, including equity-oriented investments driven by investment manager skill (e.g., private equities), has come to be known as the endowment model or Yale model. The label "Yale model" reflects the fact that the Yale University Investments Office under David Swensen pioneered the approach in the 1990s; the label "endowment model" reflects the influence of this approach among US university endowments. Swensen (2009) stated that most investors should not pursue the Yale model but should instead embrace a simpler asset allocation implemented with low-cost funds. Besides high allocations to non-traditional assets and a commitment to active management, the approach characteristically seeks to earn illiquidity premiums, which endowments with long time horizons are well positioned to capture. Exhibit 46, showing the Yale endowment asset allocation, makes these points. In the exhibit, "absolute return" indicates investment in event-driven and value-driven strategies.

Exhibit 46 Yale University Endowment Asset Allocation as of June 2014

	Yale University	US Educational Institution Mean
Absolute return	17.4%	23.3%
Domestic equity	3.9	19.3
Fixed income	4.9	9.3
Foreign equity	11.5	22.0
Natural resources	8.2	8.5
Private equity	33.0	10.0

Exhibit 46 (Continued)

	Yale University	US Educational Institution Mean
Real estate	17.6	4.2
Cash	3.5	3.5

Source: Yale University (2014, p. 13).

In almost diametrical contrast to the endowment model is the asset allocation approach of Norway's Government Pension Fund Global (Statens pensjonsfond Utland), often called the Norway model.[42] This model's asset allocation is highly committed to passive investment in publicly traded securities (subject to environmental, social, and governance [ESG] concerns), reflecting a belief in the market's informational efficiency. Since 2009, the asset allocation has followed an approximate 60/40 stock/bond mix.

Risk parity. A risk parity asset allocation is based on the notion that each asset (asset class or risk factor) should contribute equally to the total risk of the portfolio for a portfolio to be well diversified. Recall that in Section 2, we identified various criticisms and potential shortcomings of mean–variance optimization, one of which was that, while the resulting asset allocations may appear diversified across assets, the sources of risk may not be diversified. In the section on risk budgeting, Exhibit 19 contained a risk decomposition of a reverse-optimization-based asset allocation from a United Kingdom–based investor. There, we noted that the overall equity/fixed income split was approximately 54% equities and 46% fixed income, yet of the 10% standard deviation, approximately 74% of the risk came from equities while only 26% came from fixed income.

Risk parity is a relatively controversial approach. Although there are several variants, the most common risk parity approach has the following mathematical form:

$$w_i \times \mathrm{Cov}\left(r_i, r_P\right) = \frac{1}{n}\sigma_P^2 \tag{3}$$

where

w_i = the weight of asset i

$\mathrm{Cov}(r_i, r_P)$ = the covariance of asset i with the portfolio

n = the number of assets

σ_P^2 = the variance of the portfolio

In general, there is not a closed-form solution to the problem, and it must be solved using some form of optimization (mathematical programming). Prior to Markowitz's development of mean–variance optimization, which simultaneously considered both risk and return, most asset allocation approaches focused only on return *and ignored risk* (or accounted for it in an ad hoc manner). The primary criticism of risk parity is that it makes the opposite mistake: It *ignores expected returns*. In general, most of the rules-based risk approaches—such as other forms of volatility weighting, minimum volatility, and target volatility—suffer from this shortcoming.

With risk parity, the contribution to risk is highly dependent on the formation of the opportunity set. For example, if the opportunity set consists of seven equity asset classes and three fixed-income asset classes, intuitively, 70% of risk will come from the equities and 30% of risk will come from fixed income. Conversely, if the

42 See Curtis (2012).

opportunity set consists of three equity asset classes and seven fixed-income asset classes, intuitively, 70% of risk will come from fixed income and 30% of risk will come from equities. The point is that practitioners of risk parity must be very cognizant of the formation of their opportunity set.

Exhibit 47 gives a US-centric example consisting of five equity asset classes and three fixed-income asset classes. A constrained optimization routine (weights must sum to 100%) was used to determine the weight to each asset class, such that all asset classes contributed the same amount to total risk. In this case, each asset class contributed 0.8%, resulting in an asset allocation with a total standard deviation of 6.41%. In this example, 5/8 of total risk comes from equity asset classes and 3/8 comes from fixed-income asset classes. Earlier, we explained that reverse optimization can be used to infer the expected return of any set of presumed efficient weights. In Exhibit 47, based on a total market risk premium of 2.13% and a risk-free rate of 3%, we inferred the reverse-optimized total returns (final column). In this case, these seem to be relatively reasonable expected returns.

Exhibit 47 Risk Parity Portfolio Weights and Risk-Budgeting Statistics Based on Reverse-Optimized Returns

Asset Class	Weight	Marginal Contribution to Total Risk (MCTR)	ACTR	Percentage Contribution to Total Standard Deviation	Reverse-Optimized Total Returns
US large-cap equities	7.7%	10.43%	0.80%	12.50%	6.47%
US mid-cap equities	6.1	13.03	0.80	12.50	7.33
US small-cap equities	5.9	13.61	0.80	12.50	7.52
Non-US developed market equities	5.6	14.38	0.80	12.50	7.78
Emerging market equities	4.5	17.74	0.80	12.50	8.89
Non-US bonds	15.5	5.17	0.80	12.50	4.72
US TIPS	23.9	3.36	0.80	12.50	4.12
US bonds	30.8	2.60	0.80	12.50	3.86
Total	100.0%		6.41%	100.00%	5.13%

After deriving a risk parity–based asset allocation, the next step in the process is to borrow (use leverage) or to lend (save a portion of wealth, presumably in cash) so that the overall portfolio corresponds to the investor's risk appetite. Continuing with our example, the market risk premium is 2.13% (above the assumed risk-free rate of 3%) and the market variance is 0.41% (i.e., 6.41% squared); thus, the implied market trade-off of expected return (in excess of the risk-free rate) for risk is 2.13% divided by 0.41%, which equals approximately 5.2. Investors with a greater appetite for risk than the market as a whole would borrow money to lever up the risk parity portfolios, while investors with a lower appetite for risk would invest a portion of their wealth in cash.

Back tests of levered risk parity portfolios have produced promising results, although critics of these back tests argue that they suffer from look-back bias and are very dependent on the ability to use extremely large amounts of leverage at low borrow rates (which may not have been feasible); see, for example, Anderson, Bianchi, and Goldberg (2012). Proponents of risk parity have suggested that the idea of "leverage aversion" contributes to the success of the strategy. Black (1972) suggested that restrictions on leverage and a general aversion to leverage may cause return-seeking investors to pursue higher-returning assets, such as stocks. All else equal, this behavior would reduce the price of bonds, thus allowing the investor to buy bonds at a small discount, hold them to maturity, and realize the full value of the bond. Asness, Frazzini, and Pedersen (2012) have offered this idea as a potential explanation for why a levered (bond-centric) asset allocation might outperform an equity-centric asset allocation with equivalent or similar risk.

The 1/N rule. One of the simplest asset allocation heuristics involves equally weighting allocations to assets. DeMiguel, Garlappi, and Uppal (2009) define an approach in which $1/N$ of wealth is allocated to each of N assets available for investment at each rebalancing date. Calendar rebalancing to equal weighting at quarterly intervals is one common rebalancing discipline used. By treating all assets as indistinguishable in terms of mean returns, volatility, and correlations, in principle, $1/N$ rule portfolios should be dominated by methods that optimize asset class weights to exploit differences in investment characteristics. In empirical studies comparing approaches, however, the $1/N$ rule has been found to perform considerably better, based on Sharpe ratios and certainty equivalents, than theory might suggest. One possible explanation is that the $1/N$ rule sidesteps problems caused by optimizing when there is estimation error in inputs.

PORTFOLIO REBALANCING IN PRACTICE

6

The reading “Introduction to Asset Allocation” provided an introduction to rebalancing, including some detailed comments on strategic considerations. This section aims to present useful additional insight and information.

Meanings of “Rebalancing”

Rebalancing has been defined as the discipline of adjusting portfolio weights to more closely align with the strategic asset allocation. In that sense, rebalancing includes policy regarding the correction of any drift away from strategic asset allocation weights resulting from market price movements and the passage of time for finite-lived assets, such as bonds. In liability-relative asset allocation, adjusting a liability-hedging portfolio to account for changes in net duration exposures from the passage of time, for example, would fall under the rubric of rebalancing.

Some use the term “rebalancing” more expansively, to include the combined effects on asset class weights not only of rebalancing in the above sense but also of active allocation activities. In that sense, rebalancing would include tactical allocations. Although rebalancing policy can be established to accommodate tactical adjustments, tactical asset allocation per se is not covered under “rebalancing” as the term is used here.

Changes in asset allocation weights in response to changes in client circumstances, goals, or other client factors are sometimes also referred to as “rebalancing” (especially if the adjustments are minor). These activities fall under the scope of client monitoring and asset allocation review, as described elsewhere in the CFA curriculum.

An appropriate rebalancing policy involves a weighing of benefits and costs. Benefits depend on the idea that if an investor’s strategic asset allocation is optimal, then any divergence in the portfolio from that asset allocation represents an expected utility loss to the investor. Rebalancing benefits the investor by reducing the present value of expected losses from not tracking the optimum. In theory, the basic cost of not rebalancing is this present value of expected utility losses from straying from the optimum.[43]

Apart from the above considerations of trade-offs, disciplined rebalancing has tended to reduce risk while incrementally adding to returns. Several interpretations of this empirical finding have been offered, including the following:

- *Rebalancing earns a diversification return*. The compound growth rate of a portfolio is greater than the weighted average compound growth rates of the component portfolio holdings (given positive expected returns and positive asset weights). Given sufficiently low transaction costs, this effect leads to what has been called a *diversification return* to frequent rebalancing to a well-diversified portfolio.[44]
- *Rebalancing earns a return from being short volatility*. In the case of a portfolio consisting of a risky asset and a risk-free asset, the return to a rebalanced portfolio can be replicated by creating a buy-and-hold position in the portfolio, writing out-of-the-money puts and calls on the risky asset, and investing the premiums in risk-free bonds.[45] As the value of puts and calls is positively related to volatility, such a position is called being short volatility (or being short gamma, by reference to the option Greeks).

Practice appears not to have produced a consensus on the most appropriate rebalancing discipline. Introduction to Asset Allocation defined and discussed calendar rebalancing[46]—sometimes mentioned as common in portfolios managed for individual investors—and percent-range rebalancing. Calendar rebalancing involves lower overhead because of lower monitoring costs. Percent-range rebalancing is a more disciplined risk control policy, however, because it makes rebalancing contingent on market movements. Without weighing costs and benefits in the abstract, Exhibit 48 assumes percent-range rebalancing and summarizes the effects of each of several key factors on the corridor width of an asset class, holding all else equal, except for the factor of the asset class's own volatility.[47] For taxable investors, transactions trigger capital gains in jurisdictions that tax them; therefore, for such investors, higher tax rates on capital gains should also be associated with wider corridors.

Exhibit 48 Factors Affecting the Optimal Corridor Width of an Asset Class

Factor	Effect on Optimal Width of Corridor (All Else Equal)	Intuition
Factors Positively Related to Optimal Corridor Width		
Transaction costs	The higher the transaction costs, the wider the optimal corridor.	High transaction costs set a high hurdle for rebalancing benefits to overcome.
Risk tolerance	The higher the risk tolerance, the wider the optimal corridor.	Higher risk tolerance means less sensitivity to divergences from the target allocation.
Correlation with the rest of the portfolio	The higher the correlation, the wider the optimal corridor.	When asset classes move in sync, further divergence from target weights is less likely.

43 See Leland (2000).
44 See Willenbrock (2011). This phenomenon was called *rebalancing return* by Mulvey and Kim (2009). Luenberger (2013) suggests that the phenomenon could be exploited by a strategy of buying high-volatility assets and rebalancing often, a process he called *volatility pumping*.
45 As shown in Ang (2014, pp. 135–139).
46 Rebalancing a portfolio to target weights on a periodic basis—for example, monthly, quarterly, semi-annually, or annually.
47 See Masters (2003).

Exhibit 48 (Continued)

Factor	Effect on Optimal Width of Corridor (All Else Equal)	Intuition
Factors Inversely Related to Optimal Corridor Width		
Volatility of the rest of the portfolio	The higher the volatility, the narrower the optimal corridor.	Higher volatility makes large divergences from the strategic asset allocation more likely.

Among positive factors, the cases of transaction costs and risk tolerance are obvious. Transaction costs can be reduced to the extent that portfolio cash flows can be used to rebalance. The case of correlation is less obvious. Because of correlations, the rebalancing triggers among different asset classes are linked.

Consider correlation in a two–asset class scenario. Suppose one asset class is above its target weight, so the other asset class is below its target weight. A further increase in the value of the overweight asset class implies, on average, a smaller divergence in the asset mix if the asset classes' returns are more highly positively correlated (because the denominator in computing the overweight asset class's weight is the sum of the values of the two asset classes). In a multi-asset-class scenario, all pair-wise asset class correlations would need to be considered, making the interpretation of correlations complex. To expand the application of the two-asset case's intuition, one simplification involves considering the balance of a portfolio to be a single hypothetical asset and computing an asset class's correlation with it.

As indicated in Exhibit 48, the higher the volatility of the rest of the portfolio, excluding the asset class being considered, the more likely a large divergence from the strategic asset allocation becomes. That consideration should point to a narrower optimal corridor, all else being equal.

In the case of an asset class's own volatility, "holding all else equal" is not practically meaningful. If rebalancing did not involve transaction costs, then higher volatility would lead to a narrower corridor, all else equal, for a risk-averse investor.[48] Higher volatility implies that if an asset class is not brought back into the optimal range after a given move away from it, the chance of an even further divergence from optimal is greater. In other words, higher volatility makes large divergence from the strategic asset allocation more likely. However, reducing a corridor's width means more frequent rebalancing and higher transaction costs. Thus, the effect of volatility on optimal corridor width involves a trade-off between controlling transaction costs and controlling risk. Conclusions also depend on the assumptions made about asset price return dynamics.

In practice, corridor width is often specified to be proportionally greater, the higher an asset class's volatility, with a focus on transaction cost control. In *volatility-based rebalancing*, corridor width is set proportionally to the asset class's own volatility. In one variation of *equal probability rebalancing* (McCalla 1997), the manager specifies a corridor for each asset class in terms of a common multiple of the standard deviation of the asset class's returns such that, under a normal probability assumption, each asset class is equally likely to trigger rebalancing.

48 As in Masters (2003).

EXAMPLE 12

Tolerance Bands for an Asset Allocation

An investment committee is reviewing the following strategic asset allocation:

Domestic equities 50% ± 5% (i.e., 45% to 55% of portfolio value)

International equities 15% ± 1.5%

Domestic bonds 35% ± 3.5%

The market for the domestic bonds is relatively illiquid. The committee views the above corridors as appropriate *if* each asset class's risk and transaction cost characteristics remain unchanged. The committee now wants to account for differences among the asset classes in setting the corridors.

Evaluate the implications of the following sets of facts for the stated tolerance bands, given an all-else-equal assumption in each case:

1. Tax rates for international equities increase by 10 percentage points.
2. Transaction costs in international equities increase by 20% relative to domestic equities, but the correlation of international equities with domestic equities and bonds declines. What is the expected effect on the tolerance band for international equities?
3. The volatility of domestic bonds increases. What is the expected effect on their tolerance band? Assume that domestic bonds are relatively illiquid.

Solution to 1:

The tolerance band for international equities should increase if the entity is a taxable investor.

Solution to 2:

Increased transaction costs point to widening the tolerance band for international equities, but declining correlations point to narrowing it. The overall effect is indeterminate.

Solution to 3:

Given that the market for domestic bonds is relatively illiquid, the increase in volatility suggests widening the rebalancing band. Containing transaction costs is more important than the expected utility losses from allowing a larger divergence from the strategic asset allocation.

One decision involved in rebalancing policy is whether to adjust asset class holdings to their target proportions, to the limits of the corridors, or to within the corridors but not to target weights. Compared with rebalancing to target weights, rebalancing to the upper or lower limit of the allowed range results in less close alignment with target proportions but lower transaction costs—an especially important consideration in the case of relatively illiquid assets. The choice among alternatives may be influenced by judgmental tactical considerations.

Because one rebalancing decision affects later rebalancing decisions, the optimal rebalancing decisions at different points in time are linked. However, optimal rebalancing in a multi-period, multi-asset case is an unsolved problem.

The analysis of Dybvig (2005) suggests that fixed transaction costs favor rebalancing to the target weights and variable transaction costs favor rebalancing to the nearest corridor border (the interior of the corridor being therefore a "no trade zone"). A number of studies have contrasted rebalancing to target weights and rebalancing to the allowed range based on particular asset classes, time periods, and measures of the benefits of rebalancing. These studies have reached a variety of conclusions, suggesting that no simple, empirically based advice can be provided.

Rebalancing in a Goals-Based Approach

The use of probability- and horizon-adjusted discount rates to size the various goal-defeasing sub-portfolios means that portfolios will usually produce returns that are higher than assumed. Thus, as time passes, the dollars allocated to the various sub-portfolios—other than labeled-goal portfolios—may be expected to exceed the actual requirements. For example, in average markets, returns should exceed the conservative requirements of a goal associated with a 90% required probability of success. Sub-portfolios with shorter time horizons for goals with high required probabilities of success will tend to contain relatively low-risk assets, whereas riskier assets may have high allocations in longer-horizon portfolios for goals with lower required probabilities of success. Thus, there is a greater chance that the exposure to lower-risk assets will creep up before one experiences the same for riskier assets. Thus, failing to rebalance the portfolio will gradually move it down the risk axis—and the defined efficient frontier—and thus lead the client to take less risk than he or she can bear.

CONCLUSIONS 7

This reading has surveyed how appropriate asset allocations can be determined to meet the needs of a variety of investors. Among the major points made have been the following:

- The objective function of asset-only mean–variance optimization is to maximize the expected return of the asset mix minus a penalty that depends on risk aversion and the expected variance of the asset mix.
- Criticisms of MVO include the following:
  - The outputs (asset allocations) are highly sensitive to small changes in the inputs.
  - The asset allocations are highly concentrated in a subset of the available asset classes.
  - Investors are often concerned with characteristics of asset class returns such as skewness and kurtosis that are not accounted for in MVO.
  - While the asset allocations may appear diversified across assets, the sources of risk may not be diversified.
  - MVO allocations may have no direct connection to the factors affecting any liability or consumption streams.
  - MVO is a single-period framework that tends to ignore trading/rebalancing costs and taxes.
- Deriving expected returns by reverse optimization or by reverse optimization tilted toward an investor's views on asset returns (the Black–Litterman model) is one means of addressing the tendency of MVO to produce efficient portfolios that are not well diversified.

- Placing constraints on asset class weights to prevent extremely concentrated portfolios and resampling inputs are other ways of addressing the same concern.
- For some relatively illiquid asset classes, a satisfactory proxy may not be available; including such asset classes in the optimization may therefore be problematic.
- Risk budgeting is a means of making optimal use of risk in the pursuit of return. A risk budget is optimal when the ratio of excess return to marginal contribution to total risk is the same for all assets in the portfolio.
- Characteristics of liabilities that affect asset allocation in liability-relative asset allocation include the following:
  - Fixed versus contingent cash flows
  - Legal versus quasi-liabilities
  - Duration and convexity of liability cash flows
  - Value of liabilities as compared with the size of the sponsoring organization
  - Factors driving future liability cash flows (inflation, economic conditions, interest rates, risk premium)
  - Timing considerations, such longevity risk
  - Regulations affecting liability cash flow calculations
- Approaches to liability-relative asset allocation include surplus optimization, a hedging/return-seeking portfolios approach, and an integrated asset–liability approach.
  - Surplus optimization involves MVO applied to surplus returns.
  - A hedging/return-seeking portfolios approach assigns assets to one of two portfolios. The objective of the hedging portfolio is to hedge the investor's liability stream. Any remaining funds are invested in the return-seeking portfolio.
  - An integrated asset–liability approach integrates and jointly optimizes asset and liability decisions.
- A goals-based asset allocation process combines into an overall portfolio a number of sub-portfolios, each of which is designed to fund an individual goal with its own time horizon and required probability of success.
- In the implementation, there are two fundamental parts to the asset allocation process. The first centers on the creation of portfolio modules, while the second relates to the identification of client goals and the matching of these goals to the appropriate sub-portfolios to which suitable levels of capital are allocated.
- Other approaches to asset allocation include "120 minus your age," 60/40 stocks/bonds, the endowment model, risk parity, and the 1/*N* rule.
- Disciplined rebalancing has tended to reduce risk while incrementally adding to returns. Interpretations of this empirical finding include that rebalancing earns a diversification return, that rebalancing earns a return from being short volatility, and that rebalancing earns a return to supplying liquidity to the market.
- Factors positively related to optimal corridor width include transaction costs, risk tolerance, and an asset class's correlation with the rest of the portfolio. The higher the correlation, the wider the optimal corridor, because when asset classes move in sync, further divergence from target weights is less likely.

- The volatility of the rest of the portfolio (outside of the asset class under consideration) is inversely related to optimal corridor width.
- An asset class's own volatility involves a trade-off between transaction costs and risk control. The width of the optimal tolerance band increases with transaction costs for volatility-based rebalancing.

REFERENCES

Anderson, Robert M., Stephen W. Bianchi, and Lisa R. Goldberg. 2012. "Will My Risk Parity Strategy Outperform?" *Financial Analysts Journal*, vol. 68, no. 6 (November/December): 75–93.

Ang, Andrew. 2014. Asset Management. New York: Oxford University Press.

Asness, Clifford S., Andrea Frazzini, and Lasse H. Pedersen. 2012. "Leverage Aversion and Risk Parity." *Financial Analysts Journal*, vol. 68, no. 1 (January/February): 47–59.

Athayde, Gustavo M. de, and Renato G. Flôres, Jr. 2003. "Incorporating Skewness and Kurtosis in Portfolio Optimization: A Multideminsional Efficient Set." In *Advances in Portfolio Construction and Implementation*, edited by Stephen Satchell and Alan Scowcroft. Oxford, UK: Butterworth–Heinemann.

Beardsley, Xiaoxin W., Brian Field, and Mingqing Xiao. 2012. "Mean–Variance-Skewness-Kurtosis Portfolio Optimization with Return and Liquidity." *Communications in Mathematical Finance*, vol. 1, no. 1: 13–49.

Black, Fischer. 1972. "Capital market equilibrium with restricted borrowing." *Journal of Business*, vol. 45, no. 3: 444–455.

Black, Fischer, and Robert Litterman. 1990. "Asset Allocation: Combining Investors Views with Market Equilibrium." Fixed Income Research, Goldman, Sachs & Company, September.

Black, Fischer, and Robert Litterman. 1991. "Global Asset Allocation with Equities, Bonds, and Currencies." Fixed Income Research, Goldman, Sachs & Company, October.

Black, Fischer, and Robert Litterman. 1992. "Global Portfolio Optimization." *Financial Analysts Journal*, vol. 48, no. 5 (September/October): 28–43.

Blanchett, David M., and Philip U. Straehl. 2015. "No Portfolio is an Island." *Financial Analysts Journal*, vol. 71, no. 3 (May/June): 15–33.

Briec, W., K. Kerstens, and O. Jokung. 2007. "Mean–Variance-Skewness Portfolio Performance Gauging: A General Shortage Function and Dual Approach." *Management Science*, vol. 53, no. 1 (January): 135–149.

Brunel, Jean L.P. 2003. "Revisiting the Asset Allocation Challenge through a Behavioral Finance Lens." *Journal of Wealth Management*, vol. 6, no. 2 (Fall): 10–20.

Brunel, Jean L.P. 2005. "A Behavioral Finance Approach to Strategic Asset Allocation—A Case Study." *Journal of Investment Consulting*, vol. 7, no. 3 (Winter): 61–69.

Chhabra, Ashvin. 2005. "Beyond Markowitz: A Comprehensive Wealth Allocation Framework for Individual Investors." *Journal of Wealth Management*, vol. 7, no. 4 (Spring): 8–34.

Chow, George, Eric Jacquier, Mark Kritzman, and Kenneth Lowry. 1999. "Optimal Portfolios in Good Times and Bad." *Financial Analysts Journal*, vol. 55, no. 3 (May/June): 65–73.

Curtis, Gregory. 2012. "Yale versus Norway." White Paper 55, Greycourt (September).

Das, Sanjiv, Harry Markowitz, Jonathan Scheid, and Meir Statman. 2010. "Portfolio Optimization with Mental Accounts." *Journal of Financial and Quantitative Analysis*, vol. 45, no. 2 (April): 311–334.

Das, Sanjiv, Harry Markowitz, Jonathan Scheid, and Meir Statman. 2011. "Portfolios for Investors Who Want to Reach Their Goals While Staying on the Mean–Variance Efficient Frontier." *Journal of Wealth Management*, vol. 14, no. 2 (Fall): 25–31.

Davies, Ryan, Harry M. Kat, and Sa Lu. 2009. "Fund of Hedge Funds Portfolio Selection: A Multiple-Objective Approach." *Journal of Derivatives & Hedge Funds*, vol. 15, no. 2: 91–115.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. "Optimal versus Naive Diversification: How Inefficient Is the 1/*N* Portfolio Strategy?" *Review of Financial Studies*, vol. 22, no. 5: 1915–1953.

DiBartolomeo, Dan. 1993. "Portfolio Optimization: The Robust Solution." Prudential Securities Quantitative Conference. Available online at http://www.northinfo.com/documents/45.pdf

Doeswijk, Ronald, Trevin Lam, and Laurens Swinkels. 2014. "The Global Multi-Asset Market Portfolio, 1959–2012." *Financial Analysts Journal*, vol. 70, no. 2 (March/April): 26–41.

Dybvig, Philip H. 2005. "Mean-variance portfolio rebalancing with transaction costs." Working paper, Washington University in Saint Louis.

Elton, Edwin J, and Martin J. Gruber. 1992. "Optimal Investment Strategies with Investor Liabilities." *Journal of Banking & Finance*, vol. 16, no. 5: 869–890.

Gannon, James A., and Bob Collins. 2009. "Liability-Responsive Asset Allocation." Russell Research Viewpoint.

Goldberg, Lisa R., Michael Y. Hayes, and Ola Mahmoud. 2013. "Minimizing Shortfall." *Quantitative Finance*, vol. 13, no. 10: 1533–1545.

Grable, John E. 2008. "RiskCAT: A Framework for Identifying Maximum Risk Thresholds in Personal Portfolios." *Journal of Financial Planning*, vol. 21, no. 10: 52–62.

Grable, John E., and Soo-Hyun Joo. 2004. "Environmental and Biopsychosocial Factors Associated with Financial Risk Tolerance." *Financial Counseling and Planning*, vol. 15, no. 1: 73–88.

Harvey, Campbell R., John C. Liechty, Merrill W. Liechty, and Peter Müller. 2010. "Portfolio Selection with Higher Moments." *Quantitative Finance*, vol. 10, no. 5 (May): 469–485.

Idzorek, Thomas. 2008. "Lifetime Asset Allocations: Methodologies for Target Maturity Funds." Ibbotson Associates Research Report.

Idzorek, Thomas M., and Maciej Kowara. 2013. "Factor-Based Asset Allocation vs. Asset-Class-Based Asset Allocation." *Financial Analysts Journal*, vol. 69, no. 3 (May/June): 19–29.

Jobson, David J., and Bob Korkie. 1980. "Estimation for Markowitz Efficient Portfolios." *Journal of the American Statistical Association*, vol. 75, no. 371 (September): 544–554.

Jobson, David J., and Bob Korkie. 1981. "Putting Markowitz Theory to Work." *Journal of Portfolio Management*, vol. 7, no. 4 (Summer): 70–74.

Jorion, Phillipe. 1992. "Portfolio Optimization in Practice." *Financial Analysts Journal*, vol. 48, no. 1 (January/February): 68–74.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, vol. 47, no. 2: 263–292.

Kat, Harry M. 2003. "10 Things That Investors Should Know about Hedge Funds." *Journal of Wealth Management*, vol. 5, no. 4 (Spring): 72–81.

Leibowitz, Martin L., and Roy D. Henriksson. 1988. "Portfolio Optimization Within a Surplus Framework." *Financial Analysts Journal*, vol. 44, no. 2: 43–51.

Leland, Hayne. 2000. "Optimal Portfolio Implementation with Transaction Costs and Capital Gains Taxes." Working paper, University of California, Berkeley.

Luenberger, David G. 2013. *Investment Science*, 2nd ed. New York: Oxford University Press.

Markowitz, Harry M. 1952. "Portfolio Selection." *Journal of Finance*, vol. 7, no. 1 (March): 77–91.

Markowitz, Harry M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.

Maslow, A. H. 1943. "A Theory of Human Motivation." *Psychological Review*, vol. 50, no. 4: 370–396.

Masters, Seth J. 2003. "Rebalancing." *Journal of Portfolio Management*, vol. 29, no. 3: 52–57.

McCalla, Douglas B. 1997. "Enhancing the Efficient Frontier with Portfolio Rebalancing." *Journal of Pension Plan Investing*, vol. 1, no. 4: 16–32.

Michaud, Richard O. 1998. *Efficient Asset Management*. Boston: Harvard Business School Press.

Mulvey, John M. 1989. "A Surplus Optimization Perspective." *Investment Management Review*, vol. 3: 31–39.

Mulvey, John M. 1994. "An Asset-Liability System." *Interfaces*, vol. 24, no. 3: 22–33.

Mulvey, J.M., and W. Kim. 2009. "Constantly Rebalanced Portfolio—Is Mean Reversion Necessary?" in Rama Cont, ed. *Encyclopedia of Quantitative Finance*, vol 2, Hoboken, NJ: John Wiley & Sons.

Mulvey, John M., Bill Pauling, and Ron E. Madey. 2003. "Advantages of Multi-Period Portfolio Models." *Journal of Portfolio Management*, vol. 29, no. 2 (Winter): 35–45.

Nevins, Daniel. 2004. "Goal-Based Investing: Integrating Traditional and Behavioral Finance." *Journal of Wealth Management*, vol. 6, no. 4 (Spring): 8–23.

Pompian, Michael M., and John B. Longo. 2004. "A New Paradigm for Practical Application of Behavioral Finance." *Journal of Wealth Management*, vol. 7, no. 2 (Fall): 9–15.

Rockafellar, R. Tyrrell, and Stanislav Uryasev. 2000. "Optimization of Conditional Value-at-Risk." *Journal of Risk*, vol. 2, no. 3 (Spring): 21–41.

Rudd, Andrew, and Laurence B. Siegel. 2013. "Using an Economic Balance Sheet for Financial Planning." *Journal of Wealth Management*, vol. 16, no. 2 (Fall): 15–23.

Scherer, Bernd. 2002. "Portfolio Resampling: Review and Critique." *Financial Analysts Journal*, vol. 58, no. 6 (November/December): 98–109.

Sharpe, William. 1974. "Imputing Expected Security Returns from Portfolio Composition." *Journal of Financial and Quantitative Analysis*, vol. 9, no. 3 (June): 463–472.

Sharpe, William, and Lawrence G. Tint. 1990. "Liabilities: A New Approach." *Journal of Portfolio Management*, vol. 16, no. 2: 5–10.

Shefrin, H., and M. Statman. 2000. "Behavioral Portfolio Theory." *Journal of Financial and Quantitative Analysis*, vol. 35, no. 2: 127–151.

Swensen, D. 2009. *Pioneering Portfolio Management: An Unconventional Approach to Institutional Investment*, 2nd ed. New York: Free Press.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty*, vol. 5, no. 4: 297–323.

Willenbrock, Scott. 2011. "Diversification Return, Portfolio Rebalancing, and the Commodity Return Puzzle." *Financial Analysts Journal*, vol. 67, no. 4 (July/August): 42–49.

Winkelmann, Kurt. 2003. "Developing an Optimal Active Risk Budget," in *Modern Investment Management: An Equilibrium Approach* Bob Litterman, ed. New York: John Wiley & Sons.

Xiong, James X., and Thomas M. Idzorek. 2011. "The Impact of Skewness and Fat Tails on the Asset Allocation Decision." *Financial Analysts Journal*, vol. 67, no. 2 (March/April): 23–35.

Yale University. 2014. "The Yale Endowment 2014" (http://investments.yale.edu/s/Yale_Endowment_14.pdf).

PRACTICE PROBLEMS

The following information relates to questions 1–8

Megan Beade and Hanna Müller are senior analysts for a large, multi-divisional money management firm. Beade supports the institutional portfolio managers, and Müller does the same for the private wealth portfolio managers.

Beade reviews the asset allocation in Exhibit 1, derived from a mean–variance optimization (MVO) model for an institutional client, noting that details of the MVO are lacking.

Exhibit 1 Asset Allocation and Market Weights (in percent)

Asset Classes	Asset Allocation	Investable Global Market Weights
Cash	0	—
US bonds	30	17
US TIPS	0	3
Non-US bonds	0	22
Emerging market equity	25	5
Non-US developed equity	20	29
US small- and mid-cap equity	25	4
US large-cap equity	0	20

The firm's policy is to rebalance a portfolio when the asset class weight falls outside of a corridor around the target allocation. The width of each corridor is customized for each client and proportional to the target allocation. Beade recommends wider corridor widths for high-risk asset classes, narrower corridor widths for less liquid asset classes, and narrower corridor widths for taxable clients with high capital gains tax rates.

One client sponsors a defined benefit pension plan where the present value of the liabilities is \$241 million and the market value of plan assets is \$205 million. Beade expects interest rates to rise and both the present value of plan liabilities and the market value of plan assets to decrease by \$25 million, changing the pension plan's funding ratio.

Beade uses a surplus optimization approach to liability-relative asset allocation based on the objective function

$$U_m^{LR} = E\left(R_{s,m}\right) - 0.005\lambda\sigma^2\left(R_{s,m}\right)$$

where $E(R_{s,m})$ is the expected surplus return for portfolio m, λ is the risk aversion coefficient, and $\sigma^2(R_{s,m})$ is the variance of the surplus return. Beade establishes the expected surplus return and surplus variance for three different asset allocations, shown in Exhibit 2. Given $\lambda = 1.50$, she chooses the optimal asset mix.

© 2018 CFA Institute. All right reserved.

Exhibit 2 Expected Surplus Return and Volatility for Three Portfolios

	Return	Standard Deviation
Portfolio 1	13.00%	24%
Portfolio 2	12.00%	18%
Portfolio 3	11.00%	19%

Client Haunani Kealoha has a large fixed obligation due in 10 years. Beade assesses that Kealoha has substantially more funds than are required to meet the fixed obligation. The client wants to earn a competitive risk-adjusted rate of return while maintaining a high level of certainty that there will be sufficient assets to meet the fixed obligation.

In the private wealth area, the firm has designed five subportfolios with differing asset allocations that are used to fund different client goals over a five-year horizon. Exhibit 3 shows the expected returns and volatilities of the subportfolios and the probabilities that the subportfolios will exceed an expected minimum return. Client Luis Rodríguez wants to satisfy two goals. Goal 1 requires a conservative portfolio providing the highest possible minimum return that will be met at least 95% of the time. Goal 2 requires a riskier portfolio that provides the highest minimum return that will be exceeded at least 85% of the time.

Exhibit 3 Characteristics of Subportfolios

Subportfolio	A	B	C	D	E
Expected return, in percent	4.60	5.80	7.00	8.20	9.40
Expected volatility, in percent	3.46	5.51	8.08	10.80	13.59
Required Success Rate	**Minimum Expected Return for Success Rate**				
99%	1.00	0.07	–1.40	–3.04	–4.74
95%	2.05	1.75	1.06	0.25	–0.60
90%	2.62	2.64	2.37	2.01	1.61
85%	3.00	3.25	3.26	3.19	3.10
75%	3.56	4.14	4.56	4.94	5.30

Müller uses a risk parity asset allocation approach with a client's four–asset class portfolio. The expected return of the domestic bond asset class is the lowest of the asset classes, and the returns of the domestic bond asset class have the lowest covariance with other asset class returns. Müller estimates the weight that should be placed on domestic bonds.

Müller and a client discuss other approaches to asset allocation that are not based on optimization models or goals-based models. Müller makes the following comments to the client:

Comment 1 An advantage of the "120 minus your age" heuristic over the 60/40 stock/bond heuristic is that it incorporates an age-based stock/bond allocation.

Comment 2 The Yale model emphasizes traditional investments and a commitment to active management.

Comment 3 A client's asset allocation using the 1/*N* rule depends on the investment characteristics of each asset class.

1 The asset allocation in Exhibit 1 *most likely* resulted from a mean–variance optimization using:

A historical data.

B reverse optimization.

C Black–Litterman inputs.

2 Beade's suggested change in the corridor width of the rebalancing policy is correct regarding:

A high-risk asset classes.

B less liquid asset classes.

C taxable clients with high capital gains tax rates.

3 Based on Beade's interest rate expectations, the pension plan's funding ratio will:

A decrease.

B remain unchanged.

C increase.

4 Based on Exhibit 2, which portfolio provides the greatest objective function expected value?

A Portfolio 1

B Portfolio 2

C Portfolio 3

5 The asset allocation approach most appropriate for client Kealoha is *best* described as:

A a surplus optimization approach.

B an integrated asset–liability approach.

C a hedging/return-seeking portfolios approach.

6 Based on Exhibit 3, which subportfolios *best* meet the two goals expressed by client Rodríguez?

A Subportfolio A for Goal 1 and Subportfolio C for Goal 2

B Subportfolio B for Goal 1 and Subportfolio C for Goal 2

C Subportfolio E for Goal 1 and Subportfolio A for Goal 2

7 In the risk parity asset allocation approach that Müller uses, the weight that Müller places on domestic bonds should be:

A less than 25%.

B equal to 25%.

C greater than 25%.

8 Which of Müller's comments about the other approaches to asset allocation is correct?

A Comment 1

B Comment 2

C Comment 3

The following information relates to questions 9–13

Investment adviser Carl Monteo determines client asset allocations using quantitative techniques such as mean–variance optimization (MVO) and risk budgets. Monteo is reviewing the allocations of three clients. Exhibit 1 shows the expected return and standard deviation of returns for three strategic asset allocations that apply to several of Monteo's clients.

Exhibit 1 Strategic Asset Allocation Alternatives

	Adviser's Forecasts	
Asset Allocation	Expected Return (%)	Standard Deviation of Returns (%)
A	10	12.0
B	8	8.0
C	6	2.0

Monteo interviews client Mary Perkins and develops a detailed assessment of her risk preference and capacity for risk, which is needed to apply MVO to asset allocation. Monteo estimates the risk aversion coefficient (λ) for Perkins to be 8 and uses the following utility function to determine a preferred asset allocation for Perkins:

$$U_m = E(R_m) - 0.005\lambda\sigma_m^2$$

Another client, Lars Velky, represents Velky Partners (VP), a large institutional investor with $500 million in investable assets. Velky is interested in adding less liquid asset classes, such as direct real estate, infrastructure, and private equity, to VP's portfolio. Velky and Monteo discuss the considerations involved in applying many of the common asset allocation techniques, such as MVO, to these asset classes. Before making any changes to the portfolio, Monteo asks Velky about his knowledge of risk budgeting. Velky makes the following statements:

Statement 1 An optimum risk budget minimizes total risk.

Statement 2 Risk budgeting decomposes total portfolio risk into its constituent parts.

Statement 3 An asset allocation is optimal from a risk-budgeting perspective when the ratio of excess return to marginal contribution to risk is different for all assets in the portfolio.

Monteo meets with a third client, Jayanta Chaterji, an individual investor. Monteo and Chaterji discuss mean–variance optimization. Chaterji expresses concern about using the output of MVOs for two reasons:

Criticism 1: The asset allocations are highly sensitive to changes in the model inputs.

Criticism 2: The asset allocations tend to be highly dispersed across all available asset classes.

Monteo and Chaterji also discuss other approaches to asset allocation. Chaterji tells Monteo that he understands the factor-based approach to asset allocation to have two key characteristics:

Characteristic 1	The factors commonly used in the factor-based approach generally have low correlations with the market and with each other.
Characteristic 2	The factors commonly used in the factor-based approach are typically different from the fundamental or structural factors used in multifactor models.

Monteo concludes the meeting with Chaterji after sharing his views on the factor-based approach.

9 Based on Exhibit 1 and the risk aversion coefficient, the preferred asset allocation for Perkins is:

- **A** Asset Allocation A.
- **B** Asset Allocation B.
- **C** Asset Allocation C.

10 In their discussion of the asset classes that Velky is interested in adding to the VP portfolio, Monteo should tell Velky that:

- **A** these asset classes can be readily diversified to eliminate idiosyncratic risk.
- **B** indexes are available for these asset classes that do an outstanding job of representing the performance characteristics of the asset classes.
- **C** the risk and return characteristics associated with actual investment vehicles for these asset classes are typically significantly different from the characteristics of the asset classes themselves.

11 Which of Velky's statements about risk budgeting is correct?

- **A** Statement 1
- **B** Statement 2
- **C** Statement 3

12 Which of Chaterji's criticisms of MVO is/are valid?

- **A** Only Criticism 1
- **B** Only Criticism 2
- **C** Both Criticism 1 and Criticism 2

13 Which of the characteristics put forth by Chaterji to describe the factor-based approach is/are correct?

- **A** Only Characteristic 1
- **B** Only Characteristic 2
- **C** Both Characteristic 1 and Characteristic 2

SOLUTIONS

1 A is correct. The allocations in Exhibit 1 are most likely from an MVO model using historical data inputs. MVO tends to result in asset allocations that are concentrated in a subset of the available asset classes. The allocations in Exhibit 1 have heavy concentrations in four of the asset classes and no investment in the other four asset classes, and the weights differ greatly from global market weights. Compared to the use of historical inputs, the Black–Litterman and reverse-optimization models most likely would be less concentrated in a few asset classes and less distant from the global weights.

2 A is correct. Higher-risk assets should have a wider corridor to avoid frequent, costly rebalancing. Beade's other suggestions are not correct. Less liquid asset classes should have a wider, not narrower, corridor width. Less liquid assets should have a wider corridor to avoid frequent rebalancing. For taxable investors, transactions trigger capital gains in jurisdictions that tax them. For such investors, higher tax rates on capital gains should be associated with wider (not narrower) corridor widths.

3 A is correct. The original funding ratio is the market value of assets divided by the present value of liabilities. This plan's ratio is \$205 million/\$241 million = 0.8506. When the assets and liabilities both decrease by \$25 million, the funding ratio will decrease to \$180 million/\$216 million = 0.8333.

4 B is correct. The objective function expected value is $U_m^{LR} = E\left(R_{s,m}\right) - 0.005\lambda\sigma^2\left(R_{s,m}\right)$. λ is equal to 1.5, and the expected value of the objective function is shown in the rightmost column below.

Portfolio	$E(R_s,m)$	$\sigma^2(R_{s,m})$	$U_m^{LR} = E(R_{s,m}) - 0.005(1.5)\sigma^2(R_{s,m})$
1	13.00	576	8.68
2	12.00	324	9.57
3	11.00	361	8.29

Portfolio 2 generates the highest value, or utility, in the objective function.

5 C is correct. The hedging/return-seeking portfolios approach is best for this client. Beade should construct two portfolios, one that includes riskless bonds that will pay off the fixed obligation in 10 years and the other a risky portfolio that earns a competitive risk-adjusted return. This approach is a simple two-step process of hedging the fixed obligation and then investing the balance of the assets in a return-seeking portfolio.

6 A is correct. Goal 1 requires a success rate of at least 95%, and Subportfolio A has the highest minimum expected return (2.05%) meeting this requirement. Goal 2 requires the highest minimum expected return that will be achieved 85% of the time. Subportfolio C meets this requirement (and has a minimum expected return of 3.26%).

7 C is correct. A risk parity asset allocation is based on the notion that each asset class should contribute equally to the total risk of the portfolio. Bonds have the lowest risk level and must contribute 25% of the portfolio's total risk, so bonds must be overweighted (greater than 25%). The equal contribution of each asset class is calculated as:

$$w_i \times \text{Cov}\left(r_i, r_p\right) = \frac{1}{n}\sigma_p^2$$

where

w_i = weight of asset i
$\text{Cov}(r_i, r_p)$ = covariance of asset i with the portfolio
n = number of assets
σ_p^2 = variance of the portfolio

In this example, there are four asset classes, and the variance of the total portfolio is assumed to be 25%; therefore, using a risk parity approach, the allocation to each asset class is expected to contribute (1/4 × 25%) = 6.25% of the total variance. Because bonds have the lowest covariance, they must have a higher relative weight to achieve the same contribution to risk as the other asset classes.

8 A is correct. Comment 1 is correct because the "120 minus your age" rule reduces the equity allocation as the client ages, while the 60/40 rule makes no such adjustment. Comments 2 and 3 are not correct. The Yale model emphasizes investing in alternative assets (such as hedge funds, private equity, and real estate) as opposed to investing in traditional asset classes (such as stock and bonds). The 1/*N* rule allocates an equal weight to each asset without regard to its investment characteristics, treating all assets as indistinguishable in terms of mean returns, volatility, and correlations.

9 C is correct. The risk aversion coefficient (λ) for Mary Perkins is 8. The utility of each asset allocation is calculated as follows:

Asset Allocation A:

$$\begin{aligned} U_A &= 10.0\% - 0.005(8)(12\%)^2 \\ &= 4.24\% \end{aligned}$$

Asset Allocation B:

$$\begin{aligned} U_B &= 8.0\% - 0.005(8)(8\%)^2 \\ &= 5.44\% \end{aligned}$$

Asset Allocation C:

$$\begin{aligned} U_C &= 6.0\% - 0.005(8)(2\%)^2 \\ &= 5.84\% \end{aligned}$$

Therefore, the preferred strategic allocation is Asset Allocation C, which generates the highest utility given Perkins's level of risk aversion.

10 C is correct. Less liquid asset classes—such as direct real estate, infrastructure, and private equity—represent unique challenges when applying many of the common asset allocation techniques. Common illiquid asset classes cannot be readily diversified to eliminate idiosyncratic risk, so representing overall asset class performance is problematic. Furthermore, there are far fewer indexes that attempt to represent aggregate performance for these less liquid asset classes than indexes of traditional highly liquid asset classes. Finally, the risk and return characteristics associated with actual investment vehicles—such as direct real estate funds, infrastructure funds, and private equity funds—are typically significantly different from the characteristics of the asset classes themselves.

11 B is correct. The goal of risk budgeting is to maximize return per unit of risk. A risk budget identifies the total amount of risk and attributes risk to its constituent parts. An optimum risk budget allocates risk efficiently.

12 A is correct. One common criticism of MVO is that the model outputs, the asset allocations, tend to be highly sensitive to changes in the model. Another common criticism of MVO is that the resulting asset allocations tend to be highly concentrated in a subset of the available asset classes.

13 A is correct. The factors commonly used in the factor-based approach generally have low correlations with the market and with each other. This results from the fact that the factors typically represent what is referred to as a zero (dollar) investment or self-financing investment, in which the underperforming attribute is sold short to finance an offsetting long position in the better-performing attribute. Constructing factors in this manner removes most market exposure from the factors (because of the offsetting short and long positions); as a result, the factors generally have low correlations with the market and with one another. Also, the factors commonly used in the factor-based approach are typically similar to the fundamental or structural factors used in multifactor models.

PORTFOLIO MANAGEMENT

STUDY SESSION

10

Asset Allocation and Related Decisions in Portfolio Management (2)

In practice, the asset allocation decision is affected by numerous constraints that present practical challenges to asset allocation. Significant investor-based constraints include investable assets, liquidity needs, time horizon, and regulatory and tax environments.

This study session examines the effects of these constraints and presents adaptations to address them by institutional investor type. Also discussed are behavioral biases that influence the asset allocation process and ways to overcome these biases.

When the strategic asset allocation includes exposure to global markets, nondomestic currencies create additional sources of portfolio volatility and potential returns. How currency exposures can be managed to reflect a client's investment objectives and constraints is explored.

READING ASSIGNMENTS

Reading 20	Asset Allocation with Real-World Constraints by Peter Mladina, Brian J. Murphy, CFA, and Mark Ruloff, FSA, EA, CERA
Reading 21	Currency Management: An Introduction by William A. Barker, PhD, CFA

© 2018 CFA Institute. All rights reserved.

READING

20

Asset Allocation with Real-World Constraints

by Peter Mladina, Brian J. Murphy, CFA, and Mark Ruloff, FSA, EA, CERA

Peter Mladina is at Northern Trust and UCLA (USA). Brian J. Murphy, CFA, is at Willis Towers Watson (USA). Mark Ruloff, FSA, EA, CERA, is at Aon Hewitt (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** discuss asset size, liquidity needs, time horizon, and regulatory or other considerations as constraints on asset allocation;
☐	**b.** discuss tax considerations in asset allocation and rebalancing;
☐	**c.** recommend and justify revisions to an asset allocation given change(s) in investment objectives and/or constraints;
☐	**d.** discuss the use of short-term shifts in asset allocation;
☐	**e.** identify behavioral biases that arise in asset allocation and recommend methods to overcome them.

1 INTRODUCTION

This reading illustrates ways in which the asset allocation process must be adapted to accommodate specific asset owner circumstances and constraints. It addresses adaptations to the asset allocation inputs given an asset owner's asset size, liquidity, and time horizon as well as external constraints that may affect the asset allocation choice (Section 2). We also discuss the ways in which taxes influence the asset allocation process for the taxable investor (Section 3). In addition, we discuss the circumstances that should trigger a re-evaluation of the long-term strategic asset allocation (Section 4), when and how an asset owner might want to make short-term shifts in asset allocation (Section 5), and how innate investor behaviors can interfere with successful long-term planning for the investment portfolio (Section 6). Throughout the reading, we illustrate the application of these concepts using a series of hypothetical investors.

© 2017 CFA Institute. All rights reserved.

2 CONSTRAINTS IN ASSET ALLOCATION

General asset allocation principles assume that all asset owners have equal ability to access the entirety of the investment opportunity set, and that it is merely a matter of finding that combination of asset classes that best meets the wants, needs, and obligations of the asset owner. In practice, however, it is not so simple. An asset owner must consider a number of constraints when modeling and choosing among asset allocation alternatives. Some of the most important are asset size, liquidity needs, taxes, and time horizon. Moreover, regulatory and other external considerations may influence the investment opportunity set or the optimal asset allocation decision.

2.1 Asset Size

The size of an asset owner's portfolio has implications for asset allocation. It may limit the opportunity set—the asset classes accessible to the asset owner—by virtue of the scale needed to invest successfully in certain asset classes or by the availability of investment vehicles necessary to implement the asset allocation.

Economies and diseconomies of scale are perhaps the most important factors relevant to understanding asset size as a constraint. The size of an asset owner's investment pool may be too small—or too large—to capture the returns of certain asset classes or strategies efficiently. Asset owners with larger portfolios can generally consider a broader set of asset classes and strategies. On the one hand, they are more likely to have sufficient governance capacity—sophistication and staff resources—to develop the required knowledge base for the more complex asset classes and investment vehicles. They also have sufficient size to build a diversified portfolio of investment strategies, many of which have substantial minimum investment requirements. On the other hand, some asset owners may have portfolios that are *too* large; their desired minimum investment may exhaust the capacity of active external investment managers in certain asset classes and strategies. Although "too large" and "too small" are not rigidly defined, the following example illustrates the difficulty of investing a very large portfolio. Consider an asset owner with an investment portfolio of US$25 billion who is seeking to make a 5% investment in global small-cap stocks:

- The median total market capitalization of the stocks in the S&P Global SmallCap is approximately US$555 million.
- Assume a small-cap manager operates a 50-stock portfolio and is willing to own 3% of the market cap of any one of its portfolio companies. Their average position size would be US$17 million, and an effective level of assets under management (AUM) would be on the order of US$850 million. Beyond that level, the manager may be forced to expand the portfolio beyond 50 stocks or to hold position sizes greater than 3% of a company's market cap, which could then create liquidity issues for the manager.
- Now, our US$25 billion fund is looking to allocate US$1.25 billion to small-cap stocks (US$25 billion × 5%). They want to diversify this allocation across three or four active managers—a reasonable allocation of governance resources in the context of all of the fund's investment activities. The average allocation per manager is approximately US$300 to US$400 million, which would constitute between 35% and 50% of each manager's AUM. This exposes both the asset owner and the investment manager to an undesirable level of operational risk.

Although many large asset owners have found effective ways to implement a small-cap allocation, this example illustrates some of the issues associated with managing a large asset pool. These include such practical considerations as the number of investment managers that might need to be hired to fulfill an investment allocation and the ability of the asset owner to identify and monitor the required number of managers.

Research has shown that investment managers tend to incur certain disadvantages from increasing scale: Growth in AUM leads to larger trade sizes, incurring greater price impact; capital inflows may cause active investment managers to pursue ideas outside of their core investment theses; and organizational hierarchies may slow down decision making and reduce incentives.[1] Asset *owners,* however, are found to have *increasing* returns to scale, as discussed below.

A study of pension plan size and performance (using data spanning 1990–2008) found that large defined benefit plans outperformed smaller ones by 45–50 basis points per year on a risk-adjusted basis.[2] The gains are derived from a combination of cost savings related to internal management, a greater ability to negotiate fees with external managers, and the ability to support larger allocations to private equity and real estate investments. As fund size increases, the "per participant" costs of a larger governance infrastructure decline and the plan sponsor can allocate resources away from such asset classes as small-cap stocks, which are sensitive to diseconomies of scale, to such other areas as private equity funds or co-investments where they are more likely to realize scale-related benefits.

Whereas owners of large asset pools may achieve these operating efficiencies, scale may also impose obstacles related to the liquidity and trading costs of the underlying asset. Above some size, it becomes difficult to deploy capital effectively in certain active investment strategies. As illustrated in Exhibit 1, owners of very large portfolios may face size constraints in allocating to active equity strategies. The studies referenced earlier noted that these asset owners frequently choose to invest passively in developed equity markets where their size inhibits alpha potential. The asset owner's finite resources can then be allocated instead toward such strategies as private equity, hedge funds, and infrastructure, where their scale and resources provide a competitive advantage.

Exhibit 1 Asset Size and Investor Constraints

Asset Class	Investor Constraints by Size
■ Cash equivalents and money market funds	No size constraints.
■ Large-cap developed market equity ■ Small-cap developed market equity ■ Emerging market equity	Generally accessible to large and small asset owners, although the very large asset owner may be constrained in the amount of assets allocated to certain active strategies and managers.

(continued)

1 See Stein (2002); Chen, Hong, Huang, and Kubik (2004); and Pollet and Wilson (2008).

2 See Dyck and Pomorski (2011). The median plan in this study was just over US$2 billion. The 25th percentile plan was US$780 million, and the 75th percentile plan was US$6.375 billion.

Exhibit 1 (Continued)

Asset Class	Investor Constraints by Size
■ Developed market sovereign bonds ■ Investment-grade bonds ■ Non-investment-grade bonds ■ Private real estate equity	Generally accessible to large and small asset owners, although to achieve prudent diversification, smaller asset owners may need to implement via a commingled vehicle.
Alternative Investments ■ Hedge funds ■ Private debt ■ Private equity ■ Infrastructure ■ Timberland and farmland	May be accessible to large and small asset owners, although if offered as private investment vehicles, there may be legal minimum qualifications that exclude smaller asset owners. The ability to successfully invest in these asset classes may also be limited by the asset owner's level of investment understanding/expertise. Prudent diversification may require that smaller asset owners implement via a commingled vehicle, such as a fund of funds, or an ancillary access channel, such as a liquid alternatives vehicle or an alternatives ETF. For very large funds, the allocation may be constrained by the number of funds available.

Even in these strategies, very large asset owners may be constrained by scale. In smaller or less liquid markets, can a large asset owner invest enough that the exposure contributes a material benefit to the broader portfolio? For example, a sovereign wealth fund or large public pension plan may not find enough attractive hedge fund managers to fulfill their desired allocation to hedge funds. True alpha is rare, limiting the opportunity set. Asset owners who find that they have to split their mandate into many smaller pieces may end up with an index-like portfolio but with high active management fees; one manager's active bets may cancel out those of another active manager. A manager mix with no true alpha becomes index-like because the uncompensated, idiosyncratic return variation is diversified away. A much smaller allocation may be achievable, but it may be too small to meaningfully affect the risk and return characteristics of the overall portfolio. More broadly, a very large size makes it more difficult to benefit from opportunistic investments in smaller niche markets or from skilled investment managers who have a small set of unique ideas or concentrated bets. No hard and fast rules exist to determine whether a particular asset owner is too small or too large to effectively access an asset class. Greater governance resources more commonly found among owners of larger asset pools create the capacity to pursue the more complex investment opportunities, but the asset owner may still need to find creative ways to implement the desired allocation. Each asset owner has a unique set of knowledge and constraints that will influence the opportunity set.

Smaller asset owners (typically institutions with less than US$500 million in assets, and private wealth investors with less than US$25 million in assets) also find that their opportunity set may be constrained by the size of their investment portfolio. This is primarily a function of the more limited governance infrastructure typical of smaller asset owners: They may be too small to adequately diversify across the range of asset classes and investment managers or may have staffing constraints (insufficient asset size to justify a dedicated internal staff). Complex strategies may be beyond the reach of asset owners that have chosen not to develop investment expertise internally or where the oversight committee lacks individuals with sufficient investment understanding. In some asset classes and strategies, commingled investment vehicles can be used to achieve the needed diversification, provided the governing documents do not prohibit their use.

Access to other asset classes and strategies—private equity, private real estate, hedge funds, and infrastructure—may still be constrained for smaller asset owners. The commingled vehicles through which these strategies are offered typically require high minimum investments. For successful private equity and hedge fund managers, in particular, minimum investments can be in the tens of millions of (US) dollars, even for funds of funds.

Regulatory restrictions can also impose a size constraint. In the United Kingdom, for example, an asset owner in a private investment vehicle must qualify as an elective professional client, meaning they must meet two of the following three conditions:

1 The client has carried out transactions, in significant size, on the relevant market at an average frequency of 10 per quarter over the previous four quarters.
2 The size of the client's financial instrument portfolio exceeds €500,000.
3 The client works or has worked in the financial sector for at least one year in a professional position, which requires knowledge of the transactions or services envisaged.

In the United States, investors must be either accredited or qualified purchasers to invest in many private equity and hedge fund vehicles. To be a qualified purchaser, a natural person must have at least US$5 million in investments, a company must have at least US$25 million in investable assets, and an investment manager must have at least US$25 million under management. In Hong Kong, the Securities and Futures Commission requires that an investor must meet the qualifications of a "Professional Investor" to invest in certain categories of assets. A Professional Investor is generally defined as a trust with total assets of not less than HK$40 million, an individual with a portfolio not less than HK$8 million, or a corporation or partnership with a portfolio not less than HK$8 million or total assets of not less than HK$40 million. The size constraints related to these asset classes suggest that smaller asset owners have real challenges achieving an effective private equity or hedge fund allocation.

Asset size as a constraint is often a more acute issue for individual investors than institutional asset owners. Wealthy families may pool assets through such vehicles as family limited partnerships, investment companies, fund of funds, or other forms of commingled vehicles to hold their assets. These pooled vehicles can then access investment vehicles, asset classes, and strategies that individual family members may not have portfolios large enough to access on their own.

Where Asset Size Constrains Investment Opportunity

As of early 2016, the 10 largest sovereign wealth funds globally each exceed US$400 billion in assets. For a fund of this size, a 5% allocation to hedge funds (the average sovereign wealth fund allocation) would imply US$20 *billion* to be deployed. The global hedge fund industry manages approximately US$2.8 trillion in total; 73% of the funds manage less than US$100 million. The remaining 27% of the funds (roughly 3,000) manage 72% of the industry's AUM; their implied average AUM is therefore US$670 million. If we assume that the asset owner would want to be no more than 20% of a firm's AUM, we can infer that the average investment might be approximately US$130 million. With US$20 billion to deploy, the fund would need to invest with nearly 150 funds to achieve a 5% allocation to hedge funds.

Sources: Sovereign Wealth Fund Institute, BarclayHedge, Eurekahedge (2016).

EXAMPLE 1

Asset Size Constraints in Asset Allocation

1 Akkarat Aromdee is the recently retired President of Alpha Beverage, a producer and distributor of energy drinks throughout Southeast Asia. Upon retiring, the company provided a lump sum retirement payment of THB880,000,000 (equivalent to €20 million), which was rolled over to a tax-deferred individual retirement savings plan. Aside from these assets, Aromdee owns company stock worth about THB70,000,000. The stock is infrequently traded. He has consulted with an investment adviser, and they are reviewing the following asset allocation proposal:

Global equities	40%
Global high-yield bonds	15%
Domestic intermediate bonds	30%
Hedge funds	10%
Private equity	5%

Describe asset size constraints that Aromdee might encounter in implementing this asset allocation. Discuss possible means to address them.

2 The CAF$40 billion Government Petroleum Fund of Caflandia is overseen by a nine-member Investment Committee. The chief investment officer has a staff with sector heads in global equities, global bonds, real estate, hedge funds, and derivatives. The majority of assets are managed by outside investment managers. The Investment Committee, of which you are a member, approves the asset allocation policy and makes manager selection decisions. Staff has recommended an increase in the private equity allocation from its current 0% to 15%, to be implemented over the next 12 to 36 months. The head of global equities will oversee the implementation of the private equity allocation.

Given the asset size of the fund, formulate a set of questions regarding the feasibility of this recommendation that you would like staff to address at the next Investment Committee meeting.

3 The Courneuve University Endowment has US$250 million in assets. The current allocation is 65% global large-capitalization stocks and 35% high-quality bonds, with a duration target of 5.0 years. The University has adopted a 5% spending policy. University enrollment is stable and expected to remain so. A capital spending initiative of US$100 million for new science buildings in the next three to seven years is being discussed, but it has not yet been approved. The University has no dedicated investment staff and makes limited use of external resources. Investment recommendations are formulated by the University's treasurer and approved by the Investment Committee, composed entirely of external board members.

The new president of the University has stated that he feels the current policy is overly restrictive, and he would like to see a more diversified program that takes advantage of the types of investment strategies used by large endowment programs. Choosing from among the following asset classes, propose a set of asset classes to be considered in the revised asset allocation. Justify your response.

- Cash equivalents and money market funds
- Large-cap developed market equity
- Small-cap developed market equity
- Emerging market equity
- Developed market sovereign bonds
- Investment-grade bonds
- Non-investment-grade bonds
- Private real estate equity
- Hedge funds
- Private debt
- Private equity

Solution to 1:

With a THB88 million (€2 million) allocation to hedge funds and a THB44 million (€1 million) allocation to private equity funds, Aromdee may encounter restrictions on his eligibility to invest in the private investment vehicles typically used for hedge fund and private equity investment. To the extent he is eligible to invest in hedge funds and/or private equity funds, a fund-of-funds or similar commingled arrangement would be essential to achieving an appropriate level of diversification. Additionally, it is essential that he and his adviser develop the necessary level of expertise to invest in these alternative assets. To achieve a prudent level of diversification, the allocation to global high-yield bonds would most likely need to be accomplished via a commingled investment vehicle.

Solution to 2:

Questions regarding the feasibility of the recommendation include the following:

- How many private equity funds do you expect to invest in to achieve the 15% allocation to private equity?
- What is the anticipated average allocation to each fund?
- Are there a sufficient number of high-quality private equity funds willing to accept an allocation of that size?
- What expertise exists at the staff or board level to conduct due diligence on private equity investment funds?
- What resources does the staff have to oversee the increased allocation to private equity?

Solution to 3:

Asset size and limited governance resources are significant constraints on the investment opportunity set available to the Endowment. The asset allocation should emphasize large and liquid investments, such as cash equivalents, developed and emerging market equity, and sovereign and investment-grade bonds. Some small portion of assets, however, could be allocated to commingled investments in real estate, private equity, or hedge funds. Given the University's limited staff resources, it is necessary to ensure that the board members have the level of expertise necessary to select and monitor these more complex asset classes. The Endowment might also consider engaging an outside expert to advise on investment activities in these asset classes.

2.2 Liquidity

Two dimensions of liquidity must be considered when developing an asset appropriate allocation solution: the liquidity needs of the asset owner and the liquidity characteristics of the asset classes in the opportunity set. Integrating the two dimensions is an essential element of successful investment planning.

The need for liquidity in an investment portfolio will vary greatly by asset owner and by the goals the assets are set aside to achieve. For example, a bank will typically have a very large portfolio supporting its day-to-day operations. That portfolio is likely to experience very high turnover and a very high need for liquidity; therefore, the investment portfolio must hold high-quality, very short-term, and highly liquid assets.

The same bank may have another designated investment pool one level removed from operating assets. Although the liquidity requirements for this portfolio may be lower, the investments most likely feature a high degree of liquidity—a substantial allocation to investment-grade bonds, perhaps with a slight extension of maturity. For its longer-term investment portfolio, the bank may choose to allocate some portion of its portfolio to less liquid investments. The opportunity set for each portfolio will be constrained by applicable banking laws and regulations.

Long-term investors, such as sovereign wealth funds and endowment funds, can generally exploit illiquidity premiums available in such asset classes as private equity, real estate, and infrastructure investments. However, pension plans may be limited in the amount of illiquidity they can absorb. For example, a frozen pension plan may anticipate the possibility of eliminating its pension obligation completely by purchasing a group annuity and relinquishing the responsibility for making pension payments to an insurance company. If there is a significant probability that the company will take this step in the near term, liquidity of plan assets will become a primary concern; and if there is a substantial allocation to illiquid assets, the plan sponsor may be unable to execute the desired annuity purchase transaction.

Liquidity needs must also consider the particular circumstances and financial strength of the asset owner and what resources they may have beyond those held in the investment portfolio. The following examples illustrate this point:

- A university must consider its prospects for future enrollments and the extent to which it relies on tuition to meet operating needs. If the university experiences a significant drop in enrollment, perhaps because of a poor economic environment, or takes on a new capital improvement project, the asset allocation policy for the endowment should reflect the increased probability of higher outflows to support university operations.
- A foundation whose mission supports medical research in a field in which a break-through appears imminent may desire a higher level of liquidity to fund critical projects than would a foundation that supports ongoing community efforts.
- An insurance company whose business is predominantly life or auto insurance, where losses are actuarially predictable, can absorb more liquidity risk than a property/casualty reinsurer whose losses are subject to unpredictable events, such as natural disasters.
- A family with several children nearing college-age will have higher liquidity needs than a couple of the same age and circumstances with no children.

When assessing the appropriateness of any given asset class for a given asset owner, it is wise to evaluate potential liquidity needs in the context of an extreme market stress event. The market losses of the 2008–2009 global financial crisis were extreme. Simultaneously, other forces exacerbated investors' distress: Many university endowments were called upon to provide an increased level of operating support; insurers dipped into reserves to offset operating losses; community foundations found their beneficiaries in even greater need of financial support; and some individual investors experienced setbacks that caused them to move, if only temporarily, from being net contributors to net spenders of financial wealth. A successful asset allocation effort will stress the proposed allocation; it will anticipate, where possible, the likely behavior of other facets of the saving/spending equation during times of stress.

It is also important to consider the intersection of asset class and investor liquidity in the context of the asset owner's governance capacity. Although the mission of the organization or trust may allow for a certain level of illiquidity, if those responsible for the oversight of the investment program do not have the mental fortitude or discipline to maintain course through the crisis, illiquid and less liquid investments are unlikely to produce the rewards typically expected of these exposures. Although rates of return may be mean-reverting, wealth is not. Losses resulting from panic selling during times of stress become permanent losses; there are fewer assets left to earn returns in a post-crash recovery.

The Case of Vanishing Liquidity

In the global financial crisis of 2008–2009, many investors learned painful truths about liquidity. When most needed—whether to rebalance or to meet spending obligations—it can evaporate. As investors liquidated their most liquid assets to meet financial obligations (or to raise cash in fear of further market declines), the remaining less liquid assets in their portfolios became an ever-larger percentage of the portfolio. Many investors were forced to sell private partnership interests on the secondary market at steeply discounted prices. Others defaulted on outstanding private fund capital commitments.[3]

Similarly, illiquidity became a substantial problem during the Asian currency crisis of 1997–1998 and again with the Russian debt default and Long-Term Capital Management (LTCM) crisis of 1998. In the following paragraphs, we describe several "liquidity crises" that are often used in stress testing asset allocation choices.

The Asian Currency Crisis of 1997

In the spring of 1997, Thailand spent billions to defend the Thai baht against speculative attacks, finally capitulating and devaluing the baht in July 1997. This triggered a series of moves throughout the region to defend currencies against speculators. Ultimately, these efforts were unsuccessful and many countries abandoned the effort and allowed their currencies to float freely. The Philippines, Indonesia, and South Korea abandoned their pegs against the US dollar. On 27 October 1997, rattled by the currency crisis, Asian and European markets declined sharply in advance of the opening of the US markets. The S&P 500 declined nearly 7%, and trading on US stock markets was suspended.

The Russian Debt/LTCM Crisis of August 1998

On 17 August 1998, the Russian government defaulted on its short-term debt. This unprecedented default of a sovereign debtor roiled the global bond markets. A global flight-to-quality ensued, which caused credit spreads to widen and liquidity to evaporate. Highly levered investors experienced significant losses. Long-Term Capital Management, with reported notional exposure of over US$125 billion (a 25-to-1 leverage ratio), exacerbated these price declines as they faced their own liquidity crisis and were forced to liquidate large relative value, distressed, convertible arbitrage, merger arbitrage, and equity positions. Ultimately, the magnitude of the liquidity squeeze for LTCM and the risk of potential disruption to global markets caused the New York branch of the Federal Reserve Bank to orchestrate a disciplined, structured bailout of the LTCM fund.

Financial markets are increasingly linked across borders and asset classes; as a result, changes in liquidity conditions in one country can directly affect liquidity conditions elsewhere. These linkages do improve access to financing and capital markets, but they

3 As an example, Endowment Fund A (EFA) committed US$10 million to Private Equity Fund A (PEF-A) two years ago. Since then, the manager of PEF-A called US$8 million from EFA to make private equity investments. EFA still has a contractual obligation to contribute US$2 million as PEF-A finds investments it plans to add to its portfolio. EFA informs PEF-A that it will not honor the remaining US$2 million capital commitment.

also show that a liquidity problem in one part of the world can ripple across the globe—increasing volatility, creating higher execution costs for investors, and possibly leading to a reduction in credit availability and a decline in economic activity.

EXAMPLE 2

Liquidity Constraints in Asset Allocation

The Frentel Furniture Pension Fund has £200 million frozen in a defined benefit pension plan that is 85% funded. The plan has a provision that allows employees to elect a lump sum distribution of their pension benefit at retirement. The company is strong financially and is committed to fully funding the pension obligations over time. However, they also want to minimize cash contributions to the plan. Few governance resources are allocated to the pension fund, and there is no dedicated staff for pension investment activities. The current asset allocation is as shown:

Global equities	20%
Private equity	10%
Real estate	10%
Infrastructure	5%
Hedge funds	15%
Bonds	40%

The company expects to reduce their employee headcount sometime in the next three to five years, and they are tentatively planning incentives to encourage employees to retire early.

Discuss the appropriateness of the current asset allocation strategy for the pension fund, including benefits and concerns.

Solution:

In addition to the size constraints a £200 million (≈ US$250 million) plan faces when attempting to invest in real estate, private equity, infrastructure, and hedge funds, the likelihood of early retirement incentives and lump-sum distribution requests in the next three to five years indicates a need for increased sensitivity to liquidity concerns. Investments in private equity, infrastructure, and real estate may be unsuitable for the plan given their less liquid nature. Although hedge fund investments would likely be accessible via a commingled vehicle, the liquidity of the commingled vehicle should be evaluated to determine if it is consistent with the liquidity needs of the plan.

2.3 Time Horizon

An asset owner's time horizon is a critical constraint that must be considered in any asset allocation exercise. A liability to be paid at a given point in the future or a goal to be funded by a specified date each define the asset owner's horizon, thus becoming a basic input to the asset allocation solution. The changing composition of the asset owner's assets and liabilities must also be considered. As time progresses, the character of both *assets* (human capital) and *liabilities* changes.

Changing Human Capital When asset allocation considers such extended portfolio assets as human capital, the optimal allocation of financial capital can change through time (Bodie, Merton, and Samuelson 1992). Assuming no change in the investor's utility function, as human capital—with its predominately bond-like risk—declines over time, the asset allocation for financial capital would reflect an increasing allocation to bonds. This is a prime example of how time horizon can influence asset allocation.

Changing Character of Liabilities The changing character of liabilities through time will also affect the asset allocation aligned to fund those liabilities.

As an example, the term structure of liabilities changes as they approach maturity. A pension benefit program is a simple way to illustrate this point. When the employee base is young and retirements are far into the future, the liability can be hedged with long-term bonds. As the employee base ages and prospective retirements are not so far into the future, the liability is more comparable to intermediate- or even short-term bonds. When retirements are imminent, the structure of the liabilities can be characterized as cash-like, and an optimal asset allocation would also have cash-like characteristics.

Similarly, the overall profile of an individual investor's liabilities changes with the progression of time, particularly for investors with finite investment horizons. Nearer-term goals and liabilities move from partially funded to fully funded, while other, longer-term goals and liabilities move progressively closer to funding. As the relative weights of the goals to be funded shift and the time horizon associated with certain goals shortens, the aggregate asset allocation must be adapted if it is to remain aligned with the individual's goals.

Time horizon is also likely to affect the manner in which an investor prioritizes certain goals and liabilities. This will influence the desired risk profile of the assets aligned to fund them. Consider a 75-year-old retired investor with two goals:

1 Fund consumption needs through age 95
2 Fund consumption needs from age 95 through age 105

He most likely assigns a much higher priority to funding goal 1, given the lower probability that he will live beyond age 95.[4] Let's also assume that he has sufficient assets to fund goal 1 and to partially fund goal 2. The higher priority assigned to goal 1 indicates he is less willing to take risk, and this sub-portfolio will be invested more conservatively. Now consider goal 2: Given the low probability of living past 95 and the fact that he does not currently have sufficient assets to fund that goal, the sub-portfolio assigned to goal 2 is likely to have a more growth-oriented asset allocation. The priority of a given goal can change as the investor's time horizon shortens—or lengthens.

Consider the hypothetical investors Ivy and Charles Lee from the reading "Introduction to Asset Allocation." Ivy is a 54-year-old life science entrepreneur. Charles is a 55-year-old orthopedic surgeon. They have two unmarried children aged 25 (Deborah) and 18 (David). Deborah has a daughter with physical limitations. Four goals have been identified for the Lees:

1 Lifestyle/future consumption needs
2 College education for son David, 18 years old
3 Charitable gift to a local art museum in 5 years
4 Special needs trust for their granddaughter, to be funded at the death of Charles

---

4 A 75-year-old US American male has a life expectancy of 11.3 years, per the Social Security Administration's 2013 "Actuarial Life Tables," https://www.ssa.gov/oact/STATS/table4c6.html.

The lifestyle/consumption goal is split into three components: required minimum consumption requirements (a worst-case scenario of reduced lifestyle), baseline consumption needs (maintaining current standard of living), and aspirational consumption needs (an improved standard of living). At age 54, the risk preferences assigned to these goals might look something like the following:

Lifestyle Goals	Risk Preference	Asset Allocation	Sub-Portfolio as % of Total*
Required minimum	Conservative	100% bonds and cash	65%
Baseline	Moderate	60% equities/40% bonds	10%
Aspirational	Aggressive	100% equities	4%
College education	Conservative	100% bonds and cash	1%
Charitable gift (aspirational)	Aggressive	100% equities	5%
Special needs trust	Moderate	60% equities/40% bonds	15%
Aggregate portfolio		≈ 25% equities/75% bonds and cash	100%

* The present value of each goal as a proportion of the total portfolio.

The asset allocation for the total portfolio aggregates the asset allocations for each of the goal-aligned sub-portfolios, weighted by the present value of each goal. For the Lees, this is an overall asset allocation of about 25% equities and 75% bonds and cash. (Each goal is discounted to its present value by expected return of its respective goal-aligned sub-portfolio.)

Move forward 20 years. The Lees are now in their mid-70s, and their life expectancy is about 12 years. Their son has completed his college education and is successfully established in his own career. The charitable gift has been made. These two goals have been realized. The assets needed to fund the baseline consumption goal are significantly reduced because fewer future consumption years need to be funded. The special needs trust for their granddaughter remains a high priority. Although the Lee's risk preferences for these goals have not changed, the overall asset allocation *will* change because the total portfolio is an aggregated mix of the remaining goal-aligned sub-portfolios, weighted by their current present values:

Lifestyle Goals	Risk Preference	Asset Allocation	Sub-Portfolio as % of Total*
Required minimum	Conservative	100% bonds and cash	54%
Baseline	Moderate	60% equities/40% bonds	9%
Aspirational	Aggressive	100% equities	3%
Special needs trust	Moderate	60% equities/40% bonds	34%
Aggregate portfolio		≈ 30% equities/70% bonds and cash	100%

* The present value of each goal as a proportion of the total portfolio. The implied assumption is that current assets are sufficient to fund all goals, provided the Lees adopt an aggressive asset allocation strategy for the aspirational and charitable gifting goals. If the value of current assets exceeds the present value of all goals, the Lees would have greater flexibility to adopt a lower risk preference for some or all goals.

Although for ease of illustration our example assumed the Lee's risk preferences remained the same, this is not likely to be the case in the real world. Required minimum and baseline consumption goals would remain very important; there is less flexibility to withstand losses caused by either reduced earnings potential or lower likelihood of the market regaining lost ground within the shorter horizon. The aspirational lifestyle goal is likely to be a much lower priority, and it may have been eliminated altogether. The special needs trust may have a higher (or lower) priority as the needs of the granddaughter and the ability of her parents to provide for her needs after their death become more evident. The preferred asset allocation for each of these goals will shift over the course of the investor's lifetime.

As an investor's time horizon shifts, both human capital and financial market considerations, along with changes in the investor's priorities, will most likely lead to different asset allocation decisions.

EXAMPLE 3

Time Horizon Constraints in Asset Allocation

Akkarat Aromdee, the recently retired President of Alpha Beverage, is 67 years old with a remaining life expectancy of 15 years. Upon his retirement two years ago, he established a charitable foundation and funded it with THB600 million (≈ US$17.3 million). The remaining financial assets, THB350 million (≈ US$10 million), were transferred to a trust that will allow him to draw a lifetime income. The assets are invested 100% in fixed-income securities, consistent with Aromdee's desire for a high level of certainty in meeting his goals. He is a widower with no children. His consumption needs are estimated at THB20 million annually. Assets remaining in the trust at his death will pass to the charities named in the trust.

While vacationing in Ko Samui, Aromdee met and later married a 45-year-old woman with two teenage children. She has limited financial assets of her own. Upon returning from his honeymoon, Aromdee meets with his investment adviser. He intends to pay the college expenses of his new stepchildren—THB2 million annually for eight years, beginning five years from now. He would also like to ensure that his portfolio can provide a modest lifetime income for his wife after his death.

Discuss how these changed circumstances are likely to influence Aromdee's asset allocation.

Solution:

At the time Aromdee established the trust, the investment horizon was 15 years and his annual consumption expenditures could easily be funded from the trust. His desire to support his new family introduces two new horizons to be considered: In five years, the trust will begin making annual payments of THB2 million to fund college expenses, and the trust will continue to make distributions to his wife after his death, though at a reduced rate. When the trust needed to support only his consumption requirements, a conservative asset allocation was appropriate. However, the payment of college expenses will reduce his margin of safety and the lengthening of the investment horizon suggests that he should consider adding equity-oriented investments to the asset mix to provide for growth in assets over time.

Time Diversification of Risk

In practice, investors often align lower risk/lower return assets with short-term goals and liabilities and higher risk/higher return assets with long-term goals and liabilities. It is generally believed that longer-horizon goals can tolerate the higher volatility associated with higher risk/higher return assets as below average and above average returns even out over time. This is the notion of time diversification.

Mean–variance optimization, typically conducted using a multi-year time horizon, assumes that asset returns follow a random walk; returns in Year X are independent of returns in Year X – 1. Under this baseline assumption, there *is* no reduction in risk with longer time horizons.[5] Although the *probability* of reduced wealth or of a shortfall in funding a goal or liability (based on the mean of the distribution of possible outcomes) may be lower at longer time horizons, the dispersion of possible outcomes widens as the investment horizon expands. Thus, the *magnitude* of potential loss or shortfall can be greater.

Consider the choice of investing US$100,000 in an S&P 500 Index fund with a 10% expected return and 15% standard deviation versus a risk-free asset with a 3% annual return.[6] The table below compares the return of the risk-free asset over various time horizons, with the range of predicted returns for the S&P 500 Index fund at a 95% confidence interval. Although the *mean* return of the distribution of S&P 500 returns exceeds that of the risk-free asset in each time period (thus the notion that the volatility of higher risk, higher return assets evens out over time), the lower boundary of expected S&P 500 returns is less than the initial investment for all periods less than 10 years! The lower boundary of the S&P 500 outcomes does not exceed the ending wealth of the risk-free investment until the investment horizon is extended to 20 years. If the confidence interval is expanded to 99%, the lower boundary of S&P 500 outcomes falls below the initial investment up until and through 20 years!

	Ending Wealth (US$)		
	S&P 500 95% Confidence Interval		
	Lower Boundary	**Upper Boundary**	**Risk-Free Asset**
1 year	81,980	147,596	103,000
5 years	83,456	310,792	115,927
10 years	102,367	657,196	134,392
15 years	133,776	130,4376	155,797
20 years	180,651	2,565,345	180,611

Although one-year returns are largely independent, there is some evidence that risky asset returns can display mean-reverting tendencies over intermediate to longer time horizons. An assumption of mean-reverting risky asset returns would support the conventional arguments for funding long-term goals and liabilities with higher risk/higher return assets, and it would also support a reduction in the allocation to these riskier assets as the time horizon shortens.

5 See Samuelson (1963) and Samuelson (1969).
6 This example is drawn from Kritzman (2015).

2.4 Regulatory and Other External Constraints

Just as an integrated asset/liability approach to asset allocation is likely to result in a different allocation decision than what might have been selected in an asset-only context, external considerations may also influence the asset allocation decision. Local laws and regulations can have a material effect on an investor's asset allocation decisions.

Pension funds, insurance companies, sovereign wealth funds, and endowments and foundations are each subject to externally imposed constraints that are likely to tilt their asset allocation decision away from what may have been selected in a pure asset/liability context.

2.4.1 *Insurance Companies*

Unlike pension fund or endowment assets—which are legally distinct from the assets of the sponsoring entity—insurance companies' investment activities are an integral part of their day-to-day operations. Although skilled underwriting may be the focus of the firm as the key to profitability, investment returns are often a material contributor to profits or losses. Regulatory requirements and accounting treatment vary from country to country, but insurers are most often highly focused on matching assets to the projected, probabilistic cash flows of the risks they are underwriting. Fixed-income assets, therefore, are typically the largest component of an insurance company's asset base, and investing with skill in this asset class is a key to competitive pricing and success. In some regions, the relevant accounting treatment may be a book value approach, rendering variability in the market pricing of assets to be a secondary consideration as long as an asset does not have to have its book value written down as "other than temporarily impaired" ("OTTI"). Risk considerations for an insurance company include the need for capital to pay policyholder benefits and other factors that directly influence the company's financial strength ratings. Some of the key considerations are risk-based capital measures, yield, liquidity, the potential for forced liquidation of assets to fund negative claims development, and credit ratings.

Additionally, allocations to certain asset classes are often constrained by a regulator. For example, the maximum limit on equity exposure is often 10%, but it ranges as high as 30% in Switzerland and 40% in Chile. Israel and Korea impose a limit of 15% on real estate investments. Restrictions on non-publicly traded securities might also limit the allocation to such assets as private equity, for example, and there may also be limits on the allocation to high-yield bonds. Insurance regulators generally set a minimum capital level for each insurer based on that insurer's mix of assets, liabilities, and risk. Many countries are moving to Solvency II regulatory standards designed to harmonize risk-based capital requirements for insurance companies across countries.[7] Asset classes are often treated differently for purposes of determining whether an insurer meets risk-based capital requirements.

2.4.2 *Pension Funds*

Pension fund asset allocation decisions may be constrained by regulation and influenced by tax rules. Some countries regulate maximum or minimum percentages in certain asset classes. For example, Japanese pension funds must hold a certain minimum percentage of assets in Japanese bonds in order to maintain their tax-exempt status. Mexico limits a plan's investment in foreign securities to no more than 20%; Brazil limits real estate investments to 8%; and China allows real estate investments only through bonds and shares of property companies. Ukraine limits bond investments to no more than 40%.[8]

7 Solvency II is an EU legislative program implemented in all 28 member states, including the United Kingdom, in January 2016. It introduces a new, harmonized EU-wide insurance regulatory regime.

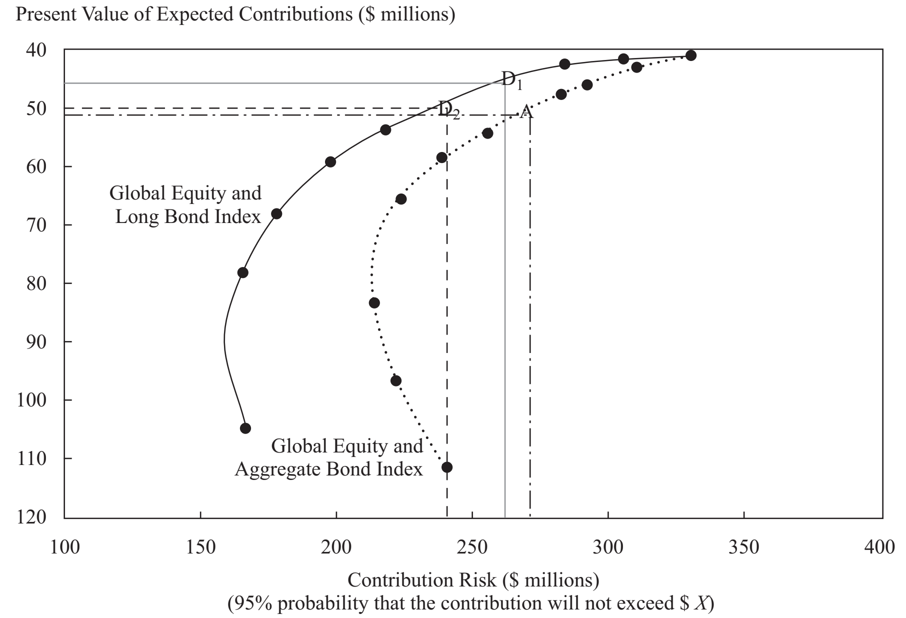
8 OECD, "Annual Survey of Investment Regulation of Pension Funds" (2014).

Pension funds are also subject to a wide array of funding, accounting, reporting, and tax constraints that may influence the asset allocation decision. (For example, US public pension funding and public and corporate accounting rules favor equity investments—higher equity allocations support a higher discount rate—and thus lower pension cost. Loss recognition is deferred until later through the smoothing mechanism.) The plan sponsor's appetite for risk is defined in part by these constraints, and the choice among asset allocation alternatives is often influenced by funding and financial statement considerations, such as the anticipated contributions, the volatility of anticipated contributions, or the forecasted pension expense or income under a given asset allocation scenario. The specific constraints vary by jurisdiction, and companies with plans in multiple jurisdictions must satisfy the rules and regulations of each jurisdiction while making sound financial decisions for the organization as a whole.

Exhibit 2 illustrates how funding considerations may affect the asset allocation decision. In this chart, risk is defined as the probability of contributions exceeding some threshold amount. In this case, the risk threshold is specified as the 95th percentile of the present value of contributions—that point on the distribution of possible contributions (using Monte Carlo simulation) where the plan sponsor can be 95% certain that contributions will not exceed that amount.

Assume that an allocation of 70% equities/30% aggregate bonds represents the most efficient portfolio for the plan sponsor's desired level of risk in an asset optimization framework. In Exhibit 2, we can see that the 70% equity/30% aggregate bond mix (Portfolio A) is associated with a present value (PV) of expected contributions of approximately US\$51 million (y-axis) and a 95% confidence level that contributions will not exceed approximately US\$275 million (x-axis)—Portfolio A in Exhibit 2. If the plan sponsor were to shift to longer-duration bonds (from aggregate to long bonds) to better match the duration of liabilities—Portfolio D_1 on Exhibit 2—the PV of expected contributions declines by approximately US\$5 million and the 95% confidence threshold improves to approximately US\$265 million. In fact, Portfolio D_1 results in nearly the lowest PV of contributions for this plan sponsor. (Note that the vertical axis is ordered from highest contributions at the bottom and lowest contributions at the top, consistent with the notion of lower contributions as a better outcome.)

Exhibit 2 Efficient Frontiers Where Risk Is Defined as the Risk of Large Contributions



Now consider Portfolio D_2, 60% equities/40% long bonds. Reducing the equity exposure from 70% to 60% lowers the contribution risk significantly, with only marginally higher expected PV of contributions than Portfolio A. (A lower equity allocation implies a lower expected rate of return, which increases the PV of contributions. However, the lower equity allocation also reduces the probability that less-than-expected returns will lead to unexpectedly large contributions.) The sponsor that wishes to reduce contribution risk is likely to give serious consideration to moving from Portfolio A to Portfolio D_2.

By iterating through various efficient frontiers using different definitions of risk, the sponsor is able to better understand the risk and reward trade-offs of alternative asset allocation choices. The regulatory or tax constraints on minimum and maximum contributions, or on minimum required funded levels, or other values that are important to the plan sponsor, can be factored into the simulations so the sponsor can better understand how these constraints might affect the risk and reward trade-offs.

2.4.3 *Endowments and Foundations*

Endowments and foundations are often established with the expectation that they will exist in perpetuity and thus can invest with a long investment horizon. In addition, the sponsoring entity often has more flexibility over payments from the fund than does a pension plan sponsor or insurance company. As a result, endowments and foundations generally can adopt a higher-risk asset allocation than other institutions. However, two categories of externally imposed constraints may influence the asset allocation decisions of an endowment or foundation: tax incentives and credit-worthiness considerations.

- *Tax incentives*. Although some endowments and foundations—US public foundations and some Austrian and Asian foundations, for example—are not required to make minimum distributions, many countries provide tax benefits tied to certain minimum spending requirements. For example, a private foundation may be subject to a requirement that it make charitable expenditures equal to at least 5% of the market value of its assets each year or risk losing its tax-favored status. These spending requirements may be relaxed if certain types of socially responsible investments are made, which can, in turn, create a bias toward socially responsible investments for some endowments and foundations, irrespective of their merits in an asset allocation context.

- *Credit considerations*. Although endowments and foundations typically have a very long investment horizon, sometimes external factors may restrict the level of risk-taking in the portfolio. For example, endowment or foundation assets are often used to support the balance sheet and borrowing capabilities of the university or the foundation organization. Lenders often require that the borrower maintain certain minimum balance sheet ratios. Therefore, the asset allocation adopted by the organization will consider the risks of breaking these bond covenants or otherwise negatively affecting the borrowing capabilities of the organization.

As an example, although a hospital foundation fund would normally have a long investment horizon and the ability to invest in less liquid asset classes, it might limit the allocation to illiquid assets in order to support certain liquidity and balance sheet metrics specified by its lender(s).

2.4.4 *Sovereign Wealth Funds*

Although every sovereign wealth fund (SWF) is unique with respect to its mission and objectives, some broad generalizations can be made with respect to the external constraints that may affect a fund's asset allocation choices. In general, SWFs are government-owned pools of capital invested on behalf of the peoples of their states or countries, investing with a long-term orientation. They are not generally seeking to defease a set of liabilities or known obligations as is common with pension funds and, to a lesser extent, endowment funds.

The governing entities adopt regulations that constrain the opportunity set for asset allocation. For example, the Chilean SWF may hold no more than 7% in equities; the Korean SWF KIC cannot invest in Korean won-denominated assets; and the Norwegian SWF NBIM is not permitted to invest in any alternative asset class other than real estate, which is limited to no more than 10% of assets. Furthermore, as publicly owned entities, SWFs are typically subject to broad public scrutiny and tend to adopt a lower-risk asset allocation than might otherwise be considered appropriate given their long-term investment horizon in order to avoid reputation risk.

In addition to the broad constraints of asset size, liquidity, time horizon, and regulations, there may be cultural or religious factors which also constrain the asset allocation choices. Environmental, social, and governance (ESG) considerations are becoming increasingly important to institutional and individual investors alike. Sharia law, for example, prohibits investment in any business that has links to pork, alcohol, tobacco, pornography, prostitution, gambling, or weaponry, and it constrains investments in most businesses that operate on interest payments (like major Western banks and mortgage providers) and in businesses that transfer risk (such as major Western insurers).[9]

ESG goals are not typically modeled during the asset allocation decision process. Instead, these goals may be achieved through the implementation of the asset allocation, or the asset owner may choose to set aside a targeted portion of the assets for these missions. The asset allocation process would treat this "set-aside" in much the same way that a concentrated stock position might be handled: The risk, return, and correlation characteristics of this holding are specified; the "set aside" asset becomes an asset class in the investor's opportunity set; and the asset allocation constraints will designate a certain minimum investment in this asset class.

9 Islamic Investment Network (www.islamicinvestmentnetwork.com/sharialaw.php).

EXAMPLE 4

External Constraints and Asset Allocation

1 An insurance company has traditionally invested its pension plan using the asset allocation strategy adopted for its insurance assets: The pension assets are 95% invested in high-quality intermediate duration bonds and 5% in global equities. The duration of pension liabilities is approximately 25 years. Until now, the company has always made contributions sufficient to maintain a fully funded status. Although the company has a strong capability to fund the plan adequately and a relatively high tolerance for variability in asset returns, as part of a refinement in corporate strategy, management is now seeking to reduce long-term expected future cash contributions. Management is willing to accept more risk in the asset return, but they would like to limit contribution risk and the risk to the plan's funded status. The Investment Committee is considering three asset allocation proposals for the pension plan:

   A Maintain the current asset allocation with the same bond portfolio duration.

   B Increase the equity allocation and lengthen the bond portfolio duration to increase the hedge of the duration risk in the liabilities.

   C Maintain the current asset allocation of 95% bonds and 5% global equities, but increase the duration of bond investments.

   Discuss the merits of each proposal.

2 A multinational corporation headquartered in Mexico has acquired a former competitor in the United States. It will maintain both the US pension plan with US$250 million in assets and the Mexican pension plan with MXN$18,600 million in assets (≈ US$1 billion). Both plans are 95% funded and have similar liability profiles. The Mexican pension trust has an asset allocation policy of 30% equities (10% invested in the Mexican equity market and 20% in equity markets outside Mexico), 10% hedge funds, 10% private equity, and 50% bonds. The treasurer has proposed that the company adopt a consistent asset allocation policy across all of the company's pension plans worldwide.

   Critique the treasurer's proposal.

Solution to 1:

Given the intermediate duration bond allocation, Proposal A fails to consider the mismatch between pension assets and liabilities and risks a reduction in the funded status and *increased* contributions if bond yields decline. (If yields decline across the curve, the shorter duration bond portfolio will fail to hedge the increase in liabilities.) To meet the objective of lower future contributions, the asset allocation must include a higher allocation to equities. Proposal B has this higher allocation, and the extension of duration in the bond portfolio in Proposal B reduces balance sheet and surplus risk relative to the pension liabilities. The net effect could be a reduction in short-term contribution risk; moreover, if the greater expected return on equities is realized, it should result in reduced contributions to the plan over the long term. Proposal C improves the hedging of the liabilities, and it may result in a modest improvement in the expected return on assets if the yield curve is upward-sloping. However, the expected return on Proposal C is likely lower than the expected return of

Proposal B and is therefore unlikely to achieve the same magnitude of reduction in future cash contributions. Proposal C would be appropriate if the goal was focused on reducing surplus risk rather than reducing long-term contributions.

Solution to 2:

The treasurer's proposal fails to consider the relative asset size of the two pension plans as well as the likelihood that plans in different jurisdictions may be subject to different funding, regulatory, and financial reporting requirements. The US pension plan may be unable to effectively access certain alternative asset classes, such as private equity, infrastructure, and hedge funds. Although economies of scale may be realized if management of the pension assets is consolidated under one team, the legal and regulatory differences of the markets in which they operate mean that the asset allocation policy must be customized to each plan.

3 ASSET ALLOCATION FOR THE TAXABLE INVESTOR

Portfolio theory developed in a frictionless world. But in the real world, taxes on income and capital gains can erode the returns achieved by taxable investors. The asset owner who ignores taxes during the asset allocation process is overlooking an economic variable that can materially alter the outcome. Although tax adjustments can be made after the asset allocation has been determined, this is a suboptimal approach because the pre-tax and after-tax risk and return characteristics of each asset class can be materially different.

Some assets are less tax efficient than others because of the character of their returns—the contribution of interest, dividends, and realized or unrealized capital gains to the total return. Interest income is usually taxed in the tax year it is received, and it often faces the highest tax rates. Therefore, assets that generate returns largely comprised of interest income tend to be less tax efficient in many countries.[10] Jurisdictional rules can also affect how the returns of certain assets are taxed. In the United States, for example, the interest income from state and local government bonds is generally exempt from federal income taxation. As a result, these bonds often constitute a large portion of a US high-net-worth investor's bond allocation. Preferred stocks, often used in lieu of bonds as an income-producing asset, are also eligible for more favorable tax treatment in many jurisdictions, where the income from preferred shares may be taxed at more favorable dividend tax rates.

The tax environment is complex. Different countries have different tax rules and rates, and these rules and rates can change frequently. However, looking across the major economies, there are some high-level commonalities in how investment returns are taxed. Interest income is taxed typically (but not always) at progressively higher income tax rates. Dividend income and capital gains are taxed typically (but not always) at lower tax rates than those applied to interest income and earned income (wages and salaries, for example). Capital losses can be used to offset capital gains (and sometimes income). Generally, interest income incurs the highest tax rate, with dividend income taxed at a lower rate in some countries, and long-term capital gains receive the most favorable tax treatment in many jurisdictions. Once we move beyond these general commonalities, however, the details of tax treatment among countries quickly diverge.

10 See Deloitte's tax guides and country highlights: https://dits.deloitte.com/#TaxGuides.

Entities and accounts can be subject to different tax rules. For example, retirement savings accounts may be tax deferred or tax exempt, with implications for the optimal asset allocation solution. These rules provide opportunities for strategic asset *location*—placing less tax-efficient assets in tax-advantaged accounts.

We will provide a general framework for considering taxes in asset allocation. We will not survey global tax regimes or incorporate all potential tax complexities into the asset allocation solution. When considering taxes in asset allocation, the objective is to model material investment-related taxes, thereby providing a closer approximation to economic reality than is represented when ignoring taxes altogether.

For simplicity, we will assume a basic tax regime that represents no single country but includes the key elements of investment-related taxes that are roughly representative of what a typical taxable asset owner in the major developed economies must contend with.

3.1 After-Tax Portfolio Optimization

After-tax portfolio optimization requires adjusting each asset class's expected return and risk for expected tax. The expected after-tax return is defined in Equation 1:

$$r_{at} = r_{pt}(1 - t) \tag{1}$$

where

r_{at} = the expected after-tax return
r_{pt} = the expected pre-tax (gross) return
t = the expected tax rate

This can be straightforward for bonds in cases where the expected return is driven by interest income. Take, for example, an investment-grade par bond with a 3% coupon expected to be held to maturity. If interest income is subject to a 40% expected tax rate, the bond has an expected after-tax return of 1.80% [0.03(1 – 0.40) = 0.018].

The expected return for equity typically includes both dividend income and price appreciation (capital gains). Equation 2 expands Equation 1 accordingly:

$$r_{at} = p_d r_{pt}(1 - t_d) + p_a r_{pt}(1 - t_{cg}) \tag{2}$$

where

p_d = the proportion of r_{pt} attributed to dividend income
p_a = the proportion of r_{pt} attributed to price appreciation
t_d = the dividend tax rate
t_{cg} = the capital gains tax rate

The treatment of the capital gains portion of equity returns can be more complex. Assuming no dividend income, a stock with an 8% expected pre-tax return that is subject to a 25% capital gains tax rate has an expected after-tax return of 6% [0.08(1 – 0.25) = 0.06]. This is an approximation satisfactory for modeling purposes.[11]

11 A more precise estimation of the expected after-tax return also takes into account the effect of the holding period on the capital gains tax. The 6% expected after-tax return in the example assumes the capital gain is realized and taxed at the end of a 1-year holding period. If the holding period is materially longer than one year, an economic benefit accrues to the investor from the compounding of the deferred capital gains tax liability. This equates to a reduction in the *implied annual* capital gains tax rate that can be used in place of t_{cg} in Equation 2. Other tax complexities can be factored into asset class assumptions—for example, various taxes and tax treatments, expenses, withholding taxes on foreign investments for otherwise tax-exempt investors, and portfolio turnover rates—but these issues are beyond the scope of this reading. For those interested in a more detailed discussion of these issues, see Mladina (2011).

Taxable assets may have existing unrealized capital gains or losses (i.e., the cost basis is below or above market value), which come with embedded tax liabilities (or tax assets). Although there is not a clear consensus on how best to deal with existing unrealized capital gains (losses), many approaches adjust the asset's current market value for the value of the embedded tax liability (asset) to create an after-tax value. Reichenstein (2006) approximates the after-tax value by subtracting the value of the embedded capital gains tax from the market value, as if the asset were sold today. Horan and Al Zaman (2008) assume the asset is sold in the future and discount the tax liability to its present value using the asset's after-tax return as the discount rate. Turvey, Basu, and Verhoeven (2013) argue that the after-tax risk-free rate is the more appropriate discount rate because the embedded tax liability is analogous to an interest-free loan from the government, where the tax liability can be arbitraged away by dynamically investing in the risk-free asset. We will discuss how to incorporate after-tax values into the portfolio optimization process in Section 3.3, where we address strategies to reduce the impact of taxes.

The ultimate purpose of an asset can be a consideration when modeling tax adjustments. In the preceding material on asset allocation, we discussed goals-based investing. If the purpose of a given pool of assets is to fund consumption in 10 years, then that 10-year holding period may influence the estimated implied annual capital gains tax rate. If the purpose of the specified pool of assets is to fund a future gift of appreciated stock to a tax-exempt charity, then capital gains tax may be ignored altogether. Through this alignment of goals with assets, goals-based investing facilitates more-precise tax adjustments.

Although correlation assumptions need not be adjusted when modeling asset allocation choices for the taxable asset owner (taxes are proportional to return, after-tax co-movements are the same as pre-tax co-movements), taxes do affect the standard deviation assumption for each asset class. The expected after-tax standard deviation is defined in Equation 3:

$$\sigma_{at} = \sigma_{pt}(1 - t) \tag{3}$$

where

σ_{at} = the expected after-tax standard deviation

σ_{pt} = the expected pre-tax standard deviation

Taxes alter the distribution of returns by both reducing the expected mean return and muting the dispersion of returns. Taxes truncate both the high and low ends of the distribution of returns, resulting in lower highs and higher lows. The effect of taxes is intuitive when considering a positive return, but the same economics apply to a negative return: Losses are muted by the same $(1 - t)$ tax adjustment. The investor is not taxed on losses but instead receives the economic benefit of a capital loss, whether realized or not. In many countries, a realized capital loss can offset a current or future realized capital gain. An unrealized capital loss captures the economic benefit of a cost basis that is above the current market value, making a portion of expected future appreciation tax free.

How does the optimal asset allocation along a pre-tax efficient frontier compare with the optimal asset allocation along an after-tax efficient frontier? Let's assume all investment assets are taxable and that cost bases equal current market values. Assume also that interest income is taxed at 40%, and dividend income and capital gains are taxed at 25%.

The asset classes we will consider include investment-grade (IG) bonds, high-yield (HY) bonds, and equity. Exhibit 3 shows the expected pre-tax returns and standard deviations for each asset class as well as the correlation matrix.[12]

Exhibit 3 Expected Pre-Tax Return and Risk

	Return	Std. Dev.	
IG bonds	3.0%	4.0%	
HY bonds	5.0%	10.0%	
Equity	8.0%	20.0%	
Correlations	**IG Bonds**	**HY Bonds**	**Equity**
IG bonds	1.0	0.2	0.0
HY bonds	0.2	1.0	0.7
Equity	0.0	0.7	1.0

Employing mean–variance portfolio optimization with these pre-tax inputs, we obtain the optimal asset allocations in Exhibit 4, which shows the allocations for portfolios P1 (lowest risk), P25, P50 (median risk), P75, and P100 (highest risk)—each on an efficient frontier comprised of 100 portfolios.

Exhibit 4 Optimal Pre-Tax Asset Mixes

	$P1_{pt}$	$P25_{pt}$	$P50_{pt}$	$P75_{pt}$	$P100_{pt}$
IG bonds	93%	52%	25%	0%	0%
HY bonds	5%	18%	26%	33%	0%
Equity	2%	30%	49%	67%	100%

Using Equations 1, 2, and 3, we calculate the expected after-tax returns and standard deviations displayed in Exhibit 5. No adjustments are made to correlations.

Exhibit 5 Expected After-Tax Return and Risk

	Return	Std. Dev.
IG bonds	1.8%	2.4%
HY bonds	3.0%	6.0%
Equity	6.0%	15.0%

12 For ease of illustration, we have assumed that the investment-grade (IG) and high-yield (HY) bond returns are comprised of 100% interest income. In practice, where bonds may be purchased at a discount or sold in the future for an expected capital gain (or loss), some portion of the expected return would be eligible for capital gains tax treatment. Equation 3 can be modified to accommodate the dual tax treatment of interest income and capital gains, although it would be necessary to make certain assumptions regarding the proportion of returns attributable to capital gains.

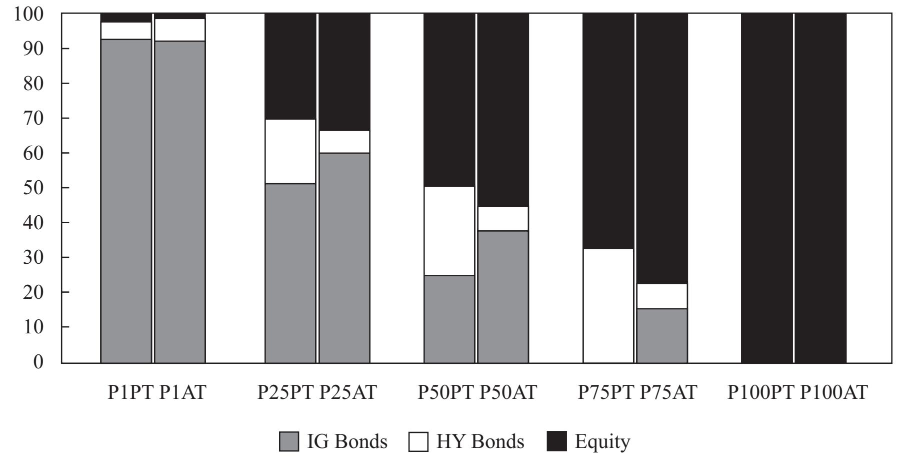
Portfolio optimization using these after-tax inputs produces the optimal asset allocations shown in Exhibit 6.

Exhibit 6 Optimal After-Tax Asset Mixes

	$P1_{at}$	$P25_{at}$	$P50_{at}$	$P75_{at}$	$P100_{at}$
IG bonds	92%	60%	38%	16%	0%
HY bonds	7%	7%	7%	7%	0%
Equity	1%	33%	55%	77%	100%

In Exhibit 7, we compare the pre-tax and after-tax efficient frontiers from these previous exhibits. Note that the portfolios at either extreme (P1 and P100) are essentially unchanged after taxes are factored into the assumptions. In portfolios P25, P50, and P75, however, you can see a significant reduction in the allocation to high-yield bonds. This is because of the heavier tax burden imposed on high-yield bonds. Although investment-grade bonds receive the same tax treatment, they are less risky than high-yield bonds and demonstrate a lower correlation with equity, so they continue to play the important role of portfolio risk reduction.

Exhibit 7 Pre-Tax and After-Tax Asset Allocation Comparisons



The optimal after-tax asset allocation depends on the interaction of after-tax returns, after-tax risk, and correlations. If an asset class or strategy is tax inefficient, it can still play a diversifying role in an optimal after-tax asset allocation if the asset or strategy offers sufficiently low correlations. After-tax portfolio optimization helps answer that question.

3.2 Taxes and Portfolio Rebalancing

Among tax-exempt institutional asset owners, periodic portfolio rebalancing—reallocating assets to return the portfolio to its target strategic asset allocation—is an integral part of sound portfolio management. This is no less true for taxable asset owners, but with the important distinction that more frequent rebalancing exposes the taxable asset owner to realized taxes that could have otherwise been deferred or even avoided. Whereas the tax burden incurred by liquidating assets to fund-required consumption cannot be avoided, rebalancing is discretionary; thus, the taxable asset owner should consider the trade-off between the benefits of tax minimization and the merits of maintaining the targeted asset allocation by rebalancing. The decision to rebalance and incur taxes is driven by each asset owner's unique circumstances.

Because after-tax volatility is less than pre-tax volatility (Equation 3) and asset class correlations remain the same, it takes larger asset class movements to materially alter the risk profile of the taxable portfolio. This suggests that rebalancing ranges for a taxable portfolio can be wider than those of a tax-exempt portfolio with a similar risk profile.

For example, consider a portfolio with a 50% allocation to equity, where equity returns are subject to a 25% tax rate. A tax-exempt investor may establish a target allocation to equities of 50%, with an acceptable range of 40% to 60% (50% plus or minus 10%). A taxable investor with the same target equity allocation can achieve a similar risk constraint with a range of 37% to 63% (50% plus or minus 13%). The equivalent rebalancing range for the taxable investor is derived by adjusting the permitted 10% deviation (up or down) by the tax rate, as shown in Equation 4:

$$R_{at}=R_{pt}/(1-t) \quad (4)$$

where

R_{at} = the after-tax rebalancing range
R_{pt} = the pre-tax rebalancing range

In our example, the 10% rebalancing range for a tax-exempt investor becomes a 13.3% rebalancing range for a taxable investor (when ranges are viewed and monitored from the same gross return perspective):

$$0.10/(1-0.25) = 13.3\%$$

Broader rebalancing ranges for the taxable investor reduce the frequency of trading and, consequently, the amount of taxable gains.

3.3 *Strategies to Reduce Tax Impact*

Additional strategies can be used to reduce taxes, including tax-loss harvesting and choices in the placement of certain types of assets in taxable or tax-exempt accounts (strategic asset location). Tax-loss harvesting is intentionally trading to realize a capital loss, which is then used to offset a current or future realized capital gain in another part of the portfolio, thereby reducing the taxes owned by the investor. It is discussed elsewhere in the curriculum, but we address strategic asset location strategies here.

Strategic asset location refers to placing (or locating) less tax-efficient assets in accounts with more favorable tax treatment, such as retirement savings accounts.

Aggregating assets across accounts with differing tax treatment requires modifying the asset value inputs to the portfolio optimization. Assets held in tax-*exempt* accounts require no tax adjustment to their market values. Assets in tax-*deferred* accounts grow tax free but are taxed upon distribution. Because these assets cannot be distributed (and consumed) without incurring the tax, the tax burden is inseparable from the economic value of the assets. Thus, the after-tax value of assets in a tax-deferred account is defined by Equation 5:

$$v_{at} = v_{pt}(1-t_i) \quad (5)$$

where

v_{at} = the after-tax value of assets
v_{pt} = the pre-tax market value of assets
t_i = the expected income tax rate upon distribution

In our earlier example, we had three asset classes: investment-grade bonds, high-yield bonds, and equities. If we assume that each of these three asset classes can be held in either of two account types—taxable or tax-deferred—then our optimization uses six different after-tax asset classes (three asset classes times two account types). The three asset classes in taxable accounts use the after-tax return and risk inputs derived earlier. The three asset classes in tax-deferred accounts (which grow tax free) use expected pre-tax return and risk inputs. The optimization adds constraints based on the after-tax value of the assets currently available in each account type and derives the optimal after-tax asset allocation and asset location simultaneously.

As a general rule, the portion of a taxable asset owner's assets that are eligible for lower tax rates and deferred capital gains tax treatment should first be allocated to the investor's taxable accounts. For example, equities should generally be held in taxable accounts, while taxable bonds and high-turnover trading strategies should generally be located in tax-exempt and tax-deferred accounts to the extent possible.

One important exception to this general rule regarding asset location applies to assets held for near-term liquidity needs. Because tax-exempt and tax-deferred accounts may not be immediately accessible without tax penalty, a portion of the bond allocation may be held in taxable accounts if its role is to fund near-term consumption requirements.

EXAMPLE 5

Asset Allocation and the Taxable Investor

1 Sarah Moreau, 45 years old, is a mid-level manager at a consumer products company. Her investment portfolio consists entirely of tax-deferred retirement savings accounts. Through careful savings and investments, she is on track to accumulate sufficient assets to retire at age 60. Her portfolio is currently allocated as indicated below:

Investment-grade bonds	20%
High-yield bonds	20%
Common stock–dividend income strategy	30%
Common stock–total return (capital gain) strategy	30%
Total portfolio	100%

The common stock–dividend income strategy focuses on income-oriented, high-dividend-paying stocks; the common stock–total return strategy focuses on stocks that represent good, long-term opportunities but pay little to no dividend. For the purposes of this example, we will assume that the expected long-term return is equivalent between the two strategies. Moreau has a high comfort level with this portfolio and the overall level of risk it entails.

Moreau has recently inherited additional monies, doubling her investable assets. She intends to use this new, taxable portfolio to support causes important to her personally over her lifetime. There is no change in her risk tolerance. She is interviewing prospective investment managers and has asked each to recommend an asset allocation strategy for the new portfolio using the same set of asset classes. She has received the following recommendations:

	Recommendation		
	A	B	C
Investment-grade bonds	20%	40%	30%
High-yield bonds	20%	0%	0%
Common stock–dividend income strategy	30%	30%	0%
Common stock–total return (capital gain) strategy	30%	30%	70%
Total portfolio	100%	100%	100%

Which asset allocation is *most* appropriate for the new portfolio? Justify your response.

2 How should Moreau distribute these investments among her taxable and tax-exempt accounts?

3 You are a member of the Investment Committee for a multinational corporation, responsible for the supervision of two portfolios. Both portfolios were established to fund retirement benefits: One is a tax-exempt defined benefit pension fund, and the other is taxable, holding assets intended to fund non-exempt retirement benefits. The pension fund has a target allocation of 70% equities and 30% fixed income, with a +/– 5% rebalancing range. There is no formal asset allocation policy for the taxable portfolio; it has simply followed the same allocation adopted by the pension portfolio. Because of recent strong equity market returns, both portfolios are now allocated 77% to equities and 23% to bonds. Management expects that the equity markets will continue to produce strong returns in the near term. Staff has offered the following options for rebalancing the portfolios:

A Do not rebalance.

B Rebalance both portfolios to the 70% equity/30% fixed-income target allocation.

C Rebalance the tax-exempt portfolio to the 70% equity/30% fixed-income target allocation, but expand the rebalancing range for the taxable portfolio.

Which recommendation is *most* appropriate? Justify your response.

Solution to 1:

Recommendation C would be the most appropriate asset allocation for the new portfolio. The high-yield bond and common stock–dividend income strategies are tax disadvantaged in a taxable portfolio. (Although investment-grade bonds are also tax disadvantaged, they maintain the role of controlling portfolio risk to maintain Moreau's risk preference.) By shifting this equity-like risk to the total return common stock strategy, Moreau should achieve a greater after-tax return. Given the lower standard deviation characteristics of after-tax equity returns when held in the taxable portfolio, a higher allocation to common stocks may be justified without exceeding Moreau's desired risk level. Recommendations A and B do not consider the negative tax implications of holding the high-yield and/or common stock–dividend income strategies in a taxable portfolio. Recommendation B also fails to consider Moreau's overall risk tolerance: The volatility of the common stock–capital gain strategy is lower when held in a taxable portfolio, thus a higher allocation to this strategy can enhance returns while remaining within Moreau's overall risk tolerance.[13]

Solution to 2:

If Moreau is willing to think of her investable portfolio as a single portfolio, rather than as independent "retirement" and "important causes" portfolios, she should hold the allocation to high-yield bonds and dividend-paying stocks in her tax-exempt retirement portfolio. In addition, subject to the overall volatility of the individual tax-exempt and taxable portfolios, it would be sensible to bear any increased stock risk in the taxable portfolio. A new optimization for *all* of Moreau's assets—using pre-tax and after-tax risk and return assumptions and subject to the constraint that half of the assets are held in a taxable portfolio and half are held in the tax-exempt portfolio—would more precisely allocate investments across portfolio (account) types.

Asset Location for Optimal Tax Efficiency

	Tax Advantaged Retirement Account	Taxable Account
Investment-grade bonds	X	
High-yield bonds	X	
Common stock–dividend income strategy	X	
Common stock–total return (capital gain) strategy		X

Solution to 3:

Recommendation C is the most appropriate course of action. Rebalancing of the tax-exempt portfolio is unencumbered by tax considerations, and rebalancing maintains the desired level of risk. The rebalancing range for the taxable portfolio can be wider than that of the tax-exempt portfolio based on the desire to minimize avoidable taxes and the lower volatility of after-tax equity returns. Recommendation A (no rebalancing) does not address the increased level of risk in the tax-exempt portfolio that results from the increase in the stock allocation. Recommendation B would create an unnecessary tax liability for the company, given that the portfolio is still operating in a reasonable range of risk when adjusted for taxes.

13 Investment-grade bonds also have lower after-tax volatility. The equivalent risk portfolios in pre-tax and after-tax environments are a function of a complex interaction of after-tax returns, standard deviations, and correlations.

Increasing Allocations to Fixed Income in Corporate Pension Plans

Increasing allocations to fixed income by defined benefit pension funds worldwide have been driven largely by a desire to better hedge plan liabilities. In some countries, accounting standards discourage de-risking. De-risking, however, is not the only argument in favor of a higher fixed-income allocation.

De-risking

There has been much discussion globally of pension plans "de-risking"—moving toward larger fixed-income allocations to better hedge liabilities, thereby reducing contribution uncertainty. Some countries' accounting rules, however—most notably those in the United States—discourage companies from moving in that direction. Under US GAAP accounting rules, for example, a higher allocation to equities allows the plan sponsor to employ a higher return assumption, thereby reducing pension cost, a non-cash expense that directly affects reported income.

For underfunded pension plans, de-risking leads to higher pension contributions. If a company has a weak core business with a higher-than-average probability of going bankrupt and makes only the minimum required contribution, it might be argued that the asset allocation decision was contrary to the interests of plan participants. If the company were to go bankrupt, the participants would get only the benefits covered by any government guaranty program. Had the company taken equity risk in the plan, there would have been a possibility of closing the funding gap, resulting in higher benefit payments.

Efficient Allocation of Risk

A higher allocation to fixed income—and a lower allocation to equity—might also be driven by corporate governance considerations. Pension investment activities are not a core competency of many companies, especially non-financial companies. Assuming that the company has a limited appetite for risk, shareholders might prefer that management allocate its risk budget to the core business of the company where they are expected to have skill, rather than to the pension fund. The rewards per unit of risk should presumably be greater in the company's core business, and the improved profitability should offset the increase in pension contributions required as a result of the lower equity allocation.

A Holistic Approach to Asset Location

Finally, some have argued that an asset allocation of 100% fixed-income securities can be justified on the premise that the company is acting as an agent for the benefit of all stakeholders, including shareholders and plan participants. This argument centers on tax-efficient asset location. A taxable investor—the shareholder and plan participant—should prefer to take his long-term equity risk in that portion of his overall portfolio where he will receive the benefit of lower capital gains rates rather than in tax-deferred accounts, the proceeds of which will be taxed at income tax rates. Consider a small business owner with US$3 million in total assets. The assets are split between a pension fund of which he is the sole participant (US$1 million) and a taxable portfolio (US$2 million). Assume that the asset allocation that represents his preferred level of risk is 67% equities and 33% fixed income. Where should this individual hold his equity exposure? As discussed, the more favorable tax treatment of equity returns argues for holding the equity exposure in his taxable account, while the investments subject to the higher tax rate should be held in the tax-deferred account—the pension plan. Theoretically, this tax efficiency argument can be extended to pension funds operated by publicly traded companies.[14]

---

14 For those interested in a more detailed discussion of this concept, see "The Case against Stock in Public Pension Funds" (Bader and Gold 2007) or the UBS Q-Series article, "Pension Fund Asset Allocation" (Cooper and Bianco 2003).

4 REVISING THE STRATEGIC ASSET ALLOCATION

An asset owner's strategic asset allocation is not a static decision. Circumstances often arise that justify revisiting the original decision, either to confirm its appropriateness or to consider a change to the current allocation strategy. It is sound financial practice to periodically re-examine the asset allocation strategy even in the absence of one of the external factors discussed next. Many institutional asset owners typically revisit the asset allocation policy at least once every five years through a formal asset allocation study, and all asset owners should affirm annually that the asset allocation remains appropriate given their needs and circumstances.

The circumstances that might trigger a special review of the asset allocation policy can generally be classified as relating to a change in *goals*, a change in *constraints*, or a change in *beliefs*. Among the reasons to review the strategic asset allocation are the following:

Goals

- Changes in business conditions affecting the organization supporting the fund and, therefore, expected changes in the cash flows
- A change in the investor's personal circumstances that may alter her risk appetite or risk capacity

Over an individual's lifespan, or throughout the course of an institutional fund's lifespan, it is unlikely that the investment goals and objectives will remain unchanged. An individual may get married, have children, or become disabled, for example, each of which may have implications for the asset allocation strategy.

Significant changes in the core business of an organization supporting or benefiting from the trust might prompt a re-examination of the asset allocation strategy. For example, an automobile manufacturer that has historically generated a significant portion of its revenues from its consumer finance activities may find that technology is disrupting this source of revenue as more online tools become available to car buyers. With greater uncertainty in its revenue stream, company management may move to reduce risk-taking in the pension fund in order to achieve a goal of reducing the variability in year-to-year contributions.

A university may embark on a long-term capital improvement plan that is reliant on the endowment fund for financial support. Or the university may be experiencing declining enrollments and must lean more heavily on the endowment fund to support its ongoing operational expenditures. The source of funds to a sovereign wealth fund may shrink considerably or even evaporate. When any of these, or similar, events occur or are anticipated, the existing asset allocation policy should be re-evaluated.

Constraints A material change in any one of the constraints mentioned earlier—time horizon, liquidity needs, asset size, or regulatory or other external constraints—is also reason to re-examine the existing asset allocation policy. Some of these changes might include the following:

- Changes in the expected payments from the fund
- A significant cash inflow or unanticipated expenditure
- Changes in regulations governing donations or contributions to the fund
- Changes in time horizon resulting from the adoption of a lump sum distribution option at retirement
- Changes in asset size as a result of the merging of pension plans

Changes in the expected payments from the fund can materially affect the asset allocation strategy. For example, a university reduces its spending policy from 5% to 4% of assets annually; an individual retires early, perhaps for health reasons or an involuntary late-career layoff; or a US corporate pension sponsor reduces or freezes pension benefits because it can no longer afford increasing Pension Benefit Guaranty Corporation[15] premiums. Faced with lower payouts, the university endowment may have greater latitude to invest in less liquid segments of the market. Decisions as to how and where to invest given this greater flexibility should be made within the framework of an asset allocation study to ensure the resulting allocation achieves the optimal trade-off of risk and return.

Similarly, a significant cash inflow has the potential to materially affect the asset allocation strategy. If a university endowment fund with £500 million in assets receives a gift of £100 million, the new monies *could* be invested in parallel with the existing assets, but that fails to consider the increased earning potential of the fund and any spending requirements associated with the donation. Pausing to formally reassess the fund's goals, objectives, constraints, and opportunities through an asset allocation study allows the asset owner to consider more broadly how best to maximize this additional wealth.

A change in regulations may also give rise to a change in asset allocation policy. Examples of regulatory changes that could trigger a re-examination of the asset allocation include the following:

- Regulatory changes in the United States in 2006 mandated a change in the liability discount rate, which resulted in larger pension contributions. With higher required contributions, there was less need to reach for higher investment returns. Many US corporate pension plans began de-risking (adopting an asset allocation strategy focused on hedging the liabilities) to reduce contribution volatility.
- UK tax incentives (30% of social impact investment costs can be deducted from income tax) and relaxed regulations for institutional investors were instituted to encourage socially responsible (impact) investing.

Again, an asset allocation study to objectively evaluate the effect of these changes on the investment opportunity set can help ensure that any new investment strategies adopted are consistent with the fund's overarching goals and objectives.

Beliefs Investment beliefs are a set of guiding principles that govern the asset owner's investment activities. Beliefs are not static, however, and changes in the economic environment and capital market expectations or a change in trustees or committee members are two factors that may lead to an altering of the principles that guide investment activities.[16]

An integral aspect of any asset allocation exercise is the forecasting of expected returns, volatilities, and correlations of the asset classes in the opportunity set. It follows, then, that a material change in the outlook for one or more of the asset classes may heavily influence the asset allocation outcome.

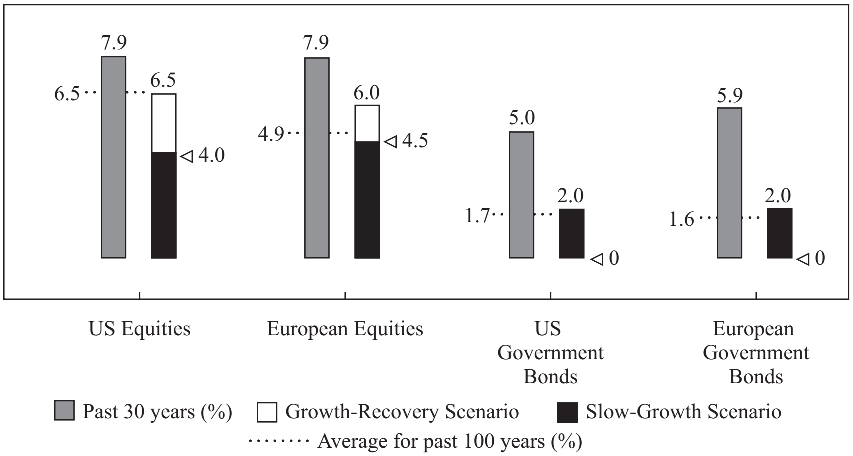
Consider the current environment relative to the environment that prevailed in 1984–2014. The 1984–2014 investing environment was characterized by declining inflation and interest rates (from the extraordinarily high levels of the 1970s and early 1980s); strong global GDP growth, aided by favorable demographics; gains in productivity; and rapid growth in China. Corporate profit growth was extremely robust, reflecting revenue growth from new markets, declining corporate taxes over

15 The Pension Benefit Guaranty Corporation insures certain US pension plan benefits.
16 For an example of an investment belief statement, see www.uss.co.uk/how-uss-invests/investment-approach/investment-beliefs-and-principles.

the period, and improved efficiencies. Despite increased market turbulence, returns on US and Western European equities and bonds during the past 30 years were considerably higher than the long-run trend.

The environment of 2015–2016 was much less favorable for investors. The dramatic decline in inflation and interest rates ended, and labor force expansion and productivity gains stalled, with negative implications for GDP growth. The largest developed-country companies that generated much of the profits of the past 30 years were faced with competitive pressures as emerging-market companies expanded and technology advances changed the competitive landscape. In April 2016, McKinsey Global Institute published a projection of stocks and bonds under two growth scenarios—a slow growth scenario and a moderate growth scenario (Exhibit 8). In neither instance do the expected returns of the next 30 years come close to the returns of the past 30 years.[17] Clearly, an asset allocation developed in 2010 built on return expectations based on the prior 26 years would look materially different than an asset allocation developed using more current, forward-looking return assumptions.

Exhibit 8 A Major Shift in Underlying Return Assumptions



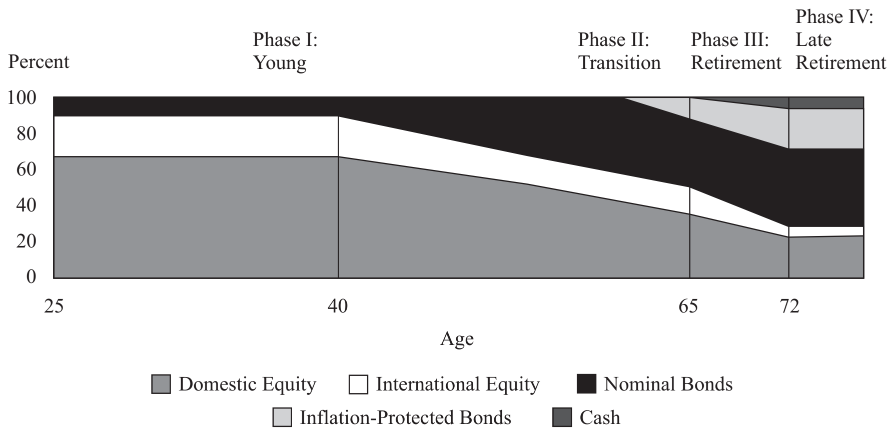
Notes:
Numbers for growth-recovery and slow-growth scenarios reflect the range between the low end of the slow-growth scenario and the high end of the growth-recovery scenario.
European equities: Weighted average real returns based on each year's Geary-Khamis purchasing power parity GDP for 14 countries in Western Europe.
US and European government bonds: Bond duration for United States is primarily 10 years; for Europe, duration varies by country but is typically 20 years.
Source: McKinsey Global Institute (www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/why-investors-may-need-to-lower-their-sights).

Finally, as new advisers or members join the Investment Committee, they bring their own beliefs and biases regarding certain investment activities. Conducting an asset allocation study to educate these new members of the oversight group and introduce them to the investment philosophy and process that has been adopted by the organization will smooth their integration into the governance system and ensure that they have a holistic view of the asset owner's goals and objectives.

17 McKinsey Global Institute, "Diminishing Returns: Why Investors May Need to Lower Their Expectations" (May 2016).

In some instances, a change to an asset allocation strategy may reasonably be implemented without a formal asset allocation study. Certain milestones are reasonable points at which to implement a change in the policy, in most instances, reducing the level of risk. (For pension funds, these "milestones" are typically related to changes in the plan's funded status.) Anticipating these milestones by putting an asset allocation policy in place that anticipates these changes allows the investor to respond more quickly to changing circumstances and in a non-reactive and objective manner. This rebalancing policy is frequently referred to as a "glide path." Target-date mutual funds common in retirement investing for individuals are one example of this approach to asset allocation. Exhibit 9 illustrates one fund company's approach to migrating the asset allocation away from equities and towards bonds as retirement approaches.

Exhibit 9 An Asset Allocation Glide Path



Source: Vanguard, "Target-Date Funds: A Solid Foundation for Retirement Investors" (May 2009): www.vanguard.com/jumppage/targetretirement/TRFCOMM.pdf.

In an institutional framework, the Investment Committee may specify certain funding levels it seeks to achieve. At the start of the period, an underfunded pension plan might adopt a higher equity allocation in an attempt to reduce the underfunding. If this is successful, the plan becomes better funded and there is less of a desire or need to take the higher level of equity risk. A pension fund may quickly implement "pre-programmed" asset allocation changes as the funded status of a pension plan improves. Typically, these planned reallocations are spelled out in an Investment Policy Statement.

EXAMPLE 6

Revising the Strategic Asset Allocation

1 Auldberg University Endowment Fund (AUE) has assets totaling CAF$200 million. The current asset allocation is as follows:

- CAF$100 million in domestic equities
- CAF$60 million in domestic government debt
- CAF$40 million in Class B office real estate

AUE has historically distributed to the University 5% of the 36-month moving average of net assets, contributing approximately CAF$10 million of Auldberg University's CAF$60 million annual operating budget. Real estate income (from the University's CAF$350 million direct investment in domestic commercial real estate assets, including office buildings and industrial parks, much of it near the campus) and provincial subsidies have been the main source of income to the University. Admission is free to all citizens who qualify academically.

Growth in the Caflandia economy has been fueled by low interest rates, encouraging excess real estate development. There is a strong probability that the economy will soon go into recession, negatively impacting both the property values and the income potential of the University's real estate holdings.

Gizi Horvath, a University alumnus, has recently announced an irrevocable CAF$200 million gift to AUE, to be paid in equal installments over the next five years. AUE employs a well-qualified staff with substantial diverse experience in equities, fixed income, and real estate. Staff has recommended that the gift from Ms. Horvath be invested using the same asset allocation policy that the endowment has been following successfully for the past five years. They suggest that the asset allocation policy should be revisited once the final installment has been received.

Critique staff's recommendation, and identify the case facts that support your critique.

2 The Government Petroleum Fund of Caflandia (GPFC) is operating under the following asset allocation policy, which was developed with a 20-year planning horizon. Target weights and actual weights are given:

	Target Asset Allocation	Current Asset Allocation
Global equities	30%	38%
Global high-yield bonds	10%	15%
Domestic intermediate bonds	30%	25%
Hedge funds	15%	15%
Private equity	15%	7%

When this asset allocation policy was adopted 5 years ago, the petroleum revenues that support the sovereign wealth fund were projected to continue to grow for at least the next 25 years and intergenerational distributions were expected to begin in 20 years. However, since the adoption of this policy, alternate fuel sources have eroded both the price and quantity of oil exports, the economy is undergoing significant restructuring, inflows to the fund have been suspended, and distributions are expected to begin within 5 years.

What are the implications of this change in the liquidity constraints for the current asset allocation policy?

3 O-Chem Corp has a defined benefit pension plan with US$1.0 billion in assets. The plan is closed, the liabilities are frozen, and the plan is currently 65% funded. The company intends to increase cash contributions to improve the funded status of the plan and then purchase annuities to fully address all of the plan's pension obligations. As part of an asset allocation analysis conducted every five years, the company has recently decided to allocate 80% of assets to liability-matching bonds and the remaining 20%

to a mix of global equities and real estate. An existing private equity portfolio is in the midst of being liquidated. This allocation reflects a desired reduction in the level of investment risk.

O-Chem has just announced an ambitious US$15 billion capital investment program to build new plants for refining and production. The CFO informed the Pension Committee that the company will be contributing to the plan only the minimum funding required by regulations for the foreseeable future. It is estimated that achieving fully funded status for the pension plan under minimum funding requirements and using the current asset allocation approach will take at least 10 years.

What are the implications of this change in funding policy for the pension plan's asset allocation strategy?

Solution to 1:

The size of the anticipated contributions will double AUE's assets over the next five years, potentially increasing the opportunity set of asset classes suitable for their investment program. Given that a typical asset allocation study encompasses a long investment horizon—10 years, 20 years, or more—staff should begin to evaluate the opportunities available to them today in *anticipation* of the future cash flows. Given the material change in the economic balance sheet along with changes in the asset size, liquidity, and time horizon constraints, AUE should plan on a regular, more frequent, formal review of the asset allocation policy until the situation stabilizes. The asset allocation study should explore the feasibility of adding new asset classes as well as the ability to improve diversification within existing categories, perhaps by including non-domestic equities and bonds. Furthermore, the forecast economic environment may materially alter the outflows from the fund in support of the University's day-to-day operations. Cash flows from the University's real estate holdings are likely to decline, as are the values of those real estate assets. Given the outlook for real estate, a strong case can be made to limit or reduce the endowment's investment in real estate; moreover, consideration should be given to the effect of declining income from the current real estate investment.

Solution to 2:

GPFC had adopted a long-range asset allocation policy under the expectation of continuing net cash inflows and no immediate liquidity constraints. With the change in circumstances, the need for liquidity in the fund has increased significantly. The current asset allocation policy allocates 40% of the fund's assets to less liquid asset classes—high-yield bonds, hedge funds, and private equity. Although the allocation to private equity has not been fully implemented, the fund is overweight high-yield bonds and at the target weight for hedge funds. These asset classes—or the size of the allocation to these asset classes—may no longer be appropriate for the fund given the change in circumstances.

Solution to 3:

The Investment Committee should conduct a new asset allocation study to address the changes in cash flow forecasts. The lower contributions imply that the pension plan will need to rely more heavily on investment returns to reach its funding objectives. A higher allocation to return-seeking assets, such as public and private equities, is warranted. The company should suspend the current private equity liquidation plan until the new asset allocation study has been completed. A liability-matching bond portfolio is still appropriate, although less than the current 80% of assets should be allocated to this portfolio.

5 SHORT-TERM SHIFTS IN ASSET ALLOCATION

Strategic asset allocation (SAA), or policy asset allocation, represents long-term investment policy targets for asset class weights, whereas tactical asset allocation (TAA) allows short-term deviations from SAA targets.[18] TAA moves might be justified based on cyclical variations within a secular trend (e.g., stage of business or monetary cycle) or temporary price dislocations in capital markets. TAA has the objective of increasing return, or risk-adjusted return, by taking advantage of short-term economic and financial market conditions that appear more favorable to certain asset classes. In seeking to capture a short-term return opportunity, TAA decisions move the investor's risk away from the targeted risk profile. TAA is predicated on a belief that investment returns, in the short run, are predictable. (This contrasts with the random walk assumption more strongly embedded in most SAA processes.) Using either short-term views or signals, the investor actively re-weights broad asset classes, sectors, or risk factor premiums. TAA is not concerned with individual security selection. In other words, generating alpha through TAA decisions is dependent on successful market or factor timing rather than security selection. TAA is an asset-only approach. Although tactical asset allocation shifts must still conform to the risk constraints outlined in the investment policy statement, they do not expressly consider liabilities (or goals in goals-based investing).

The SAA policy portfolio is the benchmark against which TAA decisions are measured. Tactical views are developed and bets are sized relative to the asset class targets of the SAA policy portfolio. The sizes of these bets are typically subject to certain risk constraints. The most common risk constraint is a pre-established allowable range around each asset class's policy target. Other risk constraints may include either a predicted tracking error budget versus the SAA or a range of targeted risk (e.g., an allowable range of predicted volatility).

The success of TAA decisions can be evaluated in a number of ways. Three of the most common are

- a comparison of the Sharpe ratio realized under the TAA relative to the Sharpe ratio that would have been realized under the SAA;
- evaluating the information ratio or the *t*-statistic of the average excess return of the TAA portfolio relative to the SAA portfolio; and
- plotting the realized return and risk of the TAA portfolio versus the realized return and risk of portfolios along the SAA's efficient frontier. This approach is particularly useful in assessing the risk-adjusted TAA return. The TAA portfolio may have produced a higher return or a higher Sharpe ratio than the SAA portfolio, but it could be less optimal than other portfolios along the investor's efficient frontier of portfolio choices.

The composition of the portfolio's excess return over the SAA portfolio return can also be examined more closely using attribution analysis, evaluating the specific overweights and underweights that led to the performance differential.

Tactical investment decisions may incur additional costs—higher trading costs and taxes (in the case of taxable investors). Tactical investment decisions can also increase the concentration of risk relative to the policy portfolio. For example, if the tactical decision is to overweight equities, not only is the portfolio risk increased but

---

18 SAA and TAA are distinct from GTAA (global tactical asset allocation), an opportunistic investment strategy that seeks to take advantage of pricing or valuation anomalies across multiple asset classes, typically equities, fixed income, and currencies.

also the diversification of risk contributions is reduced. This is particularly an issue when the SAA policy portfolio relies on uncorrelated asset classes. These costs should be weighed against the predictability of short-term returns.

There are two broad approaches to TAA. The first is discretionary, which relies on a qualitative interpretation of political, economic, and financial market conditions. The second is systematic, which relies on quantitative signals to capture documented return anomalies that may be inconsistent with market efficiency.

5.1 Discretionary TAA

Discretionary TAA is predicated on the existence of manager skill in predicting and timing short-term market moves away from the expected outcome for each asset class that is embedded in the SAA policy portfolio. In practice, discretionary TAA is typically used in an attempt to mitigate or hedge risk in distressed markets while enhancing return in positive return markets (i.e., an asymmetric return distribution).

Short-term forecasts consider a large number of data points that provide relevant information about current and expected political, economic, and financial market conditions that may affect short-term asset class returns. Data points might include valuations, term and credit spreads, central bank policy, GDP growth, earnings expectations, inflation expectations, and leading economic indicators. Price-to-earnings ratios, price-to-book ratios, and the dividend yield are commonly used valuation measures that can be compared to historical averages and across similar assets to inform short-to-intermediate-term tactical shifts. Term spreads provide information about the business cycle, inflation, and potential future interest rates. Credit spreads gauge default risk, borrowing conditions, and liquidity. Other data points are more directly related to current and expected GDP and earnings growth.

Short-term forecasts may also consider economic sentiment indicators. TAA often assumes a close relationship between the economy and capital market returns. Because consumer spending is a major driver of GDP in developed countries, consumer sentiment is a key consideration. Consumer confidence surveys provide insight as to the level of optimism regarding the economy and personal finances.

TAA also considers market sentiment—indicators of the optimism or pessimism of financial market participants. Data points considered in gauging market sentiment include margin borrowing, short interest, and a volatility index.

- Margin borrowing measures give an indication of the current level of bullishness, and the capacity for more or less margin borrowing has implications for future bullishness. Higher prices tend to inspire confidence and spur more buying; similarly, more buying on margin tends to spur higher prices. The aggregate level of margin can be an indicator that bullish sentiment is overdone, although the level of borrowing must be considered in the context of the rate of change in borrowing.
- Short interest measures give an indication of current bearish sentiment and also have implications for future bearishness. Although rising short interest indicates increasing negative sentiment, a high short interest ratio may be an indication of the extreme pessimism that often occurs at market lows.
- The volatility index, commonly known as the fear index, is a measure of market expectations of near-term volatility. VDAX-NEW in Germany, V2X in the United Kingdom, and VIX in the United States each measure the level of expected volatility of their respective indexes as implied by the bid/ask quotations of index options; it rises when put option buying increases and falls when call buying activity increases.

Different approaches to discretionary TAA may include different data points and relationships and also may prioritize and weight those data points differently depending on both the approach and the prevailing market environment. Despite the plethora of data inputs, the interpretation of this information is qualitative at its core.

5.2 Systematic TAA

Using signals, systematic TAA attempts to capture asset class level return anomalies that have been shown to have some predictability and persistence. Value and momentum, for example, are factors that have been determined to offer some level of predictability, both among securities within asset classes (for security selection) and at the asset class level (for asset class timing).

The value factor is the return of value stocks over the return of growth stocks. The momentum factor is the return of stocks with higher prior returns over the return of stocks with lower prior returns. Value and momentum (and size) factors have been determined to have some explanatory power regarding the relative returns of equity securities within the equity asset class. Value and momentum phenomena are also present at the asset class level and can be used in making tactical asset allocation decisions across asset classes.

Valuation ratios have been shown to have some explanatory power in predicting variation in future equity returns. Predictive measures for equities include dividend yield, cash flow yield, and Shiller's earnings yield (the inverse of Shiller's P/E[19]). Sometimes these yield measures are defined as the excess of the yield over the local risk-free rate or inflation.[20]

Other asset classes have their own value signals, such as yield and carry in currencies, commodities, and/or fixed income. Carry in currencies uses short-term interest rate differentials to determine which currencies (or currency-denominated assets) to overweight (or own) and which to underweight (or sell short). Carry in commodities compares positive (backwardation) and negative (contango) roll yields to determine which commodities to own or short. And for bonds, yields-to-maturity and term premiums (yields in excess of the local risk-free rate) signal the relative attractiveness of different fixed-income markets.

Asset classes can trend positively or negatively for some time before changing course. Trend following is an investment or trading strategy based on the expectation that asset class (or asset) returns will continue in the same upward or downward trend that they have most recently exhibited.[21] A basic trend signal is the most recent 12-month return: The expectation is that the direction of the most recent 12-month returns can be expected to persist for the next 12 months. Shorter time frames and different weighting schemes can also be used. For example, another trend signal is the moving-average crossover, where the moving average price of a shorter time frame is compared with the moving average price of a longer time frame. This signals an upward (downward) trend when the moving average of the shorter time frame is above (below) the moving average of the longer time frame. Trend signals are widely used in systematic TAA. Asset classes may be ranked or categorized into positive or negative buckets based on their most recent prior 12-month performance and over- or

19 A price-to-earnings ratio based on the average inflation-adjusted earnings of the previous 10 years.

20 Return predictability for equity markets is driven by historical mean-reversion, which tends to occur over the intermediate-term. These valuation measures are often used as signals for TAA, but they can also be used to shape return expectations for SAA.

21 Trend following is also called time-series momentum. Cross-sectional momentum describes the relative momentum returns of securities within the same asset class.

underweighted accordingly. More-complex signals for both momentum/trend signals (such as those that use different lookback periods or momentum signals correlated with earnings momentum) and value/carry are also used.

EXAMPLE 7

Short-Term Shifts in Asset Allocation

1 The investment policy for Alpha Beverage Corporation's pension fund allows staff to overweight or underweight asset classes, within pre-established bands, using a TAA model that has been approved by the Investment Committee. The asset allocation policy is reflected in Exhibit 10, and the output of the TAA model is given in Exhibit 11. Using the data presented in Exhibits 10 and 11, recommend a TAA strategy for the pension fund and justify your response.

Exhibit 10 Strategic Asset Allocation Policy

SAA Policy	Current Weight	Target Allocation	Upper Policy Limit	Lower Policy Limit
Investment-grade bonds	45%	40%	45%	35%
High-yield bonds	10%	10%	15%	5%
Developed markets equity	35%	40%	45%	35%
Emerging markets equity	10%	10%	15%	5%

Exhibit 11 Trend Signal (the positive or negative trailing 12-month excess return)

	12-Month Return	Risk-Free Return	Excess Return	Signal
Investment-grade bonds	4%	1%	3%	Long
High-yield bonds	–2%	1%	–3%	Short
Developed markets equity	5%	1%	4%	Long
Emerging markets equity	–10%	1%	–11%	Short

2 One year later, the Investment Committee for Alpha Beverage Corporation is conducting its year-end review of pension plan performance. Staff has prepared the following exhibits regarding the tactical asset allocation decisions taken during the past year. Assume that all investments are implemented using passively managed index funds. Evaluate the effectiveness of the TAA decisions.

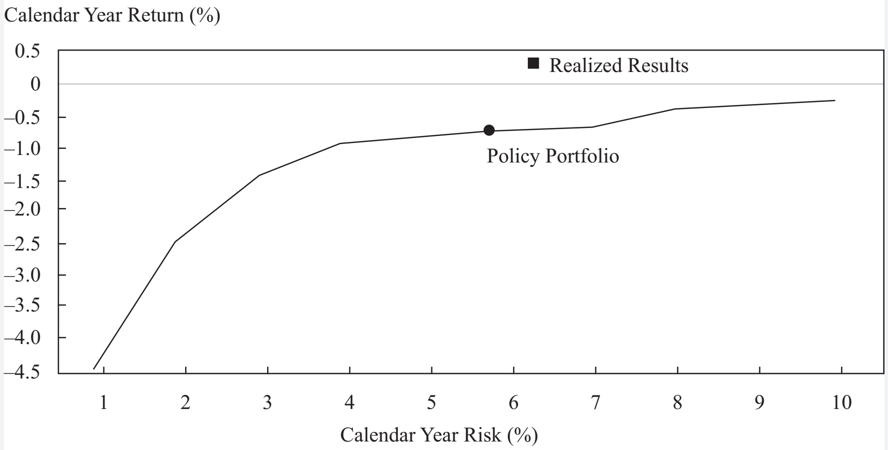
Exhibit 12A

Asset Class	Asset Allocation	Calendar Year Return
Investment-grade bonds	45%	3.45%
High-yield bonds	5%	–6.07%
Developed markets equity	45%	–0.32%
Emerging markets equity	5%	–14.60%

Exhibit 12B

	Policy Portfolio	Realized Results
12-month return	–0.82%	0.38%
Risk-free rate	0.50%	0.50%
Standard deviation	5.80%	6.20%
Sharpe ratio	–0.23	–0.02

Exhibit 12C



Solution to 1:

The TAA decision must be taken in the context of the SAA policy constraints. Thus, although the signals for high-yield bonds and emerging market equities are negative, the minimum permissible weight in each is 5%. Similarly, although the signals for investment-grade bonds and developed markets equities are positive, the maximum permissible weight in each is 45%. Asset classes can be over- or underweighted to the full extent of the policy limits. Based on the trend signals and the policy constraints, the recommended tactical asset allocation is as follows:

- Investment-grade bonds 45% *(overweight by 5%)*
- High-yield bonds 5% *(underweight by 5%)*
- Developed markets equity 45% *(overweight by 5%)*
- Emerging markets equity 5% *(underweight by 5%)*

Solution to 2:

The decision to overweight investment grade bonds and underweight emerging markets equity and high-yield bonds was a profitable one. The chosen asset allocation added approximately 120 basis points to portfolio return over the year. Although portfolio risk was elevated relative to the policy portfolio (standard deviation of 6.2% versus 5.8% for the policy portfolio), the portfolio positioning improved the fund's Sharpe ratio relative to allocations they might have selected along the efficient frontier.

A Silver Lining to the 2008–2009 Financial Crisis

Prior to 2008, corporate pension plans had begun to shift the fixed-income component of their policy portfolios from an intermediate maturity bond index to a long bond index. Despite the relatively low interest rates at the time, this move was made to better align the plans' assets with the long duration liability payment stream. The fixed-income portfolios were typically benchmarked against a long government and credit index that included both government and corporate bonds. Swaps or STRIPS* were sometimes used to extend duration.

During the financial crisis that began in 2008, these heavier and longer-duration fixed-income positions performed well relative to equities (the long government and credit index was up 8%, whereas the S&P 500 Index was down 37% in 2008), providing plan sponsors with a level of investment protection that had not been anticipated. Additionally, with its exposure to higher-returning government bonds that benefited from investors' flight to safety, this fixed-income portfolio often outperformed the liabilities. (Recall from the earlier discussion on pension regulation that pension liabilities are typically measured using corporate bond yields. Thus, liabilities rose in the face of declining corporate bond yields while the liability-hedging asset rose even further given its overall higher credit quality.) This was an unintended asset/liability mismatch that had very positive results. Subsequent to this rally in bonds, some plan sponsors made a tactical asset allocation decision—to move out of swaps and government bonds and into physical corporate bonds (non-derivative fixed-income exposure)—locking in the gains and better hedging the liability.

* Treasury STRIPS are fixed-income securities with no interest payments that are sold at a discount to face value and mature at par. STRIPS is an acronym for Separate Trading of Registered Interest and Principal of Securities.

6 DEALING WITH BEHAVIORAL BIASES IN ASSET ALLOCATION

Although global capital markets are competitive pricing engines, human behavior can be less rational than most economic models assume. Behavioral finance—the hybrid study of financial economics and psychology—has documented a number of behavioral biases that commonly arise in investing. The CFA Program reading "The Behavioral Biases of Individuals" discusses 16 common behavioral biases. The biases most relevant in asset allocation include loss aversion, the illusion of control, mental accounting, representative bias, framing, and availability bias. An effective investment program will address these decision-making risks through a formal asset allocation process with its own objective framework, governance, and controls. An important first step toward mitigating the negative effects of behavioral biases is simply acknowledging that they exist; just being aware of them can reduce their influence on decision making. It is also possible to incorporate certain behavioral biases into the investment decision-making process to produce better outcomes. This is most commonly practiced in goals-based investing. We will discuss strategies that help deal with these common biases.

6.1 Loss Aversion

Loss-aversion bias is an emotional bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains. A number of studies on loss aversion suggest that, psychologically, losses are significantly more powerful than gains. The utility derived from a gain is much lower than the utility given up with an equivalent loss. This behavior is related to the marginal utility of wealth, where each additional dollar of wealth is valued incrementally less with increasing levels of wealth.

A diversified multi-asset class portfolio is generally thought to offer an approximately symmetrical distribution of returns around a positive expected mean return. Financial market theory suggests that a rational investor would think about risk as the dispersion or uncertainty (variance) around the mean (expected) outcome. However, loss aversion suggests the investor assigns a greater weight to the negative outcomes than would be implied by the actual shape of the distribution. Looking at this another way, risk is not measured relative to the expected mean return but rather on an absolute basis, relative to a 0% return. The loss-aversion bias may interfere with an investor's ability to maintain his chosen asset allocation through periods of negative returns.

In goals-based investing, loss-aversion bias can be mitigated by framing risk in terms of shortfall probability or by funding high-priority goals with low-risk assets.

Shortfall probability is the probability that a portfolio will not achieve the return required to meet a stated goal. Where there are well-defined, discrete goals, sub-portfolios can be established for each goal and the asset allocation for that sub-portfolio would use shortfall probability as the definition of risk.

Similarly, by segregating assets into sub-portfolios aligned to goals designated by the client as high-priority and investing those assets in risk-free or low risk assets of similar duration, the adviser mitigates the loss-aversion bias associated with this particular goal—freeing up other assets to take on a more appropriate level of risk. Riskier assets can then be used to fund lower-priority and aspirational goals.

In institutional investing, loss aversion can be seen in the herding behavior among plan sponsors. Adopting an asset allocation not too different from the allocation of one's peers minimizes reputation risk.

6.2 Illusion of Control

The illusion of control is a cognitive bias—the tendency to overestimate one's ability to control events. It can be exacerbated by overconfidence, an emotional bias. If investors believe they have more or better information than what is reflected in the market, they have (excessive) confidence in their ability to generate better outcomes. They may perceive *information* in what are random price movements, which may lead to more frequent trading, greater concentration of portfolio positions, or a greater willingness to employ tactical shifts in their asset allocation. The following investor behaviors might be attributed to this illusion of control:

- Alpha-seeking behaviors, such as attempted market timing in the form of extreme tactical asset allocation shifts or all in/all out market calls—the investor who correctly anticipated a market reversal now believes he has superior insight on valuation levels.
- Alpha-seeking behaviors based on a belief of superior resources—the institutional investor who believes her internal resources give her an edge over other investors in active security selection and/or the selection of active investment managers.
- Excessive trading, use of leverage, or short selling—the long/short equity investor who moves from a normal exposure range of 65% long/20% short to 100% long/50% short.
- Reducing, eliminating, or even shorting asset classes that are a significant part of the global market portfolio based on non-consensus return and risk forecasts—the chair of a foundation's investment committee who calls for shortening the duration of the bond portfolio from six years to six months based on insights drawn from his position in the banking industry.
- Retaining a large, concentrated legacy asset that contributes diversifiable risk—the employee who fails to diversify her holding of company stock.

Hindsight bias—the tendency to perceive past investment outcomes as having been predictable—exacerbates the illusion of control.

In the asset allocation process, an investor who believes he or she has better information than others may use estimates of return and risk that produce asset allocation choices that are materially different from the market portfolio. This can result in undiversified portfolios with outsized exposures to just one or two minor asset classes, called extreme corner portfolios. Using such biased risk and return estimates results in a biased asset allocation decision—precisely what an objective asset allocation process seeks to avoid.

The illusion of control can be mitigated by using the global market portfolio as the starting point in developing the asset allocation. Building on the basic principles of CAPM, Markowitz's mean–variance theory, and efficient market theory, the global market portfolio offers a theoretically sound benchmark for asset allocation. Deviations from this baseline portfolio must be thoughtfully considered and rigorously vetted, ensuring the asset allocation process remains objective. A formal asset allocation process that employs long-term return and risk forecasts, optimization constraints anchored around asset class weights in the global market portfolio, and strict policy ranges will significantly mitigate the illusion of control bias in asset allocation.

6.3 Mental Accounting

Mental accounting is an information-processing bias in which people treat one sum of money differently from another sum based solely on the mental account the money is assigned to. Investors may separate assets or liabilities into buckets based on subjective criteria. For example, an investor may consider his retirement investment portfolio independent of the portfolio that funds his child's education, even if the combined asset allocation of the two portfolios is sub-optimal. Or an employee with significant exposure to her employer's stock through vested stock options may fail to consider this exposure alongside other assets when establishing a strategic asset allocation.

Goals-based investing incorporates mental accounting directly into the asset allocation solution. Each goal is aligned with a discrete sub-portfolio, and the investor can specify the acceptable level of risk for each goal. Provided each of the sub-portfolios lies along the same efficient frontier, the sum of the sub-portfolios will also be efficient.[22]

Concentrated stock positions also give rise to another common mental accounting issue that affects asset allocation. For example, the primary source of an entrepreneur's wealth may be a concentrated equity position in the publicly traded company he founded. The entrepreneur may prefer to retain a relatively large exposure to this one security within his broader investment portfolio despite the inherent risk. Although there may be rational reasons for this preference—including ownership control, an information advantage, and tax considerations—the desire to retain this riskier exposure is more often the result of a psychological loyalty to the asset that generated his wealth. This mental accounting bias is further reinforced by the endowment effect—the tendency to ascribe more value to an asset already owned rather than another asset one might purchase to replace it.

The concentrated stock/mental accounting bias can be accommodated in goals-based asset allocation by assigning the concentrated stock position to an aspirational goal—one that the client would *like* to achieve but to which he or she is willing to assign a lower probability of success. Whereas lifetime consumption tends to be a high-priority goal requiring a well-diversified portfolio to fund it with confidence, an aspirational goal such as a charitable gift may be an important but much less highly valued goal. It can reasonably be funded with the concentrated stock position. (This could have the additional benefit of avoiding capital gains tax altogether!)

6.4 Representative Bias

Representative, or recency, bias is the tendency to overweight the importance of the most recent observations and information relative to a longer-dated or more comprehensive set of long-term observations and information. Tactical shifts in asset allocation, those undertaken in response to recent returns or news—perhaps shifting the asset allocation toward the highest or lowest allowable ends of the policy ranges—are particularly susceptible to recency bias. Return chasing is a common manifestation of recency bias, and it results in overweighting asset classes with good recent performance.

It is believed that asset prices largely follow a random walk; past prices cannot be used to predict future returns. If this is true, then shifting the asset allocation in response to recent returns, or allowing recent returns to unduly influence the asset class assumptions used in the asset allocation process, will likely lead to sub-optimal results. *If,* however, asset class returns exhibit trending behavior, the recent past *may* contain information relevant to tactical shifts in asset allocation. And if asset class returns are mean-reverting, comparing current valuations to historical norms may signal the potential for a reversal or for above-average future returns.

Recency bias is not uniformly negative. Random walk, trending, and mean-reversion may be simultaneously relevant to the investment decision-making process, although their effect on asset prices will unfold over different time horizons. The strongest defenses against recency bias are an objective asset allocation process and a strong governance framework. It is important that the investor objectively evaluate the motivation underlying the response to recent market events. A formal asset allocation policy with pre-specified allowable ranges will constrain recency bias. A strong

22 This condition holds when the asset allocation process is unconstrained. With a long-only constraint, some efficiency is lost but the effect is much less significant than the loss of efficiency from inaccurately specifying risk aversion (which goals-based approaches to asset allocation attempt to mitigate). See Das, Markowitz, Scheid, and Statman (2010) and Das et al. (2011).

governance framework with the appropriate level of expertise and well-documented investment beliefs increases the likelihood that shifts in asset allocation are made objectively and in accordance with those beliefs.

6.5 Framing Bias

Framing bias is an information-processing bias in which a person may answer a question differently based solely on the way in which it is asked. One example of framing bias is common in committee-oriented decision-making processes. In instances where one individual frequently speaks first and speaks with great authority, the views of other committee members may be suppressed or biased toward this first position put on the table.

A more nuanced form of framing bias can be found in asset allocation. The investor's choice of an asset allocation may be influenced merely by the manner in which the risk-to-return trade-off is presented.

Risk can mean different things to different investors: volatility, tail risk, the permanent loss of capital, or a failure to meet financial goals. These definitions are all closely related, but the relative importance of each of these aspects can influence the investor's asset allocation choice. Further, the investor's perception of each of these risks can be influenced by the manner in which they are presented—gain and loss potential framed in money terms versus percentages, for example.

Investors are often asked to evaluate portfolio choices using expected return, with standard deviation as the sole measure of risk. Standard deviation measures the dispersion or volatility around the mean (expected) return. Other measures of risk may also be used. Value at risk (VaR) is a loss threshold: "If I choose this asset mix, I can be pretty sure that my losses will not exceed X, most of the time." More formally, VaR is the minimum loss that would be expected a certain percentage of the time over a certain period of time given the assumed market conditions. Conditional value at risk (CVaR) is the probability-weighted average of losses when the VaR threshold is breached. VaR and CVaR both measure downside or tail risk.

Exhibit 13 shows the expected return and risk for five portfolios that span an efficient frontier from P1 (lowest risk) to P100 (highest risk). A normal distribution of returns is assumed; therefore, the portfolio's VaR and CVaR are a direct function of the portfolio's expected return and standard deviation. In this case, standard deviation, VaR, and CVaR measure precisely the same risk but frame that risk differently. Standard deviation presents that risk as volatility, while VaR and CVaR present it as risk of loss. When dealing with a normal distribution, as this example presumes, the 5% VaR threshold is simply the point on the distribution 1.65 standard deviations below the expected mean return.

Exhibit 13 There's More Than One Way to Frame Risk

	P1	P25	P50	P75	P100
Return	3.2%	4.9%	6.0%	7.0%	8.0%
Std. Dev.	3.9%	7.8%	11.9%	15.9%	20.0%
VaR (5%)	–3.2%	–8.0%	–13.6%	–19.3%	–25.0%
CVaR (5%)	–4.8%	–11.2%	–18.5%	–25.8%	–33.2%

When viewing return and volatility alone, many investors may gravitate to P50 with its 6.0% expected return and 11.9% standard deviation. P50 represents the median risk portfolio that appeals to many investors in practice because it balances high-risk and low-risk choices with related diversification benefits. However, loss-aversion bias suggests that some investors who gravitate to the median choice might actually find the −18.5% CVaR of P50 indicative of a level of risk they find very uncomfortable. The CVaR frame intuitively communicates a different perspective of exactly the same risk that is already fully explained by standard deviation—namely, the downside or tail-risk aspects of the standard deviation and mean. With this example, you can see that how risk is framed and presented can affect the asset allocation decision.

The framing effect can be mitigated by presenting the possible asset allocation choices with multiple perspectives on the risk/reward trade-off. The most commonly used risk measure—standard deviation—can be supplemented with additional measures, such as **shortfall probability** (the probability of failing to meet a specific liability or goal)[23] and tail-risk measures (e.g., VaR and CVaR). Historical stress tests and Monte Carlo simulations can also be used to capture and communicate risk in a tangible way. These multiple perspectives of the risk and reward trade-offs among a set of asset allocation choices compel the investor to consider more carefully what outcomes are acceptable or unacceptable.

6.6 Availability Bias

Availability bias is an information-processing bias in which people take a mental shortcut when estimating the probability of an outcome based on how easily the outcome comes to mind. Easily recalled outcomes are often perceived as being more likely than those that are harder to recall or understand. For example, more recent events or events in which the investor has personally been affected are likely to be assigned a higher probability of occurring again, regardless of the objective odds of the event actually occurring.

As an example, many private equity investors experienced a liquidity squeeze during the financial crisis that began in 2008. Their equity portfolios had suffered large losses, and their private equity investments were illiquid. Worse yet, they were contractually committed to additional capital contributions to those private equity funds. At the same time, their financial obligations continued at the same or an even higher pace. Investors who personally experienced this confluence of negative events are likely to express a strong preference for liquid investments, assigning a higher probability to such an event occurring again than would an investor who had cash available to acquire the private equity interests that were sold at distressed prices.

Familiarity bias stems from availability bias: People tend to favor the familiar over the new or different because of the ease of recalling the familiar. In asset allocation, familiarity bias most commonly results in a **home bias**—a preference for securities listed on the exchanges of one's home country. However, concentrating portfolio exposure in home country securities, particularly if the home country capital markets are small, results in a less diversified, less efficient portfolio. Familiarity bias can be mitigated by using the global market portfolio as the starting point in developing the asset allocation, where deviations from this baseline portfolio must be thoughtfully considered and rigorously vetted.

Familiarity bias may also cause investors to fall into the trap of comparing their investment decisions (and performance) to others', without regard for the appropriateness of those decisions for their own specific facts and circumstances. By avoiding

---

23 Shortfall risk and shortfall probability are often used to refer to the same concept. This author prefers shortfall probability because the measure refers to the probability of shortfall, not the magnitude of the potential shortfall. For example, you may have a low probability of shortfall but the size of the shortfall could be significant. In this case, it could be misleading to say the shortfall risk is low.

comparison of investment returns or asset allocation decisions with others, an organization is more capable of identifying the asset allocation that is best tailored to their needs.

Investment decision making is subject to a wide range of potential behavioral biases. This is true in both private wealth *and* institutional investing. Employing a formal asset allocation process using the global market portfolio as the starting point for asset allocation modeling is a key component of ensuring the asset allocation decision is as objective as possible.

A strong governance structure, such as that discussed in the overview reading on asset allocation, is a necessary first step to mitigating the effect that these behavioral biases may have on the long-term success of the investment program. Bringing a diverse set of views to the deliberation process brings more tools to the table to solve any problem and leads to better and more informed decision making. A clearly stated mission—a common goal—and a commitment from committee members and other stakeholders to that mission are critically important in constraining the influence of these biases on investment decisions.

Effective Investment Governance

Six critical elements of effective investment governance are

1 clearly articulated long- and short-term investment objectives of the investment program;
2 allocation of decision rights and responsibilities among the functional units in the governance hierarchy, taking account of their knowledge, capacity, time, and position in the governance hierarchy;
3 established processes for developing and approving the investment policy statement that will govern the day-to-day operation of the investment program;
4 specified processes for developing and approving the program's strategic asset allocation;
5 a reporting framework to monitor the program's progress toward the agreed-upon goals and objectives; and
6 periodic governance audits.

EXAMPLE 8

Mitigating Behavioral Biases in Asset Allocation

Ivy Lee, the retired founder of a publicly traded company, has two primary goals for her investment assets. The first goal is to fund lifetime consumption expenditures of US$1 million per year for herself and her husband; this is a goal the Lees want to achieve with a high degree of certainty. The second goal is to provide an end-of-life gift to Auldberg University. Ivy has a diversified portfolio of stocks and bonds totaling US$5 million and a sizable position in the stock of the company she founded. The following table summarizes the facts.

Investor Profile	
Annual consumption needs	US$1,000,000
Remaining years of life expectancy	40
Diversified stock holdings	US$3,000,000
Diversified bond holdings	US$2,000,000
Concentrated stock holdings	US$15,000,000
Total portfolio	US$20,000,000

Assume that a 60% equity/40% fixed-income portfolio represents the level of risk Ivy is willing to assume with respect to her consumption goal. This 60/40 portfolio offers an expected return of 6.0%. (For simplicity, this illustration ignores inflation and taxes.)

The present value of the expected consumption expenditures is US$15,949,075. This is the amount needed on hand today, which, if invested in a portfolio of 60% equities and 40% fixed income, would fully fund 40 annual cash distributions of US$1,000,000 each.[24]

The concentrated stock has a highly uncertain expected return and comes with significant idiosyncratic (stock-specific) risk. A preliminary mean–variance optimization using three "asset classes"—stocks, bonds, and the concentrated stock—results in a zero allocation to the concentrated stock position. But Ivy prefers to retain as much concentrated stock as possible because it represents her legacy and she has a strong psychological loyalty to it.

1 Describe the behavioral biases most relevant to developing an asset allocation recommendation for Ivy.
2 Recommend and justify an asset allocation for Ivy given the facts presented above.

Solution to 1:

Two behavioral biases that the adviser must be aware of in developing an asset allocation recommendation for Ivy are illusion of control and mental accounting. Because Ivy was the founder of the company whose stock comprises 75% of her investment portfolio, she may believe she has more or better information about the return prospects for this portion of the portfolio. The belief that she has superior information may lead to a risk assessment that is not reflective of the true risk in the holding. Using a goals-based approach to asset allocation may help Ivy more fully understand the risks inherent in the concentrated stock position. The riskier, concentrated stock position can be assigned to a lower-priority goal, such as the gift to Auldberg University.

Solution to 2:

	Beginning Asset Allocation	Recommended Asset Allocation
Diversified stocks	US$3,000,000	US$9,600,000
Diversified bonds	US$2,000,000	US$6,400,000
Funding of lifestyle goal		*US$16,000,000*

24 Assumes cash distributions occur at the beginning of the year and the expected return is the geometric average.

	Beginning Asset Allocation	Recommended Asset Allocation
Concentrated stock	US$15,000,000	US$4,000,000
Total portfolio	US$20,000,000	US$20,000,000

It is recommended that Ivy fully fund her high-priority lifestyle consumption needs (US$15,949,075) with US$16 million in a diversified portfolio of stocks and bonds. To achieve this, US$11 million of the concentrated stock position should be sold and the proceeds added to the diversified portfolio that supports lifestyle consumption needs. The remaining US$4 million of concentrated stock can be retained to fund the aspirational goal of an end-of-life gift to Auldberg University. In this example, the adviser has employed the mental accounting bias to achieve a suitable outcome: By illustrating the dollar value needed to fund the high-priority lifetime consumption needs goal, the adviser was able to clarify for Ivy the risks in retaining the concentrated stock position. The adviser might also simulate portfolio returns and the associated probability of achieving Ivy's goals using a range of scenarios for the performance of the concentrated stock position. Framing the effect this one holding may have on the likelihood of achieving her goals may help Ivy agree to reduce the position size. Consideration of certain behavioral biases like mental accounting can improve investor outcomes when they are incorporated in an objective decision-making framework.

SUMMARY

- The primary constraints on an asset allocation decision are asset size, liquidity, time horizon, and other external considerations, such as taxes and regulation.
- The size of an asset owner's portfolio may limit the asset classes accessible to the asset owner. An asset owner's portfolio may be too small—or too large—to capture the returns of certain asset classes or strategies efficiently.
- Complex asset classes and investment vehicles require sufficient governance capacity.
- Large-scale asset owners may achieve operating efficiencies, but they may find it difficult to deploy capital effectively in certain active investment strategies given liquidity conditions and trading costs.
- Smaller portfolios may also be constrained by size. They may be too small to adequately diversify across the range of asset classes and investment managers, or they may have staffing constraints that prevent them from monitoring a complex investment program.
- Investors with smaller portfolios may be constrained in their ability to access private equity, private real estate, hedge funds, and infrastructure investments because of the high required minimum investments and regulatory restrictions associated with those asset classes. Wealthy families may pool assets to meet the required minimums.
- The liquidity needs of the asset owner and the liquidity characteristics of the asset classes each influence the available opportunity set.
- Liquidity needs must also take into consideration the financial strength of the investor and resources beyond those held in the investment portfolio.

- When assessing the appropriateness of any given asset class for a given investor, it is important to evaluate potential liquidity needs in the context of an extreme market stress event.
- An investor's time horizon must be considered in any asset allocation exercise. Changes in human capital and the changing character of liabilities are two important time-related constraints of asset allocation.
- External considerations—such as regulations, tax rules, funding, and financing needs—are also likely to influence the asset allocation decision.
- Taxes alter the distribution of returns by both reducing the expected mean return and muting the dispersion of returns. Asset values and asset risk and return inputs to asset allocation should be modified to reflect the tax status of the investor. Correlation assumptions do not need to be adjusted, but taxes do affect the return and the standard deviation assumptions for each asset class.
- Periodic portfolio rebalancing to return the portfolio to its target strategic asset allocation is an integral part of sound portfolio management. Taxable investors must consider the tax implications of rebalancing.
- Rebalancing thresholds may be wider for taxable portfolios because it takes larger asset class movements to materially alter the risk profile of the taxable portfolio.
- Strategic asset location is the placement of less tax-efficient assets in accounts with more-favorable tax treatment.
- An asset owner's strategic asset allocation should be re-examined periodically, even in the absence of a change in the asset owner's circumstances.
- A special review of the asset allocation policy may be triggered by a change in goals, constraints, or beliefs.
- In some situations, a change to an asset allocation strategy may be implemented without a formal asset allocation study. Anticipating key milestones that would alter the asset owner's risk appetite, and implementing pre-established changes to the asset allocation in response, is often referred to as a "glide path."
- Tactical asset allocation (TAA) allows short-term deviations from the strategic asset allocation (SAA) targets and are expected to increase risk-adjusted return. Using either short-term views or signals, the investor actively re-weights broad asset classes, sectors, or risk-factor premiums. The sizes of these deviations from the SAA are often constrained by the Investment Policy Statement.
- The success of TAA decisions is measured against the performance of the SAA policy portfolio by comparing Sharpe ratios, evaluating the information ratio or the *t*-statistic of the average excess return of the TAA portfolio relative to the SAA portfolio, or plotting outcomes versus the efficient frontier.
- TAA incurs trading and tax costs. Tactical trades can also increase the concentration of risk.
- Discretionary TAA relies on a qualitative interpretation of political, economic, and financial market conditions and is predicated on a belief of persistent manager skill in predicting and timing short-term market moves.
- Systematic TAA relies on quantitative signals to capture documented return anomalies that may be inconsistent with market efficiency.
- The behavioral biases most relevant in asset allocation include loss aversion, the illusion of control, mental accounting, recency bias, framing, and availability bias.

- An effective investment program will address behavioral biases through a formal asset allocation process with its own objective framework, governance, and controls.
- In goals-based investing, loss-aversion bias can be mitigated by framing risk in terms of shortfall probability or by funding high-priority goals with low-risk assets.
- The cognitive bias, illusion of control, and hindsight bias can all be mitigated by using a formal asset allocation process that uses long-term return and risk forecasts, optimization constraints anchored around asset class weights in the global market portfolio, and strict policy ranges.
- Goals-based investing incorporates the mental accounting bias directly into the asset allocation solution by aligning each goal with a discrete sub-portfolio.
- A formal asset allocation policy with pre-specified allowable ranges may constrain recency bias.
- The framing bias effect can be mitigated by presenting the possible asset allocation choices with multiple perspectives on the risk/reward trade-off.
- Familiarity bias, a form of availability bias, most commonly results in an overweight in home country securities and may also cause investors to inappropriately compare their investment decisions (and performance) to other organizations. Familiarity bias can be mitigated by using the global market portfolio as the starting point in developing the asset allocation and by carefully evaluating any potential deviations from this baseline portfolio.
- A strong governance framework with the appropriate level of expertise and well-documented investment beliefs increases the likelihood that shifts in asset allocation are made objectively and in accordance with those beliefs. This will help to mitigate the effect that behavioral biases may have on the long-term success of the investment program.

REFERENCES

Bader, Lawrence N., and Jeremy Gold. 2007. "The Case against Stock in Public Pension Funds." *Financial Analysts Journal*, vol. 63, no. 1 (January/February): 55–62.

Bodie, Zvi, Robert C. Merton, and William F. Samuelson. 1992. "Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model." *Journal of Economic Dynamics & Control*, vol. 16: 427–449.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey Kubik. 2004. "Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization." *American Economic Review*, vol. 94, no. 5: 1276–1302.

Cooper, Stephen, and David Bianco. 2003. "Q-SeriesTM: Pension Fund Asset Allocation." UBS Investment Research (September).

Das, Sanjiv, Harry Markowitz, Jonathan Scheid, and Meir Statman. 2010. "Portfolio Optimization with Mental Accounts." *Journal of Financial and Quantitative Analysis*, vol. 45, no. 2: 311–334.

Das, Sanjiv, Harry Markowitz, Jonathan Scheid, and Meir Statman. 2011. "Portfolios for Investors Who Want to Reach Their Goals While Staying on the Mean–Variance Efficient Frontier." *Journal of Wealth Management*, vol. 14, no. 2 (Fall): 25–31.

Dyck, Alexander, and Lukasz Pomorski. 2011. "Is Bigger Better? Size and Performance in Pension Plan Management." Rotman School of Management Working Paper No. 1690724.

Horan, Stephen, and Ashraf Al Zaman. 2008. "Tax-Adjusted Portfolio Optimization and Asset Location: Extensions and Synthesis." *Journal of Wealth Management*, vol. 11, no. 3: 56–73.

Kritzman, Mark. 2015. "What Practitioners Need to Know about Time Diversification (corrected March 2015)." *Financial Analysts Journal*, vol. 71, no. 1 (January/February): 29–34.

Mladina, Peter. 2011. "Portfolio Implications of Triple Net Returns." *Journal of Wealth Management*, vol. 13, no. 4 (Spring): 51–59.

Pollet, Joshua, and Mungo Wilson. 2008. "How Does Size Affect Mutual Fund Behavior?" *Journal of Finance*, vol. 63, no. 6 (December): 2941–2969.

Reichenstein, William. 2006. "After-Tax Asset Allocation." *Financial Analysts Journal*, vol. 62, no. 4 (July/August): 14–19.

Samuelson, Paul A. 1963. "Risk and Uncertainty: A Fallacy of Large Numbers." *Scientia*, vol. 57, no. 98: 108–113.

Samuelson, Paul A. 1969. "Lifetime Portfolio Selection by Dynamic Stochastic Programming." *Review of Economics and Statistics*, vol. 51, no. 3: 239–246.

Stein, Jeremy. 2002. "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms." *Journal of Finance*, vol. 57, no. 5 (October): 1891–1921.

Turvey, Philip, Anup Basu, and Peter Verhoeven. 2013. "Embedded Tax Liabilities and Portfolio Choice." *Journal of Portfolio Management*, vol. 39, no. 3: 93–101.

PRACTICE PROBLEMS

The following information relates to questions 1–6

Rebecca Mayer is an asset management consultant for institutions and high-net-worth individuals. Mayer meets with Sebastian Capara, the newly appointed Investment Committee chairman for the Kinkardeen University Endowment (KUE), a very large tax-exempt fund.

Capara and Mayer review KUE's current and strategic asset allocations, which are presented in Exhibit 1. Capara informs Mayer that over the last few years, Kinkardeen University has financed its operations primarily from tuition, with minimal need of financial support from KUE. Enrollment at the University has been rising in recent years, and the Board of Trustees expects enrollment growth to continue for the next five years. Consequently, the board expects very modest endowment support to be needed during that time. These expectations led the Investment Committee to approve a decrease in the endowment's annual spending rate starting in the next fiscal year.

Exhibit 1 Kinkardeen University Endowment—Strategic Asset Allocation Policy

Asset Class	Current Weight	Target Allocation	Lower Policy Limit	Upper Policy Limit
Developed markets equity	30%	30%	25%	35%
Emerging markets equity	28%	30%	25%	35%
Investment-grade bonds	15%	20%	15%	25%
Private real estate equity	15%	10%	5%	15%
Infrastructure	12%	10%	5%	15%

As an additional source of alpha, Mayer proposes tactically adjusting KUE's asset-class weights to profit from short-term return opportunities. To confirm his understanding of tactical asset allocation (TAA), Capara tells Mayer the following:

Statement 1 The Sharpe ratio is suitable for measuring the success of TAA relative to SAA.

Statement 2 Discretionary TAA attempts to capture asset-class-level return anomalies that have been shown to have some predictability and persistence.

Statement 3 TAA allows a manager to deviate from the IPS asset-class upper and lower limits if the shift is expected to produce higher expected risk-adjusted returns.

Capara asks Mayer to recommend a TAA strategy based on excess return forecasts for the asset classes in KUE's portfolio, as shown in Exhibit 2.

© 2018 CFA Institute. All rights reserved.

Exhibit 2 Short-Term Excess Return Forecast

Asset Class	Expected Excess Return
Developed markets equity	2%
Emerging markets equity	5%
Investment-grade bonds	–3%
Private real estate equity	3%
Infrastructure	–1%

Following her consultation with Capara, Mayer meets with Roger Koval, a member of a wealthy family. Although Koval's baseline needs are secured by a family trust, Koval has a personal portfolio to fund his lifestyle goals.

In Koval's country, interest income is taxed at progressively higher income tax rates. Dividend income and long-term capital gains are taxed at lower tax rates relative to interest and earned income. In taxable accounts, realized capital losses can be used to offset current or future realized capital gains. Koval is in a high tax bracket, and his taxable account currently holds, in equal weights, high-yield bonds, investment-grade bonds, and domestic equities focused on long-term capital gains.

Koval asks Mayer about adding new asset classes to the taxable portfolio. Mayer suggests emerging markets equity given its positive short-term excess return forecast. However, Koval tells Mayer he is not interested in adding emerging markets equity to the account because he is convinced it is too risky. Koval justifies this belief by referring to significant losses the family trust suffered during the recent economic crisis.

Mayer also suggests using two mean–variance portfolio optimization scenarios for the taxable account to evaluate potential asset allocations. Mayer recommends running two optimizations: one on a pre-tax basis and another on an after-tax basis.

1 The change in the annual spending rate, in conjunction with the board's expectations regarding future enrollment and the need for endowment support, could justify that KUE's target weight for:

- **A** infrastructure be increased.
- **B** investment-grade bonds be increased.
- **C** private real estate equity be decreased.

2 Which of Capara's statements regarding tactical asset allocation is correct?

- **A** Statement 1
- **B** Statement 2
- **C** Statement 3

3 Based on Exhibits 1 and 2, to attempt to profit from the short-term excess return forecast, Capara should increase KUE's portfolio allocation to:

- **A** developed markets equity and decrease its allocation to infrastructure.
- **B** emerging markets equity and decrease its allocation to investment-grade bonds.
- **C** developed markets equity and increase its allocation to private real estate equity.

4 Given Koval's current portfolio and the tax laws of the country in which he lives, Koval's portfolio would be more tax efficient if he reallocated his taxable account to hold more:

- **A** high-yield bonds.

B investment-grade bonds.

C domestic equities focused on long-term capital gain opportunities.

5 Koval's attitude toward emerging markets equity reflects which of the following behavioral biases?

A Hindsight bias

B Availability bias

C Illusion of control

6 In both of Mayer's optimization scenarios, which of the following model inputs could be used without adjustment?

A Expected returns

B Correlation of returns

C Standard deviations of returns

The following information relates to questions 7–13

Elsbeth Quinn and Dean McCall are partners at Camel Asset Management (CAM). Quinn advises high-net-worth individuals, and McCall specializes in retirement plans for institutions.

Quinn meets with Neal and Karina Martin, both age 44. The Martins plan to retire at age 62. Twenty percent of the Martins' $600,000 in financial assets is held in cash and earmarked for funding their daughter Lara's university studies, which begin in one year. Lara's education and their own retirement are the Martins' highest-priority goals. Last week, the Martins learned that Lara was awarded a four-year full scholarship for university. Quinn reviews how the scholarship might affect the Martins' asset allocation strategy.

The Martins have assets in both taxable and tax-deferred accounts. For baseline retirement needs, Quinn recommends that the Martins maintain their current overall 60% equity/40% bonds (± 8% rebalancing range) strategic asset allocation. Quinn calculates that given current financial assets and expected future earnings, the Martins could reduce future retirement savings by 15% and still comfortably retire at 62. The Martins wish to allocate that 15% to a sub-portfolio with the goal of making a charitable gift to their alma mater from their estate. Although the gift is a low-priority goal, the Martins want the sub-portfolio to earn the highest return possible. Quinn promises to recommend an asset allocation strategy for the Martins' aspirational goal.

Next, Quinn discusses taxation of investments with the Martins. Their interest income is taxed at 35%, and capital gains and dividends are taxed at 20%. The Martins want to minimize taxes. Based on personal research, Neal makes the following two statements:

Statement 1 The after-tax return volatility of assets held in taxable accounts will be less than the pre-tax return volatility.

Statement 2 Assets that receive more favorable tax treatment should be held in tax-deferred accounts.

The equity portion of the Martins' portfolios produced an annualized return of 20% for the past three years. As a result, the Martins' equity allocation in both their taxable and tax-deferred portfolios has increased to 71%, with bonds falling to 29%.

The Martins want to keep the strategic asset allocation risk levels the same in both types of retirement portfolios. Quinn discusses rebalancing; however, Neal is somewhat reluctant to take money out of stocks, expressing confidence that strong investment returns will continue.

Quinn's CAM associate, McCall, meets with Bruno Snead, the director of the Katt Company Pension Fund (KCPF). The strategic asset allocation for the fund is 65% stocks/35% bonds. Because of favorable returns during the past eight recession-free years, the KCPF is now overfunded. However, there are early signs of the economy weakening. Since Katt Company is in a cyclical industry, the Pension Committee is concerned about future market and economic risk and fears that the high-priority goal of maintaining a fully funded status may be adversely affected. McCall suggests to Snead that the KCPF might benefit from an updated IPS. Following a thorough review, McCall recommends a new IPS and strategic asset allocation.

The proposed IPS revisions include a plan for short-term deviations from strategic asset allocation targets. The goal is to benefit from equity market trends by automatically increasing (decreasing) the allocation to equities by 5% whenever the S&P 500 Index 50-day moving average crosses above (below) the 200-day moving average.

7 Given the change in funding of Lara's education, the Martins' strategic asset allocation would *most likely* decrease exposure to:

- **A** cash.
- **B** bonds.
- **C** equities.

8 The *most* appropriate asset allocation for the Martins' new charitable gift sub-portfolio is:

- **A** 40% equities/60% bonds.
- **B** 70% equities/30% bonds.
- **C** 100% equities/0% bonds.

9 Which of Neal's statements regarding the taxation of investments is correct?

- **A** Statement 1 only
- **B** Statement 2 only
- **C** Both Statement 1 and Statement 2

10 Given the Martins' risk and tax preferences, the taxable portfolio should be rebalanced:

- **A** less often than the tax-deferred portfolio.
- **B** as often as the tax-deferred portfolio.
- **C** more often than the tax-deferred portfolio.

11 During the rebalancing discussion, which behavioral bias does Neal exhibit?

- **A** Framing bias
- **B** Loss aversion
- **C** Representative bias

12 Given McCall's IPS recommendation, the *most* appropriate new strategic asset allocation for the KCPF is:

- **A** 40% stocks/60% bonds.
- **B** 65% stocks/35% bonds.
- **C** 75% stocks/25% bonds.

13 The proposal for short-term adjustments to the KCPF asset allocation strategy is known as:

A de-risking.

B systematic tactical asset allocation.

C discretionary tactical asset allocation.

SOLUTIONS

1 A is correct. A lower annual spending rate, in addition to the board's expectations of rising enrollment and minimal need for endowment support over the next five years, indicates a decreased need for liquidity. Therefore, KUE could justify an increase in the strategic allocation to less liquid asset classes (such as private real estate equity and infrastructure) and a decrease in the strategic allocation to liquid assets (such as investment-grade bonds).

2 A is correct. The Sharpe ratio is suitable for measuring the success of TAA relative to SAA. Specifically, the success of TAA decisions can be evaluated by comparing the Sharpe ratio realized under the TAA with the Sharpe ratio that would have been realized under the SAA.

3 A is correct. The forecast for expected excess returns is positive for developed markets equity and negative for infrastructure. Therefore, to attempt to profit from the short-term excess return forecast, KUE can overweight developed markets equity and underweight infrastructure. These adjustments to the asset-class weights are within KUE's lower and upper policy limits.

4 C is correct. As a general rule, the portion of a taxable asset owner's assets that are eligible for lower tax rates and deferred capital gains tax treatment should first be allocated to the investor's taxable accounts. Assets that generate returns mainly from interest income tend to be less tax efficient and in Koval's country are taxed at progressively higher rates. Also, the standard deviation (volatility) of after-tax returns is lower when equities are held in a taxable account. Therefore, Koval's taxable account would become more tax efficient if it held more domestic equities focused on long-term capital gain opportunities.

5 B is correct. Availability bias is an information-processing bias in which people take a mental shortcut when estimating the probability of an outcome based on how easily the outcome comes to mind. On the basis of the losses incurred by his family trust during the recent economic crisis, Koval expresses a strong preference for avoiding the emerging markets equity asset class. Such behavior is consistent with availability bias, where investors who personally experience an adverse event are likely to assign a higher probability to such an event occurring again.

6 B is correct. After-tax portfolio optimization requires adjusting each asset class's expected return and risk for expected taxes. The correlation of returns is not affected by taxes and does not require an adjustment when performing after-tax portfolio optimization.

7 A is correct. The changing character of liabilities through time affects the asset allocation to fund those liabilities. The Martins' investment horizon for some of their assets has changed. The amount of liquidity needed for Lara's near-term education has been greatly reduced owing to the receipt of the scholarship. The Martins will likely still have to pay for some university-related expenses; however, a large part of the $120,000 in cash that is earmarked for Lara's expenses can now be allocated to the Martins' long-term goal of early retirement. Retirement is 18 years away, much longer than the one- to five-year horizon for university expenses. Therefore, the Martins' allocation to cash would likely decrease.

8 C is correct. The Martins' sub-portfolio is aspirational and a low priority. Investors are usually willing to take more risk on lower-priority, aspirational portfolios. The charitable gift will be made from their estate, which indicates a long time horizon. In addition, the Martins want the highest return possible. Therefore, the highest allocation to equities is most appropriate.

9 A is correct. Taxes alter the distribution of returns by both reducing the expected mean return and muting the dispersion of returns. The portion of an owner's taxable assets that are eligible for lower tax rates and deferred capital gains tax treatment should first be allocated to the investor's taxable accounts.

10 A is correct. The Martins wish to maintain the same risk level for both retirement accounts based on their strategic asset allocation. However, more frequent rebalancing exposes the taxable asset owner to realized taxes that could have otherwise been deferred or even avoided. Rebalancing is discretionary, and the Martins' also wish to minimize taxes. Because after-tax return volatility is lower than pre-tax return volatility, it takes larger asset-class movements to materially alter the risk profile of a taxable portfolio. This suggests that rebalancing ranges for a taxable portfolio can be wider than those of a tax-exempt/tax-deferred portfolio with a similar risk profile; thus, rebalancing occurs less frequently.

11 C is correct. Representative, or recency, bias is the tendency to overweight the importance of the most recent observations and information relative to a longer-dated or more comprehensive set of long-term observations and information. Return chasing is a common result of this bias, and it results in overweighting asset classes with strong recent performance.

12 A is correct. McCall recommends a new IPS. Changes in the economic environment and capital market expectations or changes in the beliefs of committee members are factors that may lead to an altering of the principles that guide investment activities. Because the plan is now overfunded, there is less need to take a higher level of equity risk. The Pension Committee is concerned about the impact of future market and economic risks on the funding status of the plan. Katt Company operates in a cyclical industry and could have difficulty making pension contributions during a recession. Therefore, a substantial reduction in the allocation to stocks and an increase in bonds reduce risk. The 40% stocks/60% bonds alternative increases the allocation to bonds from 35% to 60%. Increasing the fixed-income allocation should moderate plan risk, provide a better hedge for liabilities, and reduce contribution uncertainty.

13 B is correct. Using rules-based, quantitative signals, systematic tactical asset allocation (TAA) attempts to capture asset-class-level return anomalies that have been shown to have some predictability and persistence. Trend signals are widely used in systematic TAA. A moving-average crossover is a trend signal that indicates an upward (downward) trend when the moving average of the shorter time frame, 50 days, is above (below) the moving average of the longer time frame, 200 days.

READING

21

Currency Management: An Introduction

by William A. Barker, PhD, CFA

William A. Barker, PhD, CFA (Canada).

LEARNING OUTCOMES

Mastery	The candidate should be able to:
☐	**a.** analyze the effects of currency movements on portfolio risk and return;
☐	**b.** discuss strategic choices in currency management;
☐	**c.** formulate an appropriate currency management program given financial market conditions and portfolio objectives and constraints;
☐	**d.** compare active currency trading strategies based on economic fundamentals, technical analysis, carry-trade, and volatility trading;
☐	**e.** describe how changes in factors underlying active trading strategies affect tactical trading decisions;
☐	**f.** describe how forward contracts and FX (foreign exchange) swaps are used to adjust hedge ratios;
☐	**g.** describe trading strategies used to reduce hedging costs and modify the risk–return characteristics of a foreign-currency portfolio;
☐	**h.** describe the use of cross-hedges, macro-hedges, and minimum-variance-hedge ratios in portfolios exposed to multiple foreign currencies;
☐	**i.** discuss challenges for managing emerging market currency exposures.

INTRODUCTION

1

Globalization has been one of the most persistent themes in recent history, and this theme applies equally to the world of finance. New investment products, deregulation, worldwide financial system integration, and better communication and information

© 2013 CFA Institute. All rights reserved.

networks have opened new global investment opportunities. At the same time, investors have increasingly shed their "home bias" and sought investment alternatives beyond their own borders.

The benefits of this trend for portfolio managers have been clear, both in terms of the broader availability of higher-expected-return investments as well as portfolio diversification opportunities. Nonetheless, investments denominated in foreign currencies also bring a unique set of challenges: measuring and managing foreign exchange risk. Buying foreign-currency denominated assets means bringing currency risk into the portfolio. Exchange rates are volatile and, at least in the short to medium term, can have a marked impact on investment returns and risks—*currency matters*. The key to the superior performance of global portfolios is the effective management of this currency risk.

This reading explores basic concepts and tools of currency management. Section 2 reviews some of the basic concepts of foreign exchange (FX) markets. The material in subsequent sections presumes an understanding of these concepts. Section 3 examines some of the basic mathematics involved in measuring the effects of foreign-currency investments on portfolio return and risk. Section 4 discusses the *strategic* decisions portfolio managers face in setting the target currency exposures of the portfolio. The currency exposures that the portfolio can accept range from a fully hedged position to active management of currency risk. Section 5 discusses some of the *tactical* considerations involving active currency management if the investment policy statement (IPS) extends some latitude for active currency management. A requisite to any active currency management is having a market view; so this section includes various methodologies by which a manager can form directional views on future exchange rate movements and volatility. Section 6 covers a variety of trading tools available to implement both hedging and active currency management strategies. Although the generic types of FX derivatives tools are relatively limited—spot, forward, option, and swap contracts—the number of variations within each and the number of combinations in which they can be used is vast. Section 7 examines some of the issues involved in managing the currency exposures of emerging market currencies—that is, those that are less liquid than the major currencies. Section 8 presents a summary.

2 REVIEW OF FOREIGN EXCHANGE CONCEPTS

We begin with a review of the basic trading tools of the foreign exchange market: spot, forward, FX swap, and currency option transactions. The concepts introduced in this section will be used extensively in our discussion of currency management techniques in subsequent sections.

Most people think only of spot transactions when they think of the foreign exchange market, but in fact the spot market accounts for less than 40% of the average daily turnover in currencies.[1] Although cross-border *business* may be transacted in the spot market (making and receiving foreign currency payments), the *risk management* of these flows takes place in FX derivatives markets (i.e., using forwards, FX swaps, and currency options). So does the hedging of foreign currency assets and liabilities. It is unusual for market participants to engage in any foreign currency transactions without also managing the currency risk they create. Spot transactions typically generate derivative transactions. As a result, understanding these FX derivatives markets, and their relation to the spot market, is critical for understanding the currency risk management issues examined in this reading.

1 2013 Triennial Survey, Bank for International Settlements (2013).

2.1 Spot Markets

In professional FX markets, exchange rate quotes are described in terms of the three-letter currency codes used to identify individual currencies. Exhibit 1 shows a list of some of the more common currency codes.

Exhibit 1 Currency Codes

USD	US dollar
EUR	Euro
GBP	British pound
JPY	Japanese yen
MXN	Mexican peso
CHF	Swiss franc
CAD	Canadian dollar
SEK	Swedish krona
AUD	Australian dollar
KRW	Korean won
NZD	New Zealand dollar
BRL	Brazilian real
RUB	Russian ruble
CNY	Chinese yuan
INR	Indian rupee
ZAR	South African rand

An exchange rate is the price of the base currency expressed in terms of the price currency. For example, a USD/EUR rate of 1.3650 means the euro, the base currency, costs 1.3650 US dollars. The exact notation used to represent exchange rates can vary widely between sources, and occasionally the same exchange rate notation will be used by different sources to mean completely different things. The reader should be aware that the notation used here may not be the same as that encountered elsewhere. To avoid confusion, this reading will identify exchange rates using the convention of "P/B," which refers to the price of one unit of the base currency "B" expressed in terms of the price currency "P."[2]

2 Notation is generally not standardized in global foreign exchange markets, and there are several common ways of expressing the same currency pair (e.g., JPY/USD, USD:JPY, $/¥). However, what is common in FX markets is the concept of a "base" and "price" currency when setting exchange rate prices. Later in the reading, we will sometimes switch to discussing a "domestic" and a "foreign" currency. This designation will only be used as an illustrative device for making the explanation of various theoretical concepts easier. The candidate should be aware that describing currency pairs in terms of "foreign" and "domestic" currencies is typically not done in professional FX markets because what is the "foreign" currency and what is the "domestic" currency depends on where one is located, which can lead to confusion. For example, what is "foreign" and what is "domestic" for an investor from the Middle East trading CHF against GBP with the New York branch of a European bank, with the trade ultimately booked at the bank's headquarters in Paris?

How the professional FX market quotes exchange rates—which is the base currency, and which is the price currency, in any currency pair—is not arbitrary but follows conventions that are broadly agreed on throughout the market. Generally, there is a hierarchy as to which currency will be quoted as the base currency in any given P/B currency pair:

1 Currency pairs involving the EUR will use the EUR as the base currency (for example, GBP/EUR).
2 Currency pairs involving the GBP, other than those involving the EUR, will use the GBP as the base currency (for example, CHF/GBP).
3 Currency pairs involving either the AUD or NZD, other than those involving either the EUR or GBP, will use these currencies as the base currency (for example, USD/AUD and NZD/AUD). The market convention between these two currencies is for a NZD/AUD quote.
4 All other currency quotes involving the USD will use USD as the base currency (for example, MXN/USD).

Readers are encouraged to familiarize themselves with the quoting conventions used in the professional FX market because they are the currency quotes that will be experienced in practice. Exhibit 2 lists some of the most commonly traded currency pairs in global FX markets and their market-standard quoting conventions. These market-standard conventions will be used for the balance of this reading.

Exhibit 2 Select Market-Standard Currency Pair Quotes

Quote convention	Market name
USD/EUR	Euro-dollar
GBP/EUR	Euro-sterling
USD/GBP	Sterling-dollar
JPY/USD	Dollar-yen
USD/AUD	Aussie-dollar
CHF/USD	Dollar-Swiss
CAD/USD	Dollar-Canada
JPY/EUR	Euro-yen
CHF/EUR	Euro-Swiss
JPY/GBP	Sterling-yen

Another convention used in professional FX markets is that most spot currency quotes are priced out to four decimal places: for example, a typical USD/EUR quote would be 1.3500 and not 1.35. The price point at the fourth decimal place is commonly referred to as a "pip." Professional FX traders also refer to what is called the "big figure" or the "handle," which is the integer to the left side of the decimal place as well as the first two decimal places of the quote. For example, for a USD/EUR quote of 1.3568, 1.35 is the handle and there are 68 pips.

There are exceptions to this four decimal place rule. First, forward quotes—discussed later—will often be quoted out to five and sometimes six decimal places. Second, because of the relative magnitude of some currency values, some currency quotes will only be quoted out to two decimal places. For example, because it takes many Japanese yen to buy one US dollar, the typical spot quote for JPY/USD is priced out to only two decimal places (for example, 86.35 and not 86.3500).[3]

The spot exchange rate is usually for settlement on the second business day after the trade date, referred to as $T + 2$ settlement.[4] In foreign exchange markets—as in other financial markets—market participants confront a two-sided price in the form of a bid price and an offer price (also called an ask price) being quoted by potential counterparties. The **bid price** is the price, defined in terms of the price currency, at which the counterparty providing a two-sided price quote is willing to buy one unit of the **base** currency. Similarly, **offer price** is the price, in terms of the price currency, at which that counterparty is willing to sell one unit of the base currency. For example, given a price request from a client, a dealer might quote a two-sided price on the spot USD/EUR exchange rate of 1.3648/1.3652. This quote means that the dealer is willing to pay USD1.3648 to buy one euro (bid) and that the dealer will sell one euro (offer) for USD1.3652. The market width, usually referred to as dealer's spread or the bid–offer spread, is the difference between the bid and the offer. When transacting on a dealer's bid-offer two-sided price quote, a client is said to either "hit the bid" (selling the base currency) or "pay the offer" (buying the base currency).

An easy check to see whether the bid or offer should be used for a specific transaction is that the party *asking* the dealer for a price should be on the more expensive side of the market. For example, if one wants to buy 1 EUR, 1.3652 is more USD per EUR than 1.3648. Hence, paying the offer involves paying more EUR. Similarly, when selling 1 EUR, hitting the bid at 1.3648 means less USD received than 1.3652.

2.2 Forward Markets

Forward contracts are agreements to exchange one currency for another on a future date at an exchange rate agreed on today.[5] In contrast to spot rates, forward contracts are any exchange rate transactions that occur with settlement longer than the usual $T + 2$ settlement for spot delivery.

In professional FX markets, forward exchange rates are typically quoted in terms of "points." The points on a forward rate quote are simply the difference between the forward exchange rate quote and the spot exchange rate quote; that is, the forward premium or discount, with the points scaled so that they can be related to the last decimal place in the spot quote. Forward points are adjustments to the spot price of the base currency, using our standard price/base (P/B) currency notation.

This means that forward rate quotes in professional FX markets are typically shown as the bid–offer on the spot rate and the number of forward points at each maturity.[6] For illustration purposes, assume that the bid–offer for the spot and forward points for the USD/EUR exchange rate are as shown in Exhibit 3.

3 It should be mentioned that many electronic dealing platforms in the FX market are moving to five decimal place pricing for spot quotes, using what are referred to as "deci-pips." In this case, for example, a USD/EUR spot quote might be shown as 1.37645. Spot quotes for JPY/USD on these systems will be given out to three decimal places.

4 The exception among the major currencies is CAD/USD, for which standard spot settlement is $T + 1$.

5 These are sometimes called outright forwards to distinguish them from FX swaps, which are discussed later.

6 Maturity is defined in terms of the time between spot settlement, usually $T + 2$, and the settlement of the forward contract.

Exhibit 3 Sample Spot and Forward Quotes (Bid–Offer)

Maturity	Spot Rate or Forward Points
Spot (USD/EUR)	1.3549/1.3651
One month	–5.6/–5.1
Three months	–15.9/–15.3
Six months	–37.0/–36.3
Twelve months	–94.3/–91.8

To convert any of these quoted forward points into a forward rate, one would divide the number of points by 10,000 (to scale down to the fourth decimal place, the last decimal place in the USD/EUR spot quote) and then add the result to the spot exchange rate quote.[7] But one must be careful about which side of the market (bid or offer) is being quoted. For example, suppose a market participant was *selling* the EUR forward against the USD. Given the USD/EUR quoting convention, the EUR is the base currency. This means the market participant must use the *bid* rates (i.e., the market participant will "hit the bid") given the USD/EUR quoting convention. Using the data in Exhibit 3, the three-month forward *bid* rate in this case would be based on the bid for both the spot and the forward points, and hence would be:

$$1.3549 + \left(\frac{-15.9}{10{,}000}\right) = 1.35331$$

This result means that the market participant would be selling EUR three months forward at a price of USD1.35331 per EUR. Note that the quoted points are already scaled to each maturity—they are not annualized—so there is no need to adjust them.

Although there is no cash flow on a forward contract until settlement date, it is often useful to do a mark-to-market valuation on a forward position before then to (1) judge the effectiveness of a hedge based on forward contracts (i.e., by comparing the change in the mark-to-market of the underlying asset with the change in the mark-to-market of the forward), and (2) to measure the profitability of speculative currency positions at points before contract maturity.

As with other financial instruments, the mark-to-market value of forward contracts reflects the profit (or loss) that would be realized from closing out the position at current market prices. To close out a forward position, it must be offset with an equal and opposite forward position using the spot exchange rate and forward points available in the market when the offsetting position is created. When a forward contract is initiated, the forward rate is such that no cash changes hands (i.e., the mark-to-market value of the contract at initiation is zero). From that moment onward, however, the mark-to-market value of the forward contract will change as the spot exchange rate changes as well as when interest rates change in either of the two currencies.

Consider an example. Suppose that a market participant bought GBP10,000,000 for delivery against the AUD in six months at an "all-in" forward rate of 1.6100 AUD/GBP. (The all-in forward rate is simply the sum of the spot rate and the forward points, appropriately scaled to size.) Three months later, the market participant wants to close out this forward contract. To do that would require selling GBP10,000,000 three

7 Because the JPY/USD exchange rate is only quoted to two decimal places, forward points for the dollar/yen currency pair are divided by 100.

months forward using the AUD/GBP spot exchange rate and forward points in effect at that time.[8] Assume the bid–offer for spot and forward points three months prior to the settlement date are as follows:

Spot rate (AUD/GBP)	1.6210/1.6215
Three-month points	130/140

To sell GBP (the base currency in the AUD/GBP quote) means calculating the *bid* side of the market. Hence, the appropriate all-in three-month forward rate to use is

$$1.6210 + 130/10{,}000 = 1.6340$$

Thus, the market participant originally bought GBP10,000,000 at an AUD/GBP rate of 1.6100 and subsequently sold them at a rate of 1.6340. These GBP amounts will net to zero at settlement date (GBP10 million both bought and sold), but the AUD amounts will not net to zero because the forward rate has changed. The AUD cash flow at settlement date will be equal to

$$(1.6340 - 1.6100) \times 10{,}000{,}000 = \text{AUD}240{,}000$$

This amount is a cash *inflow* because the market participant was long the GBP with the original forward position and the GBP subsequently appreciated (the AUD/GBP rate increased).

This cash flow is paid at settlement day, which is still three months away. To calculate the mark-to-market value on the dealer's position, this cash flow must be discounted to the present. The present value of this amount is found by discounting the settlement day cash flow by the three-month discount rate. Because it is an AUD amount, the three-month AUD discount rate is used. If Libor is used and the three-month AUD Libor is 4.80% (annualized), the present value of this future AUD cash flow is then

$$\frac{\text{AUD}240{,}000}{1 + 0.048\left[\frac{90}{360}\right]} = \text{AUD}237{,}154$$

This is the mark-to-market value of the original long GBP10 million six-month forward contract when it is closed out three months prior to settlement.

To summarize, the process for marking-to-market a forward position is relatively straightforward:

1. Create an equal and offsetting forward position to the original forward position. (In the example earlier, the market participant is long GBP10 million forward, so the offsetting forward contract would be to sell GBP10 million.)
2. Determine the appropriate all-in forward rate for this new, offsetting forward position. If the base currency of the exchange rate quote is being sold (bought), then use the bid (offer) side of the market.
3. Calculate the cash flow at settlement day. This calculation will be based on the original contract size times the difference between the original forward rate and the rate calculated in Step 2. If the currency the market participant was originally long (short) subsequently appreciated (depreciated), then there will be

8 Note that the offsetting forward contract is defined in terms of the original position taken. The original position in this example was long GBP10 million, so the offsetting contract is short GBP10 million. But there is an ambiguity here: to be *long* GBP10 million at 1.6100 AUD/GBP is equivalent to being *short* AUD16,100,000 (10,000,000 × 1.6100) at the same forward rate. To avoid this ambiguity, for the purposes of this reading we will state what the relevant forward position is for mark-to-market purposes. The net gain or loss from the transaction will be reflected in the alternate currency.

a cash *inflow*. Otherwise, there will be a cash outflow. (In the earlier example, the market participant was long the GBP and it subsequently appreciated; this appreciation led to a cash inflow at the settlement day.)

4 Calculate the present value of this cash flow at the future settlement date. The currency of the cash flow and the discount rate must match. (In the example earlier, the cash flow at the settlement date is in AUD, so an AUD Libor rate is used to calculate the present value.)

Finally, we note that in the example, the mark-to-market value is given in AUD. It would be possible to translate this AUD amount into any other currency value using the current spot rate for the relevant currency pair.[9]

2.3 FX Swap Markets

An FX swap transaction consists of offsetting and simultaneous spot and forward transactions, in which the base currency is being bought (sold) spot and sold (bought) forward. These two transactions are often referred to as the "legs" of the swap. The two legs of the swap can either be of equal size (a "matched" swap) or one can be larger than the other (a "mismatched" swap). FX swaps are distinct from currency swaps. Similar to currency swaps, FX swaps involve an exchange of principal amounts in different currencies at swap initiation that is reversed at swap maturity. Unlike currency swaps, FX swaps have no interim interest payments and are nearly always of much shorter term than currency swaps.

FX swaps are important for managing currency risk because they are used to "roll" forward contracts forward as they mature. For example, consider the case of a trader who *bought* GBP1,000,000 one month forward against the CHF in order to set up a currency hedge. One month later, the forward contract will expire. To maintain this long position in the GBP against the CHF, two days prior to contract maturity, given $T + 2$ settlement, the trader must (1) sell GBP1,000,000 against the CHF spot, to settle the maturing forward contract; and (2) buy GBP1,000,000 against the CHF forward. That is, the trader is engaging in an FX swap (a matched swap in this case because the GBP currency amounts are equal).

If a trader wanted to adjust the size of the currency hedge (i.e., the size of the outstanding forward position), the forward leg of the FX swap can be of a different size than the spot transaction when the hedge is rolled. Continuing the previous example, if the trader wanted to increase the size of the long-GBP position by GBP500,000 as the outstanding forward contract expires, the transactions required would be to (1) sell GBP1,000,000 against the CHF spot, to settle the maturing forward contract; and (2) buy GBP1,500,000 against the CHF forward. This would be a mismatched swap.

The pricing of swaps will differ slightly depending on whether they are matched or mismatched swaps. If the amount of the base currency involved for the spot and forward legs of the swap are equal (a matched swap), then these are exactly offsetting transactions; one is a buy, the other a sell, and both are for the same amount. Because of this equality, a common *spot* exchange rate is typically applied to both legs of the swap transaction; it is standard practice to use the mid-market spot exchange rate for a matched swap transaction. However, the *forward* points will still be based on either

9 In the context of this example, an alternative and equivalent approach to re-denominate the mark-to-market value in another currency would be to sell the AUD forward. In our example, the cash flow is +AUD 240,000 in 90 days. If one wanted to redenominate the mark-to-market in USD, for example, one would sell this amount of AUD 90-days forward against the USD at the prevailing USD/AUD 90-day forward bid rate. This will produce a USD cash flow in 90 days. This USD amount can then be present-valued at the 90-day US Libor rate to get the USD mark-to-market value of the AUD/GBP forward position. The day-count convention used here is an "actual/360" basis.

the bid or offer, depending on whether the market participant is buying or selling the base currency forward. In the earlier example, the trader is *buying* the GBP (the base currency) forward and would hence pay the *offer* side of the market for forward points.

If the FX swap is mismatched, then pricing will need to reflect the difference in trade sizes between the two legs of the transaction. Continuing the example in which the trader increased the size of the long-GBP position by GBP500,000, this mismatched swap is equivalent to (1) a matched swap for a size of GBP1,000,000, and (2) an outright forward contract buying GBP500,000. Pricing for the mismatched swap must reflect this net GBP purchase amount. Because the matched swap would already price the forward points on the offer side of the market, typically this mismatched size adjustment would be reflected in the *spot* rate quoted as the base for the FX swap. Because a net amount of GBP is being *bought*, the spot quote would now be on the *offer* side of the CHF/GBP spot rate quote. (In addition, the trader would still pay the offer side of the market for the forward points.)

We will return to these topics later in the reading when discussing in more depth the use of forward contracts and FX swaps to adjust hedge ratios. (A **hedge ratio** is the ratio of the nominal value of the derivatives contract used as a hedge to the market value of the hedged asset.)

2.4 Currency Options

The final product type within FX markets is currency options. The market for currency options is, in many ways, similar to option markets for other asset classes, such as bonds and equities. As in other markets, the most common options in FX markets are call and put options, which are widely used for both risk management and speculative purposes. However, in addition to these vanilla options, the FX market is also characterized by active trading in exotic options. ("Exotic" options have a variety of features that make them exceptionally flexible risk management tools, compared with vanilla options.)

The risk management uses of both vanilla and exotic currency options will be examined in subsequent sections. Although daily turnover in FX options market is small in *relative* terms compared with the overall daily flow in global spot currency markets, because the overall currency market is so large, the *absolute* size of the FX options market is still very considerable.

3 CURRENCY RISK AND PORTFOLIO RETURN AND RISK

In this section, we examine the effect of currency movements on asset returns and portfolio risk. We then turn to how these effects help determine construction of a foreign asset portfolio.

3.1 Return Decomposition

In this section, we examine how international exposure affects a portfolio's return. A **domestic asset** is an asset that trades in the investor's **domestic currency** (or **home currency**). From a portfolio manager's perspective, the domestic currency is the one in which portfolio valuation and returns are reported. *Domestic* refers to a relation between the currency denomination of the asset and the investor; it is not an inherent property of either the asset or the currency. An example of a domestic asset is a USD-denominated bond portfolio from the perspective of a US-domiciled investor. The return on a domestic asset is not affected by exchange rate movements of the domestic currency.

Foreign assets are assets denominated in currencies other than the investor's home currency. An example of a foreign asset is a USD-denominated bond portfolio from the perspective of a eurozone-domiciled investor (and for whom the euro is the home currency). The return on a foreign asset will be affected by exchange rate movements in the home currency against the **foreign currency**. Continuing with our example, the return to the eurozone-domiciled investor will be affected by the USD return on the USD-denominated bond as well as movements in the exchange rate between the home currency and the foreign currency, the EUR and USD respectively.

The return of the foreign asset measured in foreign-currency terms is known as the **foreign-currency return**. Extending the example, if the value of the USD-denominated bond increased by 10%, measured in USD, that increase is the foreign-currency return to the eurozone-domiciled investor. The **domestic-currency return** on a foreign asset will reflect both the foreign-currency return on that asset as well as percentage movements in the spot exchange rate between the home and foreign currencies. The domestic-currency return is multiplicative with respect to these two factors:

$$R_{DC} = (1 + R_{FC})(1 + R_{FX}) - 1 \quad \textbf{(1)}$$

where R_{DC} is the domestic-currency return (in percent), R_{FC} is the foreign-currency return, and R_{FX} is the percentage change of the foreign currency against the domestic currency.

Returning to the example, the domestic-currency return for the eurozone-domiciled investor on the USD-denominated bond will reflect both the bond's USD-denominated return as well as movements in the exchange rate between the USD and the EUR. Suppose that the foreign-currency return on the USD-denominated bond is 10% and the USD appreciates by 5% against the EUR. In this case, the domestic-currency return to the eurozone investor will be:

$$(1 + 10\%)(1 + 5\%) - 1 = (1.10)(1.05) - 1 = 0.155 = 15.5\%$$

Although the concept is seemingly straightforward, the reader should be aware that Equation 1 hides a subtlety that must be recognized. The term R_{FX} is defined as the percentage change in the foreign currency against the domestic currency. However, this change is *not* always the same thing as the percentage change in the spot rate using market standard P/B quotes (for example, as shown in Exhibit 2). In other words, it is not always the case that $R_{FX} = \%\Delta S_{P/B}$, where the term on the right side of the equal sign is defined in standard FX market convention (note that %Δ is percentage change).

This distinction is important because which currency is considered the domestic currency and whether it is either the base or the price currency in the market standard P/B quote will lead to completely different mathematical results. This happens in two ways. First, it determines the *sign*, or direction, of the exchange rate appreciation. For example, in the illustration earlier, the euro was the domestic currency. Note also that in the market standard USD/EUR quote, the euro is the base currency. In the illustration, the USD appreciated, and an appreciation of the USD means a *depreciation* of the USD/EUR quote.

Another way in which quoting conventions affect mathematical results involves the fact that the foreign exchange return, R_{FX} in Equation 1, is calculated with the investor's domestic currency as the price currency. Even if one gets the sign, or direction, of change right, it is not necessarily the case that one can simply "flip the sign" of the percentage change in market standard P/B to get the right answer. For example, in our earlier illustration, the euro was the investor's domestic currency and it is also the base currency in the market-standard USD/EUR quote. But note that a 5% move in USD/EUR in not the same as a 5% move in EUR/USD. Assume that the current USD/

EUR spot rate is 1.3510 (meaning that the EUR/USD rate is 0.7402). An appreciation of the USD, as in our illustration, is equivalent to a depreciation of the EUR. A 5% depreciation of USD/EUR leads to 1.2835. But a 5% appreciation of EUR/USD leads to 0.7772, which inverts to 1.2867—a different number. Hence, to be accurate, the foreign exchange calculation in Equation 1 must be quoted so that the "domestic" currency is always the price currency. One must be careful when using quote conventions and make adjustments as necessary to calculate the domestic-currency return properly.

With this nuance in mind, what holds for the domestic-currency return of a single foreign asset also holds for the returns on a multi-currency portfolio of foreign assets, except now the portfolio weights must be considered. More generally, the domestic-currency return on a portfolio of multiple foreign assets will be equal to

$$R_{DC} = \sum_{i=1}^{n} \omega_i \left(1 + R_{FC,i}\right)\left(1 + R_{FX,i}\right) - 1 \qquad (2)$$

where $R_{FC,i}$ is the foreign-currency return on the i-th foreign asset, $R_{FX,i}$ is the appreciation of the i-th foreign currency against the domestic currency, and ω_i are the portfolio weights of the foreign-currency assets (defined as the percentage of the aggregate domestic-currency value of the portfolio) and $\sum_{i=1}^{n}\omega_i = 1$. (Note that if short selling is allowed in the portfolio, some of the ω_i can be less than zero.) Again, it is important that the exchange rate notation in this expression (used to calculate $R_{FX,i}$) must be consistently defined with the domestic currency as the price currency.

Assume the following information for a portfolio held by an investor in India. Performance is measured in terms of the Indian rupee (INR) and the weights of the two assets in the portfolio, at the beginning of the period, are 80% for the GBP-denominated asset and 20% for the EUR-denominated asset, respectively. (Note that the portfolio weights are measured in terms of a common currency, the INR, which is the investor's domestic currency in this case.)

	One Year Ago	**Today***
INR/GBP spot rate	84.12	85.78
INR/EUR spot rate	65.36	67.81
GBP-denominated asset value, in GBP millions	43.80	50.70
EUR-denominated asset value, in EUR millions	14.08	12.17
GBP-denominated asset value, in INR millions	3,684.46	
EUR-denominated asset value, in INR millions	920.27	
GBP-denominated assets, portfolio weight (INR)	80%	
EUR-denominated assets, portfolio weight (INR)	20%	

* Today's asset values are prior to rebalancing.

The domestic-currency return (R_{DC}) is calculated as follows:

$$R_{DC} = 0.80(1 + R_{FC,GBP})(1 + R_{FX,GBP}) + 0.20(1 + R_{FC,EUR})(1 + R_{FX,EUR}) - 1$$

Note that given the exchange rate quoting convention, the INR is the price currency in the P/B quote for both currency pairs. Adding the data from the table leads to:

$$R_{DC} = 0.80\left(\frac{50.70}{43.80}\right)\left(\frac{85.78}{84.12}\right) + 0.20\left(\frac{12.17}{14.08}\right)\left(\frac{67.81}{65.36}\right) - 1$$

This solves to 0.124 or 12.4%.

To get the *expected* future return on a foreign-currency asset portfolio, based on Equation 2, the portfolio manager would need a market opinion for the expected price movement in each of the foreign assets ($R_{A,i}$) and exchange rates ($R_{FX,i}$) in the portfolio. There are typically correlations between all of these variables—correlations between the foreign asset price movements across countries, correlations between movements among various currency pairs, and correlations between exchange rate movements and foreign-currency asset returns. The portfolio manager would need to account for these correlations when forming expectations about future asset price and exchange rate movements.

3.2 Volatility Decomposition

Now we will turn to examining the effect of currency movements on the volatility of domestic-currency returns. Equation 1 can be rearranged as

$$R_{DC} = (1 + R_{FC})(1 + R_{FX}) - 1 = R_{FC} + R_{FX} + R_{FC}R_{FX}$$

When R_{FC} and R_{FX} are small, then the cross-term ($R_{FC}R_{FX}$) is small, and as a result this equation can be approximated as

$$R_{DC} \approx R_{FC} + R_{FX} \tag{3}$$

We return to the example in which the foreign-currency return on the USD-denominated bond was 10% and the USD appreciated by 5% against the EUR. In this example, the domestic-currency return for the Eurozone investor's holding in the USD-denominated bond was approximately equal to 10% + 5% = 15% (which is close to the exact value of 15.5%). We can combine the approximation of Equation 3 with the statistical rule that:

$$\sigma^2\left(\omega_x X + \omega_y Y\right) = \omega_x^2\sigma^2\left(X\right) + \omega_y^2\sigma^2\left(Y\right) + 2\omega_x\omega_y\sigma\left(X\right)\sigma\left(Y\right)\rho\left(X,Y\right) \tag{4}$$

where X and Y are random variables, ω are weights attached to X and Y, σ^2 is variance of a random variable, σ is the corresponding standard deviation, and ρ represents the correlation between two random variables. Applying this result to the domestic-currency return approximation of Equation 3 leads to:

$$\sigma^2(R_{DC}) \approx \sigma^2(R_{FC}) + \sigma^2(R_{FX}) + 2\sigma(R_{FC})\sigma(R_{FX})\rho(R_{FC},R_{FX}) \tag{5}$$

This equation is for the variance of the domestic-currency returns (R_{DC}), but risk is more typically defined in terms of standard deviation because mean and standard deviation are measured in the same units (percent, in this case). Hence, the total risk for domestic-currency returns—that is, $\sigma(R_{DC})$—is the square root of the results calculated in Equation 5.

Note as well that because Equation 5 is based on the addition of all three terms on the right side of the equal sign, exchange rate exposure will generally cause the variance of domestic-currency returns, $\sigma^2(R_{DC})$, to increase to more than that of the foreign-currency returns, $\sigma^2(R_{FC})$, considered on their own. That is, if there was no exchange rate risk, then it would be the case that $\sigma^2(R_{DC}) = \sigma^2(R_{FC})$. Using this as our base-case scenario, adding exchange rate risk exposure to the portfolio usually adds to domestic-currency return variance (the effect is indeterminate if exchange rate movements are negatively correlated with foreign asset returns).

These results on the variance of domestic-currency return can be generalized to a portfolio of foreign-currency assets. If we define the random variables X and Y in Equation 4 in terms of the domestic-currency return (R_{DC}) of two different foreign-currency investments, and the ω_i as portfolio weights that sum to one, then the result is the variance of the domestic-currency returns for the overall foreign asset portfolio:

$$\sigma^2\left(\omega_1 R_1 + \omega_2 R_2\right) \approx \omega_1^2\sigma^2\left(R_1\right) + \omega_2^2\sigma^2\left(R_2\right) + 2\omega_1\omega_2\sigma\left(R_1\right)\sigma\left(R_2\right)\rho\left(R_1,R_2\right) \tag{6}$$

where R_i is the domestic-currency return of the *i*-th foreign-currency asset. But as shown in Equation 3, the domestic-currency return of a foreign-currency asset (R_{DC}) is itself based on the sum of two random variables: R_{FC} and R_{FX}. This means that we would have to embed the variance expression shown in Equation 5 in *each* of the $\sigma^2(R_i)$ shown in Equation 6 to get the complete solution for the domestic-currency return variance of the overall portfolio. (We would also have to calculate the correlations between *all* of the R_i.) These requirements would lead to a very cumbersome mathematical expression for even a portfolio of only two foreign-currency assets; the expression would be far more complicated for a portfolio with many foreign currencies involved.

Thus, rather than attempt to give the complete mathematical formula for the variance of domestic-currency returns for a multi-currency portfolio, we will instead focus on the key intuition behind this expression. Namely, that the domestic-currency risk exposure of the overall portfolio—that is, $\sigma(R_{DC})$—will depend not only on the variances of *each* of the foreign-currency returns (R_{FC}) and exchange rate movements (R_{FX}) but also on how each of these *interacts* with the others. Generally speaking, negative correlations among these variables will help reduce the overall portfolio's risk through diversification effects.

Note as well that the overall portfolio's risk exposure will depend on the portfolio weights (ω_i) used. If short-selling is allowed in the portfolio, some of these ω_i can be negative as long as the total portfolio weights sum to one. So, for two foreign assets with a strong positive return correlation, short selling one can create considerable diversification benefits for the portfolio. (This approach is equivalent to trading movements in the price spread between these two assets.)

As before with the difference between realized and expected domestic-currency portfolio returns (R_{DC}), there is a difference between realized and expected domestic-currency portfolio risk, $\sigma(R_{DC})$. For Equation 6 to apply to the expected future volatility of the domestic-currency return of a multi-currency foreign asset portfolio, we would need to replace the observed, historical values of the variances and covariances in Equation 6 with their expected future values. This can be challenging, not only because it potentially involves a large number of variables but also because historical price patterns are not always a good guide to future price behavior. Variance and correlation measures are sensitive to the time period used to estimate them and can also vary over time. These variance and correlation measures can either drift randomly with time, or they can be subject to abrupt movements in times of market stress. It should also be clear that these observed, historical volatility and correlation measures need not be the same as the forward-looking *implied* volatility (and correlation) derived from option prices. Although sometimes various survey or consensus forecasts can be used, these too can be sensitive to sample size and composition and are not always available on a timely basis or with a consistent starting point. As with any forecast, they are also not necessarily an accurate guide to future developments; judgment must be used.

Hence, to calculate the expected future risk of the foreign asset portfolio, the portfolio manager would need a market opinion—however derived—on the variance of each of the foreign-currency asset returns (R_{FC}) over the investment horizon as well the variance of future exchange rate movements (R_{FX}) for each currency pair. The portfolio manager would also need a market opinion of how each of these future variables would interact with each other (i.e., their expected correlations). Historical price patterns can serve as a guide, and with computers and large databases, this modeling problem is daunting but not intractable. But the portfolio manager must always be mindful that historical risk patterns may not repeat going forward.

EXAMPLE 1

Portfolio Risk and Return Calculations

The following table shows current and future expected asset prices, measured in their domestic currencies, for both eurozone and Canadian assets (these can be considered "total return" indexes). The table also has the corresponding data for the CAD/EUR spot rate.

	Eurozone		Canada	
	Today	**Expected**	**Today**	**Expected**
Asset price	100.69	101.50	101.00	99.80
CAD/EUR	1.2925	1.3100		

1 What is the expected domestic-currency return for a eurozone investor holding the Canadian asset?

2 What is the expected domestic-currency return for a Canadian investor holding the eurozone asset?

3 From the perspective of the Canadian investor, assume that $\sigma(R_{FC})$ = 3% (the expected risk for the foreign-currency asset is 3%) and the $\sigma(R_{FX})$ = 2% (the expected risk of exchange rate movements is 2%). Furthermore, the expected correlation between movements in foreign-currency asset returns and movements in the CAD/EUR rate is +0.5. What is the expected risk of the domestic-currency return [$\sigma(R_{DC})$]?

Solution to 1:

For the eurozone investor, the R_{FC} = (99.80/101.00) – 1 = –1.19%. Note that, given we are considering the eurozone to be "domestic" for this investor and given the way the R_{FX} expression is defined, we will need to convert the CAD/EUR exchange rate quote so that the EUR is the *price* currency. This leads to R_{FX} = [(1/1.3100)/(1/1.2925)] – 1 = –1.34%. Hence, for the eurozone investor, R_{DC} = (1 – 1.19%)(1 – 1.34%) – 1 = –2.51%.

Solution to 2:

For the Canadian investor, the R_{FC} = (101.50/100.69) – 1 = +0.80%. Given that in the CAD/EUR quote the CAD is the price currency, for this investor the R_{FX} = (1.3100/1.2925) – 1 = +1.35%. Hence, for the Canadian investor the R_{DC} = (1 + 0.80%)(1 + 1.35%) – 1 = 2.16%.

Solution to 3:

Because this is a single foreign-currency asset we are considering (not a portfolio of such assets), we can use Equation 5:

$$\sigma^2(R_{DC}) \approx \sigma^2(R_{FC}) + \sigma^2(R_{FX}) + 2\sigma(R_{FC})\sigma(R_{FX})\rho(R_{FC},R_{FX})$$

Inserting the relevant data leads to

$$\sigma^2(R_{DC}) \approx (3\%)^2 + (2\%)^2 + 2(3\%)(2\%)(0.50) = 0.0019$$

Taking the square root of this leads to $\sigma(R_{DC}) \approx$ 4.36%. (Note that the units in these expressions are all in percent, so in this case 3% is equivalent to 0.03 for calculation purposes.)

CURRENCY MANAGEMENT: STRATEGIC DECISIONS

4

There are a variety of approaches to currency management, ranging from trying to avoid all currency risk in a portfolio to actively seeking foreign exchange risk in order to manage it and enhance portfolio returns.

There is no firm consensus—either among academics or practitioners—about the most effective way to manage currency risk. Some investment managers try to hedge all currency risk, some leave their portfolios unhedged, and others see currency risk as a potential source of incremental return to the portfolio and will actively trade foreign exchange. These widely varying management practices reflect a variety of factors including investment objectives, investment constraints, and beliefs about currency markets.

Concerning beliefs, one camp of thought holds that in the long run currency effects cancel out to zero as exchange rates revert to historical means or their fundamental values. Moreover, an efficient currency market is a zero-sum game (currency "A" cannot appreciate against currency "B" without currency "B" depreciating against currency "A"), so there should not be any long-run gains overall to speculating in currencies, especially after netting out management and transaction costs. Therefore, both currency hedging and actively trading currencies represent a cost to a portfolio with little prospect of consistently positive active returns.

At the other extreme, another camp of thought notes that currency movements can have a dramatic impact on short-run returns and return volatility and holds that there are pricing inefficiencies in currency markets. They note that much of the flow in currency markets is related to international trade or capital flows in which FX trading is being done on a need-to-do basis and these currency trades are just a spinoff of the other transactions. Moreover, some market participants are either not in the market on a purely profit-oriented basis (e.g., central banks, government agencies) or are believed to be "uninformed traders" (primarily retail accounts). Conversely, speculative capital seeking to arbitrage inefficiencies is finite. In short, marketplace diversity is believed to present the potential for "harvesting alpha" through active currency trading.

This ongoing debate does not make foreign-currency risk in portfolios go away; it still needs to managed, or at least, recognized. Ultimately, each portfolio manager or investment oversight committee will have to reach their own decisions about how to manage risk and whether to seek return enhancement through actively trading currency exposures.

Fortunately, there are a well-developed set of financial products and portfolio management techniques that help investors manage currency risk no matter what their individual objectives, views, and constraints. Indeed, the potential combinations of trading tools and strategies are almost infinite, and can shape currency exposures to custom-fit individual circumstance and market opinion. In this section, we explore various points on a spectrum reflecting currency exposure choices (a risk spectrum) and the guidance that portfolio managers use in making strategic decisions about where to locate their portfolios on this continuum. First, however, the implication of investment objectives and constraints as set forth in the investment policy statement must be recognized.

4.1 The Investment Policy Statement

The Investment Policy Statement (IPS) mandates the degree of discretionary currency management that will be allowed in the portfolio, how it will be benchmarked, and the limits on the type of trading polices and tools (e.g., such as leverage) than can be used.

The starting point for organizing the investment plan for any portfolio is the IPS, which is a statement that outlines the broad objectives and constraints of the beneficial owners of the assets. Most IPS specify many of the following points:

- the general objectives of the investment portfolio;
- the risk tolerance of the portfolio and its capacity for bearing risk;
- the time horizon over which the portfolio is to be invested;
- the ongoing income/liquidity needs (if any) of the portfolio; and
- the benchmark against which the portfolio will measure overall investment returns.

The IPS sets the guiding parameters within which more specific portfolio management policies are set, including the target asset mix; whether and to what extent leverage, short positions, and derivatives can be used; and how actively the portfolio will be allowed to trade its various risk exposures.

For most portfolios, currency management can be considered a sub-set of these more specific portfolio management policies within the IPS. The currency risk management policy will usually address such issues as the

- target proportion of currency exposure to be passively hedged;
- latitude for active currency management around this target;
- frequency of hedge rebalancing;
- currency hedge performance benchmark to be used; and
- hedging tools permitted (types of forward and option contracts, etc.).

Currency management should be conducted within these IPS-mandated parameters.

4.2 The Portfolio Optimization Problem

Having described the IPS as the guiding framework for currency management, we now examine the strategic choices that have to be made in deciding the benchmark currency exposures for the portfolio, and the degree of discretion that will be allowed around this benchmark. This process starts with a decision on the optimal foreign-currency asset and FX exposures.

Optimization of a multi-currency portfolio of foreign assets involves selecting portfolio weights that locate the portfolio on the efficient frontier of the trade-off between risk and expected return defined in terms of the investor's domestic currency. As a simplification of this process, consider the portfolio manager examining the expected return and risk of the multi-currency portfolio of foreign assets by using different combinations of portfolio weights (ω_i) that were shown in Equations 2 and 6, respectively, which are repeated here:

$$R_{DC} = \sum_{i=1}^{n} \omega_i \left(1 + R_{FC,i}\right)\left(1 + R_{FX,i}\right) - 1$$

$$\sigma^2\left(\omega_1 R_1 + \omega_2 R_2\right) \approx \omega_1^2 \sigma^2\left(R_1\right) + \omega_2^2 \sigma^2\left(R_2\right) + 2\omega_1 \sigma\left(R_1\right)\omega_2 \sigma\left(R_2\right)\rho\left(R_1, R_2\right)$$

Recall that the R_i in the equation for variance are the R_{DC} for each of the foreign-currency assets. Likewise, recall that the R_{FX} term is defined such that the investor's "domestic" currency is the price currency in the P/B exchange rate quote. In other words, this calculation may require using the algebraic reciprocal of the standard market quote convention. These two equations together show the domestic-currency return and risk for a multi-currency portfolio of foreign assets.

When deciding on an optimal investment position, these equations would be based on the *expected* returns and risks for each of the foreign-currency assets; and hence, including the *expected* returns and risks for each of the foreign-currency exposures. As we have seen earlier, the number of market parameters for which the portfolio manager would need to have a market opinion grows geometrically with the complexity (number of foreign-currency exposures) in the portfolio. That is, to calculate the expected efficient frontier, the portfolio manager must have a market opinion for *each* of the $R_{FC,i}$, $R_{FX,i}$, $\sigma(R_{FC,i})$, $\sigma(R_{FX,i})$, and $\rho(R_{FC,i}\ R_{FX,i})$, as well as for each of the $\rho(R_{FC,i}\ R_{FC,j})$ and $\rho(R_{FX,i}\ R_{FX,j})$. This would be a daunting task for even the most well-informed portfolio manager.

In a perfect world with complete (and costless) information, it would likely be optimal to *jointly* optimize all of the portfolio's exposures—over all currencies and all foreign-currency assets—simultaneously. In the real world, however, this can be a much more difficult task. Confronted with these difficulties, many portfolio managers handle asset allocation with currency risk as a two-step process: (1) portfolio optimization over fully hedged returns; and (2) selection of active currency exposure, if any. Derivative strategies can allow the various risk exposures in a portfolio to be "unbundled" from each other and managed separately. The same applies for currency risks. Because the use of derivatives allows the price risk ($R_{FC,i}$) and exchange rate risk ($R_{FX,j}$) of foreign-currency assets to be unbundled and managed separately, a starting point for the selection process of portfolio weights would be to assume a complete currency hedge. That is, the portfolio manager will choose the exposures to the foreign-currency assets first, and then decide on the appropriate currency exposures afterward (i.e., decide whether to relax the full currency hedge). These decisions are made to simplify the portfolio construction process.

If the currency exposures of foreign assets could be perfectly and costlessly hedged, the hedge would completely neutralize the effect of currency movements on the portfolio's domestic-currency return (R_{DC}).[10] In Equation 2, this would set R_{FX} = 0, meaning that the domestic-currency return is then equal to the foreign-currency return ($R_{DC} = R_{FC}$). In Equation 5, this would set $\sigma^2(R_{DC}) = \sigma^2(R_{FC})$, meaning that the domestic-currency return risk is equal to the foreign-currency return risk.

Removing the currency effects leads to a simpler, two-step process for portfolio optimization. First the portfolio manager could pick the set of portfolio weights (ω_i) for the foreign-currency assets that optimize the expected foreign-currency asset risk–return trade-off (assuming there is no currency risk). Then the portfolio manager could choose the desired currency exposures for the portfolio and decide whether and by how far to relax the constraint to a full currency hedge for each currency pair.

4.3 Choice of Currency Exposures

A natural starting point for the strategic decisions is the "currency-neutral" portfolio resulting from the two-step process described earlier. The question then becomes, How far along the risk spectrum between being fully hedged and actively trading currencies should the portfolio be positioned?

4.3.1 *Diversification Considerations*

The time horizon of the IPS is important. Many investment practitioners believe that in the long run, adding unhedged foreign-currency exposure to a portfolio does not affect expected long-run portfolio returns; hence in the long run, it would not matter if the portfolio was hedged. (Indeed, portfolio management costs would be reduced without a hedging process.) This belief is based on the view that in the long run,

10 A "costless" hedge in this sense would not only mean zero transaction costs, but also no "roll yield."

currencies “mean revert” to either some fair value equilibrium level or a historical average; that is, that the *expected* $\%\Delta S = 0$ for a sufficiently long time period. This view typically draws on the expectation that purchasing power parity (PPP) and the other international parity conditions that link movements in exchange rates, interest rates, and inflation rates will eventually hold over the long run.

Supporting this view, some studies argue that in the long-run currencies will in fact mean revert, and hence that currency risk is lower in the long run than in the short run (an early example is Froot 1993). Although much depends on how long run is defined, an investor (IPS) with a very long investment horizon and few immediate liquidity needs—which could potentially require the liquidation of foreign-currency assets at disadvantageous exchange rates—might choose to forgo currency hedging and its associated costs. Logically, this would require a portfolio benchmark index that is also unhedged against currency risk.

Although the international parity conditions may hold in the long run, it can be a *very* long time—possibly decades. Indeed, currencies can continue to drift away from the fair value mean reversion level for much longer than the time period used to judge portfolio performance. Such time periods are also typically longer than the patience of the portfolio manager’s oversight committee when portfolio performance is lagging the benchmark. If this very long-run view perspective is not the case, then the IPS will likely impose some form of currency hedging.

Diversification considerations will also depend on the *asset composition* of the foreign-currency asset portfolio. The reason is because the foreign-currency asset returns (R_{FC}) of different asset classes have different correlation patterns with foreign-currency returns (R_{FX}). If there is a negative correlation between these two sets of returns, having at least some currency exposure may help portfolio diversification and moderate the domestic-currency return risk, $\sigma(R_{DC})$. (Refer to Equation 5 in Section 3.3.)

It is often asserted that the correlation between foreign-currency returns and foreign-currency asset returns tends to be greater for fixed-income portfolios than for equity portfolios. This assertion makes intuitive sense: both bonds and currencies react strongly to movements in interest rates, whereas equities respond more to expected earnings. As a result, the implication is that currency exposures provide little diversification benefit to fixed-income portfolios and that the currency risk should be hedged. In contrast, a better argument can be made for carrying currency exposures in global equity portfolios.

To some degree, various studies have corroborated this relative advantage to currency hedging for fixed income portfolios. But the evidence seems somewhat mixed and depends on which markets are involved. One study found that the hedging advantage for fixed-income portfolios is not always large or consistent (Darnell 2004). Other studies (Campbell 2010; Martini 2010) found that the optimal hedge ratio for foreign-currency equity portfolios depended critically on the investor’s domestic currency. (Recall that the hedge ratio is defined as the ratio of the nominal value of the hedge to the market value of the underlying.) For some currencies, there was no risk-reduction advantage to hedging foreign equities (the optimal hedge ratio was close to 0%), whereas for other currencies, the optimal hedge ratio for foreign equities was close to 100%.

Other studies indicate that the optimal hedge ratio also seems to depend on *market conditions* and longer-term trends in currency pairs. For example, Campbell, Serfaty-de Medeiros, and Viceira (2007) found that there were no diversification benefits from currency exposures in foreign-currency bond portfolios, and hence to minimize the risk to domestic-currency returns these positions should be fully hedged. The authors also found, however, that during the time of their study (their data spanned 1975 to 2005), the US dollar seemed to be an exception in terms of its correlations with foreign-currency asset returns. Their study found that the US dollar tended to appreciate against foreign currencies when global bond prices fell (for example, in times of global financial stress there is a tendency for investors to shift investments into the perceived safety of reserve currencies). This finding would suggest that keeping some exposure to the US dollar in a global bond portfolio would be beneficial. For non-US investors, this would mean under-hedging the currency exposure to the USD (i.e., a hedge ratio less than 100%), whereas for US investors it would mean over-hedging their foreign-currency exposures back into the USD.[11]

Given this diversity of opinions and empirical findings, it is not surprising to see actual hedge ratios vary widely in practice among different investors. Nonetheless, it is still more likely to see currency hedging for fixed-income portfolios rather than equity portfolios, although actual hedge ratios will often vary between individual managers.

4.3.2 *Cost Considerations*

The costs of currency hedging also guide the strategic positioning of the portfolio. Currency hedges are not a "free good" and they come with a variety of expenses that must be borne by the overall portfolio. Optimal hedging decisions will need to balance the benefits of hedging against these costs.

Hedging costs come mainly in two forms: trading costs and opportunity costs. The most immediate costs of hedging involve trading expenses, and these come in several forms:

- Trading involves dealing on the bid–offer spread offered by banks. Their profit margin is based on these spreads, and the more the client trades and "pays away the spread," the more profit is generated by the dealer. Maintaining a 100% hedge and rebalancing frequently with every minor change in market conditions would be expensive. Although the bid–offer spreads on many FX-related products (especially the spot exchange rate) are quite narrow, "churning" the hedge portfolio would progressively add to hedging costs and detract from the hedge's benefits.
- Some hedges involve currency options; a long position in currency options requires the payment of up-front premiums. If the options expire out of the money (OTM), this cost is unrecoverable.
- Although forward contracts do not require the payment of up-front premiums, they do eventually mature and have to be "rolled" forward with an FX swap transaction to maintain the hedge. Rolling hedges will typically generate cash inflows or outflows. These cash flows will have to be monitored, and as necessary, cash will have to be raised to settle hedging transactions. In other words, even though the currency hedge may *reduce* the volatility of the domestic mark-to-market value of the foreign-currency asset portfolio, it will typically *increase* the volatility in the organization's cash accounts. Managing these cash flow costs can accumulate to become a significant portion of the portfolio's value, and they become more expensive (for cash outflows) the higher interest rates go.
- One of the most important trading costs is the need to maintain an administrative infrastructure for trading. Front-, middle-, and back-office operations will have to be set up, staffed with trained personnel, and provided with specialized technology systems. Settlement of foreign exchange transactions in a variety of

11 Note that since the 2008 financial crisis followed by the European sovereign debt crisis, some currencies—the USD, JPY, and CHF in particular—seem to act as a safe haven and appreciate in times of market stress. Keeping some of these currency exposures in the portfolio—having hedge ratios that are not set at 100%—can help hedge losses on riskier assets, especially for foreign-currency equity portfolios (which are more risk exposed than bond portfolios). We will return to this topic in a later section when we discuss minimum-variance hedges and macro hedges.

currencies means having to maintain cash accounts in these currencies to make and receive these foreign-currency payments. Together all of these various overhead costs can form a significant portion of the overall costs of currency trading.

A second form of costs associated with hedging are the opportunity cost of the hedge. To be 100% hedged is to forgo any possibility of favorable currency rate moves. If skillfully handled, accepting and managing currency risk—or any financial risk—can potentially add value to the portfolio, even net of management fees. (We discuss the methods by which this might be done in Section 5.)

These opportunity costs lead to another motivation for having a strategic hedge ratio of less than 100%: regret minimization. Although it is not possible to accurately predict foreign exchange movements in advance, it is certainly possible to judge after the fact the results of the decision to hedge or not. Missing out on an advantageous currency movement because of a currency hedge can cause *ex post* regret in the portfolio manager or client; so too can having a foreign-currency loss if the foreign-currency asset position was unhedged. Confronted with this *ex ante* dilemma of whether to hedge, many portfolio managers decide simply to "split the difference" and have a 50% hedge ratio (or some other rule-of-thumb number). Both survey evidence and anecdotal evidence show that there is a wide variety of hedge ratios actually used in practice by managers, and that these variations cannot be explained by more "fundamental" factors alone. Instead, many managers appear to incorporate some degree of regret minimization into hedging decisions (for example, see Michenaud and Solnik 2008).

All of these various hedging expenses—both trading and opportunity costs—will need to be managed. Hedging is a form of insurance against risk, and in purchasing any form of insurance the buyer matches their needs and budgets with the policy selected. For example, although it may be possible to buy an insurance policy with full, unlimited coverage, a zero deductible, and no co-pay arrangements, such a policy would likely be prohibitively expensive. Most insurance buyers decide that it is not necessary to insure against every outcome, no matter how minor. Some minor risks can be accepted and "self-insured" through the deductible; some major risks may be considered so unlikely that they are not seen as worth paying the extra premium. (For example, most ordinary people would likely not consider buying insurance against being kidnapped.)

These same principles apply to currency hedging. The portfolio manager (and IPS) would likely not try to hedge every minor, daily change in exchange rates or asset values, but only the larger adverse movements that can materially affect the overall domestic-currency returns (R_{DC}) of the foreign-currency asset portfolio. The portfolio manager will need to balance the benefits and costs of hedging in determining both strategic positioning of the portfolio as well as any latitude for active currency management. However, around whatever strategic positioning decision taken by the IPS in terms of the benchmark level of currency exposure, hedging cost considerations alone will often dictate a *range* of permissible exposures instead of a single point. (This discretionary range is similar to the deductible in an insurance policy.)

4.4 Locating the Portfolio Along the Currency Risk Spectrum

The strategic decisions encoded in the IPS with regard to the trade-off between the benefits and costs of hedging, as well as the potential for incremental return to the portfolio from active currency management, are the foundation for determining specific currency management strategies. These strategies are arrayed along a spectrum from very risk-averse passive hedging, to actively seeking out currency risk in order to manage it for profit. We examine each in turn.

4.4.1 *Passive Hedging*

In this approach, the goal is to keep the portfolio's currency exposures close, if not equal to, those of a benchmark portfolio used to evaluate performance. Note that the benchmark portfolio often has no foreign exchange exposure, particularly for fixed-income assets; the benchmark index is a "local currency" index based only on the foreign-currency asset return (R_{FC}). However, benchmark indexes that have some foreign exchange risk are also possible.

Passive hedging is a rules-based approach that removes almost all discretion from the portfolio manager, regardless of the manager's market opinion on future movements in exchange rates or other financial prices. In this case, the manager's job is to keep portfolio exposures as close to "neutral" as possible and to minimize tracking errors against the benchmark portfolio's performance. This approach reflects the belief that currency exposures that differ from the benchmark portfolio inject risk (return volatility) into the portfolio without any sufficiently compensatory return. Active currency management—taking positional views on future exchange rate movements—is viewed as being incapable of consistently adding incremental return to the portfolio.

But the hedge ratio has a tendency to "drift" with changes in market conditions, and even passive hedges need periodic rebalancing to realign them with investment objectives. Often the management guidance given to the portfolio manager will specify the rebalancing period—for example, monthly. There may also be allowance for intra-period rebalancing if there have been large exchange rate movements.

4.4.2 *Discretionary Hedging*

This approach is similar to passive hedging in that there is a "neutral" benchmark portfolio against which actual portfolio performance will be measured. However, in contrast to a strictly rules-based approach, the portfolio manager now has some limited discretion on how far to allow actual portfolio risk exposures to vary from the neutral position. Usually this discretion is defined in terms of percentage of foreign-currency market value (the portfolio's currency exposures are allowed to vary plus or minus x% from the benchmark). For example, a eurozone-domiciled investor may have a US Treasury bond portfolio with a mandate to keep the hedge ratio within 95% to 105%. Assuming no change in the foreign-currency return (R_{FC}), but allowing exchange rates (R_{FX}) to vary, this means the portfolio can tolerate exchange rate movements between the EUR and USD of up to 5% before the exchange rate exposures in the portfolio are considered excessive. The manager is allowed to manage currency exposures within these limits without being considered in violation of the IPS.

This discretion allows the portfolio manager at least some limited ability to express directional opinions about future currency movements—to accept risk in an attempt to earn reward—in order to add value to the portfolio performance. Of course, the portfolio manager's actual performance will be compared with that of the benchmark portfolio.

4.4.3 *Active Currency Management*

Further along the spectrum between extreme risk aversion and purely speculative trading is active currency management. In principle, this approach is really just an extension of discretionary hedging: the portfolio manager is allowed to express directional opinions on exchange rates, but is nonetheless kept within mandated risk limits. The performance of the manager—the choices of risk exposures assumed—is benchmarked against a "neutral" portfolio. But for all forms of active management (i.e., having the discretion to express directional market views), there is no allowance for unlimited speculation; there are risk management systems in place for even the most speculative investment vehicles, such as hedge funds. These controls are designed to prevent traders from taking unusually large currency exposures and risking the solvency of the firm or fund.

In many cases, the difference between discretionary hedging and active currency management is one of emphasis more than degree. The primary duty of the discretionary hedger is to protect the portfolio from currency risk. As a secondary goal, within limited bounds, there is some scope for directional opinion in an attempt to enhance overall portfolio returns. If the manager lacks any firm market conviction, the natural neutral position for the discretionary hedger is to be flat—that is, to have no meaningful currency exposures. In contrast, the active currency manager is supposed to take currency risks and manage them for profit. The primary goal is to add alpha to the portfolio through successful trading. Leaving actual portfolio exposures near zero for extended periods is typically not a viable option.

4.4.4 *Currency Overlay*

Active management of currency exposures can extend beyond limited managerial discretion within hedging boundaries. Sometimes accepting and managing currency risk for profit can be considered a portfolio objective. Active currency management is often associated with what are called **currency overlay programs**, although this term is used differently by different sources.

- In the most limited sense of the term, currency overlay simply means that the portfolio manager has outsourced managing currency exposures to a firm specializing in FX management.[12] This could imply something as limited as merely having the external party implement a fully passive approach to currency hedges. If dealing with FX markets and managing currency hedges is beyond the professional competence of the investment manager, whose focus is on managing foreign equities or some other asset class, then hiring such external professional help is an option.
- A broader view of currency overlay allows the externally hired currency overlay manager to take directional views on future currency movements (again, with the caveat that these be kept within predefined bounds). Sometimes a distinction is made between currency overlay and "foreign exchange as an asset class." In this classification, currency overlay is limited to the currency exposures already in the foreign asset portfolio. For example, if a eurozone-domiciled investor has GBP- and CHF-denominated assets, currency overlay risks are allowed only for these currencies.
- In contrast, the concept of foreign exchange as an asset class does not restrict the currency overlay manager, who is free to take FX exposures in any currency pair where there is value-added to be harvested, regardless of the underlying portfolio. In this sense, the currency overlay manager is very similar to an FX-based hedge fund. To implement this form of active currency management, the currency overlay manager would have a *joint* opinion on a range of currencies, and have market views not only on the expected movements in the spot rates but also the likelihood of these movements (the variance of the expected future spot rate distribution) as well as the expected correlation between future spot rate movements. Basically, the entire portfolio of currencies is actively managed and optimized over all of the expected returns, risks, and correlations among all of the currencies in the portfolio.

12 Typically currency overlay programs involve external managers. However, some large, sophisticated institutional accounts may have in-house currency overlay programs managed by a separate group of specialists within the firm.

We will focus on this latter form of currency overlay in this reading: active currency management conducted by external, FX-specialized sub-advisors to the portfolio.

It is quite possible to have the foreign-currency asset portfolio fully hedged (or allow some discretionary hedging internally) but then also to add an external currency overlay manager to the portfolio. This approach separates the hedging and alpha function mandates of the portfolio. Different organizations have different areas of expertise; it often makes sense to allocate managing the hedge (currency "beta") and managing the active FX exposures (currency "alpha") to those individuals with a comparative advantage in that function.

Adding this form of currency overlay to the portfolio (FX as an asset class) is similar in principle to adding any type of alternative asset class, such as private equity funds or farmland. In each case, the goal is the search for alpha. But to be most effective in adding value to the portfolio, the currency overlay program should add incremental returns (alpha) and/or greater diversification opportunities to improve the portfolio's risk–return profile. To do this, the currency alpha mandate should have minimum correlation with both the major asset classes and the other alpha sources in the portfolio.

Once this FX as an asset class approach is taken, it is not necessary to restrict the portfolio to a single overlay manager any more than it is necessary to restrict the portfolio to a single private equity fund. Different overlay managers follow different strategies (these are described in more detail in Section 5). Within the overall portfolio allocation to "currency as an alternative asset class", it may be beneficial to diversify across a range of active management styles, either by engaging several currency overlay managers with different styles or by applying a fund-of-funds approach, in which the hiring and management of individual currency overlay managers is delegated to a specialized external investment vehicle.

Whether managed internally or externally (via a fund of funds) it will be necessary to monitor, or benchmark, the performance of the currency overlay manager: Do they generate the returns expected from their stated trading strategy? Many major investment banks as well as specialized market-information firms provide a wide range of proprietary indexes that track the performance of the investible universe of currency overlay managers; sometimes they also offer sub-indexes that focus on specific trading strategies (for example, currency positioning based on macroeconomic fundamentals). However, the methodologies used to calculate these various indexes vary between suppliers. In addition, different indexes show different aspects of active currency management. Given these differences between indexes, there is no simple answer for which index is most suitable as a benchmark; much depends on the specifics of the active currency strategy.

EXAMPLE 2

Currency Overlay

Windhoek Capital Management is a South Africa-based investment manager that runs the Conservative Value Fund, which has a mandate to avoid all currency risk in the portfolio. The firm is considering engaging a currency overlay manager to help with managing the foreign exchange exposures of this investment vehicle. Windhoek does not consider itself to have the in-house expertise to manage FX risk.

Brixworth & St. Ives Asset Management is a UK-based investment manager, and runs the Aggressive Growth Fund. This fund is heavily weighted toward emerging market equities, but also has a mandate to seek out inefficiencies in the global foreign exchange market and exploit these for profit. Although Brixworth & St. Ives manages the currency hedges for all of its investment funds in-house, it is also considering engaging a currency overlay manager.

1 Using a currency overlay manager for the Conservative Value Fund is *most likely* to involve:
   A joining the alpha and hedging mandates.
   B a more active approach to managing currency risks.
   C using this manager to passively hedge their foreign exchange exposures.

2 Using a currency overlay manager for the Aggressive Growth Fund is *most likely* to involve:
   A separating the alpha and hedging mandates.
   B a less discretionary approach to managing currency hedges.
   C an IPS that limits active management to emerging market currencies.

3 Brixworth & St. Ives is *more likely* to engage multiple currency overlay managers if:
   A their returns are correlated with asset returns in the fund.
   B the currency managers' returns are correlated with each other.
   C the currency managers' use different active management strategies.

Solution to 1:

C is correct. The Conservative Value Fund wants to avoid all currency exposures in the portfolio and Windhoek believes that it lacks the currency management expertise to do this.

Solution to 2:

A is correct. Brixworth & St. Ives already does the FX hedging in house, so a currency overlay is more likely to be a pure alpha mandate. This should not change the way that Brixworth & St. Ives manages its hedges, and the fund's mandate to seek out inefficiencies in the global FX market is unlikely to lead to a restriction to actively manage only emerging market currencies.

Solution to 3:

C is correct. Different active management strategies may lead to a more diversified source of alpha generation, and hence reduced portfolio risk. Choices A and B are incorrect because a higher correlation with foreign-currency assets in the portfolio or among overlay manager returns is likely to lead to less diversification.

4.5 Formulating a Client-Appropriate Currency Management Program

We now try to bring all of these previous considerations together in describing how to formulate an appropriate currency management program given client objectives and constraints, as well as overall financial market conditions. Generally speaking, the *strategic* currency positioning of the portfolio, as encoded in the IPS, should be biased toward a more-fully hedged currency management program the more

- short term the investment objectives of the portfolio;
- risk averse the beneficial owners of the portfolio are (and impervious to *ex post* regret over missed opportunities);

- immediate the income and/or liquidity needs of the portfolio;
- fixed-income assets are held in a foreign-currency portfolio;
- cheaply a hedging program can be implemented;
- volatile (i.e., risky) financial markets are;[13] and
- skeptical the beneficial owners and/or management oversight committee are of the expected benefits of active currency management.

The relaxation of any of these conditions creates latitude to allow a more proactive currency risk posture in the portfolio, either through wider tolerance bands for discretionary hedging, or by introducing foreign currencies as a separate asset class (using currency overlay programs as an alternative asset class in the overall portfolio). In the latter case, the more currency overlay is expected to generate alpha that is uncorrelated with other asset or alpha-generation programs in the portfolio, the more it is likely to be allowed in terms of strategic portfolio positioning.

INVESTMENT POLICY STATEMENT

Kailua Kona Advisors runs a Hawaii-based hedge fund that focuses on developed market equities located outside of North America. Its investor base consists of local high-net-worth individuals who are all considered to have a long investment horizon, a high tolerance for risk, and no immediate income needs. In its prospectus to investors, Kailua Kona indicates that it actively manages both the fund's equity and foreign-currency exposures, and that the fund uses leverage through the use of loans as well as short-selling.

Exhibit 4 presents the hedge fund's currency management policy included in the IPS for this hedge fund.

Exhibit 4 Hedge Fund Currency Management Policy: An Example

Overall Portfolio Benchmark:	MSCI EAFE Index (local currency)
Currency Exposure Ranges:	Foreign-currency exposures, based on the USD market value of the equities actually held by the fund at the beginning of each month, will be hedged back into USD within the following tolerance ranges of plus or minus: ■ EUR: 20% ■ GBP: 15% ■ JPY: 10% ■ CHF: 10% ■ AUD: 10% ■ SEK: 10% Other currency exposures shall be left unhedged.
Rebalancing:	The currency hedges will be rebalanced at least monthly, to reflect changes in the USD-denominated market value of portfolio equity holdings.

(continued)

13 As we will see, this also increases hedging costs when currency options are used.

Exhibit 4 (Continued)

Hedging Instruments:	■ Forward contracts up to 12 months maturity; ■ European put and call options can be bought or written, for maturities up to 12 months; and ■ Exotic options of up to 12 months maturity can be bought or sold.
Reporting:	Management will present quarterly reports to the board detailing net foreign-currency exposures and speculative trading results. Speculative trading results will be benchmarked against a 100% hedged currency exposure.

With this policy, Kailua Kona Advisors is indicating that it is willing to accept foreign-currency exposures within the portfolio but that these exposures must be kept within pre-defined limits. For example, suppose that at the beginning of the month the portfolio held EUR10 million of EUR-denominated assets. Also suppose that this EUR10 million exposure, combined with all the other foreign-currency exposures in the portfolio, matches Kailua Kona Advisors' desired portfolio weights by currency (as a US-based fund, these desired percentage portfolio allocations across all currencies will be based in USD).

The currency-hedging guidelines indicate that the hedge (for example, using a short position in a USD/EUR forward contract) should be between EUR8 million and EUR12 million, giving some discretion to the portfolio manager on the size of the net exposure to the EUR. At the beginning of the next month, the USD values of the foreign assets in the portfolio are measured again, and the process repeats. If there has been either a large move in the foreign-currency value of the EUR-denominated assets and/or a large move in the USD/EUR exchange rate, it is possible that Kailua Kona Advisors' portfolio exposure to EUR-denominated assets will be too far away from the desired percentage allocation.[14] Kailua Kona Advisors will then need to either buy or sell EUR-denominated assets. If movements in the EUR-denominated value of the assets or in the USD/EUR exchange rate are large enough, this asset rebalancing may have to be done before month's end. Either way, once the asset rebalancing is done, it establishes the new EUR-denominated asset value on which the currency hedge will be based (i.e., plus or minus 20% of this new EUR amount).

If the portfolio is not 100% hedged—for example, continuing the Kailua Kona illustration, if the portfolio manager only hedges EUR9 million of the exposure and has a residual exposure of being long EUR1 million—the success or failure of the manager's tactical decision will be compared with a "neutral" benchmark. In this case, the comparison would be against the performance of a 100% fully hedged portfolio—that is, with a EUR10 million hedge.

14 The overall portfolio percentage allocations by currency will also depend on the price moves of all *other* foreign-currency assets and exchange rates as well, but we will simplify our example by ignoring this nuance.

CURRENCY MANAGEMENT: TACTICAL DECISIONS

5

The previous section discussed the *strategic* decisions made by the IPS on locating the currency management practices of the portfolio along a risk spectrum ranging from a very conservative approach to currency risk to very active currency management. In this section, we consider the case in which the IPS has given the portfolio manager (or currency overlay manager) at least some limited discretion for actively managing currency risk within these mandated strategic bounds. This then leads to *tactical* decisions: which FX exposures to accept and manage within these discretionary limits. In other words, tactical decisions involve active currency management.

A market view is a prerequisite to any form of active management. At the heart of the trading decision in FX (and other) markets, lies a view on future market prices and conditions. This market opinion guides all decisions with respect to currency risk exposures, including whether currency hedges should be implemented and, if so, how they should be managed.

In what follows, we will explore some of the methods used to form directional views about the FX market. However, a word of caution that cannot be emphasized enough: *There is no simple formula, model, or approach that will allow market participants to precisely forecast exchange rates (or any other financial prices) or to be able to be confident that any trading decision will be profitable.*

5.1 Active Currency Management Based on Economic Fundamentals

In this section, we lay out a framework for developing a view about future exchange rate movements based on underlying fundamentals. In contrast to other methods for developing a market view (which are discussed in subsequent sections), at the heart of this approach is the assumption that, in free markets, exchange rates are determined by logical economic relationships and that these relationships can be modeled.

The simple economic model described in this section is based on the assumption that in the long run, the real exchange rate will converge to its "fair value," but short- to medium-term factors will shape the convergence path to this equilibrium.[15]

Recall that the real exchange rate reflects the ratio of the real purchasing power between two countries; that is, the once nominal purchasing power in each country is adjusted by its respective price level as well as the spot exchange rate between the two countries. The long-run equilibrium level for the real exchange rate is determined by purchasing power parity or some other model of an exchange rate's fair value, and serves as the anchor for longer-term movements in exchange rates.

Over shorter time frames, movements in real exchange rates will also reflect movements in the real interest rate differential between countries. Recall that the real interest rate (r) is the nominal interest rate adjusted by the expected inflation rate, or $r = i - \pi^{\varepsilon}$, where i is the nominal interest rate and π^{ε} is the expected inflation rate over the same term as the nominal and real interest rates. Movements in risk premiums will also affect exchange rate movements over shorter-term horizons. The riskier a country's assets are perceived to be by investors, the more likely they are to move their investments out of that country, thereby depressing the exchange rate. Finally,

15 This model was derived and explained by Rosenberg and Barker (2017), which forms part of the Level II curriculum. Because of this prior coverage, we provide only an abbreviated summary of this simple economic model here.

our model also recognizes that there are two currencies involved in an exchange rate quote (the price and base currencies) and hence movements in exchange rates will reflect movements in the *differentials* between these various factors.

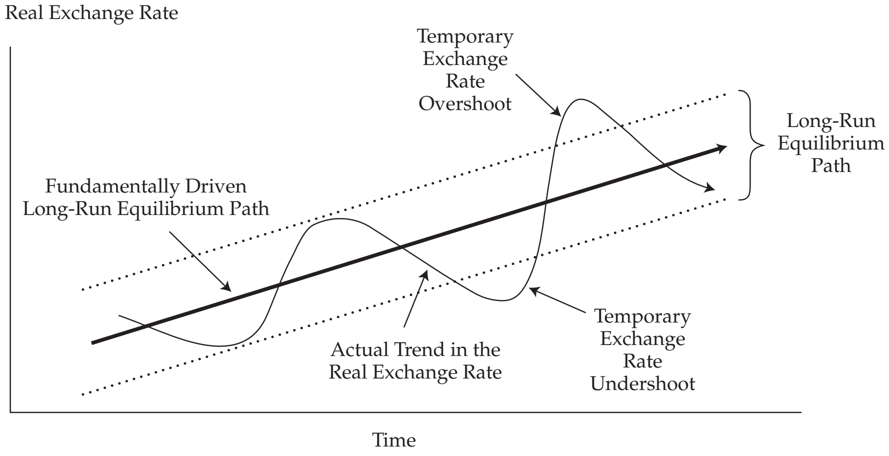
As a result, our model indicates that all else equal, the base currency's real exchange rate should appreciate if there is an upward movement in

- its long-run equilibrium real exchange rate;
- either its real or nominal interest rates, which should attract foreign capital;
- expected foreign inflation, which should cause the foreign currency to depreciate; and
- the foreign risk premium, which should make foreign assets less attractive compared with the base currency nation's domestic assets.

The real exchange rate should also increase if it is currently below its long-term equilibrium value. All of this makes intuitive sense.

In summary, our model is a mix of long-term, medium-term, and short-term factors. The long-run equilibrium real exchange rate is the anchor for exchange rates and the point of long-run convergence for exchange rate movements. Movements in the short- to medium-term factors (nominal interest rates, expected inflation) affect the timing and path of convergence to this long-run equilibrium. A stylized depiction of the price dynamics generated by this interaction between short-, medium-, and longer-term pricing factors is shown in Exhibit 5.

Exhibit 5 Interaction of Long-term and Short-term Factors in Exchange Rates



Source: Based on Rosenberg (2002), page 32.

It needs to be stressed that it can be very demanding to model how each of these separate effects—nominal interest rate, expected inflation, and risk premium differentials—change over time and affect exchange rates. It can also be challenging to model movements in the long-term equilibrium real exchange rate. A broad variety of factors, such as fiscal and monetary policy, will affect all of these variables in our simple economic model.[16]

16 A broader discussion of exchange rate economics can be found in Rosenberg and Barker (2017).

5.2 Active Currency Management Based on Technical Analysis

Another approach to forming a market view is based on technical analysis. This approach is based on quite different assumptions compared with modeling based on economic fundamentals. Whereas classical exchange rate economics tends to view market participants as rational, markets as efficient, and exchange rates as driven by underlying economic factors, technical analysis ignores economic analysis. Instead, technical analysis is based on three broad themes.[17]

First, market technicians believe that in a liquid, freely traded market the historical price data can be helpful in projecting future price movements.[18] The reason is because many traders have already used any useful data external to the market to generate their trading positions, so this information is already reflected in current prices. Therefore, it is not necessary to look outside of the market to form an opinion on future price movements. This means it is not necessary to examine interest rates, inflation rates, or risk premium differentials (the factors in our fundamentally based model) because exchange rates already incorporate these factors.

Second, market technicians believe that historical patterns in the price data have a tendency to repeat, and that this repetition provides profitable trade opportunities. These price patterns repeat because market prices reflect human behavior and human beings have a tendency to react in similar ways to similar situations, even if this repetitive behavior is not always fully rational. For example, when confronted with an upward price trend, many market participants eventually come to believe that it will extrapolate (an attitude of "irrational exuberance" or "this time it is different"). When the trend eventually breaks, a panicked position exit can cause a sharp overshoot of fair value to the downside. Broadly speaking, technical analysis can be seen as the study of market psychology and how market participant emotions—primarily greed and fear—can be read from the price data and used to predict future price moves.

Third, technical analysis does not attempt to determine where market prices *should* trade (fair value, as in fundamental analysis) but where they *will* trade. Because these price patterns reflect trader emotions, they need not reflect—at least immediately—any cool, rational assessment of the underlying economic or fundamental situation. Although market prices may eventually converge to fair value in the long run, the long run can be a very long time indeed. In the meanwhile, there are shorter-term trading opportunities available in trading the technical patterns in the price data.

Combined, these three principles of technical analysis define a discipline dedicated to identifying patterns in the historical price data, especially as it relates to identifying market trends and market turning points. (Technical analysis is less useful in a trendless market.) Technical analysis tries to identify when markets have become **overbought** or **oversold**, meaning that they have trended too far in one direction and are vulnerable to a trend reversal, or correction. Technical analysis also tries to identify what are called **support levels** and **resistance levels**, either within ongoing price trends or at their extremities (i.e., turning points). These support and resistance levels are price points on dealers' order boards where one would except to see clustering of bids and offers, respectively. At these exchange rate levels, the price action is expected to get "sticky" because it will take more order flow to pierce the wall of either bids or offers. But once these price points are breached, the price action can be expected to accelerate as **stops** are triggered. (Stops, in this sense, refer to stop-loss orders, in which traders leave resting bids or offers away from the current market price to be

17 Some material in this section is based on Sine and Strong (2012).

18 In many other asset classes, technical analysis is based on trade volume data as well as price data. But there are no reliable, timely, and high-frequency trade volume data that are publicly available for over-the-counter (OTC) FX markets.

filled if the market reaches those levels. A stop-loss order is triggered when the price action has gone against a trader's position, and it gets the trader out of that position to limit further losses.)

Technical analysis uses visual cues for market patterns as well as more quantitative technical indicators. There is a wide variety of technical indexes based on market prices that are used in this context. Some technical indicators are as simple as using moving averages of past price points. The 200-day moving average of daily exchange rates is often seen as an important indicator of likely support and resistance. Sometimes two moving averages are used to establish when a price trend is building momentum. For example, when the 50-day moving average crosses the 200-day moving average, this is sometimes seen as a price "break out" point.

Other technical indicators are based on more complex mathematical formulae. There is an extremely wide variety of these more mathematical indicators, some of them very esoteric and hard to connect intuitively with the behavior of real world financial market participants.

In summary, many FX active managers routinely use technical analysis—either alone or in conjunction with other approaches—to form a market opinion or to time position entry and exit points. Even though many technical indicators lack the intellectual underpinnings provided by formal economic modeling, they nonetheless remain a prominent feature of FX markets.

5.3 Active Currency Management Based on the Carry Trade

The **carry trade** is a trading strategy of borrowing in low-yield currencies and investing in high-yield currencies.[19] If technical analysis is based on ignoring economic fundamentals, then the carry trade is based on exploiting a well-recognized violation of one of the international parity conditions often used to describe these economic fundamentals: uncovered interest rate parity. Recall that uncovered interest rate parity asserts that, *on a longer-term average*, the return on an unhedged foreign-currency asset investment will be the same as a domestic-currency investment. Assuming that the base currency in the P/B quote is the low-yield currency, stated algebraically uncovered interest rate parity asserts that

$$\%\Delta S_{H/L} \approx i_H - i_L$$

where $\%\Delta S_{H/L}$ is the percentage change in the $S_{H/L}$ spot exchange rate (the low-yield currency is the base currency), i_H is the interest rate on the high-yield currency and i_L is the interest rate on the low-yield currency. If uncovered interest rate parity holds, the yield spread *advantage* for the high-yielding currency (the right side of the equation) will, on average, be matched by the *depreciation* of the high-yield currency (the left side of the equation; the low-yield currency is the base currency and hence a positive value for $\%\Delta S_{H/L}$ means a depreciation of the high-yield currency). According to the uncovered interest rate parity theorem, it is this offset between (1) the yield advantage and (2) the currency depreciation that equates, on average, the unhedged currency returns.

But in reality, the historical data show that there are persistent deviations from uncovered interest rate parity in FX markets, at least in the short to medium term. Indeed, high-yield countries often see their currencies *appreciate*, not depreciate, for extended periods of time. The positive returns from a combination of a favorable yield differential plus an appreciating currency can remain in place long enough to present attractive investment opportunities.

19 The term "carry" is related to what is known as the cost of carry—that is, of carrying or holding an investment. This investment has either an implicit or explicit cost (borrowing cost) but may also produce income. The net cost of carry is the difference between these two return rates.

This persistent violation of uncovered interest rate parity described by the carry trade is often referred to as the **forward rate bias**. An implication of uncovered interest rate parity is that the forward rate should be an unbiased predictor of future spot rates. The historical data, however, show that the forward rate is not the center of the distribution for future spot rates; in fact, it is a *biased* predictor (for example, see Kritzman 1999). Hence the name "forward rate bias." With the forward rate premium or discount defined as $F_{P/B} - S_{P/B}$ the "bias" in the forward rate bias is that the premium typically overstates the amount of appreciation of the base currency, and the discount overstates the amount of depreciation. Indeed, the forward discount or premium often gets even the *direction* of future spot rate movements wrong.

The carry trade strategy (borrowing in low-yield currencies, investing in high-yield currencies) is equivalent to a strategy based on trading the forward rate bias. Trading the forward rate bias involves buying currencies selling at a forward discount, and selling currencies trading at a forward premium. This makes intuitive sense: It is desirable to buy low and sell high.

To show the equivalence of the carry trade and trading the forward rate bias, recall that covered interest rate parity (which is enforced by arbitrage) is stated as

$$\frac{F_{P/B} - S_{P/B}}{S_{P/B}} = \frac{\left(i_P - i_B\right)\left(\frac{t}{360}\right)}{1 + i_B\left(\frac{t}{360}\right)}$$

This equation shows that when the base currency has a lower interest rate than the price currency (i.e., the right side of the equality is positive) the base currency will trade at a forward premium (the left side of the equality is positive). That is, being low-yield currency and trading at a forward premium is synonymous. Similarly, being a high-yield currency means trading at a forward discount. Borrowing in the low-yield currency and investing in the high-yield currency (the carry trade) is hence equivalent to selling currencies that have a forward premium and buying currencies that have a forward discount (trading the forward rate bias). We will return to these concepts in Section 6.1.2 when we discuss the roll yield in hedging with forward contracts. Exhibit 6 summarizes several key points about the carry trade.

Exhibit 6 The Carry Trade: A Summary

	Buy/Invest	Sell/Borrow
Implementing the carry trade	High-yield currency	Low-yield currency
Trading the forward rate bias	Forward discount currency	Forward premium currency

The gains that one can earn through the carry trade (or equivalently, through trading the forward rate bias) can be seen as the risk premiums earned for carrying an unhedged position—that is, for absorbing currency risk. (In efficient markets, there is no extra reward without extra risk.) Long periods of market stability can make these extra returns enticing to many investors, and the longer the yield differential persists between high-yield and low-yield currencies, the more carry trade positions will have a tendency to build up. But these high-yield currency advantages can be erased quickly, particularly if global financial markets are subject to sudden bouts of stress. This is especially true because the carry trade is a *leveraged* position: borrowing in the low-yielding currency and investing in the high-yielding currency. These occasional large losses mean that the return distribution for the carry trade has a pronounced negative skew.

This negative skew derives from the fact that the **funding currencies** of the carry trade (the low-yield currencies in which borrowing occurs) are typically the safe haven currencies, such as the USD, CHF, and JPY. In contrast, the **investment currencies** (the high-yielding currencies) are typically currencies perceived to be higher risk, such as several emerging market currencies. Any time global financial markets are under stress there is a flight to safety that causes rapid movements in exchange rates, and usually a panicked unwinding of carry trades. As a result, traders running carry trades often get caught in losing positions, with the leverage involved magnifying their losses. Because of the tendency for long periods of relatively small gains in the carry trade to be followed by brief periods of large losses, the carry trade is sometimes characterized as "picking up nickels in front of a steamroller." One guide to the riskiness of the carry trade is the volatility of spot rate movements for the currency pair; all else equal, lower volatility is better for a carry trade position.

We close this section by noting that although the carry trade can be based on borrowing in a single funding currency and investing in a single high-yield currency, it is more common for carry trades to use multiple funding and investment currencies. The number of funding currencies and investment currencies need not be equal: for example, there could be five of one and three of the other. Sometimes the portfolio weighting of exposures between the various funding and investment currencies are simply set equal to each other. But the weights can also be optimized to reflect the trader's market view of the expected movements in each of the exchange rates, as well as their individual risks ($\sigma[\%\Delta S]$) and the expected correlations between movements in the currency pairs. These trades can be dynamically rebalanced, with the relative weights among both funding and investment currencies shifting with market conditions.

5.4 Active Currency Management Based on Volatility Trading

Another type of active trading style is unique to option markets and is known as volatility trading (or simply "vol trading").[20] To explain this trading style, we will start with a quick review of some option basics.

The derivatives of the option pricing model show the sensitivity of the option's premium to changes in the factors that determine option value. These derivatives are often referred to as the "Greeks" of option pricing. There is a very large number of first, second, third, and cross-derivatives that can be taken of an option pricing formula, but the two most important Greeks that we will consider here are the following:

- **Delta:** The sensitivity of the option premium to a small change in the price of the underlying[21] of the option, typically a financial asset. This sensitivity is an indication of *price* risk.
- **Vega:** The sensitivity of the option premium to a small change in implied volatility. This sensitivity is an indication of *volatility* risk.

The most important concept to grasp in terms of volatility trading is that the use of options allows the trader, through a variety of trading strategies, to *unbundle* and isolate all of the various risk factors (the Greeks) and trade them separately. Once an initial option position is taken (either long or short), the trader has exposure to

20 In principle, this trading style can be applied to all asset classes with options, not just FX trading. But FX options are the most liquid and widely traded options in the world, so it is in FX where most of volatility trading likely takes place in global financial markets.

21 The underlying asset of a derivative is typically referred to simply as the "underlying."

all of the various Greeks/risk factors. The unwanted risk exposures, however, can then be hedged away, leaving *only* the desired risk exposure to express that specific directional view.

Delta hedging is the act of hedging away the option position's exposure to delta, the price risk of the underlying (the FX spot rate, in this case). Because delta shows the sensitivity of the option price to changes in the spot exchange rate, it thus defines the option's hedge ratio: The size of the offsetting hedge position that will set the *net* delta of the combined position (option plus delta hedge) to zero. Typically implementing this delta hedge is done using either forward contracts or a spot transaction (spot, by definition, has a delta of one, and no exposure to any other of the Greeks; forward contracts are highly correlated with the spot rate). For example, if a trader was long a call option on USD/EUR with a nominal value of EUR1 million and a delta of +0.5, the delta hedge would involve a short forward position in USD/EUR of EUR0.5 million. That is, the size of the delta hedge is equal to the option's delta times the nominal size of the contract. This hedge size would set the net delta of the overall position (option and forward) to zero.[22] Once the delta hedge has set the net delta of the position to zero, the trader then has exposure *only* to the other Greeks, and can use various trading strategies to position in these (long or short) depending on directional views.

Although one could theoretically trade *any* of the other Greeks, the most important one traded is vega; that is, the trader is expressing a view on the future movements in implied volatility, or in other words, is engaged in volatility trading. Implied volatility is not the same as realized, or observed, historical volatility, although it is heavily influenced by it. By engaging in volatility trading, the trader is expressing a view about the future volatility of exchange rates *but not their direction* (the delta hedge set the net delta of the position to zero).

One simple option strategy that implements a volatility trade is a **straddle**, which is a combination of both an at-the-money (ATM) put and an ATM call. A long straddle buys both of these options. Because their deltas are –0.5 and +0.5, respectively, the net delta of the position is zero; that is, the long straddle is delta neutral. This position is profitable in more volatile markets, when either the put or the call go sufficiently in the money to cover the upfront cost of the two option premiums paid. Similarly, a short straddle is a bet that the spot rate will stay relatively stable. In this case, the payout on any option exercise will be less than the twin premiums the seller has collected; the rest is net profit for the option seller. A similar option structure is a **strangle** position for which a long position is buying out-of-the-money (OTM) puts and calls with the same expiry date and the same degree of being out of the money (we elaborate more on this subject later). Because OTM options are being bought, the cost of the position is cheaper—but conversely, it also does not pay off until the spot rate passes the OTM strike levels. As a result, the risk–reward for a strangle is more moderate than that for a straddle.

The interesting thing to note is that by using delta-neutral trading strategies, volatility is turned into a product that can be actively traded like any other financial product or asset class, such as equities, commodities, fixed-income products, and so on. Volatility is not constant nor are its movements completely random. Instead volatility is determined by a wide variety of underlying factors—both fundamental and technical—that the trader can express an opinion on. Movements in volatility are cyclical, and typically subject to long periods of relative stability punctuated by sharp upward spikes in volatility as markets come under periodic bouts of stress (usually the result of some dramatic event, financial or otherwise). Speculative vol traders—for example, among currency overlay managers—often want to be net-short volatility.

---

22 Strictly speaking, the net delta would be *approximately* equal to zero because forward contracts do not have identical price properties to those of the spot exchange rate. But it is close enough for our purposes here, and we will ignore this small difference.

The reason is because most options expire out of the money, and the option writer then gets to keep the option premium without delivery of the underlying currency pair. The amount of the option premium can be considered the risk premium, or payment, earned by the option writer for absorbing volatility risk. It is a steady source of income under "normal" market conditions. Ideally, these traders would want to "flip" their position and be long volatility ahead of volatility spikes, but these episodes can be notoriously difficult to time. Most hedgers typically run options positions that are net-long volatility because they are buying protection from unanticipated price volatility. (Being long the option means being exposed to the time decay of the option's time value; that is similar to paying insurance premiums for the protection against exchange rate volatility.)

We can also note that just as there are *currency overlay* programs for actively trading the portfolio's currency exposures (as discussed in Section 4.4.4) there can also be *volatility overlay* programs for actively trading the portfolio's exposures to movements in currencies' implied volatility. Just as currency overlay programs manage the portfolio's exposure to currency delta (movements in spot exchange rates), volatility overlay programs manage the portfolio's exposure to currency vega. These volatility overlay programs can be focused on earning speculative profits, but can also be used to hedge the portfolio against risk (we will return to this concept in the discussion of macro hedges in Section 6.4.2.).

Enumerating all the potential strategies for trading foreign exchange volatility is beyond the scope of this reading. Instead, the reader should be aware that this dimension of trading FX volatility (not price) exists and sees a large amount of active trading. Moreover, the best traders are able to think and trade in both dimensions simultaneously. Movements in volatility are often correlated with directional movements in the price of the underlying. For example, when there is a flight to safety as carry trades unwind, there is typically a spike in volatility (and options prices) at the same time. Although pure vol trading is based on a zero-delta position, this need not always be the case; a trader can express a market opinion on volatility (vega exposure) and still have a directional exposure to the underlying spot exchange rate as well (delta exposure). That is, the overall trading position has net vega and delta exposures that reflect the *joint* market view.

We end this section by explaining how currency options are quoted in professional FX markets. (This information will be used in Section 6 when we discuss other option trading strategies.) Unlike exchanged-traded options, such as those used in equity markets, OTC options for currencies are not described in terms of specific strike levels (i.e., exchange rate levels). Instead, in the interdealer market, options are described in terms of their "delta." Deltas for puts can range from a minimum of –1 to a maximum of 0, with a delta of –0.5 being the point at which the put option is ATM; OTM puts have deltas between 0 and –0.5. For call options, delta ranges from 0 to +1, with 0.5 being the ATM point. In FX markets, these delta values are quoted both in *absolute* terms (i.e., in positive rather than negative values) and as percentages, with standard FX option quotes usually in terms of 25-delta and 10-delta options (i.e., a delta of 0.25 and 0.10, respectively; the 10-delta option is deeper OTM and hence cheaper than the 25-delta option). The FX options market is the most liquid around these standard delta quoting points (ATM, 25-delta, 10-delta), but of course, as a flexible OTC market, options of any delta/strike price can be traded. The 25-delta put option (for example) will still go in the money if the spot price dips below a *specific* exchange rate level; this *implied* strike price is *backed out* of an option pricing model once all the other pricing factors, including the current spot rate and the 25-delta of the option, are put into the option pricing model. (The specific option pricing model used is agreed on by both parties to the trade.)

These standard delta price points are often used to define option trading strategies. For example, a 25-delta strangle would be based on 25-delta put and call options. Similarly, a 10-delta strangle would be based on 10-delta options (and would cost less and have a more moderate payoff structure than a 25-delta strangle). Labeling option structures by their delta is common in FX markets.

EXAMPLE 3

Active Strategies

Annie McYelland works as an analyst at Scotland-based Kilmarnock Advisors, an investment firm that offers several investment vehicles for its clients. McYelland has been put in charge of formulating the firm's market views for some of the foreign currencies that these vehicles have exposures to. Her market views will be used to guide the hedging and discretionary positioning for some of the actively managed portfolios.

McYelland begins by examining yield spreads between various countries and the implied volatility extracted from the option pricing for several currency pairs. She collects the following data:

One-Year Yield Levels	
Switzerland	−0.103%
United States	0.162%
Poland	4.753%
Mexico	4.550%

One-Year Implied Volatility	
PLN/CHF	8.4%
MXN/CHF	15.6%
PLN/USD	20.3%
MXN/USD	16.2%

Note: PLN = Polish zloty; the Swiss yields are negative because of Swiss policy actions.

McYelland is also examining various economic indicators to shape her market views. After studying the economic prospects for both Japan and New Zealand, she expects that the inflation rate for New Zealand is about to accelerate over the next few years, whereas the inflation rate for Japan should remain relatively stable. Turning her attention to the economic situation in India, McYelland believes that the Indian authorities are about to tighten monetary policy, and that this change has not been fully priced into the market. She reconsiders her short-term view for the Indian rupee (i.e., the INR/USD spot rate) after conducting this analysis.

McYelland also examines the exchange rate volatility for several currency pairs to which the investment trusts are exposed. Based on her analysis of the situation, she believes that the exchange rate between Chilean peso and the US dollar (CLP/USD) is about to become much more volatile than usual, although she has no strong views about whether the CLP will appreciate or depreciate.

One of McYelland's colleagues, Catalina Ortega, is a market technician and offers to help McYelland time her various market position entry and exit points based on chart patterns. While examining the JPY/NZD price chart, Ortega notices that the 200-day moving average is at 62.0405 and the current spot rate is 62.0315.

1 Based on the data she collected, all else equal, McYelland's *best* option for implementing a carry trade position would be to fund in:

- **A** USD and invest in PLN.
- **B** CHF and invest in MXN.
- **C** CHF and invest in PLN.

2 Based on McYelland's inflation forecasts, all else equal, she would be *more likely* to expect a(n):

- **A** depreciation in the JPY/NZD.
- **B** increase in capital flows from Japan to New Zealand.
- **C** more accommodative monetary policy by the Reserve Bank of New Zealand.

3 Given her analysis for India, McYelland's short-term market view for the INR/USD spot rate is now *most likely* to be:

- **A** biased toward appreciation.
- **B** biased toward depreciation.
- **C** unchanged because it is only a short-run view.

4 Using CLP/USD options, what would be the *cheapest* way for McYelland to implement her market view for the CLP?

- **A** Buy a straddle
- **B** Buy a 25-delta strangle
- **C** Sell a 40-delta strangle

5 Based on Ortega's analysis, she would *most likely* expect:

- **A** support near 62.0400.
- **B** resistance near 62.0310.
- **C** resistance near 62.0400.

Solution to 1:

C is correct. The yield spread between the funding and investment currencies is the widest and the implied volatility (risk) is the lowest. The other choices have a narrower yield spread and higher risk (implied volatility).

Solution to 2:

A is correct. All else equal, an increase in New Zealand's inflation rate will decrease its real interest rate and lead to the real interest rate differential favoring Japan over New Zealand. This would likely result in a depreciation of the JPY/NZD rate over time. The shift in the relative real returns should lead to reduced capital flows from Japan to New Zealand (so Choice B is incorrect) and the RBNZ—New Zealand's central bank—is more likely to tighten monetary policy than loosen it as inflation picks up (so Choice C is incorrect).

Solution to 3:

B is correct. Tighter monetary policy in India should lead to higher real interest rates (at least in the short run). This increase will cause the INR to appreciate against the USD, but because the USD is the base currency, this will be represented as depreciation in the INR/USD rate. Choice C is incorrect because a tightening of monetary policy that is not fully priced-in to market pricing is likely to move bond yields and hence the exchange rate in the short run (given the simple economic model in Section 5.1).

Solution to 4:

B is correct. Either a long straddle or a long strangle will profit from a marked increase in volatility in the spot rate, but a 25-delta strangle would be cheaper (because it is based on OTM options). Writing a strangle—particularly one that is close to being ATM, which is what a 40-delta structure is—is likely to be exercised in favor of the counterparty if McYelland's market view is correct.

Solution to 5:

C is correct. The 200-day moving average has not been crossed yet, and it is higher than the current spot rate. Hence this technical indicator suggests that resistance lies above the current spot rate level, likely in the 62.0400 area. Choice A is incorrect because the currency has not yet appreciated to 62.0400, so it cannot be considered a "support" level. Given that the currency pair has already traded through 62.0310 and is still at least 90 pips away from the 200-day moving average, it is more likely to suspect that resistance still lies above the current spot rate.

6 TOOLS OF CURRENCY MANAGEMENT

In this section, we focus on how the portfolio manager uses financial derivatives to implement both the *strategic* positioning of the portfolio along the risk spectrum (i.e., the performance benchmark) as well as the *tactical* decisions made in regard to variations around this "neutral" position. The manager's market view—whether based on carry, fundamental, currency volatility, or technical considerations—leads to this active management of risk positioning around the strategic benchmark point. Implementing both strategic and tactical viewpoints requires the use of trading tools, which we discuss in this section.

The balance of this reading will assume that the portfolio's strategic foreign-currency asset exposures and the maximum amount of currency risk desired have already been determined by the portfolio's IPS. We begin at the conservative end of the risk spectrum by describing a passive hedge for a single currency (with a 100% hedge ratio). After discussing the costs and limitations of this approach, we move out further along the risk spectrum by describing strategies in which the basic "building blocks" of financial derivatives can be combined to implement the manager's tactical positioning and construct much more customized risk–return profiles. Not surprisingly, the basic trading tools themselves—forwards, options, FX swaps—are used for both strategic and tactical risk management and by both hedgers and speculators alike (although for different ends). Note that the instruments covered as tools of currency management are not nearly an exhaustive list. For example, exchange-traded funds for currencies are a vehicle that can be useful in managing currency risk.

6.1 Forward Contracts

In this section, we consider the most basic form of hedging: a 100% hedge ratio for a single foreign-currency exposure. Futures or forward contracts on currencies can be used to obtain full currency hedges, although most institutional investors prefer to use forward contracts for the following reasons:

1. Futures contracts are standardized in terms of settlement dates and contract sizes. These may not correspond to the portfolio's investment parameters.
2. Futures contracts may not always be available in the currency pair that the portfolio manager wants to hedge. For example, the most liquid currency futures contracts trade on the Chicago Mercantile Exchange (CME). Although there are CME futures contracts for all major exchange rates (e.g., USD/EUR, USD/GBP) and many cross rates (e.g., CAD/EUR, JPY/CHF), there are not contracts available for all possible currency pairs. Trading these cross rates would need multiple futures contracts, adding to portfolio management costs. In addition, many of the "second tier" emerging market currencies may not have liquid futures contracts available against any currency, let alone the currency pair in which the portfolio manager is interested.
3. Futures contracts require up-front margin (initial margin). They also have intra-period cash flow implications, in that the exchange will require the investor to post additional variation margin when the spot exchange rate moves against the investor's position. These initial and ongoing margin requirements tie up the investor's capital and require careful monitoring through time, adding to the portfolio management expense. Likewise, margin flows can go in the investor's favor, requiring monitoring and reinvestment.

In contrast, forward contracts do not suffer from any of these drawbacks. Major global investment dealers (such as Deutsche Bank, Royal Bank of Scotland, UBS, etc.) will quote prices on forward contracts for practically every possible currency pair, settlement date, and transaction amount. They typically do not require margin to be posted or maintained.[23]

Moreover, the daily trade volume globally for OTC currency forward and swap contracts dwarfs that for exchange-traded currency futures contracts; that is, forward contracts are more liquid than futures for trading in large sizes. Reflecting this liquidity, forward contracts are the predominant hedging instrument in use globally. For the balance of this section, we will focus only on currency forward contracts. However, separate side boxes discuss exchange-traded currency futures contracts and currency-based exchange-traded funds (ETFs).

FUTURES CONTRACTS ON THE CME

The vast majority of trading in foreign exchange markets is conducted by large money-management institutions and global currency dealers. Many public sector market participants also create large foreign exchange flows; particularly sovereign wealth funds (SWFs) and central banks managing their country's foreign-currency reserves. The minimum dealing sizes for all of these institutional players are at least one million units of the base currency; typically the dealing sizes can be much larger. For the reasons mentioned in Section 6.1, these institutional market participants almost always use OTC forward contracts rather than exchange-traded futures contracts.

23 Some forward contracts—typically between investment dealers or between dealers and some large institutional accounts—are governed by what are called Credit Support Annexes (CSAs) that do require collateral to be posted. The amount of collateral required can change over the life of the forward contract if the mark-to-market losses on the position exceed pre-specified bounds.

But this preference does not mean that exchange-traded futures are unimportant to the FX market or that they do not have their own niche of market participants. The largest complex of exchange-traded currency products—with more than 50 futures contracts and 30 option contracts on futures—is hosted by the Chicago Mercantile Exchange. Collectively these products see an average daily turnover of approximately USD120 billion, which is not that much different than the average daily turnover on the main interbank dealing platforms, Reuters D3000 and EBS, which as of this writing have an average daily turnover of approximately USD140 billion each.[24] Moreover, the CME is experiencing a higher growth rate in its turnover than the main interbank platforms.

The mix of market participants on the CME is different than that of the interbank market. Typically, the CME will appeal to those accounts that are trading in smaller dealing sizes. (For example, the most heavily traded currency contract on the CME is for the USD/EUR currency pair, and its contract size is EUR125,000.) The CME also provides market access with tight pricing and good liquidity (for these smaller deal sizes) to those accounts that may not have the creditworthiness or the credit relationship with major FX dealers in order to access the FX market through other channels. These less creditworthy accounts are able to access the exchange-traded market because the central clearing house demands margin and guarantees delivery on all contracts.

There is a diverse range of market participants who fit this description, including small hedge funds, proprietary trading firms, and active individual traders. One important class of market participants on the CME are **managed futures funds**, which are often used by small institutions, high-net-worth individuals, and other accounts that would fall under the realm of private wealth management. These managed futures funds are pools of private capital managed by **commodity trading advisers** (CTAs). To trade futures contracts on behalf of others, CTAs need to be registered with the US Commodity Futures Trading Commission (CFTC) by application through the National Futures Association, which is a self-regulatory organization (SRO) that oversees derivatives trading on US exchanges. These managed futures funds are managed on a discretionary basis by the CTAs, who use leverage and aggressive trading strategies to earn speculative profits. As such, managed futures funds can be thought of as being somewhat similar to hedge funds, and as an alternative asset class, the returns on managed futures are typically uncorrelated with more traditional asset classes. This feature provides diversification benefits to the overall portfolio.

Finally, note that FX quoting conventions for contracts quoted on the CME are frequently different from those used elsewhere in the currency market (as shown in Exhibit 2). The CME is based in the United States and, as a matter of convention, uses the USD as the price currency in the foreign exchange quote for all currency pairs.[25] As a result, many of the CME futures and options contracts are quoted in terms that are different from standard interbank market quotes in the OTC market. For example, the dollar/yen futures contract has a nominal size of JPY12,500,000 and is quoted in terms of the amount of US dollars to buy one yen—that is, in terms of USD/JPY rather than the more usual JPY/USD. Hence, a recent price quote on this contract was 0.012807 rather than 78.08. (To make this quote more readable this price is often scaled by multiplying by 100 or 10,000: the price in USD of 100 JPY and 10,000 JPY, respectively.)

6.1.1 *Hedge Ratios with Forward Contracts*

In principle, setting up a full currency hedge is relatively straight forward: match the current market value of the foreign-currency exposure in the portfolio with an equal and offsetting position in a forward contract. In practice, of course, it is not that simple because the market value of the foreign-currency assets will change with market

24 It might seem odd that the CME transacts almost as much as the major interbank trading platforms, EBS and Reuters. However, these interbank platforms deal almost exclusively in *spot* transactions, which are only a fraction of overall FX market trading. Most of the approximately USD4 trillion in daily turnover in FX markets is accounted for by FX swap and outright forward transactions.

25 That is, for all currency pairs that involve the USD. The CME also has contracts for JPY/CAD and other cross rates.

conditions. This means that the actual hedge ratio will typically *drift* away from the desired hedge ratio as market conditions change. A **static hedge** (i.e., unchanging hedge) will avoid transaction costs, but will also tend to accumulate unwanted currency exposures as the value of the foreign-currency assets change. This characteristic will cause a mismatch between the market value of the foreign-currency asset portfolio and the nominal size of the forward contract used for the currency hedge; this is pure currency risk. For this reason, the portfolio manager will typically need to implement a **dynamic hedge** by rebalancing the portfolio periodically. This hedge rebalancing will mean adjusting some combination of the size, number, and maturities of the forward currency contracts.

A simple example will illustrate this rebalancing process. Suppose that an investor domiciled in Switzerland has a EUR-denominated portfolio that, at the start of the period, is worth EUR1,000,000. Assume a monthly hedge-rebalancing cycle. To hedge this portfolio, the investor would sell EUR1,000,000 one month forward against the CHF. Assume that one month later, the EUR-denominated investment portfolio is then actually worth only EUR950,000. To roll the hedge forward for the next month, the investor will engage in a mismatched FX swap. (Recall that a "matched" swap means that both the spot and forward transactions—the near and far "legs" of the swap, respectively—are of equal size). For the near leg of the swap, EUR1 million will be bought at spot to settle the expiring forward contract. (The euro amounts will then net to zero, but a Swiss franc cash flow will be generated, either a loss or a gain for the investor, depending on how the CHF/EUR rate has changed over the month). For the far leg of the swap, the investor will sell EUR950,000 forward for one month.

Another way to view this rebalancing process is to consider the case in which the original short forward contract has a three-month maturity. In this case, rebalancing after one month would mean that the manager would have to *buy* 50,000 CHF/EUR two months forward. There is no cash flow at the time this second forward contract is entered, but the *net* amount of euro for delivery at contract settlement two months into the future is now the euro hedge amount desired (i.e., EUR950,000). There will be a net cash flow (denominated in CHF) calculated over these two forward contracts on the settlement date two months hence.

Although rebalancing a dynamic hedge will keep the actual hedge ratio close to the target hedge ratio, it will also lead to increased transaction costs compared with a static hedge. The manager will have to assess the cost–benefit trade-offs of how frequently to dynamically rebalance the hedge. These will depend on a variety of idiosyncratic factors (manager risk aversion, market view, IPS guidelines, etc.), and so there is no single "correct" answer—different managers will likely make different decisions.

However, we can observe that the higher the degree of risk aversion, the more frequently the hedge is likely to be rebalanced back to the "neutral" hedge ratio. Similarly, the greater the tolerance for active trading, and the stronger the commitment to a particular market view, the more likely it is that the actual hedge ratio will be allowed to vary from a "neutral" setting, possibly through entering into new forward contracts. (For example, if the P/B spot rate was seen to be oversold and likely to rebound higher, an actively traded portfolio might buy the base currency through forward contracts to lock in this perceived low price—and thus change the actual hedge ratio accordingly.) The sidebar on executing a hedge illustrates the concepts of rolling hedges, FX swaps and their pricing (bid–offer), and adjusting hedges for market views and changes in market values.

EXECUTING A HEDGE

Jiao Yang works at Hong Kong-based Kwun Tong Investment Advisors; its reporting currency is the Hong Kong Dollar (HKD). She has been put in charge of managing the firm's foreign-currency hedges. Forward contracts for two of these hedges are coming due for settlement, and Yang will need to use FX swaps to roll these hedges forward three months.

Hedge #1: Kwun Tong has a short position of JPY800,000,000 coming due on a JPY/HKD forward contract. The market value of the underlying foreign-currency assets has not changed over the life of the contract, and Yang does not have a firm opinion on the expected future movement in the JPY/HKD spot rate.

Hedge #2: Kwun Tong has a short position of EUR8,000,000 coming due on a HKD/EUR forward contract. The market value of the EUR-denominated assets has increased (measured in EUR). Yang expects the HKD/EUR spot rate to decrease.

The following spot exchange rates and three-month forward points are in effect when Yang transacts the FX swaps necessary to roll the hedges forward:

	Spot Rate	Three-Month Forward Points
JPY/HKD	10.80/10.82	–20/–14
HKD/EUR	10.0200/10.0210	125/135

Note: The JPY/HKD forward points will be scaled by 100; the HKD/EUR forward points will be scaled by 10,000

As a result, Yang undertakes the following transactions:

For **Hedge #1**, the foreign-currency value of the underlying assets has not changed, and she does not have a market view that would lead her to want to either over- or under-hedge the foreign-currency exposure. Therefore, to roll these hedges forward, she uses a matched swap. For matched swaps (see Section 2.3), the convention is to base pricing on the mid-market spot exchange rate. Thus, the spot leg of the swap would be to buy JPY800,000,000 at the mid-market rate of 10.81 JPY/HKD. The forward leg of the swap would require selling JPY800,000,000 forward three months. Selling JPY (the price currency in the JPY/HKD quote) is equivalent to buying HKD (the base currency). Therefore, she uses the offer-side forward points, and the all-in forward rate for the forward leg of the swap is as follows:

$$10.81 + \frac{-14}{100} = 10.67$$

For **Hedge #2**, the foreign-currency value of the underlying assets has increased; Yang recognizes that this implies that she should increase the size of the hedge greater than EUR8,000,000. She also believes that the HKD/EUR spot rate will decrease, and recognizes that this implies a hedge ratio of more than 100% (Kwun Tong Advisors has given her discretion to over- or under-hedge based on her market views). This too means that the size of the hedge should be increased more than EUR8,000,000, because Yang will want a larger short position in the EUR to take advantage of its expected depreciation. Hence, Yang uses a mismatched swap, buying EUR8,000,000 at spot rate against the HKD, to settle the maturing forward contract and then *selling* an amount *more* than EUR8,000,000 forward to increase the hedge size. Because the EUR is the base currency in the HKD/EUR quote, this means using the *bid* side for both the spot rate and the forward points when calculating the all-in forward rate:

$$10.0200 + \frac{125}{10{,}000} = 10.0325$$

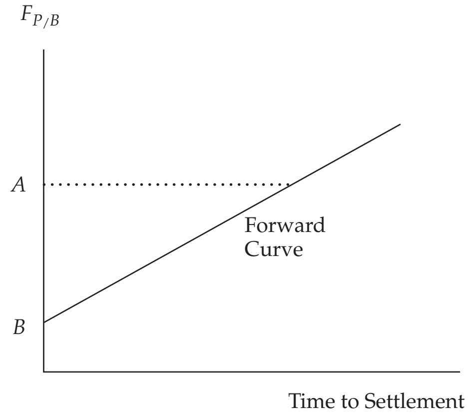
The spot leg of the swap—buying back EUR8,000,000 to settle the outstanding forward transaction—is also based on the bid rate of 10.0200. This is because Yang is selling an amount larger than EUR8,000,000 forward, and the all-in forward rate of the swap is already using the bid side of the market (as it would for a matched swap). Hence, to pick up the net increase in forward EUR sales, the dealer Yang is transacting with would price the swap so that Yang also has to use bid side of the *spot* quote for the spot transaction used to settle the maturing forward contract.

6.1.2 *Roll Yield*

The roll yield (also called the roll return) on a hedge results from the fact that forward contracts are priced at the spot rate adjusted for the number of forward points at that maturity (see the example shown in Exhibit 3). This forward point adjustment can either benefit or detract from portfolio returns (positive and negative roll yield, respectively) depending on whether the forward points are at a premium or discount, and what side of the market (buying or selling) the portfolio manager is on.

The concept of roll yield is illustrated with the simplified example shown in Exhibit 7.

Exhibit 7 The Forward Curve and Roll Yield



The magnitude of roll yield is given by $|(F_{P/B} - S_{P/B})/S_{P/B}|$ where "||" indicates absolute value. The sign depends on whether the investor needs to buy or to sell the base currency forward in order to maintain the hedge. A *positive* roll yield results from buying the base currency at a forward discount or selling it at a forward premium (the intuition here is that it is profitable to "buy low and sell high"). Otherwise, the roll yield is negative (i.e., a positive cost). Examining the case of negative roll yield, assume that to implement the hedge requires buying the base currency in the P/B quote, and that the base currency is trading at a forward premium (as shown in Exhibit 7). By using a long position in a forward contract to implement this hedge, it means paying the forward price of A. All else equal, as time passes the price of the forward contract will "roll down the curve" toward Price B as the forward contract's settlement date approaches. (Note that in reality the curve is not always linear.) At the settlement date of the forward contract, it is necessary to roll the hedging position forward to extend the currency hedge. This rolling forward will involve selling the base currency at the then-current spot exchange rate to settle the forward contract, and then going long another far-dated forward contract (i.e., an FX swap transaction). Note that the portfolio manager originally bought the base currency at Price A and then subsequently sold it at a lower Price B—and that buying high and selling low will be a cost to the portfolio. Or put differently, all else equal, the roll yield would be negative in this case.[26]

The concept of roll yield is very similar to the concept of forward rate bias (and the carry trade) introduced in Section 5. Indeed, a negative roll yield typically indicates that the hedger was trading *against* the forward rate bias by buying a currency at a forward premium (as in Exhibit 7) or selling a currency at a forward discount. This is the exact opposite of trading the forward rate bias, which is to buy at a discount and sell at a premium. Given the equivalence between the forward rate bias and the carry trade, by trading against the forward rate bias the hedger with a negative roll yield is also essentially entering into *negative* carry trade, in effect borrowing at high rates and investing at low rates. On average, this will not be a winning strategy. Given the equivalence between implementing a carry trade, trading the forward rate bias, and earning positive roll yield, we can now complete Exhibit 6 introduced in Section 5.3 on the carry trade:

Exhibit 8 The Carry Trade and Roll Yield

	Buy/Invest	Sell/Borrow	
Implementing the carry trade	High-yield currency	Low-yield currency	Earning a positive roll yield
Trading the forward rate bias	Forward discount currency	Forward premium currency	

Note as well that this concept of roll yield applies to forward and futures contracts used to trade *any* asset class, not just currencies: It applies equally well to forwards and futures on equities, fixed-income securities, commodities, and indeed, any financial product. For example, consider the case of a commodity processor that hedges the costs of its production process by going long corn futures contracts. If the futures curve for corn futures contracts is in contango (upward sloping, as in Exhibit 7), then this hedging position will also face the potential for negative roll yield.

To be fair, it is also possible for the level of, and movement in, forward points to be in the portfolio manager's favor. Extending our previous example, consider the case of a portfolio manager that has to *sell* the base currency to implement the currency hedge. In this case, the manager would be selling the base currency forward at Price A in Exhibit 7 and, all else equal and through entering an FX swap at settlement date, buying the currency back at the lower Price B—essentially, short selling a financial product with a declining price. In this case, the roll yield is positive.

Because the level of and movements in forward points can either enhance or reduce currency-hedged returns, it explains an observed tendency in foreign exchange markets for the amount of currency hedging to generally vary with movements in forward points. As forward points move against the hedger, the amount of hedging activity typically declines as the cost/benefit ratio of the currency hedge deteriorates. The opposite occurs when movements in forward points reduce hedging costs. Essentially the tendency to hedge will vary depending on whether implementing the hedge happens to be trading in the same direction of the forward rate bias strategy

26 The "all else equal" qualification refers to the fact that the all-in price of the forward contract consists of the spot rate and forward points, and both are likely to change over the life of the forward contract. It is possible that at the settlement date the spot rate would have moved higher than A, in which case the roll yield would be positive. But the larger the gap between A and B at contract initiation, the less likely this is to occur.

or against it. It is easier to sell a currency forward if there is a "cushion" when it is selling at a forward premium. Likewise, it is more attractive to buy a currency when it is trading at a forward discount. This swings the forward rate bias (and carry trade advantage) in favor of the hedge.

Combined with the manager's market view of future spot rate movements, what this concept implies is that, when setting the hedge ratio, the portfolio manager must balance the effect of expected future exchange rate movements on portfolio returns against the expected effect of the roll yield (i.e., the expected cost of the hedge).

A simple example can illustrate this effect. Consider a portfolio manager that needs to sell forward the base currency of a currency pair (P/B) to implement a currency hedge. Clearly, the manager would prefer to sell this currency at as high a price as possible. Assume that given the forward points for this currency pair and the time horizon for the hedge, the expected roll yield (cost of the hedge) is –3%. Suppose the portfolio manager had a market view that the base currency would depreciate by 4%. In this case, the hedge makes sense: It is better to pay 3% for the hedge to avoid an expected 4% loss.

Now, suppose that with a movement in forward points the new forward discount on the base currency is 6% away from the current spot rate. If the manager's market view is unchanged (an expected depreciation of the base currency of 4%), then now the use of the hedge is less clear: Does it make sense to pay 6% for the hedge to avoid an expected 4% loss? A *risk-neutral* manager would not hedge under these circumstances because the net expected value of the hedge is negative. But a *risk-averse* manager might still implement the hedge regardless of the negative net expected value. The reason is because it is possible that the market forecast is wrong and that the *actual* depreciation of the base currency (and realized loss to the portfolio) may be higher than the 6% cost of the hedge. The risk-averse manager must then weigh the *certainty* of a hedge that costs 6% against the *risk* that actual unhedged currency losses might be much higher than that.

Clearly, the cost/benefit analysis has shifted against hedging in this case, but many risk-averse investors would still undertake the hedge anyway. The risk-averse manager would likely only take an unhedged currency position if the difference between the expected cost of the hedge and the expected return on an unhedged position was so great as to make the risk acceptable. Balancing these two considerations would depend on the type of market view the manager held and the degree of conviction in it, as well as the manager's degree of risk aversion. The decision taken will vary among investors, so no definitive answer can be given as to what would be the appropriate hedging choice (different portfolio managers will make different choices given the same opportunity set). But hedging costs will vary with market conditions and the higher the expected cost of the hedge (negative roll yield) the more the cost/benefit calculation moves against using a fully hedged position. Or put another way, if setting up the hedge involves selling the low-yield currency and buying the high-yield currency in the P/B pair (i.e., an implicit carry position), then the more likely the portfolio will be fully hedged or even over-hedged. The opposite is also true: Trading against the forward rate bias is likely to lead to lower hedge ratios, all else equal.

EXAMPLE 4

The Hedging Decision

The reporting currency of Hong Kong-based Kwun Tong Investment Advisors is the Hong Kong dollar (HKD). The investment committee is examining whether it should implement a currency hedge for the firm's exposures to the GBP and the ZAR (the firm has long exposures to both of these foreign currencies). The hedge would use forward contracts. The following data relevant to assessing the expected cost of the hedge and the expected move in the spot exchange rate has been developed by the firm's market strategist.

	Current Spot Rate	Six-Month Forward Rate	Six-Month Forecast Spot Rate
HKD/GBP	12.4610	12.6550	12.3000
HKD/ZAR	0.9510	0.9275	0.9300

1 Recommend whether to hedge the firm's long GBP exposure. Justify your recommendation.

2 Discuss the trade-offs in hedging the firm's long ZAR exposure.

Solution to 1:

Kwun Tong is long the GBP against the HKD, and HKD/GBP is selling at a forward premium of +1.6% compared with the current spot rate. All else equal, this is the expected roll yield—which is in the firm's favor, in this case, because to implement the hedge Kwun Tong would be *selling* GBP, the base currency in the quote, at a price *higher* than the current spot rate. Moreover, the firm's market strategist expects the GBP to *depreciate* by 1.3% against the HKD. Both of these considerations argue for hedging this exposure.

Solution to 2:

Kwun Tong is long the ZAR against the HKD, and HKD/ZAR is selling at a forward discount of –2.5% compared with the current spot rate. Implementing the hedge would require the firm to *sell* the base currency in the quote, the ZAR, at a price *lower* than the current spot rate. This would imply that, all else equal, the roll yield would go against the firm; that is, the expected cost of the hedge would be 2.5%. But the firm's strategist also forecasts that the ZAR will depreciate against the HKD by 2.2%. This makes the decision to hedge less certain. A risk-neutral investor would not hedge because the expected cost of the hedge is more than the expected depreciation of the ZAR. But this is only a point forecast and comes with a degree of uncertainty—there is a risk that the HKD/ZAR spot rate might depreciate by more than the 2.5% cost of the hedge. In this case, the decision to hedge the currency risk would depend on the trade-offs between (1) the level of risk aversion of the firm; and (2) the conviction the firm held in the currency forecast—that is, the level of certainty that the ZAR would not depreciate by more than 2.5%.

6.2 Currency Options

One of the costs of forward contracts is the opportunity cost. Once fully hedged, the portfolio manager forgoes any upside potential for future currency moves in the portfolio's favor. Currency options remove this opportunity cost because they provide the manager the right, but not the obligation, to buy or sell foreign exchange at a future date at a rate agreed on today. The manager will only exercise the option at the expiry date if it is favorable to do so.[27]

27 Almost all options in the FX market are European-style options, which only allow for exercise at the expiry date.

Consider the case of a portfolio manager who is long the base currency in the P/B quote and needs to sell this currency to implement the hedge. One approach is to simply buy an at-the-money put option on the P/B currency pair. Matching a long position in the underlying with a put option is known as a **protective put** strategy. Suppose the current spot rate is 1.3650 and the strike price on the put option bought is 1.3650. If the P/B rate subsequently goes down (P appreciates and B depreciates) by the expiry date, the manager can exercise the option, implement the hedge, and guarantee a selling price of 1.3650. But if the P/B rate increases (P depreciates and B appreciates), the manager can simply let the option expire and collect the currency gains.

Unfortunately, like forward contracts, currency options are not "free goods" and, like any form of insurance, there is always a price to be paid for it. Buying an option means paying an upfront premium. This premium is determined, first, by its **intrinsic value**, which is the difference between the spot exchange rate and the strike price of the option (i.e., whether the option is in the money, at the money, or out of the money, respectively). ATM options are more expensive than OTM options, and frequently these relatively expensive options expire worthless.[28]

The second determinant of an option's premium is its **time value**, which in turn is heavily influenced by the volatility in exchange rates. Regardless of exchange rate volatility, however, options are always moving toward expiry. In general, the time value of the option is always declining. This is the time decay of the option's value (theta, one of the "Greeks" of option prices, describes this effect) and is similar in concept to that of negative roll yield on forward contracts described earlier. Time decay always works against the owner of an option.

As with forward contracts, a portfolio manager will have to make judgments about the cost/benefit trade-offs of options-based strategies. Although options do allow the portfolio upside potential from favorable currency movements, options can also be a very expensive form of insurance. The manager will have to balance any market view of potential currency gains against hedging costs and the degree of risk aversion. There is no "right" answer; different managers will make different decisions about the cost/benefit trade-offs when given the same opportunity set.

EXAMPLE 5

Hedging Problems

Brixworth & St. Ives Asset Management is a UK-based firm managing a dynamic hedging program for the currency exposures in its Aggressive Growth Fund. One of the fund's foreign-currency asset holdings is denominated in the Mexican peso (MXN), and one month ago Brixworth & St. Ives fully hedged this exposure using a two-month MXN/GBP forward contract. The following table provides the relevant information.

28 This is nothing to be unduly concerned about. Most insurance contracts expire out of the money in that the insured event did not occur (e.g., your house did not catch fire). This does not mean the insurance was a mistake; the insurance buyer had downside protection, which is worth something. Moreover, the (absolute value of the) option's delta is *approximately* equal to the probability of exercise, which means that, on average, an ATM option will expire worthless about half the time.

	One Month Ago	Today
Value of assets (in MXN)	10,000,000	9,500,000
MXN/GBP spot rate (bid–offer)	20.0500/20.0580	19.5985/20.0065
One-month forward points (bid–offer)	625/640	650/665
Two-month forward points (bid–offer)	875/900	900/950

The Aggressive Growth Fund also has an unhedged foreign-currency asset exposure denominated in the South African rand (ZAR). The current mid-market spot rate in the ZAR/GBP currency pair is 5.1050.

1 One month ago, Brixworth & St. Ives *most likely* sold:

A MXN9,500,000 forward at an all-in forward rate of MXN/GBP 19.6635.

B MXN10,000,000 forward at an all-in forward rate of MXN/GBP 20.1375.

C MXN10,000,000 forward at an all-in forward rate of MXN/GBP 20.1480.

2 To rebalance the hedge today, the firm would *most likely* need to:

A buy MXN500,000 spot.

B buy MXN500,000 forward.

C sell MXN500,000 forward.

3 Given the data in the table, the roll yield on this hedge at the forward contracts' maturity date is *most likely* to be:

A zero.

B negative.

C positive.

4 Assuming that all ZAR/GBP options considered have the same notional amount and maturity, the *most* expensive hedge that Brixworth & St. Ives could use to hedge its ZAR exposure is a long position in a(n):

A ATM call.

B 25-delta call.

C put with a strike of 5.1050.

Solution to 1:

C is correct. Brixworth & St. Ives is long the MXN and hence must sell the MXN forward against the GBP. Selling MXN against the GBP means buying GBP, the base currency in the MXN/GBP quote. Therefore, the offer side of the market must be used. This means the all-in rate used one month ago would have been 20.0580 + 900/10,000, which equals 20.1480. Choice A is incorrect because it uses today's asset value and the bid side of the spot and one-month forward quotes and Choice B is incorrect because it uses the wrong side of the market (the bid side).

Solution to 2:

B is correct. The foreign investment went down in value in MXN terms. Therefore Brixworth & St. Ives must reduce the size of the hedge. Previously it had sold MXN10,000,000 forward against the GBP, and this amount must be reduced to MXN9,500,000 by buying MXN500,000 forward. Choice A is incorrect because hedging is done with forward contracts not spot deals. Choice C is incorrect because selling MXN forward would increase the size of the hedge, not decrease it.

Solution to 3:

B is correct. To implement the hedge, Brixworth & St. Ives must sell MXN against the GBP, or equivalently, buy GBP (the base currency in the P/B quote) against the MXN. The base currency is selling forward at a premium, and—all else equal—its price would "roll down the curve" as contract maturity approached. Having to settle the forward contract means then selling the GBP spot at a lower price. Buying high and selling low will define a negative roll yield. Moreover, the GBP has depreciated against the MXN, because the MXN/GBP spot rate declined between one month ago and now, which will also add to the negative roll yield.

Solution to 4:

A is correct. The Aggressive Growth Fund is long the ZAR through its foreign-currency assets, and to hedge this exposure it must sell the ZAR against the GBP, or equivalently, buy GBP—the base currency in the P/B quote—against the ZAR. Hedging a required purchase means a long position in a call option (not a put, which is used to hedge a required sale of the base currency in the P/B quote). An ATM call option is more expensive than a 25-delta call option.

6.3 Strategies to Reduce Hedging Costs and Modify a Portfolio's Risk Profile

In the previous sections, we showed that completely hedging currency risk is possible—but can also be expensive. It can be even more expensive when trying to avoid all downside risk while keeping the full upside potential for favorable currency movements (i.e., a protective put strategy with ATM options). Hedging can be seen as a form of insurance, but it is possible to overpay for insurance. Judgments have to be made to determine at what point the costs outweigh the benefits.

As with any form of insurance, there are always steps that can be taken to reduce hedging costs. For most typical insurance products, these cost-reduction measures include such things as higher deductibles, co-pay arrangements, and lower maximum payouts. The same sorts of measures exist in the FX derivatives market; we will explore these various alterative measures in this section. The key point to keep in mind is that all of these various cost-reduction measures invariably involve some combination of *less downside protection* and/or *less upside potential* for the hedge. In efficient markets, lower insurance premiums mean lower insurance.

These cost-reduction measures also start moving the portfolio away from a passively managed 100% hedge ratio toward discretionary hedging in which the manager is allowed to take directional positions. Once the possibility of accepting some downside risk, and some upside potential, is introduced into the portfolio, the manager is moving away from a rules-based approach to hedging toward a more active style of trading. The portfolio manager can then use the trading tools and strategies described in the following sections to express a market view and/or cut hedging costs.

The variety of trading strategies—involving various combinations of forwards, options, and swaps—that can be deployed to this end is almost infinite. We will not attempt to explore all of them in this reading, but rather to give a sense of the range of trading tools and strategies available for managing currency risk. We begin with Exhibit 9, which gives a high-level description of some of these various trading strategies that will then be explained in more detail in subsequent sections. Note that as this section progresses, we will be describing strategies at different points along the risk spectrum described in Section 4, moving in turn from passive hedge-based approaches to strategies used in more active currency management schemes.

Exhibit 9 Select Currency Management Strategies

Forward Contracts	Over-/under-hedging	Profit from market view
Option Contracts	OTM options	Cheaper than ATM
	Risk reversals	Write options to earn premiums
	Put/call spreads	Write options to earn premiums
	Seagull spreads	Write options to earn premiums
Exotic Options	Knock-in/out features	Reduced downside/upside exposure
	Digital options	Extreme payoff strategies

We will make one simplifying assumption for the following sections. Currency management strategies will differ fundamentally depending on whether the base currency of the P/B price quote must be bought or sold to decrease the foreign-currency exposure. To simplify the material and impose consistency on the discussions that follow, we will assume that the portfolio manager must sell the base currency in the P/B quote to reduce currency risk. In addition, unless otherwise noted, the notional amounts and expiration dates on all forward and options contracts are the same.[29]

6.3.1 *Over-/Under-Hedging Using Forward Contracts*

When the IPS gives the manager discretion either to over- or under-hedge the portfolio, relative to the "neutral" benchmark, there is the possibility to add incremental value based on the manager's market view. Profits from successful tactical positioning help reduce net hedging costs. For example, if the neutral benchmark hedge ratio is 100% for the base currency being hedged, and the portfolio manager has a market opinion that the base currency is likely to depreciate, then *over*-hedging through a short position in P/B forward contracts might be implemented—that is, the manager might use a hedge ratio higher than 100%. Similarly, if the manager's market opinion is that the base currency is likely to appreciate, the currency exposure might be *under*-hedged.

A variant of this approach would be to adjust the hedge ratio based on exchange rate movements: to increase the hedge ratio if the base currency depreciated, but decrease the hedge ratio if the base currency appreciated. Essentially, this approach is a form of "delta hedging" that tries to mimic the payoff function of a put option on the base currency. That is, this form of dynamic hedging with forward contracts tries to increasingly participate in any upside moves of the base currency, but increasingly hedge any downside moves. Doing so adds "convexity" to the portfolio, meaning that the hedge's payoff function will be a convex curve when this function is graphed with profit on the vertical axis and the spot rate on the horizontal axis. (Note that this concept of convexity is identical in intent to the concept of convexity describing bonds; as convexity increases the price of a bond rises more quickly in a declining yield environment and drops more slowly in a rising yield environment. Convexity is a desirable characteristic in both the fixed-income and currency-hedging contexts.)

29 Examples of implementing a hedge by *buying* the base currency will be provided in some of the practice examples.

6.3.2 *Protective Put Using OTM Options*

In the previous section, we examined a dynamic hedging strategy using forward contracts that tries to mimic the payoff function of an option and put convexity into the hedge's payoff function. The payoff functions for options are naturally convex to begin with. However, this can be a costly form of convexity (relatively high option premiums), and fully hedging a currency position with a protective put strategy using an ATM option is the most expensive means of all to buy convexity.

One way to reduce the cost of using options is to accept some downside risk by using an OTM option, such as a 25- or 10-delta option. These options will be less costly, but also do not fully protect the portfolio from adverse currency movements. Conversely, it makes sense to insure against larger risks but accept some smaller day-to-day price movements in currencies. As an analogy, it may be possible to buy a home or car insurance policy with a zero deductible—but the premiums would be exorbitant. It may be more rational to have a cheaper insurance policy and accept responsibility for minor events, but insure against extreme damage.

6.3.3 *Risk Reversal (or Collar)*

Another set of option strategies involves *selling* options (also known as writing options) to earn income that can be used to offset the cost of buying a put option, which forms the "core" of the hedge. Recall that in this section, we are using the simplifying convention that the manager is long the base currency in the P/B quote; hence puts and not calls would be used for hedging in this case.

One strategy to obtain downside protection at a lower cost than a straight protective put position is to *buy* an OTM put option and *write* an OTM call option. Essentially, the portfolio manager is selling some of the upside potential for movements in the base currency (writing a call) and using the option's premiums to help pay the cost of the long put option being purchased. This approach is similar to creating a collar in fixed-income markets. The portfolio is protected against downside movements, but its upside is limited to the strike price on the OTM call option; the exchange rate risk is confined to a corridor or "collar."

In professional FX markets, having a long position in a call option and a short position in a put option is called a **risk reversal**. For example, buying a 25-delta call and writing a 25-delta put is referred to as a *long* position in a 25-delta risk reversal. The position used to create the collar position we just described (buying a put, writing a call) would be a *short* position in a risk reversal.

The majority of currency hedging for foreign-currency asset portfolios and corporate accounts is based on the use of forward contracts and simple option strategies (protective puts/covered calls and risk reversals/collars). We now begin to transition to more active trading strategies that are designed to express market views for speculative profit.

6.3.4 *Put Spread*

A variation of the short risk reversal position is a **put spread**, which is also used to reduce the upfront cost of buying a protective put. The short risk reversal is structured by buying a put option and writing a call option: the premiums received by writing the call help cover the cost of the put. Similarly, the put spread position involves buying a put option and writing another put option to help cover the cost of the long put's premiums. This position is typically structured by buying an OTM put, and writing a deeper-OTM put to gain income from premiums; both options involved have the same maturity.

To continue our previous example, with the current spot rate at 1.3550, the portfolio manager might set up the following put spread: buy a put with a strike of 1.3500 and write a put with a strike of 1.3450. The payoff on the put spread position will then be as follows: there is no hedge protection between 1.3550 and 1.3500; the portfolio is hedged from 1.3500 down to 1.3450; at spot rates below 1.3450, the portfolio becomes unhedged again. The put spread reduces the cost of the hedge, but at the cost of more limited downside protection. The portfolio manager would then use this spread only for cases in which a modest decline in the spot exchange rate was expected, and this position would have to be closely monitored against adverse exchange rate movements.[30]

Note that the put spread structure will not be zero-cost because the deeper-OTM put (1.3450) being written will be cheaper than the less-OTM put (1.3500) being bought. However, there are approaches that will make the put spread (or almost any other option spread position) cheaper or possibly zero-cost: the manager could alter: (a) the strike prices of the options; (b) the notional amounts of the options; or (c) some combination of these two measures.

Altering the strike prices of the put options would mean moving them closer together (and hence more equal in cost). However, this would reduce the downside protection on the hedge. Instead, the portfolio manager could write a larger notional amount for the deeper-OTM option; for example, the ratio for the notionals for the options written versus bought might be 1:2. (In standard FX market notation, this would be a 1 × 2 put spread—the option with exercise price closest to being ATM is given first. However, to avoid confusion it is good practice to specify explicitly in the price quote which is the long and short positions, and what their deltas/strike prices are.) Although this structure may now be (approximately) zero-cost it is not without risks: for spot rates below 1.3450 the portfolio has now seen its exposure to the base currency double-up (because of the 1:2 proportion of notionals) and at a worse spot exchange rate for the portfolio on top of it. Creating a zero-cost structure with a 1 × 2 put spread is equivalent to adding leverage to the options position, because you are selling more options than you are buying. This means that this put spread position will have to be carefully managed. For example, the portfolio manager might choose to close out the short position in the deep-OTM put (by going long/buying an equivalent put option) before the base currency depreciates to the 1.3450 strike level. This may be a costly position exit, however, as the market moves against the manager's original positioning. Because of this, this sort of 1 × 2 structure may be more appropriate for expressing directional opinions rather than as a pure hedging strategy.

6.3.5 *Seagull Spread*

An alternative, and somewhat safer approach, would be to combine the original put spread position (1:1 proportion of notionals) with a covered call position. This is simply an extension of the concept behind risk reversals and put spreads. The "core" of the hedge (for a manager long the base currency) is the long position in a put option. This is expensive. To reduce the cost, a short risk reversal position writes a call option while a put spread writes a deep-OTM put option. Of course, the manager can always do both: that is, be long a protective put and then write *both* a call and a deep-OTM put. This option structure is sometimes referred to as a **seagull spread**.

As with the names for other option strategies based on winged creatures, the "seagull" indicates an option structure with at least three individual options, and in which the options at the most distant strikes—the wings—are on the opposite side of the market from the middle strike(s)—the body. For example, if the current spot price is 1.3550, a seagull could be constructed by going *long* an ATM put at 1.3550 (the middle strike is the "body"), *short* an OTM put at 1.3500, and *short* an OTM call

30 Another approach could be to fully hedge the underlying currency exposure with a short position in a forward contract, and then overlay this hedge with a put spread as a *tactical* position designed to profit from modest depreciation of the base currency. The underlying currency exposure is hedged by the forward contract, and the portfolio manager earns a speculative gain if the base currency depreciates enough below the strike price on the long put position to cover the net cost of this option structure.

at 1.3600 (the latter two options are the “wings”). Because the options in the “wings” are being written (sold) this is called a *short* seagull position. The risk/return profile of this structure gives full downside protection from 1.3550 to 1.3500 (at which point the short put position neutralizes the hedge) and participation in the upside potential in spot rate movements to 1.3600 (the strike level for the short call option).

Note that because *two* options are now being written to gain premiums instead of one, this approach allows the strike price of the long put position to be ATM, increasing the downside protection. The various strikes and/or notional sizes of these options (and hence their premiums) can always be adjusted up or down until a zero-cost structure is obtained. However, note that this particular seagull structure gives away some upside potential (the short call position) as well as takes on some downside risk (if the short put position is triggered, it will disable the hedge coverage coming from the long put position). As always, lower structure costs come with some combination of lower downside protection and/or less upside potential.

There are many variants of these seagull strategies, each of which provides a different risk–reward profile (and net cost). For example, for the portfolio manager wishing to hedge a long position in the base currency in the P/B quote when the current spot rate is 1.3550, another seagull structure would be to write an ATM call at 1.3550 and use the proceeds to buy an OTM put option at 1.3500 and an OTM call option at 1.3600. Note that in this seagull structure, the “body” is now a *short* option position, not a long position as in the previous example, and the “wings” are the long position. Hence, it is a *long* seagull spread. This option structure provides cheap downside protection (the hedge kicks in at the put’s 1.3500 strike) while providing the portfolio manager with unlimited participation in any rally in the base currency beyond the 1.3600 strike of the OTM call option. As before, the various option strikes and/or notional sizes on the options bought and written can be adjusted so that a zero-cost structure is obtained.

6.3.6 *Exotic Options*

In this section, we move even further away from derivatives and trading strategies used mainly for hedging, and toward the more speculative end of the risk spectrum dominated by active currency management. Exotic options are often used by more sophisticated players in the professional trading market—for example, currency overlay managers—and are less frequently used by investment funds or corporations for hedging purposes.[31]

In general, the term “exotic” refers to all options that are not “vanilla.” In FX, vanilla refers essentially to European-style put and call options. The full range of exotic options is both very broad and constantly evolving; many are extraordinarily complex both to price and even to understand. However, all exotics, no matter how complex, typically share one defining feature in common: They are designed to customize the risk exposures desired by the client and provide them at the lowest possible price.[32] Much like the trading strategies described previously, they usually involve some combination of lower downside protection and/or lower upside potential while providing the client with the specific risk exposures they are prepared to manage, and to do so at what is generally a lower cost than vanilla options.

The two most common type of exotic options encountered in foreign exchange markets are those with **knock-in/knock-out** features and digital options.

---

31 There are several reasons for this relatively light usage of “exotics” for hedging purposes, some related to the fact that many smaller entities lack familiarity with these products. Another reason involves the difficulty of getting hedge accounting treatment in many jurisdictions, which is more advantageous for financial reporting reasons. Finally, the specialized terms of such instruments make them difficult to value for regulatory and accounting purposes.

32 Although the price is low for the client compared with vanilla options, exotics are typically nonetheless high profit margin items for investment dealers.

An option with a knock-in feature is essentially a vanilla option that is created only when the spot exchange rate touches a pre-specified level (this trigger level, called the "barrier," is not the same as the strike price). Similarly a knock-out option is a vanilla option that ceases to exist when the spot exchange rate touches some pre-specified barrier level. Because these options only exist (i.e., get knocked-in or knocked-out) under certain circumstances, they are more restrictive than vanilla options and hence are cheaper. But again, the knock-in/out features provide less upside potential and/or downside protection.

Digital options are also called binary options, or all-or-nothing options. They are called this because they pay a *fixed* amount if they "touch" their exercise level at any time before expiry (even if by a single pip). This characteristic of "extreme payoff" options makes them almost akin to a lottery ticket. Because of these large payoffs, digital options usually cost more than vanilla options with the same strike price. But digitals also provide highly leveraged exposure to movements in the spot rate. This makes these exotic products more appropriate as trading tools for active currency management, rather than as hedging tools. In practice, digital options are typically used by more sophisticated speculative accounts in the FX market to express directional views on exchange rates.

A full exposition of exotic options is beyond the scope of this reading, but the reader should be aware of their existence and why they exist.

6.3.7 *Section Summary*

Clearly, loosening the constraint of a fully hedged portfolio begins to introduce complicated active currency management decisions. The following steps can be helpful to sort things out:

a First, identify the *base* currency in the P/B quote (currency pair) you are dealing with. Derivatives are typically quoted in terms of either buying or selling the *base* currency when the option is exercised. A move upward in the P/B quote is an appreciation of the base currency.

b Then, identify whether the base currency must be *bought* or *sold* to establish the hedge. These are the price movements you will be protecting against.

c If *buying* the base currency is required to implement the hedge, then the core hedge structure will be based on some combination of a long call option and/or a long forward contract. The cost of this core hedge can be reduced by buying an OTM call option or writing options to earn premiums. (But keep in mind, lower hedging costs equate to less downside protection and/or upside potential.)

d If *selling* the base currency is required to implement the hedge, then the core hedge structure will be based on some combination of a long put option and/or a short forward contract. The cost of this core hedge can be reduced by buying an OTM put option or writing options to earn premiums.

e The higher the allowed discretion for active management, the lower the risk aversion; and the firmer a particular market view is held, the more the hedge is likely to be structured to allow risk exposures in the portfolio. This approach involves positioning in derivatives that "lean the same way" as the market view. (For example, a market view that the base currency will depreciate would use some combination of short forward contracts, writing call options, buying put options, and using "bearish" exotic strategies.) This directional bias to the trading position would be superimposed on the core hedge position described in steps "c" and "d," creating an active-trading "tilt" in the portfolio.

f For these active strategies, varying the strike prices and notional amounts of the options involved can move the trading position toward a zero-cost structure. But as with hedges, keep in mind that lower cost implies less downside protection and/or upside potential for the portfolio.

A lot of different hedging tools and strategies have been named and covered in this section. Rather than attempting to absorb all of them by rote memorization (a put spread is "X" and a seagull is "Y"), the reader is encouraged instead to focus on the intuition behind a hedge, and how and why it is constructed. It matters less what name (if any) is given to any specific approach; what is important is understanding how all the moving parts fit together. The reader should focus on a "building blocks" approach in understanding how and why the parts of the currency hedge are assembled in a given manner.

EXAMPLE 6

Alternative Hedging Strategies

Brixworth & St. Ives Asset Management, the UK-based investment firm, has hedged the exposure of its Aggressive Growth Fund to the MXN with a long position in a MXN/GBP forward contract. The fund's foreign-currency asset exposure to the ZAR is hedged by buying an ATM call option on the ZAR/GBP currency pair. The portfolio managers at Brixworth & St. Ives are looking at ways to modify the risk–reward trade-offs and net costs of their currency hedges.

Jasmine Khan, one of the analysts at Brixworth & St. Ives, proposes an option-based hedge structure for the long-ZAR exposure that would replace the hedge based on the ATM call option with either long or short positions in the following three options on ZAR/GBP:

a ATM put option

b 25-delta put option

c 25-delta call option

Khan argues that these three options can be combined into a hedge structure that will have some limited downside risk, but provide complete hedge protection starting at the relevant 25-delta strike level. The structure will also have unlimited upside potential, although this will not start until the ZAR/GBP exchange rate moves to the relevant 25-delta strike level. Finally, this structure can be created at a relatively low cost because it involves option writing.

1 The *best* method for Brixworth & St. Ives to gain some upside potential for the hedge on the Aggressive Growth Fund's MXN exposure using MXN/GBP options is to replace the forward contract with a:

A long position in an OTM put.

B short position in an ATM call.

C long position in a 25-delta risk reversal.

2 While keeping the ATM call option in the ZAR/GBP, the method that would lead to *greatest* cost reduction on the hedge would be to:

A buy a 25-delta put.

B write a 10-delta call.

C write a 25-delta call.

3 Setting up Khan's proposed hedge structure would *most likely* involve being:

A long the 25-delta options and short the ATM option.

B long the 25-delta call, and short both the ATM and 25-delta put options.

C short the 25-delta call, and long both the ATM and 25-delta put options.

Solution to 1:

C is correct. The Aggressive Growth Fund has a long foreign-currency exposure to the MXN in its asset portfolio, which is hedged by selling the MXN against the GBP, or equivalently, buying the GBP—the base currency in the P/B quote—against the MXN. This need to protect against an *appreciation* in the GBP is why the hedge is using a *long* position in the forward contract. To set a collar around the MXN/GBP rate, Brixworth & St. Ives would want a long call option position with a strike greater than the current spot rate (this gives upside potential to the hedge) and a short put position with a strike less than the current spot rate (this reduces net cost of the hedge). A long call and a short put defines a long position in a risk reversal.

Choice A is incorrect because, if exercised, buying a put option would increase the fund's exposure to the MXN (sell GBP, buy MXN). Similarly, Choice B is incorrect because, if exercised, the ATM call option would increase the MXN exposure (the GBP is "called" away from the fund at the strike price with MXN delivered). Moreover, although writing the ATM call option would gain some income from premiums, writing options (on their own) is never considered the "best" hedge because the premium income earned is fixed but the potential losses on adverse currency moves are potentially unlimited.

Solution to 2:

C is correct. As before, the hedge is implemented in protecting against an *appreciation* of the base currency of the P/B quote, the GBP. The hedge is established with an ATM call option (a long position in the GBP). Writing an OTM call option (i.e., with a strike that is more than the current spot rate of 5.1050) establishes a call spread (although hedge protection is lost if ZAR/GBP expires at or above the strike level). Writing a 25-delta call earns more income from premiums than a deeper-OTM 10-delta call (although the 25-delta call has less hedge protection). Buying an option would increase the cost of the hedge, and a put option on the ZAR/GBP would increase the fund's ZAR exposure if exercised (the GBP is "put" to the counterparty at the strike price and ZAR received).

Solution to 3:

A is correct. Once again, the hedge is based on hedging the need to sell ZAR/buy GBP, and GBP is the base currency in the ZAR/GBP quote. This means the hedge needs to protect against an *appreciation* of the GBP (an appreciation of the ZAR/GBP rate). Based on Khan's description, the hedge provides protection after a certain loss point, which would be a long 25-delta call. Unlimited upside potential after favorable (i.e., down) moves in the ZAR/GBP past a certain level means a long 25-delta put. Getting the low net cost that Khan refers to means that the cost of these two long positions is financed by selling the ATM option. (Together these three positions define a long seagull spread). Choice B is incorrect because although the first two legs of the position are right, a short position in the put does not provide any unlimited upside potential (from a down-move in

ZAR/GBP). Choice C is incorrect because any option-based hedge, given the need to hedge against an up-move in the ZAR/GBP rate, is going to be based on a long call position. C does not contain any of these.

6.4 Hedging Multiple Foreign Currencies

We now expand our discussion to hedging a portfolio with multiple foreign-currency assets. The hedging tools and strategies are very similar to those discussed for hedging a single foreign-currency asset, except now the currency hedge must consider the *correlation* between the various foreign-currency risk exposures.

For example, consider the case of a US-domiciled investor who has exposures to foreign-currency assets in Australia and New Zealand. These two economies are roughly similar in that they are resource-based and closely tied to the regional economy of the Western Pacific, especially the large emerging markets in Asia. As a result, the movements in their currencies are often closely correlated; the USD/AUD and USD/NZD currency pairs will tend to move together. If the portfolio manager has the discretion to take short positions, the portfolio may (for example) possibly have a net long position in the Australian foreign-currency asset and a net short position in the New Zealand foreign-currency asset. In this case, there may be less need to hedge away the AUD and NZD currency exposures separately because the portfolio's long exposure to the AUD is diversified by the short position on the NZD.

6.4.1 *Cross Hedges and Macro Hedges*

A **cross hedge**—also referred to a **proxy hedge**—occurs when a position in one asset (or a derivative based on the asset) is used to hedge the risk exposures of a different asset (or a derivative based on it).[33] Normally, cross hedges are not needed because, as we mentioned earlier, forward contracts and other derivatives are widely available in almost every conceivable currency pair. However, if the portfolio already has "natural" cross hedges in the form of negatively correlated residual currency exposures—as in the long-AUD/short-NZD example in Section 6.4—this helps moderate portfolio risk ($\sigma[R_{DC}]$) without having to use a direct hedge on the currency exposure.

33 Sometimes a distinction is made between a "proxy" hedge and a "cross" hedge. When this distinction is made, a *proxy hed*ge removes the foreign currency risk by hedging it back to the investor's domestic currency—such as in the example with USD/AUD and USD/NZD discussed in the text. In contrast, a *cross hed*ge moves the currency risk from one foreign currency to another foreign currency. For example, a US-domiciled investor may have an exposure to both the Indonesian rupiah (IDR) and the Thai baht (THB), but based on a certain market view, may only want exposure to the THB. In this context, the manager might use currency derivatives as a cross hedge to convert the IDR/USD exposure to a THB/USD exposure. But not all market participants make this sharp of a distinction between proxy hedges and cross hedges, and these terms are often used interchangeably. The most common term found among practitioners in most asset classes is simply a cross hedge, as we are using the term here: hedging an exposure with a closely correlated product (i.e., a proxy hedge when this distinction is made). The cross hedge of moving currency exposures between various *foreign* currencies is more of a special-case application of this concept. In our example, a US investor wanting to shift currency exposures between the IDR and THB would only need to shift the relative size of the IDR/USD and THB/USD forward contracts *already* being used. As mentioned earlier, forwards are available on almost every currency pair, so a cross hedge from foreign currency "A" to foreign currency "B" would be a special case when derivatives on one of the currencies are not available.

EXAMPLE 7

Cross Hedges

Mai Nguyen works at Cape Henlopen Advisors, which runs a US-domiciled fund that invests in foreign-currency assets of Australia and New Zealand. The fund currently has equally weighted exposure to one-year Australian and New Zealand treasury bills (i.e., both of the portfolio weights, $\omega_i = 0.5$). Because the foreign-currency return on these treasury bill assets is risk-free and known in advance, their expected $\sigma(R_{FC})$ is equal to zero.

Nguyen wants to calculate the USD-denominated returns on this portfolio as well as the cross hedging effects of these investments. She collects the following information:

Expected Values	Australia	New Zealand
Foreign-currency asset return R_{FC}	4.0%	6.0%
Foreign-currency return R_{FX}	5.0%	5.0%
Asset risk $\sigma(R_{FC})$	0%	0%
Currency risk $\sigma(R_{FX})$	8.0%	10.0%
Correlation (USD/AUD; USD/NZD)	+0.85	

Using Equation 1, Nguyen calculates that the expected domestic-currency return for the Australian asset is

$$(1.04)(1.05) - 1 = 0.092$$

or 9.2%. Likewise, she determines that the expected domestic-currency return for the New Zealand asset is

$$(1.06)(1.05) - 1 = 0.113$$

or 11.3%. Together, the result is that the expected domestic-currency return (R_{DC}) on the equally weighted foreign-currency asset portfolio is the weighted average of these two individual country returns, or

$$R_{DC} = 0.5(9.2\%) + 0.5(11.3\%) = 10.3\%$$

Nguyen now turns her attention to calculating the portfolio's investment risk $[\sigma(R_{DC})]$. To calculate the expected risk for the domestic-currency return, the currency risk of R_{FX} needs to be multiplied by the *known* return on the treasury bills. The portfolio's investment risk, $\sigma(R_{DC})$, is found by calculating the standard deviation of the right-hand-side of:

$$R_{DC} = (1 + R_{FC})(1 + R_{FX}) - 1$$

Although R_{FX} is a random variable—it is not known in advance—the R_{FC} term is in fact known in advance because the asset return is risk-free. Because of this Nguyen can make use of the statistical rules that, first, $\sigma(kX) = k\sigma(X)$, where X is a random variable and k is a constant; and second, that the correlation between a random variable and a constant is zero. These results greatly simplify the calculations because, in this case, she does not need to consider the correlation between exchange rate movements and foreign-currency asset returns. Instead, Nguyen needs to calculate the risk only on the currency side. Applying these statistical rules to the above formula leads to the following results:

A The expected risk (i.e., standard deviation) of the domestic-currency return for the Australian asset is equal to $(1.04) \times 8\% = 8.3\%$.

B The expected risk (i.e., standard deviation) of the domestic-currency return for the New Zealand asset is equal to $(1.06) \times 10\% = 10.6\%$.

Adding all of these numerical values into Equation 4 leads Nguyen to calculate:

$$\sigma^2(R_{DC}) = (0.5)^2(8.3\%)^2 + (0.5)^2(10.6\%)^2 + [(2)0.5(8.3\%)0.5(10.6\%)0.85]$$
$$= 0.8\%$$

The standard deviation of this amount—that is, $\sigma(R_{DC})$—is 9.1%. Note that in the expression, all of the units are in percent, so for example, 8.3% is equivalent to 0.083 for calculation purposes. The careful reader may also note that Nguyen is able to use an exact expression for calculating the variance of the portfolio returns, rather than the approximate expressions shown in Equations 3 and 5. This is because, with risk-free foreign-currency assets, the variance of these foreign-currency returns $\sigma^2(R_{FC})$ is equal to zero.

Nguyen now considers an alternative scenario in which, instead of an equally weighted portfolio (where the $\omega_i = 0.5$), the fund has a long exposure to the New Zealand asset and a short exposure to the Australian asset (i.e., the ω_i are +1 and −1, respectively; this is similar to a highly leveraged carry trade position). Putting these weights into Equations 2 and 4 leads to

$$R_{DC} = -1.0(9.2\%) + 1.0(11.3\%) = 2.1\%$$

$$\sigma^2(R_{DC}) = (1.0)^2(8.3\%)^2 + (1.0)^2(10.6\%)^2 + [-2.0(8.3\%)(10.6\%)0.85]$$
$$= 0.3\%$$

The standard deviation—that is, $\sigma(R_{DC})$—is now 5.6%, less than either of the expected risks for foreign-currency asset returns (results A and B). Nguyen concludes that having long and short positions in positively correlated currencies can lead to much lower portfolio risk, through the benefits of cross hedging. (Nguyen goes on to calculate that if the expected correlation between USD/AUD and USD/NZD increases to 0.95, with all else equal, the expected domestic-currency return risk on the long–short portfolio drops to 3.8%.)

Some types of cross hedges are often referred to as macro hedges. The reason is because the hedge is more focused on the entire portfolio, particularly when individual asset price movements are highly correlated, rather than on individual assets or currency pairs. Another way of viewing a macro hedge is to see the portfolio not just as a collection of financial assets, but as a collection of risk exposures. These various risk exposures are typically defined in categories, such as term risk, credit risk, and liquidity risk. These risks can also be defined in terms of the potential financial scenarios the portfolio is exposed to, such as recession, financial sector stress, or inflation. Often macro hedges are defined in terms of the financial scenario they are designed to protect the portfolio from.

Putting gold in the portfolio sometimes serves this purpose by helping to provide broad portfolio protection against extreme market events. Using a volatility overlay program can also hedge the portfolio against such risks because financial stress is typically associated with a spike in exchange rates' implied volatility. Using a derivative product based on an index, rather than specific assets or currencies, can also define a macro hedge. One macro hedge specific to foreign exchange markets uses derivatives based on fixed-weight baskets of currencies (such derivatives are available in both exchange-traded and OTC form). In a multi-currency portfolio, it may not always be cost efficient to hedge each single currency separately, and in these situations a macro hedge using currency basket derivatives is an alternative approach.

6.4.2 *Minimum-Variance Hedge Ratio*

A mathematical approach to determining the optimal cross hedging ratio is known as the **minimum-variance hedge ratio**. Recall that regression analysis based on ordinary least squares (OLS) is used to minimize the variance of $\hat{\varepsilon}$, the residual between actual and fitted values of the regression

$$y_t = \alpha + \beta x_t + \varepsilon_t \text{ where } \hat{\varepsilon}_t = y_t - \left(\hat{\alpha} + \hat{\beta} x_t\right)$$

This same principle can be used to minimize the tracking error between the value of the hedged asset and the hedging instrument. In the regression formula, we substitute the percentage change in the value of the asset to be hedged for y_t, and the percentage change in value of the hedging instrument for x_t (both of these values are measured in terms of the investor's domestic currency). The calculated coefficient in this regression ($\hat{\beta}$) gives the optimal hedging ratio, which means it minimizes the variance of $\hat{\varepsilon}$ and minimizes the tracking error between changes in the value of the hedge and changes in the value of the asset it is hedging. It can be shown that the formula for the minimum-variance hedge ratio—the formula for calculating the $\hat{\beta}$ coefficient in the regression—is mathematically equal to:

$$\frac{\text{covariance } (y,x)}{\text{variance } (x)} = \text{correlation } (y,x) \times \left[\frac{\text{std. dev. } (y)}{\text{std. dev. } (x)}\right]$$

where y and x are defined as before, the change in the domestic-currency value of the asset and the hedge, respectively.

Calculating the minimum-variance hedge ratio typically applies only for "indirect" hedges based on cross hedging or macro hedges; it is not typically applied to a "direct" hedge in which exposure to a spot rate is hedged with a forward contract in that same currency pair. This is because the correlation between movements in the spot rate and its forward contract is likely to be very close to +1. Likewise, the variance in spot price movements and movements in the price of the forward contract are also likely to be approximately equal. Therefore, calculating the minimum-variance hedge ratio by regressing changes in the spot exchange rate against changes in the forward rate will almost always result is a $\hat{\beta}$ regression estimate very close to 1, and hence a minimum-variance hedge ratio close to 100%. So, undertaking the regression analysis is superfluous.

But the minimum-variance hedge ratio can be quite different from 100% when the hedge is *jointly* optimized over *both* exchange rate movements R_{FX} and changes in the foreign-currency value of the asset R_{FC}. A sidebar discusses this case.

There can also be cases when the optimal hedge ratio may not be 100% because of the market characteristics of a specific currency pair. For example, a currency pair may not have a (liquid) forward contract available and hence an alternative cross hedging instrument or a macro hedge must be used instead. We examine when such situations might come up in Section 7.

6.4.3 *Basis Risk*

The portfolio manager must be aware that any time a direct currency hedge (i.e., a spot rate hedged against its own forward contract) is replaced with an indirect hedge (cross hedge, macro hedge), **basis risk** is brought into the portfolio. This risk reflects the fact that the price movements in the exposure being hedged and the price movements in the cross hedge instrument are not perfectly correlated, and that the correlation will change with time—and sometimes both dramatically and unexpectedly. For a minimum-variance hedge ratio, this risk is expressed as instability in the $\hat{\beta}$ coefficient estimate as more data become available.

For an example of basis risk, return to the illustration earlier of the foreign-currency asset portfolio that cross hedged a long USD/AUD exposure with a short USD/NZD exposure. It is not only possible, but highly likely, that the correlation between movements in the USD/AUD and the USD/NZD spot rates will vary with time. This varying correlation would reflect movements in the NZD/AUD spot rate. Another example of basis risk would be that the correlation between a multi-currency portfolio's domestic-currency market value and the value of currency basket derivatives being used as a macro hedge will neither be perfect nor constant.

At a minimum, this means that all cross hedges and macro hedges will have to be carefully monitored and, as needed, rebalanced to account for the drift in correlations. It also means that minimum-variance hedge ratios will have to be re-estimated as more data become available. The portfolio manager should beware that sudden, unexpected spikes in basis risk can sometimes turn what was once a minimum-variance hedge or an effective cross hedge into a position that is highly correlated with the underlying assets being hedged—the opposite of a hedge.[34]

OPTIMAL MINIMUM-VARIANCE HEDGES

For simple foreign-currency asset portfolios, it may be possible to use the single-variable OLS regression technique to do a *joint* optimization of the hedge over both the foreign-currency value of the asset R_{FC} and the foreign-currency risk exposure R_{FX}. This approach will reduce the variance of the all-in domestic-currency return R_{DC}, which is the risk that matters most to the investor, not just reducing the variance of the foreign exchange risk R_{FX}.

Calculating the minimum-variance hedge for the foreign exchange risk R_{FX} proceeds by regressing changes in the spot rate against changes in the value of the hedging instrument (i.e., the forward contract). But as indicated in the text, performing this regression is typically unnecessary; for all intents and purposes, the minimum-variance hedge for a spot exchange rate using a forward contract will be close to 100%.

But when there is only a *single* foreign-currency asset involved, one can perform a joint optimization over both of the foreign-currency risks (i.e., both R_{FC} and R_{FX}) by regressing changes in the domestic-currency return (R_{DC}) against percentage changes in the value of the hedging instrument. Basing the optimal hedge ratio on the OLS estimate for β in this regression will minimize the variance of the domestic-currency return $\sigma^2(R_{DC})$. The result will be a better hedge ratio than just basing the regression on R_{FX} alone because this joint approach will also pick up any *correlations* between R_{FX} and R_{FC}. (Recall from Section 4.3.1 that the asset mix in the portfolio, and hence the correlations between R_{FX} and R_{FC}, can affect the optimal hedge ratio.) This single-variable OLS approach, however, will only work if there is a single foreign-currency asset in the portfolio.

Work by Campbell (2010) has shown that the optimal hedge ratio based jointly on movements in R_{FC} and R_{FX} for international *bond* portfolios is almost always close to 100%. However, the optimal hedge ratio for single-country foreign *equity* portfolios varies widely between currencies, and will depend on *both* the investor's domestic currency and the currency of the foreign investment. For example, the optimal hedge ratio for a US equity portfolio will be different for UK and eurozone-based investors; and for

34 Basis risk is also used in the context of forward and futures contracts because the price movements of these derivatives products do not always correspond exactly with those of the underlying currency. This is because the price of the forward contract also reflects the interest rate differential between the two countries in the currency pair as well as the term to contract maturity. But with futures and forwards, the derivatives price converges to the price of the underlying as maturity approaches, which is enforced by arbitrage. This convergence is not the case with cross hedges, which potentially can go disastrously wrong with sudden movements in market risk (price correlations), credit risk, or liquidity risk. We will also make passing reference to a term that one may sometimes come across in financial markets: a "Texas hedge." This is a derogatory term used to refer to a position that doubles-up an exposure rather than hedges it; it is the direct opposite of a hedge. For example, someone who buys both call options and their underlying may have put on a "Texas hedge."

eurozone investors, the optimal hedge ratio for a US equity portfolio can be different from that of a Canadian equity portfolio. The study found that the optimal hedge ratio for foreign equity exposures can vary widely from 100% between countries. But as the author cautions, these optimal hedge ratios are calculated on historical data that may not be representative of future price dynamics.

Minimum-Variance Hedge Ratio Example

Annie McYelland is an analyst at Scotland- based Kilmarnock Capital. Her firm is considering an investment in an equity index fund based on the Swiss Stock Market Index (SMI). The SMI is a market-cap weighted average of the twenty largest and most liquid Swiss companies, and captures about 85% of the overall market capitalization of the Swiss equity market.

McYelland is asked to formulate a currency-hedging strategy. Because this investment involves only one currency pair and one investment (the SMI), she decides to calculate the minimum-variance hedge ratio for the entire risk exposure, not just the currency exposure. McYelland collects 10 years of monthly data on the CHF/GBP spot exchange rate and movements in the Swiss Market Index.

McYelland notes that the GBP is the base currency in the CHF/GBP quote and that the formula for domestic-currency returns (R_{DC}) shown in Equation 1 requires that the domestic currency be the price currency. Accordingly, she starts by inverting the CHF/GBP quote to a GBP/CHF quote ($S_{GBP/CHF}$). Then she calculates the monthly percentage changes for this adjusted currency series ($\%\Delta S_{GBP/CHF}$) as well as for the SMI (%ΔSMI). This allows her to calculate the monthly returns of an unhedged investment in the SMI with these unhedged returns measured in the "domestic" currency, the GBP:

$$R_{DC} = (1 + R_{FC})(1 + R_{FX}) - 1$$

where R_{FC} = %ΔSMI and $R_{FX} = \%\Delta S_{GBP/CHF}$. Because McYelland wants to minimize the variance of these unhedged domestic-currency returns, she calculates the minimum-variance hedge ratio with the following OLS regression:

$$R_{DC} = \alpha + \beta(\%\Delta S_{GBP/CHF}) + \varepsilon$$

The calculated regression coefficients show that $\hat{\alpha} = -0.21$ and $\hat{\beta} = 1.35$. McYelland interprets these results to mean that the estimated $\hat{\beta}$-coefficient is the minimum-variance hedge ratio. This conclusion makes sense because $\hat{\beta}$ represents the sensitivity of the domestic-currency return on the portfolio to percentage changes in the spot rate. In this case, the return on the SMI seems very sensitive to the appreciation of the CHF. Indeed, over the 10 years of data she collected, McYelland notices that the correlation between %ΔSMI and $\%\Delta S_{GBP/CHF}$ is equal to +0.6.

On the basis of these calculations, she recommends that the minimum-variance hedge ratio for Kilmarnock Capital's exposure to the SMI be set at approximately 135%. This recommendation means that a *long* CHF1,000,000 exposure to the SMI should be hedged with a *short* position in CHF against the GBP of approximately CHF1,350,000. Because forward contracts in professional FX markets are quoted in terms of CHF/GBP for this currency pair, this would mean a *long* position in the forward contract ($F_{CHF/GBP}$)—that is, *selling* the CHF means *buying* the base currency GBP.

McYelland cautions the Investment Committee at Kilmarnock Capital that this minimum-variance hedge ratio is only approximate and must be closely monitored because it is estimated over historical data that may not be representative of future price dynamics. For example, the +0.6 correlation estimated between %ΔSMI and $\%\Delta S_{GBP/CHF}$ is the 10-year *average* correlation; future market conditions may not correspond to this historical average.[35]

35 Another reason why this hedge ratio would only be approximate is because it is based on the SMI, which is not a total-return index (i.e., including reinvested dividends). Additionally, some approaches to the minimum-variance hedge ratio incorporate the forward premium or discount in the calculation as well. Given the approximate nature of the minimum-variance hedge ratio to begin with, we ignore these other factors as second-order events.

6.5 Basic Intuitions for Using Currency Management Tools

This section has covered only some of the most common currency management tools and strategies used in FX markets—there are a great many other derivatives products and strategies that have not been covered. The key points are that there are *many* different hedging and active trading strategies, there are many possible *variations* within each of these strategies, and these strategies can be used in *combination* with each other. There is no need to cover all of what would be a very large number of possible permutations and combinations. Instead, we will close this section with a key thought: Each of these many approaches to either hedging or expressing a directional view on currency movements has its advantages and disadvantages, its risks, and its costs.

As a result, there is no single "correct" approach to initiating and managing currency exposures. Instead, at the strategic level, the IPS of the portfolio sets guidelines for risk exposures, permissible hedging tools, and strategies, which will vary among investors. At the tactical level, at which the portfolio manager has discretion on risk exposures, currency strategy will depend on the manager's management style, market view, and risk tolerance. It will also depend on the manager's perceptions of the relative costs and benefit of any given strategy. Market conditions will affect the cost/benefit calculations behind the hedging decision, as movements in forward points (expected roll yield) or exchange rate volatility (option premiums) affect the expected cost of the hedge; the same hedge structure can be "rich" or "cheap" depending on current market conditions.

Reflecting all of these considerations, different managers will likely make different decisions when confronted with the same opportunity set; and each manager will likely have a good reason for their individual decision. The most important point is that the portfolio manager be aware of all the benefits, costs, and risks of the chosen strategy and be comfortable that any remaining residual currency risks in the hedge are acceptable.

To summarize the key insights of Section 6—and continuing our example of a portfolio manager who is long the base currency in the P/B quote and wants to hedge that price risk—the manager needs to understand the following:

1 Because the portfolio has a *long* exposure to base currency, to neutralize this risk the hedge will attempt to build a *short* exposure out of that currency's derivatives using some combination of forward and/or option contracts.

2 A currency hedge is not a free good, particularly a complete hedge. The hedge cost, real or implied, will consist of some combination of lost upside potential, potentially negative roll yield (forward points at a discount or time decay on long option positions), and upfront payments of option premiums.

3 The cost of any given hedge structure will vary depending on market conditions (i.e., forward points and implied volatility).

4 The cost of the hedge is focused on its "core." For a manager with a long exposure to a currency, the cost of this "core" hedge will be the implicit costs of a short position in a forward contract (no upside potential, possible negative roll yield) or the upfront premium on a long position in a put option. Either of these two forms of insurance can be expensive. However, there are various cost mitigation methods that can be used alone or in combination to reduce these core hedging costs:

   a Writing options to gain upfront premiums.

   b Varying the strike prices of the options written or bought.

   c Varying the notional amounts of the derivative contracts.

   d Using various "exotic" features, such as knock-ins or knock-outs.

5 There is nothing inherently wrong with any of these cost mitigation approaches—but the manager *must* understand that these invariably involve some combination of reduced upside potential and/or reduced downside protection. A reduced cost (or even a zero-cost) hedge structure is perfectly acceptable, but only as long as the portfolio manager fully understands all of the residual risks in the hedge structure and is prepared to accept and manage them.

6 There are often “natural” hedges within the portfolio, in which some residual risk exposures are uncorrelated with each other and offer portfolio diversification effects. Cross hedges and macro hedges bring basis risk into the portfolio, which will have to be monitored and managed.

7 There is no single or “best” way to hedge currency risk. The portfolio manager will have to perform a due diligence examination of potential hedge structures and make a rational decision on a cost/benefit basis.

EXAMPLE 8

Hedging Strategies

Ireland-based Old Galway Capital runs several investment trusts for its clients. Fiona Doyle has just finished rebalancing the dynamic currency hedge for Overseas Investment Trust III, which has an IPS mandate to be fully hedged using forward contracts. Shortly after the rebalancing, Old Galway receives notice that one of its largest investors in the Overseas Investment Trust III has served notice of a large withdrawal from the fund.

Padma Bhattathiri works at Malabar Coast Capital, an India-based investment company. Her mandate is to seek out any alpha opportunities in global FX markets and aggressively manage these for speculative profit. The Reserve Bank of New Zealand (RBNZ) is New Zealand’s central bank, and is scheduled to announce its policy rate decision within the week. The consensus forecast among economists is that the RBNZ will leave rates unchanged, but Bhattathiri believes that the RBNZ will surprise the markets with a rate hike.

Jasmine Khan, analyst at UK-based Brixworth & St. Ives Asset Management, has been instructed by the management team to reduce hedging costs for the firm’s Aggressive Growth Fund, and that more currency exposure—both downside risk and upside potential—will have to be accepted and managed. Currently, the fund’s ZAR-denominated foreign-currency asset exposures are being hedged with a 25-delta risk reversal (on the ZAR/GBP cross rate). The current ZAR/GBP spot rate is 13.1350.

Bao Zhang is a market analyst at South Korea–based Kwangju Capital, an investment firm that offers several actively managed investment trusts for its clients. She notices that the exchange rate for the Philippines Peso (PHP/USD) is increasing (PHP is depreciating) toward its 200-day moving average located in the 42.2500 area (the current spot rate is 42.2475). She mentions this to Akiko Takahashi, a portfolio manager for one of the firm’s investment vehicles. Takahashi’s view, based on studying economic fundamentals, is that the PHP/USD rate should continue to increase, but after speaking with Zhang she is less sure. After further conversation, Zhang and Takahashi come to the view that the PHP/USD spot rate will either break through the 42.2500 level and gain upward momentum through the 42.2600 level, or stall at the 42.2500 level and then drop down through the 42.2400 level as frustrated long positions exit the market. They decide that either scenario has equal probability over the next month.

Annie McYelland is an analyst at Scotland-based Kilmarnock Capital. The firm is considering a USD10,000,000 investment in an S&P 500 Index fund. McYelland is asked to calculate the minimum-variance hedge ratio. She collects the following statistics based on 10 years of monthly data:

$\sigma(\%\Delta_{SGBP/USD})$	$\sigma(R_{DC})$	$\rho(R_{DC};\%\Delta S_{GBP/USD})$
2.7%	4.4%	0.2

Source: Data are from Bloomberg.

1 Given the sudden liquidity need announced, Doyle's *best* course of action with regard to the currency hedge is to:

A do nothing.

B reduce the hedge ratio.

C over-hedge by using currency options.

2 Given her market view, Bhattathiri would *most likely* choose which of the following long positions?

A 5-delta put option on NZD/AUD

B 10-delta put option on USD/NZD

C Put spread on JPY/NZD using 10-delta and 25-delta options

3 Among the following, replacing the current risk reversal hedge with a long position in which of the following would *best* meet Khan's instructions? (All use the ZAR/GBP.)

A 10-delta risk reversal

B Put option with a 13.1300 strike

C Call option with a 13.1350 strike

4 Which of the following positions would *best* implement Zhang's and Takahashi's market view?

A Long a 42.2450 put and long a 42.2550 call

B Long a 42.2450 put and short a 42.2400 put

C Long a 42.2450 put and short a 42.2550 call

5 Which of the following positions would *best* implement Kilmarnock Capital's minimum-variance hedge?

A Long a USD/GBP forward contract with a notional size of USD1.2 million

B Long a USD/GBP forward contract with a notional size of USD3.3 million

C Short a USD/GBP forward contract with a notional size of USD2.0 million

Solution to 1:

A is correct. After rebalancing, the Overseas Investment Trust III is fully hedged; currency risk is at a minimum, which is desirable if liquidity needs have increased. Choices B and C are incorrect because they increase the currency risk exposures.

Solution to 2:

A is correct. The surprise rate hike should cause the NZD to appreciate against most currencies. This appreciation would mean a depreciation of the NZD/AUD rate, which a put option can profit from. A 5-delta option is deep-OTM, but the price reaction on the option premiums will be more extreme than a higher-delta option. That is to say, the *percentage* change in the premiums for a 5-delta option for a given percentage change in the spot exchange rate will be higher than the percentage change in premiums for a 25-delta option. In a sense, a very low delta option is like a highly leveraged lottery ticket on the event occurring. With a surprise rate hike, the odds would swing in Bhattathiri's favor. Choice B is incorrect because the price reaction in the USD/NZD spot rate after the surprise rate hike would likely cause the NZD to appreciate; so Bhattathiri would want a call option on the USD/NZD currency pair. Choice C is incorrect because an appreciation of the NZD after the surprise rate hike would best be captured by a call spread on the JPY/NZD rate, which will likely increase (the NZD is the base currency).

Solution to 3:

A is correct. Moving to a 10-delta risk reversal will be cheaper (these options are deeper-OTM than 25-delta options) and widen the bands in the corridor being created for the ZAR/GBP rate. Choice B is incorrect because a long put provides no protection against an upside movement in the ZAR/GBP rate, which Brixworth & St. Ives is trying to hedge (recall that the fund is long ZAR in its foreign-currency asset exposure and hence needs to sell ZAR/buy GBP to hedge). Also, if Brixworth & St. Ives exercises the option, they would "put" GBP to the counterparty at the strike price and receive ZAR in return. Although this option position may be considered profitable in its own right, it nonetheless causes the firm to double-up its ZAR exposure. Choice C is incorrect because although an ATM call option on ZAR/GBP will provide complete hedge protection, it will be expensive and clearly more expensive than the current 25-delta risk reversal.

Solution to 4:

A is correct. Zhang's and Takahashi's market view is that, over the next month, a move in PHP/USD to either 42.2400 or 42.2600 is equally likely. A strangle would express this view of heightened volatility but without a directional bias, and would require a long put and a long call positions. Choice B is incorrect because it is a put spread; it will profit by a move in PHP/USD between 42.2450 and 42.2400. If it moves below 42.2400 the short put gets exercised by the counterparty and neutralizes the long put. Although less costly than an outright long put position, this structure is not positioned to profit from a move higher in PHP/USD. Choice C is incorrect because it is a short risk reversal position. It provides relatively cheap protection for a down-move in PHP/USD but is not positioned to profit from an up-move in PHP/USD.

Solution to 5:

B is correct. The formula for the minimum-variance hedge ratio (h) is:

$$h = \rho\left(R_{DC};R_{FX}\right) \times \left[\frac{\sigma\left(R_{DC}\right)}{\sigma\left(R_{FX}\right)}\right]$$

After inputting the data from the table, this equation solves to 0.33. This means that for a USD10 million investment in the S&P 500 (long position), Kilmarnock Capital would want to be *short* approximately USD3.3 million in a forward contract. Because the standard market quote for this currency pair is USD/GBP, to be short the USD means one would have to buy the GBP; that is, a *long* position in a USD/GBP forward contract. Choice A is incorrect because it inverts the ratio in the formula. Choice C is incorrect because it shows a short position in the USD/GBP forward, and because it only uses the correlation to set the contract size.

7 CURRENCY MANAGEMENT FOR EMERGING MARKET CURRENCIES

Most of the material in this reading has focused on what might be described as the major currencies, such as the EUR, GBP, or JPY. This focus is not a coincidence: The vast majority of daily flow in global FX markets is accounted for by the top half dozen currencies. Moreover, the vast majority of investable assets globally, as measured by market capitalization, are denominated in the major currencies. Nonetheless, more investors are looking at emerging markets, as well as "frontier markets," for potential investment opportunities. And many developing economies are beginning to emerge as major forces in the global economy. In the following sections, we survey the challenges for currency management and the use of non-deliverable forwards as one tool to address them.

7.1 Special Considerations in Managing Emerging Market Currency Exposures

Managing emerging market currency exposure involves unique challenges. Perhaps the two most important considerations are (1) higher trading costs than the major currencies under "normal" market conditions, and (2) the increased likelihood of extreme market events and severe illiquidity under stressed market conditions.

Many emerging market currencies are thinly traded, causing higher transaction costs (bid–offer spreads). There may also be fewer derivatives products to choose from, especially exchange-traded products. Although many global investment banks will quote spot rates and OTC derivatives for almost any conceivable currency pair, many of these are often seen as "specialty" products and often come with relatively high mark-ups. This mark-up increases trading and hedging costs. (In addition, the underlying foreign-currency asset in emerging markets can be illiquid and lack the full array of derivatives products.)

These higher currency trading costs would especially be the case for "crosses" in these currency pairs. For example, there is no reason why an investor in Chile (which uses the Chilean peso, currency code CLP) could not have an investment in assets denominated in the Thai baht (THB). But the CLP/THB cross is likely to be very thinly traded; there simply are not enough trade or capital flows between these two countries. Typically, any trade between these two currencies would go through a major intermediary currency, usually the USD. Hence, the trade would be broken into two legs: a trade in the CLP/USD pair and another in the THB/USD pair. These trades might go through different traders or trading desks at the same bank; or perhaps one leg of the trade would be done at one bank and the other leg through a different bank. There may also be time zone issues affecting liquidity; one leg of the trade may be relatively liquid at the same time as the other leg of the trade may be more thinly traded. The reason is because liquidity in most emerging market currencies is typically deepest in their domestic time zones. In any event, there are two bid–offer spreads—one for each leg of the trade—to be covered.[36]

The liquidity issue is especially important when trades in these less-liquid currencies get "crowded," for example, through an excessive build-up of carry trades or through a fad-like popularity among investors for investing in a particular region or trading theme. Trades can be much easier to gradually enter into than to quickly exit, particularly under stressed market conditions. For example, after a long period of slow build-up, carry trades into these currencies can occasionally be subject to panicked unwinds as market conditions suddenly turn. This situation typically causes market liquidity to evaporate and leaves traders locked into positions that continue to accumulate losses.

The investment return probability distributions for currency (and other) trades subject to such relatively frequent extreme events have fatter tails than the normal distribution as well as a pronounced negative skew.[37] Risk measurement and control tools (such as value at risk, or VaR) that depend on normal distributions can be misleading under these circumstances and greatly understate the risks the portfolio is actually exposed to. Many investment performance measures are also based on the normal distribution. Historical investment performance measured by such indexes as the Sharpe ratio can look very attractive during times of relative tranquility; but this seeming outperformance can disappear into deep losses faster than most investors can react (investors typically do a poor job of timing crises). As mentioned in the prior section on volatility trading, price volatility in financial markets is very cyclical and implied volatility can be subject to sharp spikes. These volatility spikes can severely affect both option prices and hedging strategies based on options.[38] Even if the initial option protection is in place, it will eventually have to be rolled as options expire—but then at much higher prices for the option buyer.

The occurrence of currency crises can also affect hedging strategies based on forward contracts. Recall that hedging a long exposure to a foreign currency typically involves selling the foreign currency forward. However, when currencies are under severe downward pressure, central banks often react by hiking the policy rate to support the domestic currency. But recall that the higher interest rates go in a country, then, all else equal, the deeper the forward discount for its currency (enforced by the arbitrage conditions of covered interest rate parity). Having to sell the currency forward at increasingly deep discounts will cause losses through negative roll yield and undermine the cost effectiveness of the hedging program.

Extreme price movements in financial markets can also undermine many hedging strategies based on presumed diversification. Crises not only affect the volatility in asset prices but also their correlations, primarily through "contagion" effects. The history of financial markets (circa 2012) has been characterized by a "risk-on, risk-off" environment dominated by swings in investor sentiment between speculative enthusiasm and pronounced flight-to-safety flows. In the process, there is often little differentiation between individual currencies, which tend to get traded together in broader baskets (such as "haven currencies"—USD, JPY, and CHF—and "commodity currencies"—AUD, NZD, and ZAR). Investors who may have believed that they had diversified their portfolio through a broad array of exposures in emerging markets

36 This is often the case for many of the cross rate currency pairs among developed market currencies as well. However, the bid–offer spreads are usually tighter for the major currency pairs.

37 Relatively frequent at least compared with that contained in the tails of a normal distributions whose variance is estimated over recent historical data. One is struck by how often "once in a lifetime" events seem to have occurred in the past 20 years.

38 Option pricing models are based on some underlying process of price diffusion for the underlying. In simple option pricing models, this diffusion process is usually assumed to be relatively "smooth." It is more difficult to model option pricing when there are extreme events in the price diffusion process.

may find instead in crises that they doubled-up their currency exposures. (Likewise, there can be correlated and extreme movements in the underlying assets of these foreign-currency exposures.)

Another potential factor affecting currency management in these "exotic" markets is government involvement in setting the exchange rate through such measures as foreign exchange market intervention, capital controls, and pegged (or at least tightly managed) exchange rates. These measures too can lead to occasional extreme events in markets; for example, when central banks intervene or when currency pegs change or get broken. Short-term stability in these government-influenced markets can lull traders into a false sense of overconfidence and over-positioning. When currency pegs break, the break can happen quickly. Assuming that investment returns will be normally distributed according to parameters estimated on recent historical data, or that correlation factors and liquidity will not change suddenly, can be lethal.

It bears noting that currency crises and government involvement in FX markets is not limited to emerging market currencies, but often occur among the major currencies as well. The central banks of major currencies will, on occasion, intervene in their own currencies or use other polices (such as sharp movements in policy rates) to influence exchange rate levels. These too can lead to extreme events in currency markets.

7.2 Non-Deliverable Forwards

Currencies of many emerging market countries trade with some form of capital controls. Where capital controls exist and delivery in the controlled currency is limited by the local government, it is often possible to use what are known as **non-deliverable forwards** (NDFs). These are similar to regular forward contracts, but they are cash settled (in the non-controlled currency of the currency pair) rather than physically settled (the controlled currency is neither delivered nor received). The non-controlled currency for NDFs is usually the USD or some other major currency. A partial list of some of the most important currencies with NDFs would include the Chinese yuan (CNY), Korean won (KRW), Russian ruble (RUB), Indian rupee (INR), and Brazilian real (BRL). The NDF is essentially a cash-settled "bet" on the movement in the spot rate of these currencies.

For example, a trader could enter into a long position in a three-month NDF for the BRL/USD. Note that the BRL—the currency with capital controls—is the price currency and the base currency, the USD, is the currency that settlement of the NDF will be made in. Assume that the current all-in rate for the NDF is 2.0280 and the trader uses an NDF with a notional size of USD1,000,000. Suppose that three months later the BRL/USD spot rate is 2.0300 and the trader closes out the existing NDF contract with an equal and offsetting spot transaction at this rate. Settlement proceeds by noting that the USD amounts net to zero (USD1,000,000 both bought and sold on settlement date), so the net cash flow generated would normally be in BRL if this was an ordinary forward contract. The net cash flow to the long position in this case would be calculated as

$$(2.0300 - 2.0280) \times 1{,}000{,}000 = \text{BRL}2{,}000$$

But with an NDF, there is no delivery in the controlled currency (hence the name *non-deliverable* forward). Settlement must be in USD, so this BRL amount is converted to USD at the then-current spot rate of 2.0300. This leads to a USD cash inflow for the long position in the NDF of

$$\text{BRL}2{,}000 \div 2.0300 \text{ BRL/USD} = \text{USD}985.22$$

The credit risk of an NDF is typically lower than for the outright forward because the principal sums in the NDF do not move, unlike with an outright "vanilla" forward contract. For example, in the illustration the cash pay-off to the "bet" was the relatively small amount of USD985.22—there was no delivery of USD1,000,000 against receipt of BRL2,028,000. Conversely, as noted previously, NDFs exist because of some form of government involvement in foreign exchange markets. Sudden changes in government policy can lead to sharp movements in spot and NDF rates, often reversing any investment gains earned during long periods of seeming (but artificial) market calm. The implicit market risk of the NDF embodies an element of "tail risk."

Finally, we note that when capital controls exist, the free cross-border flow of capital that enforces the arbitrage condition underlying covered interest rate parity no longer functions consistently. Therefore, the pricing on NDFs need not be exactly in accord with the covered interest rate parity theorem. Instead, NDF pricing will reflect the individual supply and demand conditions (and risk premia) in the offshore market, which need not be the same as the onshore market of the specific emerging market country. Some of the most active participants in the NDF market are offshore hedge funds and proprietary traders making directional bets on the emerging market currency, rather than corporate or institutional portfolio managers hedging currency exposures. Volatility in the net speculative demand for emerging market exposure can affect the level of forward points. We also note that the type and strictness of capital controls can vary among emerging markets; hence, the need for knowledge of local market regulations is another factor influencing currency risk management in these markets.

SUMMARY

In this reading, we have examined the basic principles of managing foreign exchange risk within the broader investment process. International financial markets create a wide range of opportunities for investors, but they also create the need to recognize, measure, and control exchange rate risk. The management of this risk starts with setting the overall mandate for the portfolio, encoding the investors' investment objectives and constraints into the investment policy statement and providing strategic guidance on how currency risk will be managed in the portfolio. It extends to tactical positioning when portfolio managers translate market views into specific trading strategies within the overall risk management guidelines set by the IPS. We have examined some of these trading strategies, and how a range of portfolio management tools—positions in spot, forward, option, and FX swap contracts—can be used either to hedge away currency risk, or to express a market opinion on future exchange rate movements.

What we have emphasized throughout this reading is that there is no simple or single answer for the "best" currency management strategies. Different investors will have different strategic mandates (IPS), and different portfolio managers will have different market opinions and risk tolerances. There is a near-infinite number of possible currency trading strategies, each with its own benefits, costs, and risks. Currency risk management—both at the strategic and tactical levels—means having to manage the trade-offs between all of these various considerations.

Some of the main points covered in this reading are as follows:

- In professional FX markets, currencies are identified by standard three-letter codes, and quoted in terms of a price and a base currency (P/B).

- The spot exchange rate is typically for $T + 2$ delivery, and forward rates are for delivery for later periods. Both spot and forward rates are quoted in terms of a bid–offer price. Forward rates are quoted in terms of the spot rate plus forward points.
- An FX swap is a simultaneous spot and forward transaction; one leg of the swap is buying the base currency and the other is selling it. FX swaps are used to renew outstanding forward contracts once they mature, to "roll them forward."
- The domestic-currency return on foreign-currency assets can be broken into the foreign-currency asset return and the return on the foreign currency (the percentage appreciation or depreciation of the foreign currency against the domestic currency). These two components of the domestic-currency return are multiplicative.
- When there are several foreign-currency assets, the portfolio domestic-currency return is the weighted average of the individual domestic-currency returns (i.e., using the portfolio weights, which should sum to one)
- The risk of domestic-currency returns (its standard deviation) can be approximated by using a variance formula that recognizes the individual variances and covariances (correlations) among the foreign-currency asset returns and exchange rate movements.
- The calculation of the domestic-currency risk involves a large number of variables that must be estimated: the risks and correlations between all of the foreign-currency asset returns and their exchange rate risks.
- Guidance on where to target the portfolio along the risk spectrum is part of the IPS, which makes this a *strategic* decision based on the investment goals and constraints of the beneficial owners of the portfolio.
- If the IPS allows currency risk in the portfolio, the amount of desired currency exposure will depend on both portfolio diversification considerations and cost considerations.
  - Views on the diversifying effects of foreign-currency exposures depend on the time horizon involved, the type of foreign-currency asset, and market conditions.
  - Cost considerations also affect the hedging decision. Hedging is not free: It has both direct transactional costs as well as opportunity costs (the potential for favorable outcomes is foregone). Cost considerations make a perfect hedge difficult to maintain.
- Currency management strategies can be located along a spectrum stretching from:
  - passive, rules-based, complete hedging of currency exposures;
  - discretionary hedging, which allows the portfolio manager some latitude on managing currency exposures;
  - active currency management, which seeks out currency risk in order to manage it for profit; and to
  - currency overlay programs that aggressively manage currency "alpha."
- There are a variety of methods for forming market views.
  - The use of macroeconomic fundamentals to predict future currency movements is based on estimating the "fair value" for a currency with the expectation that spot rates will eventually converge on this equilibrium value.

- Technical market indicators assume that, based on market psychology, historical price patterns in the data have a tendency to repeat. Technical indicators can be used to predict support and resistance levels in the market, as well as to confirm market trends and turning points.
- The carry trade is based on violations of uncovered interest rate parity, and is also based on selling low-yield currencies in order to invest in high-yield currencies. This approach is equivalent to trading the forward rate bias, which means selling currencies trading at a forward premium and buying currencies trading at a forward discount.
- Volatility trading uses the option market to express views on the distribution of future exchange rates, not their levels.

- Passive hedging will typically use forward contracts (rather than futures contracts) because they are more flexible. However, currency futures contracts are an option for smaller trading sizes and are frequently used in private wealth management.
- Forward contracts have the possibility of negative roll yield (the forward points embedded in the forward price can work for or against the hedge). The portfolio manager will have to balance the advantages and costs of hedging with forward contracts.
- Foreign-currency options can reduce opportunity costs (they allow the upside potential for favorable foreign-currency movements). However, the upfront option premiums must be paid.
- There are a variety of means to reduce the cost of the hedging with either forward or option contracts, but these cost-reduction measures always involve some combination of less downside protection and/or less upside potential.
- Hedging multiple foreign currencies uses the same tools and strategies used in hedging a single foreign-currency exposure; except now the correlation between residual currency exposures in the portfolio should be considered.
- Cross hedges introduce basis risk into the portfolio, which is the risk that the correlation between exposure and its cross hedging instrument may change in unexpected ways. Forward contracts typically have very little basis risk compared with movements in the underlying spot rate.
- The number of trading strategies that can be used, for hedging or speculative purposes, either for a single foreign currency or multiple foreign currencies, is near infinite. The manager must assess the costs, benefits, and risks of each in the context of the investment goals and constraints of the portfolio. There is no single "correct" approach.

REFERENCES

Bank for International Settlements. 2013. “Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity.”

Campbell, John Y. 2010. “Global Currency Hedging: What Role Should Foreign Currency Play in a Diversified Investment Portfolio?” *CFA Institute Conference Proceedings Quarterly*, vol. 27, no. 4 (December):8–18.

Campbell, John Y., Karine Serfaty-de Medeiros, and Luis. M. Viceira. 2007. “Global Currency Hedging,” NBER Working Paper 13088 (May).

Darnell, R. Max. 2004. “Currency Strategies to Enhance Returns.” In *Fixed-Income Tools for Enhancing Return and Meeting Client Objectives*. Charlottesville, VA: Association for Investment Management and Research.

Froot, Kenneth A. 1993. “Currency Hedging Over Long Horizons.” NBER Working Paper 4355 (April): www.people.hbs.edu/kfroot/oldwebsite/cvpaperlinks/currency_hedging.pdf.

Hnatkovska, Viktoria, and Martin Evans. 2005. “International Capital Flows in a World of Greater Financial Integration.” NBER Working Paper 11701 (October).

Kritzman, Mark P. 1999. “The Forward-Rate Bias.” In *Currency Risk in Investment Portfolios*. Charlottesville, VA: Association for Investment Management and Research.

Martini, Giulio. 2010. “The Continuum from Passive to Active Currency Management.” *CFA Institute Conference Proceedings Quarterly*, vol. 27, no. 1 (March):1–11.

Michenaud, Sébastien, and Bruno Solnik. 2008. “Applying Regret Theory to Investment Choices: Currency Hedging Decisions.” *Journal of International Money and Finance*, vol. 27, no. 5:677–694.

Rosenberg, Michael R. 2002. *Deutsche Bank Guide to Exchange-Rate Determination*. London: Irwin Professional Publishing (May).

Rosenberg, Michael R., and William A. Barker. 2017. “Currency Exchange Rates: Understanding Equilibrium Value.” CFA Program Curriculum, Level II.

Sine, Barry M., and Robert A. Strong. 2012. “Technical Analysis.” CFA Program Curriculum, Level I.

US Department of the Treasury. 2007. “Semiannual Report on International Economic and Exchange Rate Policies, Appendix I.” (December).

PRACTICE PROBLEMS

The following information relates to Questions 1–9

Kamala Gupta, a currency management consultant, is hired to evaluate the performance of two portfolios. Portfolio A and Portfolio B are managed in the United States and performance is measured in terms of the US dollar (USD). Portfolio A consists of British pound (GBP) denominated bonds and Portfolio B holds euro (EUR) denominated bonds.

Gupta calculates a 19.5% domestic-currency return for Portfolio A and 0% domestic-currency return for Portfolio B.

1 **Analyze** the movement of the USD against the foreign currency for Portfolio A. **Justify** your choice.

Template for Question 1

Asset	Foreign-Currency Portfolio Return	USD Relative to Foreign-Currency (circle one)
Portfolio A	15%	appreciated depreciated

Justification

2 **Analyze** the foreign-currency return for Portfolio B. **Justify** your choice.

Template for Question 2

Asset	Percentage Movement in the Spot Exchange Rate	Foreign-Currency Portfolio Return (circle one)
Portfolio B	EUR appreciated 5% against the USD	positive negative

(continued)

© 2013 CFA Institute. All rights reserved.

(Continued)

Justification

The fund manager of Portfolio B is evaluating an internally-managed 100% foreign-currency hedged strategy.

3 **Discuss** *two* forms of trading costs associated with this currency management strategy.

Gupta tells the fund manager of Portfolio B:

> "We need to seriously consider the potential costs associated with favorable currency rate movements, given that a 100% hedge-ratio strategy is being applied to this portfolio."

4 **Explain** Gupta's statement in light of the strategic choices in currency management available to the portfolio manager.

The investment policy statement (IPS) for Portfolio A provides the manager with discretionary authority to take directional views on future currency movements. The fund manager believes the foreign currency assets of the portfolio could be fully hedged internally. However, the manager also believes existing firm personnel lack the expertise to actively manage foreign-currency movements to generate currency alpha.

5 **Recommend** a solution that will provide the fund manager the opportunity to earn currency alpha through active foreign exchange management.

Gupta and the fund manager of Portfolio A discuss the differences among several active currency management methods.

6 **Evaluate** each statement independently and select the active currency approach it *best* describes. **Justify** each choice.

Template for Question 6

Gupta's Statements	Active Currency Approach (circle one)	Justification
"Many traders believe that it is not necessary to examine factors like the current account deficit, inflation, and interest rates because current exchange rates already reflect the market view on how these factors will affect future exchange rates."	carry trade technical analysis economic fundamental	
"The six-month interest rate in India is 8% compared to 1% in the United States. This presents a yield pick-up opportunity."	carry trade technical analysis economic fundamental	
"The currency overlay manager will estimate the fair value of the currencies with the expectation that observed spot rates will converge to long-run equilibrium values described by parity conditions."	carry trade technical analysis economic fundamental	

The following information is used for Question 7

Gupta interviews a currency overlay manager on behalf of Portfolio A. The foreign currency overlay manager describes volatility-based trading, compares volatility-based trading strategies and explains how the firm uses currency options to establish positions in the foreign exchange market. The overlay manager states:

Statement 1 "Given the current stability in financial markets, several traders at our firm take advantage of the fact that most options expire out-of-the money and therefore are net-short volatility."

Statement 2 "Traders that want to minimize the impact of unanticipated price volatility are net-long volatility."

7 **Compare** Statement 1 and Statement 2 and **identify** which *best* explains the view of a speculative volatility trader and which best explains the view of a hedger of volatility. **Justify** your response.

The following information is used for Questions 8 and 9

The fund manager of Portfolio B believes that setting up a full currency hedge requires a simple matching of the *current* market value of the foreign-currency exposure in the portfolio with an equal and offsetting position in a forward contract.

8 **Explain** how the hedge, as described by the fund manager, will eventually expose the portfolio to currency risk.

9 **Recommend** an alternative hedging strategy that will keep the hedge ratio close to the target hedge ratio. **Identify** the main disadvantage of implementing such a strategy.

The following information relates to Questions 10–15

Guten Investments GmbH, based in Germany and using the EUR as its reporting currency, is an asset management firm providing investment services for local high net worth and institutional investors seeking international exposures. The firm invests in the Swiss, UK, and US markets, after conducting fundamental research in order to select individual investments. Exhibit 1 presents recent information for exchange rates in these foreign markets.

Exhibit 1 Exchange Rate Data

	One Year Ago	Today
Euro-dollar (USD/EUR)*	1.2730	1.2950
Euro-sterling (GBP/EUR)	0.7945	0.8050
Euro-Swiss (CHF/EUR)	1.2175	1.2080

* The amount of USD required to buy one EUR

In prior years, the correlation between movements in the foreign-currency asset returns for the USD-denominated assets and movements in the exchange rate was estimated to be +0.50. After analyzing global financial markets, Konstanze Ostermann, a portfolio manager at Guten Investments, now expects that this correlation will increase to +0.80, although her forecast for foreign-currency asset returns is unchanged.

Ostermann believes that currency markets are efficient and hence that long-run gains cannot be achieved from active currency management, especially after netting out management and transaction costs. She uses this philosophy to guide hedging decisions for her discretionary accounts, unless instructed otherwise by the client.

Ostermann is aware, however, that some investors hold an alternative view on the merits of active currency management. Accordingly, their portfolios have different investment guidelines. For these accounts, Guten Investments employs a currency specialist firm, Umlauf Management, to provide currency overlay programs specific to each client's investment objectives. For most hedging strategies, Umlauf Management develops a market view based on underlying fundamentals in exchange rates. However, when directed by clients, Umlauf Management uses options and a variety of trading strategies to unbundle all of the various risk factors (the "Greeks") and trade them separately.

Ostermann conducts an annual review for three of her clients and gathers the summary information presented in Exhibit 2.

Exhibit 2 Select Clients at Guten Investments

Client	Currency Management Objectives
Adele Kastner – A high net worth individual with a low risk tolerance.	Keep the portfolio's currency exposures close, if not equal to, the benchmark so that the domestic-currency return is equal to the foreign-currency return.
Braunt Pensionskasse – A large private-company pension fund with a moderate risk tolerance.	Limited discretion which allows the actual portfolio currency risk exposures to vary plus-or-minus 5% from the neutral position.
Franz Trading GmbH – An exporting company with a high risk tolerance.	Discretion with respect to currency exposure is allowed in order to add alpha to the portfolio.

10 Based on Exhibit 1, the domestic-currency return over the last year (measured in EUR terms) was *higher* than the foreign-currency return for:

A USD-denominated assets.

B GBP-denominated assets.

C CHF-denominated assets.

11 Based on Ostermann's correlation forecast, the expected domestic-currency return (measured in EUR terms) on USD-denominated assets will *most* likely:

A increase.

B decrease.

C remain unchanged.

12 Based on Ostermann's views regarding active currency management, the percentage of currency exposure in her discretionary accounts that is hedged is *most likely:*

A 0%.

B 50%.

C 100%.

13 The active currency management approach that Umlauf Management is *least* likely to employ is based on:

A volatility trading.

B technical analysis.

C economic fundamentals.

14 Based on Exhibit 2, the currency overlay program *most* appropriate for Braunt Pensionskasse would:

A be fully passive.

B allow limited directional views.

C actively manage foreign exchange as an asset class.

15 Based on Exhibit 2, the client *most likely* to benefit from the introduction of an additional overlay manager is:

A Adele Kastner.

B Braunt Pensionskasse.

C Franz Trading GmbH.

The following information relates to Questions 16–19

Li Jiang is an international economist operating a subscription website through which she offers financial advice on currency issues to retail investors. One morning she receives four subscriber e-mails seeking guidance.

Subscriber 1	"As a French national now working in the United States, I hold US dollar-denominated assets currently valued at USD 700,000. The USD/EUR exchange rate has been quite volatile and now appears oversold based on historical price trends. With my American job ending soon, I will return to Europe. I want to protect the value of my USD holdings, measured in EUR terms, before I repatriate these funds back to France. To reduce my currency exposure I am going to use currency futures contracts. Can you explain the factors most relevant to implementing this strategy?"
Subscriber 2	"I have observed that many of the overseas markets for Korean export goods are slowing, while the United States is experiencing a rise in exports. Both trends can combine to possibly affect the value of the won (KRW) relative to the US dollar. As a result, I am considering a speculative currency trade on the KRW/USD exchange rate. I also expect the volatility in this exchange rate to increase."
Subscriber 3	"India has relatively high interest rates compared to the United States and my market view is that this situation is likely to persist. As a retail investor actively trading currencies, I am considering borrowing in USD and converting to the Indian rupee (INR). I then intend to invest these funds in INR-denominated bonds, but without using a currency hedge."
Subscriber 4	"I was wondering if trading in emerging market currencies provides the more opportunities for superior returns through active management than trading in Developed Market currencies."

16 For Subscriber 1, the *most* significant factor to consider would be:

A margin requirements.

B transaction costs of using futures contracts.

C different quoting conventions for future contracts.

17 For Subscriber 2, and assuming all of the choices relate to the KRW/USD exchange rate, the *best* way to implement the trading strategy would be to:

A write a straddle.

B buy a put option.

C use a long NDF position.

18 Which of the following market developments would be *most* favorable for Subscriber 3's trading plan?

A A narrower interest rate differential.

B A higher forward premium for INR/USD.

C Higher volatility in INR/USD spot rate movements.

19 Jiang's *best* response to Subscriber 4 would be that active trading in trading in emerging market currencies:

A typically leads to return distributions that are positively skewed.

B should not lead to higher returns because FX markets are efficient.

C often leads to higher returns through carry trades, but comes with higher risks and trading costs.

The following information relates to Questions 20–23

Rika Björk runs the currency overlay program at a large Scandinavian investment fund, which uses the Swedish krona (SEK) as its reporting currency. She is managing the fund's exposure to GBP-denominated assets, which are currently hedged with a GBP 100,000,000 forward contract (on the SEK/GBP cross rate, which is currently at 10.6875 spot). The maturity for the forward contract is December 1, which is still several months away. However, since the contract was initiated the value of the fund's assets has declined by GBP 7,000,000. As a result, Björk wants to rebalance the hedge immediately.

Next Björk turns her attention to the fund's Swiss franc (CHF) exposures. In order to maintain some profit potential Björk wants to hedge the exposure using a currency option, but at the same time, she wants to reduce hedging costs. She believes that there is limited upside for the SEK/CHF cross rate.

Björk then examines the fund's EUR-denominated exposures. Due to recent monetary tightening by the Riksbank (the Swedish central bank) forward points for the SEK/EUR rate have swung to a premium. The fund's EUR-denominated exposures are hedged with forward contracts.

Finally Björk turns her attention to the fund's currency exposures in several emerging markets. The fund has large positions in several Latin American bond markets, but Björk does not feel that there is sufficient liquidity in the related foreign exchange derivatives to easily hedge the fund's Latin American bond markets exposures. However, the exchange rates for these countries, measured against the SEK, are correlated with the MXN/SEK exchange rate. (The MXN is the Mexican peso, which is considered to be among the most liquid Latin American currencies). Björk considers using forward positions in the MXN to cross-hedge the fund's Latin American currency exposures.

20 To rebalance the SEK/GBP hedge, and assuming all instruments are based on SEK/GBP, Björk would buy:

A GBP 7,000,000 spot.

B GBP 7,000,000 forward to December 1.

C SEK 74,812,500 forward to December 1.

21 Given her investment goals and market view, and assuming all options are based on SEK/CHF, the *best* strategy for Björk to manage the fund's CHF exposure would be to buy an:

A ATM call option.

B ITM call option and write an OTM call option.

C OTM put option and write an OTM call option.

22 Given the recent movement in the forward premium for the SEK/EUR rate, Björk can expect that the hedge will experience higher:

A basis risk.

B roll yield.

C premia income.

23 The *most* important risk to Björk's Latin American currency hedge would be changes in:

A forward points.

B exchange rate volatility.

C cross-currency correlations.

The following information relates to Question 24

Kalila Al-Khalili has been hired as a consultant to a Middle Eastern sovereign wealth fund. The fund's oversight committee has asked her to examine the fund's financial characteristics and recommend an appropriate currency management strategy given the fund's Investment Policy Statement. After a thorough study of the fund and its finances, Al-Khalili reaches the following conclusions:

- The fund's mandate is focused on the long-term development of the country, and the royal family (who are very influential on the fund's oversight committee) are prepared to take a long-term perspective on the fund's investments.
- The fund's strategic asset allocation is tilted towards equity rather than fixed-income assets.
- Both its fixed-income and equity portfolios have a sizeable exposure to emerging market assets.

- Currently, about 90% of exchange rate exposures are hedged although the IPS allows a range of hedge ratios.
- Liquidity needs of the fund are minimal, since the government is running a balanced budget and is unlikely to need to dip into the fund in the near term to cover fiscal deficits. Indeed, the expected lifetime of country's large oil reserves has been greatly extended by recent discoveries, and substantial oil royalties are expected to persist into the future.

24 Based on her investigation, Al-Khalili would *most* likely recommend:

A active currency management.

B a hedging ratio closer to 100%.

C a narrow discretionary band for currency exposures.

SOLUTIONS

1

Template for Question 1

Asset	Foreign-Currency Portfolio Return	USD Relative to Foreign-Currency (circle one)
Portfolio A	**15%**	appreciated depreciated

Justification

The 19.5% domestic-currency return for Portfolio A is higher than the 15% foreign-currency portfolio return in GBP; therefore, the USD necessarily depreciated relative to the GBP.

The domestic-currency return on a foreign portfolio will reflect both the foreign-currency return on the portfolio and the percentage movements in the spot exchange rate between the domestic and foreign currency. The domestic-currency return is multiplicative with respect to these two factors:

$$R_{DC} = (1 + R_{FC})(1 + R_{FX}) - 1$$

where R_{DC} is the domestic-currency return (in percent), R_{FC} is the foreign-currency return of the asset (portfolio), and R_{FX} is the percentage change of the foreign currency against the domestic currency. (Note that in the R_{FX} expression the domestic currency—the USD in this case—is the price currency.)

Solving for R_{FX}: $(1 + 15\%)(1 + R_{FX}) - 1 = 19.50\%$; $R_{FX} = 3.91\%$

Thus, the USD depreciated relative to the GBP. That is, the GBP appreciated against the USD because R_{FX} is quoted in terms of USD/GBP, with the USD as the price currency and GBP as the base currency, and in this example R_{FX} is a positive number (3.91%).

2

Template for Question 2

Asset	Percentage Movement in the Spot Exchange Rate	Foreign-Currency Portfolio Return (circle one)
Portfolio B	EUR appreciated 5% against the USD	positive negative

Justification

The domestic-currency return for Portfolio B is 0%, and the EUR appreciated 5% against the USD; therefore, the foreign-currency return for Portfolio B is necessarily negative.

The domestic-currency return on a foreign portfolio will reflect both the foreign-currency return on the portfolio and the percentage movements in the spot exchange rate between the domestic and foreign currency. The domestic-currency return is multiplicative with respect to these two factors:

$$R_{DC} = (1 + R_{FC})(1 + R_{FX}) - 1$$

where R_{DC} is the domestic-currency return (in percent), R_{FC} is the foreign-currency return of the asset (portfolio), and R_{FX} is the percentage change of the foreign currency against the domestic currency. (Note that once again, the domestic currency—the USD—is the price currency in the USD/EUR quote for R_{FX}.)

Solving for R_{FC}: $(1 + R_{FC})(1 + 5\%) - 1 = 0\%$; $R_{FC} = -4.76\%$

3 Any *two* of the following four points is acceptable:

- Trading requires dealing on the bid/offer spread offered by dealers. Dealer profit margin is based on these spreads. Maintaining a 100% hedge will require frequent rebalancing of minor changes in currency movements and could prove to be expensive. “Churning” the hedge portfolio would progressively add to hedging costs and reduce the hedge’s benefits.
- A long position in currency options involves an upfront payment. If the options expire out-of-the-money, this is an unrecoverable cost.
- Forward contracts have a maturity date and need to be “rolled” forward with an FX swap transaction to maintain the hedge. Rolling hedges typically generate cash inflows and outflows, based on movements in the spot rate as well as roll yield. Cash may have to be raised to settle the hedging transactions (increases the volatility in the organization’s cash accounts). The management of these cash flow costs can accumulate and become a large portion of the portfolio’s value, and they become more expensive for cash outflows as interest rates increase.
- Hedging requires maintaining the necessary administrative infrastructure for trading (personnel and technology systems). These overhead costs can become a significant portion of the overall costs of currency trading.

4 Optimal hedging decisions require balancing the benefits of hedging against the costs of hedging. Hedging costs come mainly in two forms: trading costs and opportunity costs. Gupta is referring to the opportunity cost of the 100% hedge strategy. The opportunity cost of the 100% hedge strategy for Portfolio B is the forgone opportunity of benefiting from favorable currency rate movements. Gupta is implying that accepting some currency risk has the potential to enhance portfolio return. A complete hedge eliminates this possibility.

5 A solution is to put in place a currency overlay program for active currency management. Because internal resources for active management are lacking, the fund manager would outsource currency exposure management to a sub-advisor that specializes in foreign exchange management. This approach would allow the fund manager of Portfolio A to separate the currency hedging function (currency beta), which can be done effectively internally, and the active currency management function (currency alpha) which can be managed externally by a foreign currency specialist.

6

Template for Question 6

Gupta's Statements	Active Currency Approach (circle one)	Justification
"Many traders believe that it is not necessary to examine factors like the current account deficit, inflation, and interest rates because current exchange rates already reflect the market view on how these factors will affect future exchange rates."	carry trade technical analysis economic fundamentals	Gupta is describing active currency management based on market technicals. Market technicians believe that in a liquid, freely-traded market the historical price data already incorporates all relevant information on future price movements. Technicians believe that it is not necessary to look outside the market at data like the current account deficit, inflation and interest rates because current exchange rates already reflect the market consensus view on how these factors will affect future exchange rates.
"The six-month interest rate in India is 8% compared to 1% in the United States. This presents a yield pick-up opportunity."	carry trade technical analysis economic fundamentals	Gupta is describing active currency management based on the carry trade. This strategy is implemented by borrowing in low-yield currencies (USD at 1%) and investing in high-yield currencies (INR at 8%).
"The currency overlay manager will estimate the fair value of the currencies with the expectation that observed spot rates will converge to long-run equilibrium values described by parity conditions."	carry trade technical analysis economic fundamentals	Gupta is describing active currency management based on economic fundamentals. This approach assumes that, in free markets, exchange rates are determined by logical economic relationships that can be modeled. A fundamentals-based approach estimates the "fair value" of the currency, with the expectation that observed spot rates will converge to long-run equilibrium values described by parity conditions.

7 Statements 1 and 2 compare differences between speculative volatility traders and hedgers of volatility. Statement 1 best explains the view of a speculative volatility trader. Speculative volatility traders often want to be net-short volatility, if they believe that market conditions will remains stable. The reason for this is that most options expire out-of-the money, and the option writer can then keep the option premium as a payment earned for accepting volatility risk. (Speculative volatility traders would want to be long volatility if they thought volatility was likely to increase.) Statement 2 best describes the view of a hedger of volatility. Most hedgers are net-long volatility since they want to buy protection from unanticipated price volatility. Buying currency risk protection generally means a long option position. This can be thought of as paying an insurance premium for protection against exchange rate volatility.

8 In practice, matching the *current* market value of the foreign-currency exposure in the portfolio with an equal and offsetting position in a forward contract is likely to be ineffective over time because the market value of foreign-currency assets will change with market conditions. A static hedge (i.e., an unchanging hedge) will tend to accumulate unwanted currency exposures as the value of

the foreign-currency assets change. This will result in a mismatch between the market value of the foreign-currency asset portfolio and the nominal size of the forward contract used for the currency hedge (resulting in currency risk). For this reason, the portfolio manager will generally need to implement a dynamic hedge by rebalancing the portfolio periodically.

9 The fund manager should implement a dynamic hedging approach. Dynamic hedging requires rebalancing the portfolio periodically. The rebalancing would require adjusting some combination of the size, number, and maturities of the foreign-currency contracts.

Although rebalancing a dynamic hedge will keep the actual hedge ratio close to the target hedge ratio, it has the disadvantage of increased transaction costs compared to a static hedge.

10 C is correct. The domestic-currency return is a function of the foreign-currency return and the percentage change of the foreign currency against the domestic currency. Mathematically, the domestic-currency return is expressed as:

$$R_{DC} = (1 + R_{FC})(1 + R_{FX}) - 1$$

where R_{DC} is the domestic-currency return (in percent), R_{FC} is the foreign-currency return, and R_{FX} is the percentage change of the foreign currency against the domestic currency. Note that this R_{FX} expression is calculated using the investor's domestic currency (the EUR in this case) as the *price* currency in the P/B quote. This is different than the market-standard currency quotes in Exhibit 1, where the EUR is the *base* currency in each of these quotes. Therefore, for the foreign currency (USD, GBP, or CHF) to *appreciate* against the EUR, the market-standard quote (USD/EUR, GBP/EUR, or CHF/EUR, respectively) must *decrease*; i.e. the EUR must depreciate.

The Euro-Swiss (CHF/EUR) is the only spot rate with a negative change (from 1.2175 to 1.2080), meaning the EUR depreciated against the CHF (the CHF/EUR rate decreased). Or put differently, the CHF appreciated against the EUR, adding to the EUR-denominated return for the German investor holding CHF-denominated assets. This would result in a higher domestic-currency return (R_{DC}), for the German investor, relative to the foreign-currency return (R_{FC}) for the CHF-denominated assets. Both the Euro-dollar (USD/EUR) and Euro-sterling (GBP/EUR) experienced a positive change in the spot rate, meaning the EUR appreciated against these two currencies (the USD/EUR rate and the GBP/EUR rate both increased). This would result in a lower domestic-currency return (R_{DC}) for the German investor relative to the foreign-currency return (R_{FC}) for the USD- and GBP-denominated assets.

A is incorrect because the Euro-dollar (USD/EUR) experienced a positive change in the spot rate, meaning the EUR appreciated against the USD (the USD/EUR rate increased). This would result in a lower domestic-currency (i.e. EUR-denominated) return relative to the foreign-currency return for the USD-denominated assets, since the USD has depreciated against the EUR.

B is incorrect because the Euro-sterling (GBP/EUR) experienced a positive change in the spot rate, meaning the EUR appreciated against the GPB (the GBP/EUR rate increased). This would result in a lower domestic-currency (i.e. EUR-denominated) return relative to the foreign-currency return for the GBP-denominated assets, since the GBP has depreciated against the EUR.

11 C is correct. An increase in the expected correlation between movements in the foreign-currency asset returns and movements in the spot exchange rates from 0.50 to 0.80 would increase the domestic-currency return *risk* but would not change the *level* of expected domestic-currency return. The domestic-currency return risk is a function of the foreign-currency return risk [$\sigma(R_{FC})$] the exchange rate risk [$\sigma(R_{FX})$] and the correlation between the foreign-currency returns and exchange rate movements. Mathematically, this is expressed as:

$$\sigma^2(R_{DC}) \approx \sigma^2(R_{FC}) + \sigma^2(R_{FX}) + 2\sigma(R_{FC})\sigma(R_{FX})\rho(R_{FC}R_{FX})$$

If the correlation increases from +0.50 to +0.80, then the *variance* of the expected domestic-currency return will increase—but this will not affect the *level* of the expected domestic-currency return (R_{DC}). Refer to the equation shown for the answer in Question 1 and note that Ostermann's expected R_{FC} has not changed. (Once again, note as well that R_{FX} is defined with the domestic currency as the price currency.)

A and B are incorrect. An increase in the expected correlation between movements in the foreign-currency asset returns and movements in the spot rates from 0.50 to 0.80 would increase the domestic-currency return risk but would not impact the expected domestic-currency return.

12 A is correct. Guten believes that, due to efficient currency markets, there should not be any long-run gains for speculating (or active management) in currencies, especially after netting out management and transaction costs. Therefore, both currency hedging and actively trading currencies represent a cost to the portfolio with little prospect of consistently positive active returns. Given a long investment horizon and few immediate liquidity needs, Guten is most likely to choose to forgo currency hedging and its associated costs.

B and C are incorrect because given a long investment horizon and little immediate liquidity needs, Guten is most likely to choose to forgo currency hedging and its associated costs. Guten believes that due to efficient currency markets there should not be any long-run gains when speculating in currencies, especially after netting out management and transaction costs.

13 B is correct. Umlauf develops a market view based on underlying fundamentals in exchange rates (an economic fundamental approach). When directed by clients, Umlauf uses options and a variety of trading strategies to *unbundle* all of the various risk factors and trades them separately (a volatility trading approach). A market technical approach would entail forming a market view based on technical analysis (i.e., a belief that historical prices incorporate all relevant information on future price movements and that such movements have a tendency to repeat).

A is incorrect because, in using options and a variety of trading strategies to *unbundle* all of the various risk factors and trade them separately, Umlauf is likely to periodically employ volatility trading-based currency strategies.

C is incorrect because, in developing a market view based on underlying fundamentals in exchange rates, Umlauf does utilize an economic fundamentals approach.

14 B is correct. Braunt Pensionskasse provides the manager with limited discretion in managing the portfolio's currency risk exposures. This would be most consistent with allowing the currency overlay manager to take directional views on future currency movements (within predefined bounds) where the currency overlay is limited to the currency exposures already in the foreign asset portfolio. It would not be appropriate to use a fully-passive hedging approach since it would eliminate any alpha from currency movements. Further, a currency overlay program, which considers "foreign exchange as an asset class", would likely expose Braunt's portfolio to more currency risk than desired given the given primary performance objectives.

A is incorrect because a directional view currency overlay program is most appropriate given the limited discretion Braunt Pensionskasse has given the manager. A fully passive currency overlay program is more likely to be used when a client seeks to hedge all the currency risk.

C is incorrect because a directional view currency overlay program is most appropriate given the limited discretion Braunt Pensionskasse has given the manager. In contrast, the concept of "foreign exchange as an asset class" allows the currency overlay manager to take currency exposure positions in any currency pair where there is value-added to be harvested.

15 C is correct. The primary performance objective of Franz Trading GmbH is to add alpha to the portfolio, and thus has given the manager discretion in trading currencies. This is essentially a "foreign exchange as an asset class" approach. Braunt Pensionskasse and Kastner have more conservative currency strategies, and thus are less likely to benefit from the different strategies that a new overlay manager might employ.

A is incorrect because Franz Trading GmbH is more likely to benefit from the introduction of an additional overlay manager. Kastner is more likely to have a fully passive currency overlay program.

B is incorrect because Franz Trading GmbH is more likely to benefit from the introduction of an additional overlay manager. Braunt is more likely to have a currency overlay program where the manager takes a directional view on future currency movements.

16 A is correct. Exchange-traded futures contract not only have initial margin requirements, they also have daily mark-to-market and, as a result, can be subject to daily margin calls. Market participants must have sufficient liquidity to meet margin calls, or have their positions involuntarily liquidated by their brokers. Note that the risk of daily margin calls is not a feature of most forwards contracts; nor is initial margin. (However, this is changing among the largest institutional players in FX markets as many forward contracts now come with what are known as Collateral Support Annexes—CSAs—in which margin can be posted. Posting additional margin would typically not be a daily event, however, except in the case of extreme market moves.)

B is incorrect because futures contracts have low transactions costs. C is incorrect because whether the EUR is the price or the base currency in the quote will not affect the hedging process. In fact, on the CME the quote would be the market-standard USD/EUR quote, with the EUR as the base currency.

17 C is correct. Based on predicted export trends, Subscriber 2 most likely expects the KRW/USD rate to increase (i.e., the won—the price currency—to depreciate relative to the USD). This would require a long forward position in a forward contract, but as a country with capital controls, a NDF would be used instead. (Note: While forward contracts offered by banks are generally an institutional product, not retail, the retail version of a non-deliverable forward contract is known as a "contract for differences" (CFD) and is available at several retail FX brokers.)

A is incorrect because Subscriber 2 expects the KRW/USD rate to increase. A short straddle position would be used when the direction of exchange rate movement is unknown and volatility is expected to remain low.

B is incorrect because a put option would profit from a decrease of the KRW/USD rate, not an increase (as expected). Higher volatility would also make buying a put option more expensive.

18 B is correct. Subscriber 3's carry trade strategy is equivalent to trading the forward rate bias, based on the historical evidence that the forward rate is not the center of the distribution for the spot rate. Applying this bias involves buying currencies selling at a forward discount and selling currencies trading at a forward premium. So a higher forward premium on the lower yielding currency—the USD, the base currency in the INR/USD quote—would effectively reflect a more profitable trading opportunity. That is, a higher premium for buying or selling the USD forward is associated with a lower US interest rate compared to India. This would mean a wider interest rate differential in favor of Indian instruments, and hence potentially more carry trade profits.

A is incorrect because Subscriber 3's carry trade strategy depends on a wide interest rate differential between the high-yield country (India) and the low-yield country (the United States). The differential should be wide enough to compensate for the unhedged currency risk exposure.

C is incorrect because a guide to the carry trade's riskiness is the volatility of spot rates on the involved currencies, with rapid movements in exchange rates often associated with a panicked unwinding of carry trades. All things being equal, higher volatility is worse for carry trades.

19 C is correct. Emerging market currencies are often the investment currencies in the carry trade. This reflects the higher yields often available in their money markets compared to Developed Market economies (funding currencies are typically low-yield currencies such as the JPY). This can lead to higher holding returns, but these higher returns can also come with higher risks: carry trades are occasionally subject to panicked unwinds in stressed market conditions. When this occurs, position exit can be made more difficult by market illiquidity and higher trading costs (wider bid/offer spreads). The leverage involved in the carry trade can magnify trading losses under these circumstances.

A is incorrect because return distributions are often *negatively* skewed, reflecting the higher event risk (panicked carry trade unwinds, currency pegs being re-set, etc.) associated with the carry trade.

B is incorrect because although FX markets are typically efficient (or very close thereto) this does not mean that higher returns are not available. The key question is whether these are abnormally high *risk-adjusted* returns. Higher return in an efficient market comes with higher risk. The higher (short-term) return in the carry trade reflects the risk premia for holding unhedged currency risk, in the context of a favorable interest rate differential.

20 B is correct. The GBP value of the assets has declined, and hence the hedge needs to be *reduced* by GBP 7,000,000. This would require buying the GBP forward to net the outstanding (short) forward contract to an amount less than GBP 100,000,000.

A is incorrect because to rebalance the hedge (reduce the net size of the short forward position) the GBP must be bought *forward*, not with a spot transaction.

C is incorrect because the GBP must be bought, not sold. Buying SEK against the GBP is equivalent to selling GBP. Moreover, the amount of SEK that would be sold forward (to buy GBP 7,000,000 forward) would be determined by the *forward* rate, not the spot rate (7,000,000 × 10.6875 = 74,812,500).

21 C is correct. The fund holds CHF-denominated assets and hence Björk wants to protect against a depreciation of the CHF against the SEK, which would be a down-move in the SEK/CHF cross rate. An OTM put option provides some downside protection against such a move, while writing an OTM call option helps reduce the cost of this option structure. Note that Björk does not expect

that the SEK/CHF rate will increase, so this option (in her view) will likely expire OTM and allow her to keep the premia. This hedging structure is known as a short risk reversal (or a collar) and is a popular hedging strategy.

A is incorrect because the ATM call option will not protect against a decrease in the SEK/CHF rate. An ATM option is also expensive (compared to an OTM option). Note that Björk does not expect the SEK/CHF rate to increase, so would not want a long call option position for this rate.

B is incorrect because this structure is expensive (via the long ITM call option) and does not protect against a decrease in the SEK/CHF rate.

22 B is correct. To hedge the EUR-denominated assets Björk will be selling forward contracts on the SEK/EUR cross rate. A higher forward premium will result in higher roll return as Björk is selling the EUR forward at a higher all-in forward rate, and closing out the contract at a lower rate (all else equal), given that the forward curve is in contango.

A is incorrect because Björk is hedging EUR-denominated assets with a EUR-denominated forward contract. While it is true that the gap between spot and forward rates will be higher the higher the interest rate differential between countries, this gap (basis) converges to zero near maturity date, when the forward contracts would be rolled.

C is incorrect because forward contracts do not generate premia income; writing options does.

23 C is correct. A cross hedge exposes the fund to basis risk; that is, the risk that the hedge fails to protect against adverse currency movements because the correlations between the value of the assets being hedged and the hedging instrument change.

A is incorrect because movements in forward points (and hence roll yield) would be of secondary importance compared to the basis risk of a cross hedge.

B is incorrect because exchange rate volatility would not necessarily affect a hedge based on forward contracts, as long as the correlations between the underlying assets and the hedge remained stable. Although relevant, volatility in itself is not the "most" important risk to consider for a cross-hedge. (However, movements in volatility would affect hedges based on currency options.)

24 A is correct. The fund has a long-term perspective, few immediate liquidity needs, and a lower weight in fixed income that in equities (bond portfolios are typically associated with hedge ratios closer to 100% than equity portfolios). The emerging market exposure would also support active management, given these countries' typically higher yields (carry trade) and often volatile exchange rates.

B is incorrect because the characteristics of the fund and the beneficial investor (in this case, the royal family) do not argue for a conservative currency strategy.

C is incorrect because a more active currency management strategy would be more suitable for this fund.

Glossary

Absolute return vehicles Investments that have no direct benchmark portfolio.

Accounting defeasance Also called in-substance defeasance, accounting defeasance is a way of extinguishing a debt obligation by setting aside sufficient high-quality securities to repay the liability.

Accounting risk The risk associated with accounting standards that vary from country to country or with any uncertainty about how certain transactions should be recorded.

Accumulated benefit obligation (ABO) The present value of pension benefits, assuming the pension plan terminated immediately such that it had to provide retirement income to all beneficiaries for their years of service up to that date.

Accumulated service Years of service of a pension plan participant as of a specified date.

Active-lives The portion of a pension fund's liabilities associated with active workers.

Active management An approach to investing in which the portfolio manager seeks to outperform a given benchmark portfolio.

Active return The portfolio's return in excess of the return on the portfolio's benchmark.

Active risk The annualized standard deviation of active returns, also referred to as *tracking error* (also sometimes called *tracking risk*).

Active risk budgeting Risk budgeting that concerns active risk (risk relative to a portfolio's benchmark).

Active Share A measure of how similar a portfolio is to its benchmark. A manager who precisely replicates the benchmark will have an Active Share of zero; a manager with no holdings in common with the benchmark will have an Active Share of one.

Actual extreme events A type of scenario analysis used in stress testing. It involves evaluating how a portfolio would have performed given movements in interest rates, exchange rates, stock prices, or commodity prices at magnitudes such as occurred during past extreme market events (e.g., the stock market crash of October 1987).

Adaptive markets hypothesis (AMH) A hypothesis that applies principles of evolution—such as competition, adaptation, and natural selection—to financial markets in an attempt to reconcile efficient market theories with behavioral alternatives.

Adverse selection risk The risk associated with information asymmetry; in the context of trading, the risk of trading with a more informed trader.

Algorithmic trading Automated electronic trading subject to quantitative rules and user-specified benchmarks and constraints.

Allocation/selection interaction return A measure of the joint effect of weights assigned to both sectors and individual securities; the difference between the weight of the portfolio in a given sector and the portfolio's benchmark for that sector, times the difference between the portfolio's and the benchmark's returns in that sector, summed across all sectors.

Alpha research Research related to capturing excess risk-adjusted returns by a particular strategy; a way investment research is organized in some investment management firms.

Alternative investments Groups of investments with risk and return characteristics that differ markedly from those of traditional stock and bond investments.

Anchoring and adjustment An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Anchoring and adjustment bias An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Anchoring trap The tendency of the mind to give disproportionate weight to the first information it receives on a topic.

Angel investor An accredited individual investing chiefly in seed- and early-stage companies.

Anomalies Apparent deviations from market efficiency.

Ask price The price at which a dealer will sell a specified quantity of a security. Also called *ask*, *offer price*, or *offer*.

Ask size The quantity associated with the ask price.

Aspirational risk bucket In goal-based portfolio planning, that part of wealth allocated to investments that have the potential to increase a client's wealth substantially.

Asset covariance matrix The covariance matrix for the asset classes or markets under consideration.

Asset/liability management The management of financial risks created by the interaction of assets and liabilities.

Asset location The type of account an asset is held within, e.g., taxable or tax deferred.

Asset-only With respect to asset allocation, an approach that focuses directly on the characteristics of the assets without explicitly modeling the liabilities.

Assurity of completion In the context of trading, confidence that trades will settle without problems under all market conditions.

Assurity of the contract In the context of trading, confidence that the parties to trades will be held to fulfilling their obligations.

Asynchronism A discrepancy in the dating of observations that occurs because stale (out-of-date) data may be used in the absence of current data.

AUM fee A fee based on assets under management; an *ad valorem* fee.

Authorized participants Broker/dealers who enter into an agreement with the distributor of the fund.

Automated trading Any form of trading that is not manual, including trading based on algorithms.

Availability bias An information-processing bias in which people take a heuristic approach to estimating the probability of an outcome based on how easily the outcome comes to mind.

Average effective spread A measure of the liquidity of a security's market. The mean effective spread (sometimes dollar weighted) over all transactions in the stock in the period under study.

Back office Administrative functions at an investment firm such as those pertaining to transaction processing, record keeping, and regulatory compliance.

Backtesting A method for gaining information about a model using past data. As used in reference to VaR, it is the process of comparing the number of violations of VaR thresholds over a time period with the figure implied by the user-selected probability level.

Backwardation With respect to the futures curve in a commodity futures market, the condition of being downward sloping.

Balance of payments An accounting of all cash flows between residents and nonresidents of a country.

Bancassurance The sale of insurance by banks.

Barbell A fixed income portfolio combining securities concentrated in short and long maturities relative to the benchmark.

Base With respect to a foreign exchange quotation of the price of one unit of a currency, the currency referred to in "one unit of a currency."

Base-rate neglect A type of representativeness bias in which the base rate or probability of the categorization is not adequately considered.

Basis point value (BPV) The change in the bond price for a 1 basis point change in yield. Also called *present value of a basis point* or *price value of a basis point (PVBP)*.

Basis risk The risk resulting from using a hedging instrument that is imperfectly matched to the investment being hedged; in general, the risk that the basis will change in an unpredictable way.

Batch auction markets Auction markets where multilateral trading occurs at a single price at a prespecified point in time.

Bayes' formula A mathematical rule explaining how existing probability beliefs should be changed given new information; it is essentially an application of conditional probabilities.

Bear spread An option strategy that involves selling a put with a lower exercise price and buying a put with a higher exercise price. It can also be executed with calls.

Behavioral finance An approach to finance based on the observation that psychological variables affect and often distort individuals' investment decision making.

Behavioral finance macro A focus on market level behavior that considers market anomalies that distinguish markets from the efficient markets of traditional finance.

Behavioral finance micro A focus on individual level behavior that examines the behavioral biases that distinguish individual investors from the rational decision makers of traditional finance.

Benchmark spread The yield on a credit security over the yield on a security with little or no credit risk (benchmark bond) and with a similar duration.

Best efforts order A type of order that gives the trader's agent discretion to execute the order only when the agent judges market conditions to be favorable.

Best-in-class An ESG implementation approach that seeks to identify the most favorable companies and sectors based on ESG considerations. Also called *positive screening*.

Beta A measure of the sensitivity of a given investment or portfolio to movements in the overall market.

Beta research Research related to systematic (market) risk and return; a way investment research is organized in some investment management firms.

Bid The price at which a dealer will buy a specified quantity of a security. Also called *bid price*.

Bid–ask spread The difference between the current bid price and the current ask price of a security.

Bid price In a price quotation, the price at which the party making the quotation is willing to buy a specified quantity of an asset or security.

Bid size The quantity associated with the bid price.

Block order An order to sell or buy in a quantity that is large relative to the liquidity ordinarily available from dealers in the security or in other markets.

Bond-yield-plus-risk-premium method An approach to estimating the required return on equity which specifies that required return as a bond yield plus a risk premium.

Bottom-up approach A credit strategy approach that involves selecting the individual bonds or issuers that the investor views as having the best relative value from among a set of bonds or issuers with similar features.

Bounded rationality The notion that people have informational and cognitive limitations when making decisions and do not necessarily optimize when arriving at their decisions.

Box spread An option strategy that combines a bull spread and a bear spread having two different exercise prices, which produces a risk-free payoff of the difference in the exercise prices.

Breadth The number of truly independent decisions made each year.

Broad market indexes An index that is intended to measure the performance of an entire asset class. For example, the S&P 500 Index, Wilshire 5000, and Russell 3000 indexes for US common stocks.

Broker An agent of a trader in executing trades.

Brokered markets Markets in which transactions are largely effected through a search-brokerage mechanism away from public markets.

Bubbles Episodes in which asset market prices move to extremely high levels in relation to estimated intrinsic value.

Buffering Establishing ranges around breakpoints that define whether a stock belongs in one index or another.

Build-up approach Synonym for the risk premium approach.

Bull spread An option strategy that involves buying a call with a lower exercise price and selling a call with a higher exercise price. It can also be executed with puts.

Bullet A fixed income portfolio made up of securities targeting a single segment of the curve.

Business cycle Fluctuations in GDP in relation to long-term trend growth, usually lasting 9–11 years.

Business risk The equity risk that comes from the nature of the firm's operating activities.

Butterfly spread An option strategy that combines two bull or bear spreads and has three exercise prices, or, a measure of yield curve curvature.

Buy-side traders Professional traders that are employed by investment managers and institutional investors.

Calendar rebalancing Rebalancing a portfolio to target weights on a periodic basis; for example, monthly, quarterly, semiannually, or annually.

Calmar ratio The compound annualized rate of return over a specified time period divided by the absolute value of maximum drawdown over the same time period.

Cap A combination of interest rate call options designed to hedge a borrower against rate increases on a floating-rate loan.

Capital adequacy ratio A measure of the adequacy of capital in relation to assets.

Capital flows forecasting approach An exchange rate forecasting approach that focuses on expected capital flows, particularly long-term flows such as equity investment and foreign direct investment.

Capital market expectations (CME) Expectations concerning the risk and return prospects of asset classes.

Caplet Each component call option in a cap.

Carried interest A private equity fund manager's incentive fee; the share of the private equity fund's profits that the fund manager is due once the fund has returned the outside investors' capital.

Carry trade A trading strategy that involves buying a security and financing it at a rate that is lower than the yield on that security.

Cash balance plan A defined-benefit plan whose benefits are displayed in individual recordkeeping accounts.

Cash drag Tracking error caused by temporarily uninvested cash.

Cash flow at risk A variation of VaR that measures the risk to a company's cash flow, instead of its market value; the minimum cash flow loss expected to be exceeded with a given probability over a specified time period.

Cash flow matching Immunization approach that attempts to ensure that all future liability payouts are matched precisely by cash flows from bonds or fixed-income derivatives, such as interest rate futures, options, or swaps.

Cell approach See *stratified sampling*.

Certainty equivalent The maximum sum of money a person would pay to participate or the minimum sum of money a person would accept to not participate in an opportunity.

Chain-linking A process for combining periodic returns to produce an overall time-weighted rate of return.

Cheapest-to-deliver A bond in which the amount received for delivering the bond is largest compared with the amount paid in the market for the bond.

Civil law A legal system derived from Roman law, in which judges apply general, abstract rules or concepts to particular cases. In civil systems, law is developed primarily through legislative statutes or executive action.

Clawback provision With respect to the compensation of private equity fund managers, a provision that specifies that money paid to the fund manager be returned to the investors if, at the end of a fund's life, the investors have not received back their capital contributions and contractual share of profits.

Closed-book markets Markets in which a trader does not have real-time access to all quotes in a security.

Closeout netting In a bankruptcy, a process by which multiple obligations between two counterparties are consolidated into a single overall value owed by one of the counterparties to the other.

Closet indexer A fund that advertises itself as being actively managed but is substantially similar to an index fund in its exposures.

Cobb-Douglas model A production function (model for economic output) based on factors of labor and capital that exhibits constant returns to scale.

Cobb-Douglas production function A production function (model for economic output) based on factors of labor and capital that exhibits constant returns to scale.

Cognitive dissonance The mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions.

Cognitive errors Behavioral biases resulting from faulty reasoning; cognitive errors stem from basic statistical, information processing, or memory errors.

Collar An option strategy involving the purchase of a put and sale of a call in which the holder of an asset gains protection below a certain level, the exercise price of the put, and pays for it by giving up gains above a certain level, the exercise price of the call. Collars also can be used to provide protection against rising interest rates on a floating-rate loan by giving up gains from lower interest rates.

Commingled real estate funds (CREFs) Professionally managed vehicles for substantial commingled (i.e., pooled) investment in real estate properties.

Commitment period The period of time over which committed funds are advanced to a private equity fund.

Commodities Articles of commerce such as agricultural goods, metals, and petroleum; tangible assets that are typically relatively homogeneous in nature.

Commodity trading advisers (CTAs) Registered advisers who manage futures funds.

Common law A legal system which draws abstract rules from specific cases. In common law systems, law is developed primarily through decisions of the courts.

Community property regime A marital property regime under which each spouse has an indivisible one-half interest in property received during marriage.

Company-specific risk The non-systematic or idiosyncratic risk specific to a particular company's operations, reputation, and business environment.

Completion overlay A type of overlay that addresses an indexed portfolio that has diverged from its proper exposure.

Confidence band With reference to a quality control chart for performance evaluation, a range in which the manager's value-added returns are anticipated to fall a specified percentage of the time.

Confirmation bias A belief perseverance bias in which people tend to look for and notice what confirms their beliefs, to ignore or undervalue what contradicts their beliefs, and to misinterpret information as support for their beliefs.

Confirming evidence trap The bias that leads individuals to give greater weight to information that supports an existing or preferred point of view than to evidence that contradicts it.

Conjunction fallacy An inappropriate combining of probabilities of independent events to support a belief. In fact, the probability of two independent events occurring in conjunction is never greater than the probability of either event occurring alone; the probability of two independent events occurring together is equal to the multiplication of the probabilities of the independent events.

Conservatism bias A belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new information.

Constant returns to scale A characteristic of a production function such that a given percentage increase in capital stock and labor input results in an equal percentage increase in output.

Contingent immunization Hybrid approach that combines immunization with an active management approach when the asset portfolio's value exceeds the present value of the liability portfolio.

Continuous auction markets Auction markets where orders can be executed at any time during the trading day.

Controlled foreign corporation A company located outside a taxpayer's home country and in which the taxpayer has a controlling interest as defined under the home country law.

Convexity A measure of how interest rate sensitivity changes with a change in interest rates.

Core capital The amount of capital required to fund spending to maintain a given lifestyle, fund goals, and provide adequate reserves for unexpected commitments.

Core–satellite A way of thinking about allocating money that seeks to define each investment's place in the portfolio in relation to specific investment objectives or roles.

Corporate governance The system of internal controls and procedures used to define and protect the rights and responsibilities of various stakeholders.

Corporate venturing Investments by companies in promising young companies in the same or a related industry.

Covered call An option strategy involving the holding of an asset and sale of a call on the asset.

Creation units Large blocks of ETF shares often traded against a basket of underlying securities.

Credit default swap A swap used to transfer credit risk to another party. A protection buyer pays the protection seller in return for the right to receive a payment from the seller in the event of a specified credit event.

Credit derivative A contract in which one party has the right to claim a payment from another party in the event that a specific credit event occurs over the life of the contract.

Credit method When the residence country reduces its taxpayers' domestic tax liability by the amount of taxes paid to a foreign country that exercises source jurisdiction.

Credit risk The risk of loss caused by a counterparty's or debtor's failure to make a timely payment or by the change in value of a financial instrument based on changes in default risk. Also called *default risk*.

Credit spread forward A forward contract used to transfer credit risk to another party; a forward contract on a yield spread.

Credit spread option An option based on the yield spread between two securities that is used to transfer credit risk.

Credit VaR A variation of VaR related to credit risk; it reflects the minimum loss due to credit exposure with a given probability during a period of time.

Credited rates Rates of interest credited to a policyholder's reserve account.

Cross-default provision A provision stipulating that if a borrower defaults on any outstanding credit obligations, the borrower is considered to be in default on all obligations.

Cross hedge A hedge involving a hedging instrument that is imperfectly correlated with the asset being hedged; an example is hedging a bond investment with futures on a non-identical bond.

Currency overlay A type of overlay that helps hedge the returns of securities held in foreign currency back to the home country's currency.

Currency overlay programs A currency overlay program is a program to manage a portfolio's currency exposures for the case in which those exposures are managed separately from the management of the portfolio itself.

Current credit risk The risk of credit-related events happening in the immediate future; it relates to the risk that a payment currently due will not be paid. Also called *jump-to-default risk*.

Custom security-based benchmark Benchmarks that are custom built to accurately reflect the investment discipline of a particular investment manager. Also called *strategy benchmarks* because they reflect a manager's particular strategy.

Cyclical stocks The shares of companies whose earnings have above-average sensitivity to the business cycle.

Cyclically Adjusted P/E Ratio (CAPE) A price-to-earnings ratio in which the numerator (in a US context) is defined as the real S&P 500 price index and the denominator as the moving average of the preceding 10 years of real reported earnings on the S&P 500.

Day traders Traders that rapidly buy and sell stocks in the hope that the stocks will continue to rise or fall in value for the seconds or minutes they are prepared to hold a position. Day traders hold a position open somewhat longer than a scalper but closing all positions at the end of the day.

Dealer A business entity that is ready to buy an asset for inventory or sell an asset from inventory to provide the other side of an order. Also called *market maker*.

Decision price The prevailing price when the decision to trade is made. Also called *arrival price* or *strike price*.

Decision-reversal risk The risk of reversing a chosen course of action at the point of maximum loss.

Decision risk The risk of changing strategies at the point of maximum loss.

Deduction method When the residence country allows taxpayers to reduce their taxable income by the amount of taxes paid to foreign governments in respect of foreign-source income.

Deemed dispositions Tax treatment that assumes property is sold. It is sometimes seen as an alternative to estate or inheritance tax.

Deemed distribution When shareholders of a controlled foreign corporation are taxed as if the earnings were distributed to shareholders, even though no distribution has been made.

Default risk The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security.

Default risk premium Compensation for the possibility that the issue of a debt instrument will fail to make a promised payment at the contracted time and in the contracted amount.

Defaultable debt Debt with some meaningful amount of credit risk.

Deferred annuity An annuity that enables an individual to purchase an income stream that will begin at a later date.

Defined-benefit plan A pension plan that specifies the plan sponsor's obligations in terms of the benefit to plan participants.

Defined-contribution plan A pension plan that specifies the sponsor's obligations in terms of contributions to the pension fund rather than benefits to plan participants.

Deflation A decrease in the general level of prices; an increase in the purchasing power of a unit of currency.

Delay costs Implicit trading costs that arise from the inability to complete desired trades immediately due to order size or market liquidity. Also called *slippage*.

Delta The relationship between the option price and the underlying price, which reflects the sensitivity of the price of the option to changes in the price of the underlying.

Delta hedge An option strategy in which a position in an asset is converted to a risk-free position with a position in a specific number of options. The number of options per unit of the underlying changes through time, and the position must be revised to maintain the hedge.

Delta hedging Hedging that involves matching the price response of the position being hedged over a narrow range of prices.

Delta-normal method A measure of VaR equivalent to the analytical method but that refers to the use of delta to estimate the option's price sensitivity.

Demand deposit A deposit that can be drawn upon without prior notice, such as a checking account.

Demutualizing The process of converting an insurance company from mutual form to stock.

Descriptive statistics Methods for effectively summarizing data to describe important aspects of a dataset.

Differential returns Returns that deviate from a manager's benchmark.

Diffusion index An index that measures how many indicators are pointing up and how many are pointing down.

Direct commodity investment Commodity investment that involves the cash market purchase of physical commodities or exposure to changes in spot market values via derivatives, such as futures.

Direct market access Platforms sponsored by brokers that permit buy-side traders to directly access equities, fixed income, futures, and foreign exchange markets, clearing via the broker.

Disability income insurance A type of insurance designed to mitigate earnings risk as a result of a disability in which an individual becomes less than fully employed.

Discounted cash flow models (DCF models) Valuation models that express the idea that an asset's value is the present value of its (expected) cash flows.

Discretionary trust A trust structure in which the trustee determines whether and how much to distribute in the sole discretion of the trustee.

Disintermediation To withdraw funds from financial intermediaries for placement with other financial intermediaries offering a higher return or yield. Or, to withdraw funds from a financial intermediary for the purposes of direct investment, such as withdrawing from a mutual fund to make direct stock investments.

Dispersion The weighted *variance* of the times to receipt of cash flow; it measures the extent to which the payments are spread out around the duration.

Disposition effect As a result of loss aversion, an emotional bias whereby investors are reluctant to dispose of losers. This results in an inefficient and gradual adjustment to deterioration in fundamental value.

Distressed debt arbitrage A distressed securities investment discipline that involves purchasing the traded bonds of bankrupt companies and selling the common equity short.

Distressed securities Securities of companies that are in financial distress or near bankruptcy; the name given to various investment disciplines employing the securities of companies in distress.

Diversification effect In reference to VaR across several portfolios (for example, across an entire firm), this effect equals the difference between the sum of the individual VaRs and total VaR.

Dividend capture A trading strategy whereby an equity portfolio manager purchases stocks just before their ex-dividend dates, holds these stocks through the ex-dividend date to earn the right to receive the dividend, and subsequently sells the shares.

Dividend recapitalization A method by which a buyout fund can realize the value of a holding; involves the issuance of debt by the holding to finance a special dividend to owners.

Domestic asset An asset that trades in the investor's domestic currency (or home currency).

Domestic currency The currency of the investor, i.e., the currency in which he or she typically makes consumption purchases, e.g., the Swiss franc for an investor domiciled in Switzerland.

Domestic-currency return A rate of return stated in domestic currency terms from the perspective of the investor; reflects both the foreign-currency return on an asset as well as percentage movement in the spot exchange rate between the domestic and foreign currencies.

Donor-advised fund A fund administered by a tax-exempt entity in which the donor advises on where to grant the money that he or she has donated.

Double inflection utility function A utility function that changes based on levels of wealth.

Downside deviation A measure of volatility using only rate of return data points below the investor's minimum acceptable return.

Due diligence Investigation and analysis in support of an investment action or recommendation, such as the scrutiny of operations and management and the verification of material facts.

Duration A measure of the approximate sensitivity of a security to a change in interest rates (i.e., a measure of interest rate risk).

Duration matching Immunization approach based on the duration of assets and liabilities. Ideally, the liabilities being matched (the liability portfolio) and the portfolio of assets (the bond portfolio) should be affected similarly by a change in interest rates.

Dynamic asset allocation A strategy incorporating deviations from the strategic asset allocation that are motivated by longer-term valuation signals or economic views than usually associated with tactical asset allocation.

Dynamic hedge A hedge requiring adjustment as the price of the hedged asset changes.

Earnings at risk (EAR) A variation of VaR that reflects the risk of a company's earnings instead of its market value.

Earnings risk The risk associated with the earning potential of an individual.

Econometrics The application of quantitative modeling and analysis grounded in economic theory to the analysis of economic data.

Economic balance sheet A balance sheet that provides an individual's total wealth portfolio, supplementing traditional balance sheet assets with human capital and pension wealth, and expanding liabilities to include consumption and bequest goals. Also known as *holistic balance sheet*.

Economic exposure The risk associated with changes in the relative attractiveness of products and services offered for sale, arising out of the competitive effects of changes in exchange rates.

Economic indicators Economic statistics provided by government and established private organizations that contain information on an economy's recent past activity or its current or future position in the business cycle.

Effective convexity A second-order effect that describes how a bond's interest rate sensitivity changes with changes in yield. Effective convexity is used when the bond has cash flows that change when yields change (as in the case of callable bonds or mortgage-backed securities).

Effective duration Duration adjusted to account for embedded options.

Effective spread Two times the distance between the actual execution price and the midpoint of the market quote at the time an order is entered; a measure of execution costs that captures the effects of price improvement and market impact.

Electronic communications networks (ECNs) Computer-based auctions that operate continuously within the day using a specified set of rules to execute orders.

Emerging market debt The sovereign debt of nondeveloped countries.

Emotional biases Behavioral biases resulting from reasoning influenced by feelings; emotional biases stem from impulse or intuition.

Empirical duration A measure of interest rate sensitivity that is determined from market data.

Endogenous variable A variable whose values are determined within the system.

Endowment bias An emotional bias in which people value an asset more when they hold rights to it than when they do not.

Endowments Long-term funds generally owned by operating nonprofit institutions such as universities and colleges, museums, hospitals, and other organizations involved in charitable activities.

Enhanced derivatives products companies A type of subsidiary separate from an entity's other activities and not liable for the parent's debts. They are often used by derivatives dealers to control exposure to ratings downgrades. Also called *special purpose vehicles.*

Enhanced indexing strategy Method investors use to match an underlying market index in which the investor purchases fewer securities than the full set of index constituents but matches primary risk factors reflected in the index.

Enterprise risk management An overall assessment of a company's risk position. A centralized approach to risk management sometimes called firmwide risk management.

Environmental, social, and corporate governance (ESG) Also called socially responsible investing, refers to the explicit inclusion of ethical, environmental, or social criteria when selecting a portfolio.

Equity forward sale contract A private contract for the forward sale of an equity position.

Equity monetization The realization of cash for an equity position through a manner other than an outright sale.

Equity *q* The ratio of a company's equity market capitalization divided by net worth measured at replacement cost.

Equity risk premium Compensation for the additional risk of equity compared with debt.

ESG risk The risk to a company's market valuation resulting from environmental, social, and governance factors.

Estate All of the property a person owns or controls; may consist of financial assets, tangible personal assets, immovable property, or intellectual property.

Estate planning The process of preparing for the disposition of one's estate (e.g., the transfer of property) upon death and during one's lifetime.

Estate tax freeze A plan usually involving a corporation, partnership, or limited liability company with the goal to transfer *future* appreciation to the next generation at little or no gift or estate tax cost.

Eurozone The region of countries using the euro as a currency.

Evaluated pricing See *matrix pricing.*

***Ex post* alpha** (or Jensen's alpha) The average return achieved in a portfolio in excess of what would have been predicted by CAPM given the portfolio's risk level; an after-the-fact measure of excess risk-adjusted return.

Excess capital An investor's capital over and above that which is necessary to fund their lifestyle and reserves.

Exchange A regulated venue for the trading of investment instruments.

Exchange-traded fund Exchange-traded Funds or ETFs are hybrid investment products with many features of mutual funds combined with the trading features of common stocks or bonds. Essentially, ETFs are typically portfolios of stocks or bonds or commodities that trade throughout the day like common stocks.

Execution uncertainty Uncertainty pertaining to the timing of execution, or if execution will even occur at all.

Exemption method When the residence country imposes no tax on foreign-source income by providing taxpayers with an exemption, in effect having only one jurisdiction impose tax.

Exhaustive An index construction strategy that selects every constituent of a universe.

Exogenous shocks Events from outside the economic system that affect its course. These could be short-lived political events, changes in government policy, or natural disasters, for example.

Exogenous variable A variable whose values are determined outside the system.

Expected shortfall The average loss conditional on exceeding the VaR cutoff; sometimes referred to as *conditional VaR* or *expected tail loss.*

Expected tail loss See *expected shortfall*.

Extended portfolio assets and liabilities Assets and liabilities beyond those shown on a conventional balance sheet that are relevant in making asset allocation decisions; an example of an extended asset is human capital.

Externality Those consequences of a transaction (or process) that do not fall on the parties to the transaction (or process).

Factor covariance matrix The covariance matrix of factors.

Factor-model-based benchmark Benchmarks constructed by examining a portfolio's sensitivity to a set of factors, such as the return for a broad market index, company earnings growth, industry, or financial leverage.

Factor push A simple stress test that involves pushing prices and risk factors of an underlying model in the most disadvantageous way to estimate the impact of factor extremes on the portfolio's value.

Factor sensitivities In a multifactor model, the responsiveness of the dependent variable to factor movements. Also called *factor betas* or *factor loadings*.

Fallen angels Debt that has crossed the threshold from investment grade to high yield.

Fed model An equity valuation model that relates the earnings yield on the S&P 500 to the yield to maturity on 10-year US Treasury bonds.

Federal funds rate The interest rate on overnight loans of reserves (deposits) between US Federal Reserve System member banks.

Fiduciary A person or entity standing in a special relation of trust and responsibility with respect to other parties.

Financial buyers Buyers who lack a strategic motive.

Financial capital The tangible and intangible assets (excluding human capital) owned by an individual or household.

Financial equilibrium models Models describing relationships between expected return and risk in which supply and demand are in balance.

Financial risk Risks derived from events in the external financial markets, such as changes in equity prices, interest rates, or currency exchange rates.

Fiscal policy Government activity concerning taxation and governmental spending.

Fixed trust A trust structure in which distributions to beneficiaries are prescribed in the trust document to occur at certain times or in certain amounts.

Flexible-premium variable life A type of life insurance policy that combines the flexibility of universal life with the investment choice flexibility of variable life. Also called *variable universal life*.

Floor A combination of interest rate options designed to provide protection against interest rate decreases.

Floor broker An agent of the broker who, for certain exchanges, physically represents the trade on the exchange floor.

Floorlet Each component put option in a floor.

Forced heirship rules Legal ownership principles whereby children have the right to a fixed share of a parent's estate.

Foreign assets Assets denominated in currencies other than the investor's home currency.

Foreign currency Currency that is not the currency in which an investor makes consumption purchases, e.g., the US dollar from the perspective of a Swiss investor.

Foreign-currency return The return of the foreign asset measured in foreign-currency terms.

Formal tools Established research methods amenable to precise definition and independent replication of results.

Forward conversion with options The construction of a synthetic short forward position against the asset held long.

Forward rate bias Persistent violation of uncovered interest rate parity that is exploited by the carry trade.

Foundations Typically, grant-making institutions funded by gifts and investment assets.

Fourth market A term occasionally used for direct trading of securities between institutional investors; the fourth market would include trading on electronic crossing networks.

Framing An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Framing bias An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Front office The revenue generating functions at an investment firm such as those pertaining to trading and sales.

Front-run To trade ahead of the initiator, exploiting privileged information about the initiator's trading intentions.

Full replication approach When every issue in an index is represented in the portfolio, and each portfolio position has approximately the same weight in the fund as in the index.

Fully funded plan A pension plan in which the ratio of the value of plan assets to the present value of plan liabilities is 100 percent or greater.

Fund of funds A fund that invests in a number of underlying funds.

Funded status The relationship between the value of a plan's assets and the present value of its liabilities.

Funding currencies The low-yield currencies in which borrowing occurs in a carry trade.

Funding risk The risk that liabilities funding long asset positions cannot be rolled over at reasonable cost.

G-spread The yield on a credit security over the yield of an actual or interpolated government bond.

Gain-to-loss ratio The ratio of positive returns to negative returns over a specified period of time.

Gamblers' fallacy A misunderstanding of probabilities in which people wrongly project reversal to a long-term mean.

Gamma A numerical measure of the sensitivity of delta to a change in the underlying's value.

Global custodian An entity that effects trade settlement, safekeeping of assets, and the allocation of trades to individual custody accounts.

Global investable market A practical proxy for the world market portfolio consisting of traditional and alternative asset classes with sufficient capacity to absorb meaningful investment.

Goals-based With respect to asset allocation or investing, an approach that focuses on achieving an investor's goals (for example, related to supporting lifestyle needs or aspirations) based typically on constructing sub-portfolios aligned with those goals.

Goals-based investing An investment industry term for approaches to investing for individuals and families focused on aligning investments with goals (parallel to liability-driven investing for institutional investors).

Gordon (constant) growth model A version of the dividend discount model for common share value that assumes a constant growth rate in dividends.

Government structural policies Government policies that affect the limits of economic growth and incentives within the private sector.

Grinold–Kroner model An expression for the expected return on a share as the sum of an expected income return, an expected nominal earnings growth return, and an expected repricing return.

Growth in total factor productivity A component of trend growth in GDP that results from increased efficiency in using capital inputs; also known as technical progress.

H-model A variant of the two-stage dividend discount model in which growth begins at a high rate and declines linearly throughout the supernormal growth period until it reaches a normal growth rate that holds in perpetuity.

Hague Conference on Private International Law An intergovernmental organization working toward the convergence of private international law. Its 69 members consist of countries and regional economic integration organizations.

Halo effect An emotional bias that extends a favorable evaluation of some characteristics to other characteristics.

Health insurance A type of insurance used to cover health care and medical costs.

Health risk The risk associated with illness or injury.

Hedge funds A historically loosely regulated, pooled investment vehicle that may implement various investment strategies.

Hedge ratio The relationship of the quantity of an asset being hedged to the quantity of the derivative used for hedging.

Hedging A general strategy usually thought of as reducing, if not eliminating, risk.

Herding When a group of investors trade on the same side of the market in the same securities, or when investors ignore their own private information and act as other investors do.

High-water mark A specified net asset value level that a fund must exceed before performance fees are paid to the hedge fund manager.

High-yield investing A distressed securities investment discipline that involves investment in high-yield bonds perceived to be undervalued.

Hindsight bias A bias with selective perception and retention aspects in which people may see past events as having been predictable and reasonable to expect.

Historical simulation method The application of historical price changes to the current portfolio.

Holistic balance sheet See *economic balance sheet*.

Home bias A preference for securities listed on the exchanges of one's home country.

Home-country bias The favoring of domestic over non-domestic investments relative to global market value weights.

Home currency See *domestic currency*.

Horizon matching Hybrid approach that combines cash flow and duration matching approaches. Under this approach, liabilities are categorized as short-and long-term liabilities.

Human capital An implied asset; the net present value of an investor's future expected labor income weighted by the probability of surviving to each future age. Also called *net employment capital*.

Hybrid markets Combinations of market types, which offer elements of batch auction markets and continuous auction markets, as well as quote-driven markets.

Hypothetical events A type of scenario analysis used in stress testing that involves the evaluation of performance given events that have never happened in the markets or market outcomes to which we attach a small probability.

I-spread The yield on a credit security over the swap rate (denominated in the same currency as the credit security). Also known as interpolated spread.

Illiquidity premium Compensation for the risk of loss relative to an investment's fair value if an investment needs to be converted to cash quickly.

Illusion of control A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot. Illusion of knowledge and self-attribution biases contribute to the overconfidence bias.

Illusion of control bias A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot. Illusion of knowledge and self-attribution biases contribute to the overconfidence bias.

Immediate annuity An annuity that provides a guarantee of specified future monthly payments over a specified period of time.

Immunization An asset/liability management approach that structures investments in bonds to match (offset) liabilities' weighted-average duration; a type of dedication strategy.

Impact investing Investment approach that seeks to achieve targeted social or environmental objectives along with measurable financial returns through engagement with a company or by direct investment in projects or companies.

Implementation shortfall The difference between the money return on a notional or paper portfolio and the actual portfolio return.

Implementation shortfall strategy A strategy that attempts to minimize trading costs as measured by the implementation shortfall method. Also called *arrival price strategy*.

Implied yield A measure of the yield on the underlying bond of a futures contract implied by pricing it as though the underlying will be delivered at the futures expiration.

Incremental VaR A measure of the incremental effect of an asset on the VaR of a portfolio by measuring the difference between the portfolio's VaR while including a specified asset and the portfolio's VaR with that asset eliminated.

Indexing A common passive approach to investing that involves holding a portfolio of securities designed to replicate the returns on a specified index of securities.

Indifference curve analysis A decision-making approach whereby curves of consumption bundles, among which the decision-maker is indifferent, are constructed to identify and choose the curve within budget constraints that generates the highest utility.

Indirect commodity investment Commodity investment that involves the acquisition of indirect claims on commodities, such as equity in companies specializing in commodity production.

Inferential statistics Methods for making estimates or forecasts about a larger group from a smaller group actually observed.

Inflation An increase in the general level of prices; a decrease in the purchasing power of a unit of currency.

Inflation premium Compensation for expected inflation.

Information coefficient Formally defined as the correlation between forecast return and actual return. In essence, it measures the effectiveness of investment insight.

Information-motivated traders Traders that seek to trade on information that has limited value if not quickly acted upon.

Information ratio The mean excess return of the account over the benchmark (i.e., mean active return) relative to the variability of that excess return (i.e., tracking risk); a measure of risk-adjusted performance.

Infrastructure funds Funds that make private investment in public infrastructure projects in return for rights to specified revenue streams over a contracted period.

Initial public offering The initial issuance of common stock registered for public trading by a formerly private corporation.

Input uncertainty Uncertainty concerning whether the inputs are correct.

Inside ask The lowest available ask price. Also called *market ask*.

Inside bid The highest available bid price. Also called *market bid*.

Inside bid–ask spread Market ask price minus market bid price. Also called *market bid–ask spread*, *inside spread*, or *market spread*.

Inside quote Combination of the highest available bid price with the lowest available ask price. Also called *market quote*.

Inside spread Market ask price minus market bid price. Also called *market bid–ask spread*, *inside bid–ask spread*, or *market spread*.

Institutional investors Corporations or other legal entities that ultimately serve as financial intermediaries between individuals and investment markets.

Interest rate management effect With respect to fixed-income attribution analysis, a return component reflecting how well a manager predicts interest rate changes.

Interest spread With respect to banks, the average yield on earning assets minus the average percent cost of interest-bearing liabilities.

Internal rate of return The growth rate that will link the ending value of the account to its beginning value plus all intermediate cash flows; money-weighted rate of return is a synonym.

Intestate Having made no valid will; a decedent without a valid will or with a will that does not dispose of their property is considered to have died intestate.

Intrinsic value The difference between the spot exchange rate and the strike price of a currency option.

Inventory cycle A cycle measured in terms of fluctuations in inventories, typically lasting 2–4 years.

Inverse floater A floating-rate note or bond in which the coupon is adjusted to move opposite to a benchmark interest rate.

Investment currencies The high-yielding currencies in a carry trade.

Investment skill The ability to outperform an appropriate benchmark consistently over time.

Investment style indexes Indexes that represent specific portions of an asset category. For example, subgroups within the US common stock asset category such as large-capitalization growth stocks.

Irrevocable trust A trust arrangement wherein the settlor has no ability to revoke the trust relationship.

J factor risk The risk associated with a judge's track record in adjudicating bankruptcies and restructurings.

Joint ownership with right of survivorship Jointly owned; assets held in joint ownership with right of survivorship automatically transfer to the surviving joint owner or owners outside the probate process.

Justified P/E The price-to-earnings ratio that is fair, warranted, or justified on the basis of forecasted fundamentals.

Key rate duration A method of measuring the interest rate sensitivities of a fixed-income instrument or portfolio to shifts in key points along the yield curve.

Knock-in/knock-out Features of a vanilla option that is created (or ceases to exist) when the spot exchange rate touches a pre-specified level.

Lagging economic indicators A set of economic variables whose values correlate with recent past economic activity.

Leading economic indicators A set of economic variables whose values vary with the business cycle but at a fairly consistent time interval before a turn in the business cycle.

Legal/contract risk The possibility of loss arising from the legal system's failure to enforce a contract in which an enterprise has a financial stake; for example, if a contract is voided through litigation.

Leverage-adjusted duration gap A leverage-adjusted measure of the difference between the durations of assets and liabilities which measures a bank's overall interest rate exposure.

Leveraged floating-rate note (leveraged floater) A floating-rate note or bond in which the coupon is adjusted at a multiple of a benchmark interest rate.

Leveraged recapitalization A leveraging of a company's balance sheet, usually accomplished by working with a private equity firm.

Liability-driven investing An investment industry term that generally encompasses asset allocation that is focused on funding an investor's liabilities in institutional contexts.

Liability glide path A specification of desired proportions of liability-hedging assets and return-seeking assets and the duration of the liability hedge as funded status changes and contributions are made.

Liability insurance A type of insurance used to manage liability risk.

Liability-relative With respect to asset allocation, an approach that focuses directly only on funding liabilities as an investment objective.

Liability risk The possibility that an individual or household may be held legally liable for the financial costs associated with property damage or physical injury.

Life-cycle finance A concept in finance that recognizes as an investor ages, the fundamental nature of wealth and risk evolves.

Life insurance A type of insurance that protects against the loss of human capital for those who depend on an individual's future earnings.

Lifetime gratuitous transfer A lifetime gift made during the lifetime of the donor; also known as *inter vivos* transfers.

Limit order An instruction to execute an order when the best price available is at least as good as the limit price specified in the order.

Liquidity The ability to trade without delay at relatively low cost and in relatively large quantities.

Liquidity-motivated traders Traders that are motivated to trade based upon reasons other than an information advantage. For example, to release cash proceeds to facilitate the purchase of another security, adjust market exposure, or fund cash needs.

Liquidity risk Any risk of economic loss because of the need to sell relatively less liquid assets to meet liquidity requirements; the risk that a financial instrument cannot be purchased or sold without a significant concession in price because of the market's potential inability to efficiently accommodate the desired trading size.

Lock-up period A minimum initial holding period for investments during which no part of the investment can be withdrawn.

Locked up Said of investments that cannot be traded at all for some time.

Logical participation strategies Protocols for breaking up an order for execution over time. Typically used by institutional traders to participate in overall market volumes without being unduly visible.

Longevity risk The risk associated with living to an advanced age in retirement, including the uncertainty surrounding how long retirement will last; the risk of outliving one's financial resources.

Loss-aversion bias A bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains.

Loss given default See *loss severity*.
Loss severity The amount of loss if a default occurs. Also called *loss given default*.
***M*²** A measure of what a portfolio would have returned if it had taken on the same total risk as the market index.
Macaulay duration The percentage change in price for a percentage change in yield. The term, named for one of the economists who first derived it, is used to distinguish the calculation from modified duration. (See also *modified duration*).
Macro attribution Performance attribution analysis conducted on the fund sponsor level.
Macro expectations Expectations concerning classes of assets.
Managed futures Pooled investment vehicles, frequently structured as limited partnerships, that invest in futures and options on futures and other instruments.
Managed futures funds Pools of private capital managed by commodity trading advisers.
Manager continuation policies Policies adopted to guide the manager evaluations conducted by fund sponsors. The goal of manager continuation policies is to reduce the costs of manager turnover while systematically acting on indications of future poor performance.
Manager monitoring A formal, documented procedure that assists fund sponsors in consistently collecting information relevant to evaluating the state of their managers' operations; used to identify warning signs of adverse changes in existing managers' organizations.
Manager review A detailed examination of a manager that currently exists within a plan sponsor's program. The manager review closely resembles the manager selection process, in both the information considered and the comprehensiveness of the analysis. The staff should review all phases of the manager's operations, just as if the manager were being initially hired.
Market-adjusted implementation shortfall The difference between the money return on a notional or paper portfolio and the actual portfolio return, adjusted using beta to remove the effect of the return on the market.
Market ask The lowest available ask price.
Market bid The best available bid; highest price any buyer is currently willing to pay.
Market bid–ask spread Market ask price minus market bid price. Also called *inside bid–ask spread*, *inside spread*, or *market spread*.
Market fragmentation A condition whereby a market contains no dominant group of sellers (or buyers) that are large enough to unduly influence the market.
Market impact The effect of the trade on transaction prices. Also called *price impact*.
Market integration The degree to which there are no impediments or barriers to capital mobility across markets.
Market microstructure The market structures and processes that affect how the manager's interest in buying or selling an asset is translated into executed trades (represented by trade prices and volumes).
Market model A regression equation that specifies a linear relationship between the return on a security (or portfolio) and the return on a broad market index.
Market-not-held order A variation of the market order designed to give the agent greater discretion than a simple market order would allow. "Not held" means that the floor broker is not required to trade at any specific price or in any specific time interval.
Market on close order A market order to be executed at the closing of the market.
Market on open order A market order to be executed at the opening of the market.
Market order An instruction to execute an order as soon as possible in the public markets at the best price available.
Market quote Combination of the highest available bid price with the lowest available ask price. Also called *inside quote*.
Market risk The risk associated with interest rates, exchange rates, and equity prices.
Market risk bucket In goal-based portfolio planning, that part of wealth allocated to investments intended to maintain the client's current standard of living.
Market segmentation The degree to which there are some meaningful impediments to capital movement across markets.
Market spread Market ask price minus market bid price. Also called *market bid–ask spread*, *inside spread*, or *inside bid–ask spread*.
Marking to market A procedure used primarily in futures markets in which the parties to a contract settle the amount owed daily. Also known as the *daily settlement*.
Mass affluent An industry term for a segment of the private wealth marketplace that is not sufficiently wealthy to command certain individualized services.
Matrix pricing An approach for estimating the prices of thinly traded securities based on the prices of securities with similar attributions, such as similar credit rating, maturity, or economic sector. Also called *evaluated pricing*.
Maturity premium Compensation for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended.
Maximum loss optimization A stress test in which we would try to optimize mathematically the risk variable that would produce the maximum loss.
Mega-cap buyout funds A class of buyout funds that take public companies private.
Mental accounting bias An information-processing bias in which people treat one sum of money differently from another equal-sized sum based on which mental account the money is assigned to.
Micro attribution Performance attribution analysis carried out on the investment manager level.
Micro expectations Expectations concerning individual assets.
Middle-market buyout funds A class of buyout funds that purchase private companies whose revenues and profits are too small for them to access capital from the public equity markets.
Midquote The halfway point between the market bid and ask prices.
Minimum-variance hedge ratio A mathematical approach to determining the optimal cross hedging ratio.
Mismatch in character The potential tax inefficiency that can result if the instrument being hedged, and the tool that is being used to hedge it, produce income and loss of a different character.
Missed trade opportunity costs Unrealized profit/loss arising from the failure to execute a trade in a timely manner.
Model risk The risk that a model is incorrect or misapplied; in investments, it often refers to valuation models.
Model uncertainty Uncertainty concerning whether a selected model is correct.
Modified duration An adjustment of the duration for the level of the yield. Contrast with *Macaulay duration*.

Monetary policy Government activity concerning interest rates and the money supply.
Monetize To access an item's cash value without transferring ownership of it.
Money duration A measure of the price change in units of the currency in which the bond is denominated given a change in its yield-to-maturity.
Money markets Markets for fixed-income securities with maturities of one year or less.
Money-weighted rate of return Same as the internal rate of return; the growth rate that will link the ending value of the account to its beginning value plus all intermediate cash flows.
Multifactor model A model that explains a variable in terms of the values of a set of factors.
Multiperiod Sharpe ratio A Sharpe ratio based on the investment's multiperiod wealth in excess of the wealth generated by the risk-free investment.
Mutual funds A professionally managed investment pool in which investors in the fund typically each have a pro-rata claim on the income and value of the fund.
Mutuals With respect to insurance companies, companies that are owned by their policyholders, who share in the company's surplus earnings.
Natural liquidity An extensive pool of investors who are aware of and have a potential interest in buying and/or selling a security.
Negative screening An ESG implementation approach that excludes certain sectors or companies that deviate from an investor's accepted standards.
Net asset value Value established at the end of each trading day based on the fund's valuation of all existing assets minus liabilities, divided by the total number of shares outstanding.
Net employment capital See *human capital*.
Net interest margin With respect to banks, net interest income (interest income minus interest expense) divided by average earning assets.
Net interest spread With respect to the operations of insurers, the difference between interest earned and interest credited to policyholders.
Net wealth The difference between an individual's assets and liabilities; extends traditional financial assets and liabilities to include human capital and future consumption needs.
Net worth tax or net wealth tax A tax based on a person's assets, less liabilities.
Nominal default-free bonds Conventional bonds that have no (or minimal) default risk.
Nominal gross domestic product A money measure of the goods and services produced within a country's borders. Also called *nominal GDP*.
Nominal risk-free interest rate The sum of the real risk-free interest rate and the inflation premium.
Non-deliverable forwards Forward contracts that are cash settled (in the non-controlled currency of the currency pair) rather than physically settled (the controlled currency is neither delivered nor received).
Nonfinancial risk Risks that arise from sources other than the external financial markets, such as changes in accounting rules, legal environment, or tax rates.
Nonparametric Involving minimal probability-distribution assumptions.
Nonstationarity A property of a data series that reflects more than one set of underlying statistical properties.
Normal portfolio A portfolio with exposure to sources of systematic risk that are typical for a manager, using the manager's past portfolios as a guide.
Objective function A quantitative expression of the objective or goal of a process.
Offer price The price at which a counterparty is willing to sell one unit of the base currency.
Open market operations The purchase or sale by a central bank of government securities, which are settled using reserves, to influence interest rates and the supply of credit by banks.
Open outcry auction market Public auction where representatives of buyers and sellers meet at a specified location and place verbal bids and offers.
Operational risk The risk of loss from failures in a company's systems and procedures (for example, due to computer failures or human failures) or events completely outside of the control of organizations (which would include "acts of God" and terrorist actions).
Opportunistic participation strategies Passive trading combined with the opportunistic seizing of liquidity.
Option-adjusted spread The constant spread that, when added to all the one-period forward rates on the interest rate tree, makes the arbitrage-free value of the bond equal to its market price.
Optional stock dividends A type of dividend in which shareholders may elect to receive either cash or new shares.
Order-driven markets Markets in which transaction prices are established by public limit orders to buy or sell a security at specified prices.
Ordinary life insurance A type of life insurance policy that involves coverage for the whole of the insured's life. Also called *whole life insurance.*
Orphan equities investing A distressed securities investment discipline that involves investment in orphan equities that are perceived to be undervalued.
Orphan equity The newly issued equity of a company emerging from reorganization.
Output gap The difference between the value of GDP estimated as if the economy were on its trend growth path (potential output) and the actual value of GDP.
Overall trade balance The sum of the current account (reflecting exports and imports) and the financial account (consisting of portfolio flows).
Overbought When a market has trended too far in one direction and is vulnerable to a trend reversal, or correction.
Overconfidence bias A bias in which people demonstrate unwarranted faith in their own intuitive reasoning, judgments, and/or cognitive abilities.
Overconfidence trap The tendency of individuals to overestimate the accuracy of their forecasts.
Overlay A derivative position (or positions) used to adjust a pre-existing portfolio closer to its objectives.
Oversold The opposite of overbought; see *overbought*.
Packeting Splitting stock positions into multiple parts.
Panel method A method of capital market expectations setting that involves using the viewpoints of a panel of experts.
Partial correlation In multivariate problems, the correlation between two variables after controlling for the effects of the other variables in the system.
Partial fill Execution of a purchase or sale for fewer shares than was stipulated in the order.

Participate (do not initiate) order A variant of the market-not-held order. The broker is deliberately low-key and waits for and responds to the initiatives of more active traders.

Passive investment Investment that seeks to mimic the prevailing characteristics of the overall investments available in terms of credit quality, type of borrower, maturity, and duration rather than express a specific market view.

Passive management A buy-and-hold approach to investing in which an investor does not make portfolio changes based upon short-term expectations of changing market or security performance.

Passive traders Traders that seek liquidity in their rebalancing transactions, but are much more concerned with the cost of trading.

Payment netting A means of settling payments in which the amount owed by the first party to the second is netted with the amount owed by the second party to the first; only the net difference is paid.

Pension funds Funds consisting of assets set aside to support a promise of retirement income.

Pension surplus Pension assets at market value minus the present value of pension liabilities.

Percent-range rebalancing An approach to rebalancing that involves setting rebalancing thresholds or trigger points, stated as a percentage of the portfolio's value, around target values.

Percentage-of-volume strategy A logical participation strategy in which trading takes place in proportion to overall market volume (typically at a rate of 5–20 percent) until the order is completed.

Perfect markets Markets without any frictional costs.

Performance appraisal The evaluation of portfolio performance; a quantitative assessment of a manager's investment skill.

Performance attribution A comparison of an account's performance with that of a designated benchmark and the identification and quantification of sources of differential returns.

Performance evaluation The measurement and assessment of the outcomes of investment management decisions.

Performance measurement A component of performance evaluation; the relatively simple procedure of calculating an asset's or portfolio's rate of return.

Performance netting risk For entities that fund more than one strategy and have asymmetric incentive fee arrangements with the portfolio managers, the potential for loss in cases where the net performance of the group of managers generates insufficient fee revenue to fully cover contractual payout obligations to all portfolio managers with positive performance.

Periodic auction markets Auction markets where multilateral trading occurs at a single price at a prespecified point in time.

Permanent income hypothesis The hypothesis that consumers' spending behavior is largely determined by their long-run income expectations.

Permanent life insurance A type of life insurance that provides lifetime coverage.

Personal risk bucket In goal-based portfolio planning, that part of wealth allocated to investments intended to protect the client from a drastic decrease in lifestyle.

Plan sponsor The trustee, company, or employer responsible for a public or private institutional investment plan.

Pledging requirement With respect to banks, a required collateral use of assets.

Point estimate A single-valued estimate of a quantity, as opposed to an estimate in terms of a range of values.

Policyholder reserves With respect to an insurance company, an amount representing the estimated payments to policyholders, as determined by actuaries, based on the types and terms of the various insurance policies issued by the company.

Political risk The risk of war, government collapse, political instability, expropriation, confiscation, or adverse changes in taxation. Also called *geopolitical risk*.

Portable Moveable. With reference to a pension plan, one in which a plan participant can move his or her share of plan assets to a new plan, subject to certain rules, vesting schedules, and possible tax penalties and payments.

Portfolio overlay An array of derivative positions managed separately from the securities portfolio to achieve overall intended portfolio characteristics.

Portfolio trade A trade in which a number of securities are traded as a single unit. Also called *program trade* or *basket trade*.

Position a trade To take the other side of a trade, acting as a principal with capital at risk.

Positive screening An ESG implementation approach that seeks to identify the most favorable companies and sectors based on ESG considerations. Also called *best-in-class*.

Post-trade transparency Degree to which completed trades are quickly and accurately reported to the public.

Potential output The value of GDP if the economy were on its trend growth path.

Preferred return With respect to the compensation of private equity fund managers, a hurdle rate.

Premature death risk The risk of an individual dying earlier than anticipated; sometimes referred to as *mortality risk*.

Premium Regarding life insurance, the asset paid by the policy holder to an insurer who, in turn, has a contractual obligation to pay death benefit proceeds to the beneficiary named in the policy.

Prepackaged bankruptcy A bankruptcy in which the debtor seeks agreement from creditors on the terms of a reorganization before the reorganization filing.

Prepaid variable forward A collar and loan combined within a single instrument.

Present value of a basis point (PVBP) The change in the bond price for a 1 basis point change in yield. Also called *basis point value* (BPV).

Present value of distribution of cash flows methodology Method used to address a portfolio's sensitivity to rate changes along the yield curve, this approach seeks to approximate and match the yield curve risk of an index over discrete time periods.

Pretrade transparency Ability of individuals to quickly, easily, and inexpensively obtain accurate information about quotes and trades.

Price discovery Adjustment of transaction prices to balance supply and demand.

Price improvement Execution at a price that is better than the price quoted at the time of order placement.

Price uncertainty Uncertainty about the price at which an order will execute.

Price value of a basis point (PVBP) The change in the bond price for a 1 basis point change in yield. Also called *basis point value* (BPV).

Priced risk Risk for which investors demand compensation.

Primary capital Assets held outside a concentrated position that are at least sufficient to provide for the owner's lifetime spending needs.

Prime brokerage A suite of services that is often specified to include support in accounting and reporting, leveraged trade execution, financing, securities lending (related to short-selling activities), and start-up advice (for new entities).

Principal trade A trade with a broker in which the broker commits capital to facilitate the prompt execution of the trader's order to buy or sell.

Private equity Ownership interests in non-publicly traded companies.

Private equity funds Pooled investment vehicles that generally invest in highly illiquid assets; include venture capital funds and buyout funds.

Private placement memorandum A document used to raise venture capital financing when funds are raised through an agent.

Probate The legal process to confirm the validity of a will so that executors, heirs, and other interested parties can rely on its authenticity.

Profession An occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.

Profit-sharing plans A defined-contribution plan in which contributions are based, at least in part, on the plan sponsor's profits.

Program trading A strategy of buying or selling many stocks simultaneously.

Projected benefit obligation (PBO) A measure of a pension plan's liability that reflects accumulated service in the same manner as the ABO but also projects future variables, such as compensation increases.

Property insurance A type of insurance used by individuals to manage property risk.

Property risk The possibility that a person's property may be damaged, destroyed, stolen, or lost.

Prospect theory An alternative to expected utility theory, it assigns value to gains and losses (changes in wealth) rather than to final wealth, and probabilities are replaced by decision weights. In prospect theory, the shape of a decision maker's value function is assumed to differ between the domain of gains and the domain of losses.

Protective put An option strategy in which a long position in an asset is combined with a long position in a put.

Proxy hedge See *cross hedge*.

Prudence trap The tendency to temper forecasts so that they do not appear extreme; the tendency to be overly cautious in forecasting.

Public good A good that is not divisible and not excludable (a consumer cannot be denied it).

Purchasing power parity The theory that movements in an exchange rate should offset any difference in the inflation rates between two countries.

Pure indexing Method investors use to match an underlying market index in which the investor aims to replicate an existing market index by purchasing all of the constituent securities in the index to minimize tracking risk.

Pure sector allocation return A component of attribution analysis that relates relative returns to the manager's sector-weighting decisions. Calculated as the difference between the allocation (weight) of the portfolio to a given sector and the portfolio's benchmark weight for that sector, multiplied by the difference between the sector benchmark's return and the overall portfolio's benchmark return, summed across all sectors.

Put spread A strategy used to reduce the upfront cost of buying a protective put, it involves buying a put option and writing another put option.

Quality control charts A graphical means of presenting performance appraisal data; charts illustrating the performance of an actively managed account versus a selected benchmark.

Quantitative easing A policy measure in which a central bank buys financial assets to inject a predetermined quantity of money in the financial system.

Quote-driven markets Markets that rely on dealers to establish firm prices at which securities can be bought and sold. Also called *dealer markets*.

Quoted depth The number of shares available for purchase or sale at the quoted bid and ask prices.

Ratio spread An option strategy in which a long position in a certain number of options is offset by a short position in a certain number of other options on the same underlying, resulting in a risk-free position.

Rational economic man A self-interested, risk-averse individual who has the ability to make judgments using all available information in order to maximize his/her expected utility.

Re-base With reference to index construction, to change the time period used as the base of the index.

Real estate Interests in land or structures attached to land.

Real estate investment trusts (REITs) Publicly traded equities representing pools of money invested in real estate properties and/or real estate debt.

Real option An option involving decisions related to tangible assets or processes.

Real risk-free interest rate The single-period interest rate for a completely risk-free security if no inflation were expected.

Rebalancing In the context of asset allocation, a discipline for adjusting the portfolio to align with the strategic asset allocation.

Rebalancing overlay A type of overlay that addresses a portfolio's need to sell certain constituent securities and buy others.

Rebalancing range A range of values for asset class weights defined by trigger points above and below target weights, such that if the portfolio value passes through a trigger point, rebalancing occurs. Also known as a corridor.

Rebate rate The portion of the collateral earnings rate that is repaid to the security borrower by the security lender.

Recallability trap The tendency of forecasts to be overly influenced by events that have left a strong impression on a person's memory.

Recession A broad-based economic downturn, conventionally defined as two successive quarterly declines in GDP.

Regime A distinct governing set of relationships.

Regret The feeling that an opportunity has been missed; typically an expression of *hindsight bias*.

Regret-aversion bias An emotional bias in which people tend to avoid making decisions that will result in action out of fear that the decision will turn out poorly.

Regulatory risk The risk associated with the uncertainty of how a transaction will be regulated or with the potential for regulations to change.

Reinvestment risk The risk of reinvesting coupon income or principal at a rate less than the original coupon or purchase rate.

Relative economic strength forecasting approach An exchange rate forecasting approach that suggests that a strong pace of economic growth in a country creates attractive investment opportunities, increasing the demand for the country's currency and causing it to appreciate.

Repo rate The interest rate on a repurchase agreement.

Representativeness bias A belief perseverance bias in which people tend to classify new information based on past experiences and classifications.

Repurchase agreements (repos) In a repurchase agreement, a security owner agrees to sell a security for a specific cash amount, while simultaneously agreeing to repurchase the security at a specified future date (typically one day later) and price.

Repurchase yield The negative of the expected percent change in number of shares outstanding, in the Grinold–Kroner model.

Residence jurisdiction A framework used by a country to determine the basis for taxing income, based on residency.

Residence–residence conflict When two countries claim residence of the same individual, subjecting the individual's income to taxation by both countries.

Residence–source conflict When tax jurisdiction is claimed by an individual's country of residence and the country where some of their assets are sourced; the most common source of double taxation.

Resistance levels Price points on dealers' order boards where one would expect to see a clustering of offers.

Retired-lives The portion of a pension fund's liabilities associated with retired workers.

Returns-based benchmarks Benchmarks constructed by examining a portfolio's sensitivity to a set of factors, such as the returns for various style indexes (e.g., small-cap value, small-cap growth, large-cap value, and large-cap growth).

Reverse repos Repurchase agreement from the standpoint of the lender.

Revocable trust A trust arrangement wherein the settlor (who originally transfers assets to fund the trust) retains the right to rescind the trust relationship and regain title to the trust assets.

Risk budgeting The establishment of objectives for individuals, groups, or divisions of an organization that takes into account the allocation of an acceptable level of risk.

Risk exposure A source of risk. Also, the state of being exposed or vulnerable to a risk.

Risk premium approach An approach to forecasting the return of a risky asset that views its expected return as the sum of the risk-free rate of interest and one or more risk premiums.

Risk reversal With respect to foreign exchange option strategies, one involving a long position in a call option and a short position in a put option.

Risk tolerance The capacity to accept risk; the level of risk an investor (or organization) is willing and able to bear.

Rolling return The moving average of the holding-period returns for a specified period (e.g., a calendar year) that matches the investor's time horizon.

Sale and leaseback A transaction wherein the owner of a property sells that property and then immediately leases it back from the buyer at a rate and term acceptable to the new owner and on financial terms consistent with the marketplace.

Sample estimator A formula for assigning a unique value (a point estimate) to a population parameter.

Sample-size neglect A type of representativeness bias in which financial market participants incorrectly assume that small sample sizes are representative of populations (or "real" data).

Sandwich spread An option strategy that is equivalent to a short butterfly spread.

Satisfice A combination of "satisfy" and "suffice" describing decisions, actions, and outcomes that may not be optimal, but are adequate.

Savings–investment imbalances forecasting approach An exchange rate forecasting approach that explains currency movements in terms of the effects of domestic savings–investment imbalances on the exchange rate.

Scenario analysis A risk assessment technique involving the examination of the performance of a portfolio under specified situations.

Seagull spread An extension of the risk reversal foreign exchange option strategy that limits downside risk.

Sector/quality effect In a fixed-income attribution analysis, a measure of a manager's ability to select the "right" issuing sector and quality group.

Securities lending A form of collateralized lending that may be used to generate income for portfolios.

Security selection effect In a fixed-income attribution analysis, the residual of the security's total return after other effects are accounted for; a measure of the return due to ability in security selection.

Segmentation With respect to the management of insurance company portfolios, the notional subdivision of the overall portfolio into sub-portfolios each of which is associated with a specified group of insurance contracts.

Selective An index construction methodology that targets only those securities with certain characteristics.

Self-attribution bias A bias in which people take personal credit for successes and attribute failures to external factors outside the individual's control.

Self-control bias A bias in which people fail to act in pursuit of their long-term, overarching goals because of a lack of self-discipline.

Sell side Broker/dealers that sell securities and make recommendations for various customers, such as investment managers and institutional investors.

Separate property regime A marital property regime under which each spouse is able to own and control property as an individual.

Settlement netting risk The risk that a liquidator of a counterparty in default could challenge a netting arrangement so that profitable transactions are realized for the benefit of creditors.

Settlement risk When settling a contract, the risk that one party could be in the process of paying the counterparty while the counterparty is declaring bankruptcy.

Settlor (or grantor) An entity that transfers assets to a trustee, to be held and managed for the benefit of the trust beneficiaries.

Shari'a The law of Islam. In addition to the law of the land, some follow guidance provided by Shari'a or Islamic law.

Sharpe ratio A measure of risk-adjusted performance that compares excess returns to the total risk of the account, where total risk is measured by the account's standard deviation of returns. Also called *reward-to-variability*.

Short sale against the box Shorting a security that is held long.

Shortfall probability The probability of failing to meet a specific liability or goal.

Shortfall risk The risk that portfolio value will fall below some minimum acceptable level during a stated time horizon; the risk of not achieving a specified return target.

Shrinkage estimation Estimation that involves taking a weighted average of a historical estimate of a parameter and some other parameter estimate, where the weights reflect the analyst's relative belief in the estimates.

Shrinkage estimator The formula used in shrinkage estimation of a parameter.

Smart beta Involves the use of simple, transparent, rules-based strategies as a basis for investment decisions.

Smart routing The use of algorithms to intelligently route an order to the most liquid venue.

Smoothing rule With respect to spending rates, a rule that averages asset values over a period of time in order to dampen the spending rate's response to asset value fluctuation.

Social proof A bias in which individuals tend to follow the beliefs of a group.

Soft dollars The use of commissions to buy services other than execution services. Also called *soft dollar arrangements* or *soft commissions*.

Sole ownership Owned by one person; assets held in sole ownership are typically considered part of a decedent's estate. The transfer of their ownership is dictated by the decedent's will through the probate process.

Solow residual A measure of the growth in total factor productivity that is based on an economic growth model developed by economist Robert M. Solow.

Sortino ratio A performance appraisal ratio that divides the difference between a portfolio's return and a minimum acceptable return by downside deviation.

Source jurisdiction A framework used by a country to determine the basis for taxing income or transfers. A country that taxes income as a source within its borders imposes source jurisdiction.

Source–source conflict When two countries claim source jurisdiction of the same asset; both countries may claim that the income is derived from their jurisdiction.

Sovereign risk A form of credit risk in which the borrower is the government of a sovereign nation.

Special dividends A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Spread curve The fitted curve of credit spreads for each bond of an issuer plotted against either the maturity or duration of each of those bonds.

Spread duration A measure used in determining a portfolio's sensitivity to changes in credit spreads.

Stale price bias Bias that arises from using prices that are stale because of infrequent trading.

Static hedge A hedge that is not sensitive to changes in the price of the asset hedged.

Stationary A series of data for which the parameters that describe a return-generating process are stable.

Status quo bias An emotional bias in which people do nothing (i.e., maintain the "status quo") instead of making a change.

Status quo trap The tendency for forecasts to perpetuate recent observations—that is, to predict no change from the recent past.

Sterling ratio The compound annualized rate of return over a specified time period divided by the average yearly maximum drawdown over the same time period less an arbitrary 10%.

Stock companies With respect to insurance companies, companies that have issued common equity shares.

Stock lending Securities lending involving the transfer of equities.

Stops Stop-loss orders involve leaving bids or offers away from the current market price to be filled if the market reaches those levels.

Straddle An option strategy involving the purchase of a put and a call on the same underlying with the same exercise price and expiration date. If the put and call are held long, it is a long straddle; if they are held short, it is a short straddle.

Straight-through processing Systems that simplify transaction processing through the minimization of manual and/or duplicative intervention in the process from trade placement to settlement.

Strangle A variation on a straddle in which the put and call have different exercise prices; if the put and call are held long, it is a long strangle; if they are held short, it is a short strangle.

Strap An option strategy involving the purchase of two calls and one put.

Strategic asset allocation 1) The process of allocating money to IPS-permissible asset classes that integrates the investor's return objectives, risk tolerance, and investment constraints with long-run capital market expectations. 2) The result of the above process, also known as the policy portfolio.

Strategic buyers Buyers who have a strategic motive (e.g., realization of synergies) for seeking to buy a company.

Stratified sampling A sampling method that guarantees that subpopulations of interest are represented in the sample. Also called *representative sampling* or *cell approach*.

Strip An option strategy involving the purchase of two puts and one call.

Structural level of unemployment The level of unemployment resulting from scarcity of a factor of production.

Structural risk Risk that arises from portfolio design, particularly the choice of the portfolio allocations.

Structured note A variation of a floating-rate note that has some type of unusual characteristic such as a leverage factor or in which the rate moves opposite to interest rates.

Stylized scenario A type of analysis often used in stress testing. It involves simulating the movement in at least one interest rate, exchange rate, stock price, or commodity price relevant to the portfolio.

Sunshine trades Public display of a transaction (usually high-volume) in advance of the actual order.

Support levels Price points on dealers' order boards where one would expect to see a clustering of bids.

Surplus The difference between the value of assets and the present value of liabilities. With respect to an insurance company, the net difference between the total assets and total liabilities (equivalent to policyholders' surplus for a mutual insurance company and stockholders' equity for a stock company).

Surplus capital Capital that is in excess of primary capital.
Survey method A method of capital market expectations setting that involves surveying experts.
Survival probability The probability an individual survives in a given year; used to determine expected cash flow required in retirement.
Survivorship bias Bias that arises in a data series when managers with poor track records exit the business and are dropped from the database whereas managers with good records remain; when a data series as of a given date reflects only entities that have survived to that date.
Tactical asset allocation Asset allocation that involves making short-term adjustments to asset class weights based on short-term predictions of relative performance among asset classes.
Tail risk The risk that there are more actual events in the tail of a probability distribution than would be predicted by probability models.
Tail value at risk (or conditional tail expectation) The VaR plus the expected loss in excess of VaR, when such excess loss occurs.
Target covariance matrix A component of shrinkage estimation; allows the analyst to model factors that are believed to influence the data over periods longer than observed in the historical sample.
Tax avoidance Developing strategies that minimize tax, while conforming to both the spirit and the letter of the tax codes of jurisdictions with taxing authority.
Tax evasion The practice of circumventing tax obligations by illegal means such as misreporting or not reporting relevant information to tax authorities.
Tax premium Compensation for the effect of taxes on the after-tax return of an asset.
Tax risk The uncertainty associated with tax laws.
Taylor rule A rule linking a central bank's target short-term interest rate to the rate of growth of the economy and inflation.
Temporary life insurance A type of life insurance that covers a certain period of time, specified at purchase. Commonly referred to as "term" life insurance.
Term life insurance A type of life insurance policy that provides coverage for a specified length of time and accumulates little or no cash values.
Territorial tax system A framework used by a country to determine the basis for taxing income or transfers. A country that taxes income as a source within its borders imposes source jurisdiction.
Testamentary gratuitous transfer The bequeathing or transfer of assets upon one's death. From a recipient's perspective, it is called an inheritance.
Testator A person who makes a will.
Thematic investing An investment approach that focuses on companies within a specific sector or following a specific theme, such as energy efficiency or climate change.
Theta The change in price of an option associated with a one-day reduction in its time to expiration; the rate at which an option's time value decays.
Time deposit A deposit requiring advance notice prior to a withdrawal.
Time-series estimators Estimators that are based on lagged values of the variable being forecast; often consist of lagged values of other selected variables.
Time to expiration The time remaining in the life of a derivative, typically expressed in years.
Time value The difference between the market price of an option and its intrinsic value, determined by the uncertainty of the underlying over the remaining life of the option.
Time-weighted average price (TWAP) strategy A logical participation strategy that assumes a flat volume profile and trades in proportion to time.
Time-weighted rate of return The compound rate of growth over a stated evaluation period of one unit of money initially invested in the account.
Tobin's *q* An asset-based valuation measure that is equal to the ratio of the market value of debt and equity to the replacement cost of total assets.
Top-down approach A credit strategy approach that involves formulating a view on major macroeconomic trends and then selecting the bonds that the investor expects to perform best in the expected environment.
Total factor productivity (TFP) A variable which accounts for that part of Y not directly accounted for by the levels of the production factors (K and L).
Total future liability With respect to defined-benefit pension plans, the present value of accumulated and projected future service benefits, including the effects of projected future compensation increases.
Total rate of return A measure of the increase in the investor's wealth due to both investment income (for example, dividends and interest) and capital gains (both realized and unrealized).
Total return equity swap A swap contract that involves a series of exchanges of the total return on a specified asset or equity index in return for specified fixed or floating rate payments.
Total return payer Party responsible for paying the reference obligation cash flows and return to the receiver, but will also be compensated by the receiver for any depreciation in the index or default losses incurred on the portfolio.
Total return receiver Party receives both the cash flows from the underlying index as well as any appreciation in the index over the period in exchange for paying Libor plus a pre-determined spread.
Total return swap A swap in which one party agrees to pay the total return on a security. Often used as a credit derivative, in which the underlying is a bond.
Tracking error The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk. Also called *tracking risk*.
Tracking risk The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk. Also called *tracking error*.
Trade blotter A device for entering and tracking trade executions and orders to trade.
Trade settlement Completion of a trade wherein purchased financial instruments are transferred to the buyer and the buyer transfers money to the seller.
Trading activity In fixed-income attribution analysis, the effect of sales and purchases of bonds over a given period; the total portfolio return minus the other components determining the management effect in an attribution analysis.
Transaction exposure The risk associated with a foreign exchange rate on a specific business transaction such as a purchase or sale.
Transfer coefficient The ability to translate portfolio insights into investment decisions without constraint.

Translation exposure The risk associated with the conversion of foreign financial statements into domestic currency.

Transparency Availability of timely and accurate market and trade information.

Trigger points In the context of portfolio rebalancing, the endpoints of a rebalancing range (corridor).

Twist With respect to the yield curve, a movement in contrary directions of interest rates at two maturities; a nonparallel movement in the yield curve.

Type I error With respect to manager selection, keeping (or hiring) managers with zero value-added. (Rejecting the null hypothesis when it is correct).

Type II error With respect to manager selection, firing (or not hiring) managers with positive value-added. (Not rejecting the null hypothesis when it is incorrect).

Underfunded plan A pension plan in which the ratio of the value of plan assets to the present value of plan liabilities is less than 100 percent.

Underwriting (profitability) cycle A cycle affecting the profitability of insurance companies' underwriting operations.

Unit-linked life insurance A type of ordinary life insurance in which death benefits and cash values are linked to the investment performance of a policyholder-selected pool of investments held in a so-called separate account. Also called *variable life insurance*.

Universal life insurance A type of life insurance policy that provides for premium flexibility, an adjustable face amount of death benefits, and current market interest rates on the savings element.

Unrelated business income With respect to the US tax code, income that is not substantially related to a foundation's charitable purposes.

Unstructured modeling Modeling without a theory on the underlying structure.

Urgency of the trade The importance of certainty of execution.

Utility The level of relative satisfaction received from the consumption of goods and services.

Utility theory Theory whereby people maximize the present value of utility subject to a present value budget constraint.

Valuation reserve With respect to insurance companies, an allowance, created by a charge against earnings, to provide for losses in the value of the assets.

Value at risk (VaR) A probability-based measure of loss potential for a company, a fund, a portfolio, a transaction, or a strategy over a specified period of time.

Value-motivated traders Traders that act on value judgments based on careful, sometimes painstaking research. They trade only when the price moves into their value range.

Variable life insurance A type of ordinary life insurance in which death benefits and cash values are linked to the investment performance of a policyholder-selected pool of investments held in a so-called separate account. Also called *unit-linked life insurance*.

Variable universal life A type of life insurance policy that combines the flexibility of universal life with the investment choice flexibility of variable life. Also called *flexible-premium variable life*.

Vega A measure of the sensitivity of an option's price to changes in the underlying's volatility.

Venture capital The equity financing of new or growing private companies.

Venture capital firms Firms representing dedicated pools of capital for providing equity or equity-linked financing to privately held companies.

Venture capital fund A pooled investment vehicle for venture capital investing.

Venture capital trust An exchange-traded, closed-end vehicle for venture capital investing.

Venture capitalists Specialists who seek to identify companies that have good business opportunities but need financial, managerial, and strategic support.

Vested With respect to pension benefits or assets, said of an unconditional ownership interest.

Vintage year With reference to a private equity fund, the year it closed.

Vintage year effects The effects on returns shared by private equity funds closed in the same year.

Volatility Represented by the Greek letter sigma (σ), the standard deviation of price outcomes associated with an underlying asset.

Volatility clustering The tendency for large (small) swings in prices to be followed by large (small) swings of random direction.

Volume-weighted average price (VWAP) The average price at which a security is traded during the day, where each trade price is weighted by the fraction of the day's volume associated with the trade.

Volume-weighted average price (VWAP) strategy A logical participation strategy that involves breaking up an order over time according to a prespecified volume profile.

Wealth relative The ending value of one unit of money invested at specified rates of return.

Whole life insurance A type of life insurance policy that involves coverage for the whole of the insured's life. Also called *ordinary life insurance*.

Will A document associated with estate planning that outlines the rights others will have over one's property after death. Also called *testament*.

Within-sector selection return In attribution analysis, a measure of the impact of a manager's security selection decisions relative to the holdings of the sector benchmark.

Worst-case scenario analysis A stress test in which we examine the worst case that we actually expect to occur.

Yardeni model An equity valuation model, more complex than the Fed model, that incorporates the expected growth rate in earnings.

Yield beta A measure of the sensitivity of a bond's yield to a general measure of bond yields in the market that is used to refine the hedge ratio.

Yield curve The relationship between yield and time to maturity.

Yield to worst The yield on a callable bond that assumes a bond is called at the earliest opportunity.

Z-spread The yield spread that must be added to each point of the implied spot yield curve to make the present value of a bond's cash flows equal its current market price. Also known as zero-volatility spread.

Zero-cost collar A transaction in which a position in the underlying is protected by buying a put and selling a call with the premium from the sale of the call offsetting the premium from the purchase of the put. It can also be used to protect a floating-rate borrower against interest rate increases with the premium on a long cap offsetting the premium on a short floor.

Index

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.