2024

FRM®

EXAM PART I

Financial Markets and Products

GARP®

FRM® | Financial Risk Manager

GARP®

FRM® | Financial Risk Manager

2024

FRM®

EXAM PART I

Financial Markets and Products

Pearson

Contents

PREFACE

I want to thank you on behalf of GARP's Board of Trustees and our professional certification program staff for your support of the Financial Risk Manager (FRM®) program.

It's gratifying to see that in the 26 years since the first FRM examination, the FRM program has become the global standard for educating and credentialing financial risk management professionals. Its worldwide effects in furthering the understanding and acceptance of financial risk management have been highly positive and, in many ways, transformative.

COVID is thankfully in the rearview mirror. We now can be much more flexible in expanding—and in certain instances re-focusing and updating—the FRM program to address the many new challenges encountered by financial institutions globally.

Our FRM program advisory committee, consisting of senior risk professionals from around the world, that meets regularly to debate and settle the FRM program's subject coverage, has found no shortage of subjects for inclusion in the FRM curriculum.

One of the advisory committee's more-material challenges is to understand and assess where the global financial services industry is headed, and then identify issues and subjects most important for risk management professionals.

The FRM advisory committee also recommends how the FRM program covers subject matter. Its objective is to ensure that candidates who complete the FRM program successfully can be confident that their skills have been assessed objectively, and that they possess the requisite knowledge to succeed as a risk management professional anywhere in the world.

The FRM program's coverage is dynamic. The advisory committee reacts to and tries to anticipate market changes, global economic trends, technological advances, and regulatory adjustments; and assesses how these will affect the necessary knowledge and skill sets of a risk management professional.

The biggest change to the program's coverage for 2024 revolves around credit risk measurement and management. About two-thirds of the subject readings in *Credit Risk Measurement and Management* were updated for 2024.

Notably in 2023, GARP expanded the FRM program's coverage of operational resilience, an issue of rapidly growing importance around the world. Materials deal with structural vulnerabilities and areas of the financial system that may be under stress. The transmission of shocks to the financial system, and the assessment, modeling, and measurement of potential points of failure are other important covered concepts.

Also notable in 2023, GARP added two chapters on machine learning (ML) in the FRM Part I *Quantitative Analysis* book. These chapters not only introduce the ML methods risk managers need to understand, but also address key issues associated with artificial intelligence (AI) and ML, including transparency, interpretability, and explainability; data considerations; and risks that arise from the use of AI/ML, including the potential for bias, discrimination, and unethical behavior.

Throughout the FRM curriculum, GARP aims, wherever possible, to present lessons learned from noteworthy current events to contextualize program content and give FRM candidates critical insight.

As you will see from reviewing the program's coverage and readings, it keeps up with a world that is becoming more interconnected and complex by the day.

GARP is committed to offering a program that is dynamic, sophisticated, and responsive to the needs of financial institutions and risk professionals around the world.

We wish you the very best as you study for the FRM exams. And much success in your career as a risk-management professional.

Yours truly,

Richard Apostolik

President & CEO

FRM® COMMITTEE

ATTRIBUTIONS

Author

John C. Hull, PhD, University Professor, Maple Financial Group Chair in Derivatives and Risk Management, University of Toronto

Reviewers

Erick W. Rengifo, PhD, Associate Professor of Economics, Fordham University

David W. Wiley, MBA, CFA, President, WHW Investments, LLC

Sam Wong, PhD, CEO and Chief Quant, 5Lattice Securities Limited

Tao Pang, PhD, FRM, CFA, Director of the Financial Mathematics Program, North Carolina State University

Masao Matsuda, PhD, FRM, CAIA, President and CEO, Crossgates Investment and Risk Management

Colin White, PhD, FRM, Economist, Canadian Securities Transition Office

Charles Currat, PhD, FRM, SVP, Head of Investment Risk Methodology, Wells Fargo Bank

Claus-Peter Mueller, PhD, FRM, BCS, Head of Group Organization, Transformation, and IT, Heta Asset Resolution

Daniel Homolya, PhD, FRM, ERP, Treasury Risk and Control Manager, Diageo

Jeffrey Goodwin, FRM, Founder and President, Crescendo Investments, LLC

Patrick Steiner, FRM, Large Institution Supervision Coordinating Committee, Federal Reserve Bank of New York

1

Banks

Learning Objectives

After completing this reading, you should be able to:

- Identify the major risks faced by banks and explain how these risks can arise.
- Distinguish between economic capital and regulatory capital.
- Summarize the Basel committee regulations for regulatory capital and their motivations.
- Explain how deposit insurance gives rise to a moral hazard problem.
- Describe investment banking financing arrangements, including private placement, public offering, best efforts, firm commitment, and Dutch auction approaches.
- Describe the potential conflicts of interest among commercial banking, securities services, and investment banking divisions of a bank, and recommend solutions to these conflict of interest problems.
- Describe the distinctions between the banking book and the trading book of a bank.
- Explain the originate-to-distribute banking model and discuss its benefits and drawbacks.

Banks are the cornerstone of the world's financial system. The activities of banks in many countries can be subdivided into *commercial banking* and *investment banking*.

Commercial banking involves the traditional activities of receiving deposits and making loans. These activities can be categorized as either retail or wholesale. Retail banking involves transacting with private individuals and small businesses. Wholesale banking involves transacting with large corporations. Loans and deposits are much larger in wholesale banking than in retail banking. As a result, the administrative costs per dollar of deposits (or loans) are lower. The spread between the rates paid on deposits and the rates charged on loans is lower for wholesale banking as well.

Meanwhile, investment banking involves a variety of activities such as:

- Raising debt or equity capital for companies;
- Providing advice to companies on mergers, acquisitions, and financing decisions; and
- Acting as a broker–dealer for trading debt, equity, and other securities.

In some countries, the commercial banking and investment banking sectors are strictly separated. The U.S., for example, once limited the extent to which a single corporation could engage in both commercial and investment banking. Until the repeal of the Glass–Steagall Act in 1999, investment banks were not allowed to take deposits and make loans while commercial banks were not allowed to arrange equity issuances for other companies.

Following the financial crisis of 2007–2008, policymakers in some countries prohibited banks from putting depositors' funds at risk by engaging in proprietary trading (often referred to as *prop trading*). This is the speculative trading that an investment bank does in the hope of increasing its profitability.

1.1 THE RISKS IN BANKING

In this section, we explain three major risks that banks face. In the next section, we will outline the way in which banks are regulated to ensure that they can survive these risks.

Market Risks

Market risks are the risks arising from a bank's exposure to movements in market variables (e.g., exchange rates, interest rates, commodity prices, and equity prices). These market variables are often referred to as *risk factors*. The value of a market variable is determined by trading in the financial markets.

Consider the exchange rate between the U.S. dollar (USD) and the British pound (GBP). If the demand to buy GBP using USD is greater than the demand to sell GBP for USD, the value of the exchange rate (USD per GBP) will increase. Similarly, if the demand to sell GBP is greater than the demand to buy GBP, the exchange rate will decrease.

The values of market variables can be affected by many different events. For example, the value of the British pound decreased in June 2016 after the United Kingdom voted to leave the European Union (an event that market participants viewed as bad news for the British economy). Another example can be seen with the reinstatement of sanctions by the U.S. government on oil-producer Iran in May 2018. This event led to an increase in the price of oil because market participants thought that it could reduce the supply of oil in global markets.

A bank's exposure to movements in the values of market variables arises primarily from its trading operations. As previously explained, proprietary trading by banks is not currently allowed in the U.S. However, banks provide corporate clients and institutional investors with a wide range of products whose values depend on the prices of market variables. Consider again the USD per GBP exchange rate. Among the transactions a corporate client may request are as follows.

- *Spot transactions:* where GBP is bought or sold for almost immediate delivery.
- *Forward contracts:* where an exchange rate for the purchase or sale of a certain amount of GBP on a future date is agreed.
- *Options:* where one side has the right (but not the obligation) to buy or sell GBP at a pre-arranged price (i.e., the exercise price) at a certain future time.

For many of these contracts, banks act as market makers by quoting both a bid (i.e., the price at which they are prepared to buy) and an ask (i.e., the price at which they are prepared to sell). Banks typically ensure that their exposures to market variables are kept within certain limits, but they do not (usually) eliminate those exposures entirely. As a result, banks are always exposed to some market risk.

Credit Risks

Credit risk arises from the possibility that borrowers will fail to repay their debts. For banks, loans to corporations and individuals are a major source of credit risk. If a borrower defaults, a loss is usually incurred. In a bankruptcy, the size of the loss depends on whether assets have been pledged as collateral and how the bank's claims rank compared with those of other creditors.[1]

[1] This is discussed in more detail in Chapter 17.

A bank builds expected losses into the interest rate it charges on loans. For example, suppose the bank's cost of funds (the average interest rate paid on deposits and on the bank's debt) is 1.5%. The average interest rate charged on loans might be 4%. The difference between the two interest rates (2.5% in our example) is referred to as the *net interest margin*. If a bank expects to lose 0.8% of what it lends, it will be left with 1.7% to cover administrative/operational costs and contribute to profits.

In this example, 0.8% is the bank's expected (or average) loan losses. However, loan losses show significant variation from year to year. During stressed economic conditions, a bank might experience loan losses as high as 4%, while during good economic times these losses might be as low as 0.2%.[2] Current regulations require banks to maintain enough capital to cover losses that are estimated to occur only once every thousand years.[3]

Other bank contracts also give rise to credit risk. For example, banks trade a variety of derivatives (e.g., forward contracts and options). As already indicated, these give rise to market risk because the value of a derivatives contract depends on the underlying market variables. Derivatives also give rise to credit risk. This comes from the possibility that the counterparty to a derivatives transaction will default when the transaction has a positive value to the bank (and therefore a negative value to the counterparty).

Banks typically account for expected losses on transactions as soon as they are initiated. An accounting rule known as IFRS 9, which covers most bank-issued loans, requires banks to show the outstanding principal net of estimated expected losses over the following 12 months on their balance sheet.[4] In the case of derivatives, banks calculate a credit value adjustment (CVA) reflecting the amount they expect to lose over the life of the derivatives due to counterparty default. This is subtracted from the balance sheet value of the outstanding derivatives. In both cases, expected losses, even though they have not (yet) been incurred, are charged to income.

[2] Statistics published by the credit rating agency S&P show that the default rate per year on all rated corporate debt varied between 0.14% and 4.19% between 1981 and 2018. The worst default rate (4.19%) was in 2009, following the financial crisis. Other years with defaults rates greater than 3% were 1991 (3.25%), 2001 (3.79%), and 2002 (3.63%). In 2018, the default rate was 1.03%.

Source: S&P Global. (2019, April 19). 2018 Annual Global Corporate Default And Rating Transition Study. https://www.spratings.com/documents/20184/774196/2018AnnualGlobalCorporateDefaultAndRatingTransitionStudy.pdf

[3] This is explained in more detail in Chapter 6 of *Valuation and Risk Models*.

[4] The Financial Accounting Standards Board (FASB), the accounting standard-setting body in the United States, has a rule similar to IFRS 9. In the case of the FASB rules, it is losses over the whole life of the loan which are subtracted from the net principal. In the case of IFRS 9, losses over the whole life of the loan are considered only when there is a significant increase in credit risk.

Operational Risks

Operational risk is defined by bank regulators as:[5]

> *The risk of loss resulting from inadequate or failed internal processes, people, and systems or from external events.*

This definition includes all risks that are not market or credit risks (with the exception of strategic and reputational risks). Operational risk is harder to quantify than either market risk or credit risk. Examples from seven categories of operational risk identified by regulators include the following.

- *Internal fraud:* Rogue traders intentionally misreporting positions or employees stealing from the bank by creating loans to fictitious companies.
- *External fraud:* Cyberattacks, bank robberies, forgery, and check kiting.
- *Employment practices and work place safety:* Worker compensation claims, employee discrimination claims, and litigation arising from personal injury claims at bank branches.
- *Clients, products, and business practices:* Money laundering and other actions that are either unlawful or prohibited by regulators.
- *Damage to physical assets:* Terrorism, vandalism, earthquakes, fires, and floods.
- *Business disruption and system failures:* Hardware and software failures, telecommunication problems, and utility outages.
- *Execution, delivery, and process management:* Data entry errors, collateral management issues, and inadequate legal documentation.

Operational risk is regarded by many to be a greater challenge for banks than either market risk or credit risk. Since 2008, banks in North America and Europe have been fined hundreds of billions of dollars for operational risk violations such as money laundering, market manipulation, terrorist financing, and inappropriate activities in the mortgage market.

Significant sources of operational risk in banking include cyber risk, legal risk, and compliance risk (i.e., failure to comply with rules and regulations, either accidentally or intentionally). These risks are discussed further in Chapter 7 of *Valuation and Risk Models*.

1.2 BANK REGULATION

Banks are subject to regulations designed to protect depositors as well as maintain confidence and stability in the financial system. In this section, we outline the development of the global banking regulatory environment.

[5] See Bank for International Settlements, "Working Paper on the Regulatory Treatment of Operational Risk," September 2001.

Capital

It is important for banks to keep sufficient capital for the risks they are taking. The most important capital is equity capital. Because losses have the effect of reducing equity capital, banks must try to maintain enough equity capital to cover potential losses and remain solvent (i.e., have a positive amount of equity capital). Debt capital is the other main category of capital, and it is usually subordinate to assets held for depositors (therefore providing an extra degree of protection for depositors).

Equity capital is sometimes referred to as *going concern capital* because it absorbs losses while the bank is a going concern (i.e., it remains in business). Debt capital is referred to as *gone concern capital* because it is only affected by losses once a bank has failed. In theory, depositors are at risk only when losses are sufficiently large to wipe out both equity and debt capital.

We can distinguish between *regulatory capital* and *economic capital*. Regulatory capital is the minimum capital that regulators require banks to keep. Economic capital is a bank's own estimate of the capital it requires. In both cases, capital can be thought of as funds that are available to absorb unexpected losses. A common objective in calculating economic capital is to maintain a high credit rating (as will be described in later chapters). Economic capital is allocated to a bank's business units so that they can be compared using a *return on allocated economic capital* metric.

The amount of capital that is necessary depends on the size of possible losses. If a bank's equity capital is USD 4 billion and there is a 1% chance that the bank will incur a loss higher than USD 4 billion over a year, the equity capital will be considered insufficient by both regulators and the bank itself. This is because even a 1% chance that the bank will become insolvent is unacceptable. As mentioned earlier, the regulatory capital for credit risk is designed to be sufficient to cover a loss that is expected to be exceeded only once every thousand years.[6]

The Basel Committee

The Basel Committee for Banking Supervision was established in 1974 to provide a forum where the bank regulators from different countries could exchange ideas.[7] Prior to 1988, bank regulation and enforcement varied from country to country. In 1988, there was an international agreement (which became known as Basel I) that required regulators in all signatory countries to calculate capital requirements in the same manner. Initially, these capital requirements were designed to cover losses arising from defaults on loans and derivatives contracts.

By the 1990s, however, bank trading activities had significantly increased. In response, the Basel Committee agreed that banks should keep capital for both market risk and credit risk. This modification to Basel I, known as the *Market Risk Amendment*, was implemented in 1998.

In 1999, the Basel Committee proposed what has become known as Basel II. This agreement revised the procedure for calculating credit risk capital and introduced a capital requirement for operational risk. It took about eight years for the final Basel II rules to be worked out and implemented. The total capital requirement was then the sum of amounts for (a) credit risk, (b) market risk, and (c) operational risk.

The 2007–2008 crisis led to several bank failures and bailouts. Global bank regulators subsequently determined that the rules for calculating market risk capital were inadequate. Thus, the rules were revised in what is referred to as Basel II.5

The Basel Committee also decided that equity capital requirements needed to be increased. This latest set of regulations, called Basel III, includes a large increase in the amount of equity capital that banks are required to keep and is expected to be fully implemented by 2027.[8]

Meanwhile, the rules for market risk have been revised yet again with the *Fundamental Review of the Trading Book*, which is due to be implemented in 2022.

Standardized Models versus Internal Models

Models are necessary to determine bank capital. Some models are standardized tools developed by the Basel Committee, while others are internal models developed by the banks themselves. Generally, banks need approval from regulators before they can use a specific internal model.

The models for credit risk that were introduced in Basel I were standardized models developed by the Basel Committee. This means that two banks, when presented with the same portfolio, should calculate the same capital requirements. The Market Risk Amendment included a standardized model approach and an internal model approach. Banks could determine market risk capital using an internal model provided that the model satisfied the requirements laid down by the Basel Committee and was approved by national regulators. Basel II allowed internal

[6] See Chapter 6 of *Valuation and Risk Models* for further discussion.

[7] The Basel Committee for Banking Supervision is based at the Bank for International Settlements in Basel, Switzerland.

[8] Bank for International Settlements. (2019, March 20). *Basel III monitoring results published by the Basel Committee* [Press release]. Retrieved from https://www.bis.org/press/p190320.htm

Many bankers refer to the second half of the Basel III rules, which were agreed in 2016 and 2017, as Basel IV. These rules include limits on the extent to which internal models can be used and will be described in the next section.

models to be used to determine both credit risk capital and operational risk capital.

Since the crisis, the Basel Committee has decided to reduce bank reliance on internal models. The committee felt that it had given banks too much freedom to choose internal models that would produce the lowest capital requirements. It now requires that banks use a standardized model for determining operational risk capital. For credit risk and market risk, banks must calculate capital using a standardized model and can (if they receive approval from their national regulators) also calculate capital using an internal model. However, these internal models cannot reduce total capital requirements below a minimum level that is set equal to a certain percentage of the capital given by the standardized approach. By 2027, this percentage will be 72.5%. This means that:

$$\text{Required Capital} = \max(\text{IMC}, 0.725 \times \text{SMC})$$

where IMC is the capital given by the internal models and SMC is the capital given by the standardized (Basel Committee) models.

Trading Book versus Banking Book

When calculating regulatory capital, it is important to distinguish between the trading book and the banking book. The trading book (as its name implies) consists of assets and liabilities that are held to trade. The banking book consists of assets and liabilities that are expected to be held until maturity. Items in the trading book are subject to market risk capital calculations, whereas items in the banking book are subject to credit risk capital calculations. These calculations are quite different. In the past, there had sometimes been ambiguity as to whether a transaction (e.g., a credit derivative) should be in the banking book or the trading book. Banks tended to take advantage of this ambiguity by putting each transaction in the book that would lead to the lowest capital requirement (usually this was the trading book).

The *Fundamental Review of the Trading Book* mentioned earlier attempts to clarify the Basel Committee's rules concerning whether an instrument should be in the banking book or the trading book. If a bank has a desk for trading a specific instrument, that instrument will normally be considered to be part of the trading book. Otherwise, it will be part of the banking book.

Liquidity Ratios

Many of the problems experienced during the financial crisis were a result of a lack of liquidity, rather than a shortage of capital. Consider a bank that wants to fund five-year loans. One possibility is to issue five-year bonds so that the maturities of its assets and liabilities are matched. A tempting alternative that could lead to lower funding costs (in many interest rate environments) is to issue a three-month commercial paper. At the end of the three months, a new three-month commercial paper is issued and used to repay the first issuance. At the end of a further three months, there is a third issuance of a three-month commercial paper, which would be used to repay the second issuance, and so on.

A risk with this strategy comes when the commercial paper cannot be rolled over in the way we have described. If the market (rightly or wrongly) loses confidence in the bank, it is likely that the maturing commercial paper cannot be replaced (or must be replaced at much higher interest rates). Unless the bank has other guaranteed lines of credit, it could default on its debt and go bankrupt. Note that if the five-year loans had been financed with five-year debt, this problem would have been avoided because the loan repayments could have been used to repay the debt.

The failure of Northern Rock in the United Kingdom can be traced to this type of liquidity problem. The British bank had a mortgage portfolio that it was partly funding with commercial paper. While this mortgage portfolio was not unduly risky, problems in the U.S. mortgage market made investors nervous, and the commercial paper could not be rolled over. Lehman's demise in 2008 was also largely a result of liquidity problems of this type.

As a result of the liquidity problems encountered during the crisis, the Basel Committee has (as part of Basel III) developed two liquidity ratios to which banks are required to adhere. *The Liquidity Coverage Ratio* is a requirement designed to ensure that banks have sufficient sources of funding to survive a 30-day period of acute stress (e.g., where it is downgraded, loses deposits, or has drawdowns on its lines of credit). The *Net Stable Funding Ratio* is a requirement that limits the size of mismatches between the maturity of assets and the maturity of liabilities.

1.3 DEPOSIT INSURANCE

To maintain confidence in the banking system, many countries have introduced deposit insurance. This typically provides a certain amount of protection to a depositor against losses arising from a bank failure. In the U.S., the amount is currently USD 250,000. In some jurisdictions, all banks pay the same insurance premium per year per dollar of deposit insured. In other jurisdictions (such as the U.S.), the insurance premium is based on an assessment of each bank's risk.

If deposit insurance were provided to a bank without any other measures being taken, the insurance might encourage banks to take on more risks than they would otherwise. For example, banks could offer slightly above average interest rates to depositors and then use the funds to make risky loans at relatively high interest rates to borrowers. Without deposit insurance, this would not be possible because depositors would withdraw their money when the risks being taken became apparent. With deposit insurance, the strategy might be feasible because depositors know that they are protected in the event of bank failure and will appreciate the above average interest rates they are receiving.

This argument is an example of what is known as a *moral hazard*, which can be defined as the risk that the behavior of an insured party will change because of the mere existence of the insurance, and thus the insurance contract will become riskier. It is a serious consideration in deposit insurance, because governments certainly do not want to set up a program that encourages a bank to take *larger* risks.

Risk-based deposit insurance premiums reduce the moral hazard to some extent. The moral hazard is also lessened by regulations that ensure that a bank's required capital increases with the risks it takes (see Section 1.2).

1.4 INVESTMENT BANKING

A major activity of a bank's investment banking arm is raising capital for companies in the form of debt, equity, or more complicated securities (e.g., convertible debt). This process is referred to as *underwriting*. Typically, a company will approach the investment bank to discuss its plans to issue securities. Once the plans have been agreed upon, the securities are originated along with documentation itemizing the rights of investors who purchase the securities. A prospectus detailing the company's past performance and future prospects is also produced. This includes a discussion of risks, any outstanding lawsuits, and other relevant information. There is usually a *road show* in which senior management from the issuing company and executives from the investment bank attempt to persuade investors to buy the securities. Finally, a price for the securities is agreed upon between the bank and the issuing company, and the bank then proceeds to market the securities.

There are two types of offerings.

1. *Private placements:* where the securities are sold (or placed) with a small number of large institutional investors (e.g., pension plans and life insurance companies).
2. *Public offerings:* where securities are offered for sale to the general public.

In the case of a private placement, the investment bank receives an agreed upon fee. In the case of a public offering, the agreement between the investment bank and the issuing company can be on *a best efforts* or a *firm commitment* basis. As its name implies, best efforts means that the bank will do its best to sell the securities for the agreed upon price. However, there are no guarantees. The bank is paid a fee that usually depends (to some extent) on how successful it has been in selling the securities for the agreed upon price.

In the case of a firm commitment, the bank does guarantee that the securities will be sold for an agreed upon price. The bank buys the securities at the agreed upon price and then attempts to sell them for a higher price. Its profit is the difference between the two prices. If it misjudges the market and is unable to sell the securities for more than the agreed upon price, it will incur a loss. A firm commitment is sometimes referred to as a *bought deal*.

A firm commitment arrangement is riskier for an investment bank (but less risky for the issuing company) than a best efforts arrangement. Suppose that a company wants to issue 10 million new shares. It is currently publicly traded, and its share price (which has risen recently) is USD 58. In negotiations with its investment bank, there are two offers on the table:

1. A best efforts arrangement where shares will be sold at the best possible price and the bank will be paid USD 1.50 per share sold (to keep the example simple, we assume that the bank's fee does not depend on the price at which the shares are sold); and
2. A firm commitment arrangement where the bank guarantees that the shares can be sold for USD 50.

Table 1.1 summarizes these alternatives from the perspective of the investment bank and considers two outcomes. In the first one, the shares can be sold for USD 55; in the second one, they can be sold for USD 48. The best efforts alternative is certain to give the bank a gross profit of USD 15 million. On the other hand, the firm commitment alternative is much riskier. If the shares can be sold for USD 55, the bank will make USD 50 million. If the shares can only be sold for USD 48, however, the bank will lose USD 20 million.

The decision taken by the bank will depend on the subjective probabilities it assigns to different outcomes in conjunction with its risk appetite. For the company, the risks are less with a firm commitment because it knows it will realize USD 500 million (regardless of the final market price). Under the best efforts arrangement, the maximum amount realized (after considering the bank's fee) would be USD 535 million for the first scenario in Table 1.1 and a maximum of USD 465 million for the second scenario.[9]

[9] Table 1.1 may understate the risks of a best efforts arrangement to the company. If there is a dramatic market downturn, the issue may be withdrawn so that the company raises no capital.

Table 1.1 Profit to Bank from Best Efforts and Firm Commitment Alternatives to Sell 10 Million Shares (USD Million)

	Best Efforts, Fee Equals USD 1.50 per Share Sold	Firm Commitment, Bank Buys Shares for USD 50
Price Realized = USD 55	+15	+50
Price Realized = USD 48	+15	−20

IPOs

An IPO (initial public offering) is a first-time offering of a company's shares to the public. Prior to an IPO, shares are typically held by the company's founders, venture capitalists, and others who have provided early stage funding. The shares being sold can be a mixture of existing and new shares, which can provide additional capital for the company. Sometimes the founders retain control by arranging for the shares they keep in the company to have better voting rights than other shares.

Because the company's shares do not yet trade on an exchange, it is difficult for an investment bank to accurately assess what the share price will be after the IPO.

For example, suppose the company wishes to raise USD 100 million. The investment bank must try and estimate

$$\frac{\text{Value of Company After USD 100 Million Cash Injection}}{\text{Number of Shares Post IPO}}$$

Typically, the investment bank sets the offering price below its best estimate to make it more likely that it can sell the issue at the offering price.

There is often a substantial increase in the share price after an IPO. This means that the company could have probably issued shares at a higher price, thereby raising more money. It also indicates that IPOs tend to be good investments. Unfortunately, it is often difficult for small investors to buy IPOs.

Dutch Auctions

The advantage of using investment banks to handle an IPO is that they have the necessary expertise as well as relationships with potential investors. However, some issuers feel that they would prefer for the market to decide the right price for their company. One way they can do this is through a Dutch auction. This is a procedure where all investors (not just clients of an investment bank) are invited to submit bids indicating how many shares they would like to purchase and at what price.

As a simple example of how a Dutch auction works, suppose that a company wants to sell 500,000 shares and has received the bids presented in Table 1.2. To evaluate the bids, it is necessary to sort bidders from the highest to the lowest. This has been done in Table 1.3. We then search for the maximum price at which 500,000 shares or more can be sold. From Table 1.3, we see that 30,000 shares have been bid for at USD 70 or more, 130,000 have been bid for at USD 65 or more, 170,000 have been bid for at USD 63 or more, and so on. Furthermore, 480,000 shares have been bid for at USD 56 or more and 680,000 have been bid for at USD 55 or more. The maximum price at which 500,000 shares can be sold is therefore USD 55. All successful bidders pay this price. The seven highest bidders in Table 1.3 get the full amount of the shares for which they bid. Bidder D gets 20,000 shares (the difference between the 500,000 being sold and the 480,000 for which a higher price than USD 55 has been bid).

An advantage of Dutch auctions is that (if all potential investors in a company bid) the price charged is the one that balances supply and demand in the market. In theory, the post-IPO price should be similar to the pre-IPO price.

One high profile IPO that used the Dutch auction approach was that of Google in 2004. This auction was a little different from

Table 1.2 Bids for Ten Participants in a Dutch Auction when 500,000 Shares are Being Sold

Bidder	Number of Shares Requested	Price Bid (USD)
A	100,000	65
B	50,000	60
C	30,000	70
D	200,000	55
E	70,000	58
F	150,000	61
G	40,000	63
H	40,000	56
I	80,000	54
J	100,000	50

Table 1.3 Bids in Table 1.2 Sorted from Highest to Lowest

Bidder	Number of Shares Requested	Cumulative Number of Shares Requested	Price Bid (USD)
C	30,000	30,000	70
A	100,000	130,000	65
G	40,000	170,000	63
F	150,000	320,000	61
B	50,000	370,000	60
E	70,000	440,000	58
H	40,000	480,000	56
D	200,000	680,000	55
I	80,000	760,000	54
J	100,000	860,000	50

the "plain vanilla" Dutch auction we have described. Instead, Google reserved the right to change (at the last minute) the number of shares that would be offered and the percentage of the requested amount allocated to each bidder. When it saw the bids, it decided that the number of shares being offered would be 19,605,052 at a price of USD 85. The total value of the offering was therefore USD 1.67 billion, and investors who had bid USD 85 or more got 74.2% of the shares for which they had bid. This was a surprising decision. Google could have raised USD 2.25 billion instead of USD 1.67 billion with a more usual Dutch auction (where investors bidding USD 85 or more got 100% of the shares for which they had bid). Perhaps founders Sergei Brin and Larry Page anticipated (correctly as it turned out) that they would be able to issue more shares at a much higher price later on.

On the first day of the new issuance, Google's shares closed at USD 100.34 (i.e., 18% above the issue price). This was followed by a further 7% increase on the second day. In this example, the use of a Dutch auction did not eliminate the IPO underpricing problem we mentioned earlier. Google did use two investment banks (Morgan Stanley and Credit Suisse First Boston) to assist in the issuance. However, the fee paid was less than it would have been for a regular IPO.

Advisory Services

In addition to handling securities issuances, investment banks also offer advice to corporations on decisions involving mergers and acquisitions, divestments, and restructurings. Specifically, they assist companies in finding acquisition partners and in finding buyers for divisions that are to be divested. Investment bankers will also advise companies that are the subject of a takeover attempt by another company.

In advising Company A on a potential takeover of Company B, it is necessary for an investment bank to value Company B and to assess any potential synergies (i.e., cost savings, economies of scale, market share, or other benefits from merging the two companies). It must also consider the type of offer that should be made. This could be a:

- *Cash offer:* where the existing shares of Company B are purchased for cash,
- *Share-for-share exchange:* where newly issued shares of Company A are exchanged for those of Company B so that Company B's shareholders become shareholders of Company A, or
- Combination of a cash offer and a share-for-share exchange.

In a cash offer, the acquisition's risk and uncertainties are borne by the acquiring company. In a share-for-share exchange, they are shared between the two companies.

The initial offer is not usually the final offer, and the investment bank must use its experience to develop a reasonable plan for the price negotiations. The investment bank must assess the best way to approach the management of the target company. The takeover may be hostile (i.e., opposed by existing management) or friendly (i.e., supported by management). In some instances, it may be necessary for investment bankers to consider antitrust concerns and whether regulatory approval for the merger will be necessary.

The companies targeted by takeover attempts are also advised by investment bankers. Sometimes a company (often with the advice of an investment bank) will take steps to avoid being taken over. These are known as poison pills. Examples of poison pills are

- Granting key employees attractive stock options that can be exercised in the event of a takeover—this could dissuade a potential acquirer from proceeding because the key employees will almost certainly leave;
- Adding a provision to the company's charter making it impossible for a new owner to fire the existing directors for a period of time;
- Issuing preferred shares that automatically get converted to regular shares in the event of a takeover;
- Allowing existing shareholders to purchase shares at a discount in the event of a takeover;
- Changing the voting structure so that management has sufficient votes to block a takeover; and
- Adding a provision that allows remaining shareholders to sell shares to the new owner at a 50% premium in the event of a successful takeover.

Poison pills are illegal in many countries, but they are permitted in the U.S. However, they must be approved by the majority of shareholders. If shareholders feel that the poison pills will benefit management at the expense of shareholders, they are likely to vote against them. One argument in favor of poison pills is that they can benefit shareholders by improving the negotiating position of management in the event of a takeover attempt (potentially resulting in a price more favorable to existing shareholders).

Trading

Another activity of investment banking is trading. As discussed earlier, regulations implemented since the 2007–2008 financial crisis have limited the extent to which banks can do speculative trading. In the U.S., the Volcker rule (part of the Dodd–Frank Act) does not allow U.S. banks to engage in proprietary trading.[10] In some other countries, such as the UK, proprietary trading must be *ring fenced* to ensure that depositors are not adversely affected by any losses.[11]

Banks offer market-making services where they quote bids and offers on a wide range of different products to meet the needs of corporate treasurers and institutional investors. These products typically depend on exchange rates, interest rates, precious metal prices, and so on. Some of the products involve options, forward contracts, and more complex derivatives that will be explained in later chapters. Typically, a bank will enter into a contract with a corporate end user and then hedge all or part of its risk by trading with another financial institution.

Banks also offer brokerage services to retail clients. They will take orders from a client and arrange for them to be executed on an exchange. A full-service broker offers investment research and advice. Others offer low-cost services where little or no advice is given to clients. Sometimes a broker manages discretionary accounts where clients have entrusted them to make investment decisions on their behalf.

1.5 CONFLICTS OF INTEREST

Banking gives rise to several potential conflicts of interest. The following are examples.

- Suppose that an investment banker is handling a new equity issue for a client and is finding it difficult to persuade investors to buy the shares. The investment banker might ask the bank's brokers to advise clients to buy the securities and to buy them for clients' discretionary accounts.
- Suppose an investment banker is advising a client about a possible acquisition. It knows that the target company is a client of the commercial banking arm of the bank. Investment bankers might ask the commercial bankers for their impressions of the target and thereby gain confidential information that can be passed on to the acquiring company.
- Suppose an investment bank is hoping for lucrative business from a company. It might contact researchers that work for the brokerage end of the bank and ask them to come up with a "buy" recommendation for the company's stock to please the company's management.
- Suppose that a bank's commercial banking arm has a large loan outstanding to a client where there is a high probability of a loss. It might suggest to the client that it replace the loan with a bond issue handled by the bank's investment banking arm. If the investment banking arm agrees to this, the debt is likely to be taken over by investors who are less informed than the commercial bankers.

Conflicts of interest are handled in part by what are termed *Chinese walls*. These are rules within a bank preventing the transfer of information from one part of the bank to another. A cynic might argue that a bank will not (in practice) enforce Chinese walls if they reduce profits. However, it is in the bank's interest to enforce Chinese walls. Big fines can be (and have been) levied for violations of conflict of interest rules.

[10] The Dodd-Frank Act is a US law passed in 2010 designed to provide greater oversight of financial institutions and reduce the risks they take. In 2018, another act was passed that exempted some smaller banks from certain Dodd-Frank provisions.

[11] Financial Conduct Authority. Ring-fencing. Retrieved from https://www.fca.org.uk/consumers/ring-fencing.

Additionally, a bank's reputation is its most valuable asset. A bank that is seen to be ignoring conflict of interest rules can lose business. These fines and reputational costs are generally much greater than any gains arising from conflict of interest violations.

The types of conflicts of interest mentioned earlier are the reason why U.S. regulators have in the past separated investment banking from commercial banking. Under the Glass–Steagall Act of 1933, commercial banks could assist with the issues of Treasury and municipal bonds and handle private placements, but they were not allowed to handle public offerings and other investment banking activities. Similarly, investment banks were not allowed to take deposits and make commercial loans.

In 2007, there were five large investment banks in the U.S.: Goldman Sachs, Morgan Stanley, Merrill Lynch, Bear Stearns, and Lehman Brothers. The 2007–2008 financial crisis led to a big upheaval. Lehman declared bankruptcy, Bear Stearns was taken over by JPMorgan Chase, and Merrill Lynch was taken over by Bank of America. Goldman Sachs and Morgan Stanley became banking holding companies with both commercial and investment banking interests.

1.6 THE ORIGINATE TO DISTRIBUTE MODEL

Traditionally, banks have originated loans and kept them on their balance sheet. An alternative to this is what has become known as the *originate-to-distribute model*. Under this model, banks use their expertise to originate loans and then sell them (directly or indirectly) to investors.

Originate-to-distribute arrangements have been used in the U.S. mortgage market for many years. The U.S. government has sponsored the creation of three entities:

1. Government National Mortgage Association (GNMA) or Ginnie Mae,
2. Federal National Mortgage Association (FNMA) or Fannie Mae, and
3. Federal Home Loan Mortgage Corporation (FHLMC) or Freddie Mac.

These agencies buy mortgage portfolios from banks and other mortgage originators, package the cash flows into securities, and sell the securities to investors. These investors are not subject to credit risk because the respective agency guarantees the mortgage payments. However, the securities are subject to prepayment risk. This is the risk that the mortgage principal will be

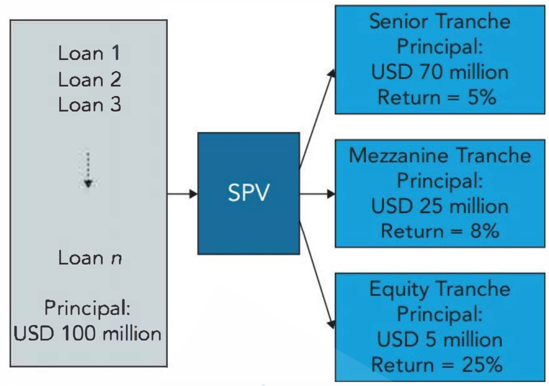


Figure 1.1 **Securitization of loans.**

paid off by the borrower earlier than expected.[12] Prior to 1999, the agencies only handled mortgages that had a low probability of default. In 1999, they started to accept subprime mortgages, which are much riskier.

Since the 1990s, banks have used the originate-to-distribute model for a wide range of loans without the help of a government agency (and in most cases without payment guarantees). This means that the loans originated by banks are converted into securities and the investors who buy the securities bear the credit risk. This process is known as *securitization*. It enables banks to remove loans from their balance sheets and frees up funds so that more loans can be made. The bank earns a fee for originating the loan and a further fee if it services the loan after it has been originated.

Tranches are often created from loan portfolios so that each tranche contains different exposures to losses on the portfolio. A simplified example is shown in Figure 1.1. A loan portfolio with a total principal of USD 100 million is sold by the bank to a special purpose vehicle (SPV), which arranges for the cash flows from the loans to be passed to three tranches. The senior, mezzanine, and equity tranches fund 70%, 25%, and 5% of the loan portfolio, respectively. The returns received by tranche holders if they do not bear any losses are 5% for the senior tranche, 8% for the mezzanine tranche, and 25% for the equity tranche.[13]

Repayments of principal flow first to the senior tranche. When that has been repaid, they flow to the mezzanine tranche, and when the mezzanine tranche has been repaid, they flow to the

[12] This tends to happen when interest rates have decreased, and the mortgage can be refinanced at a lower interest rate (but it may also happen simply because a house has been sold). The investor then must reinvest his or her funds at a lower interest rate.

[13] As will be explained, losses can reduce these returns. The reduction can be particularly severe for the equity tranche because it bears the first USD 5 million of losses.

equity tranche. Interest payments flow to the senior tranche until it has received its promised return of 5% on the outstanding principal. They then flow to the mezzanine tranche until it has received its promised return of 8% on outstanding principal. Finally, they flow to the equity tranche.

From this description, we see that the equity tranche bears the first 5% of losses, the mezzanine tranche bears the next 25% of losses, and the senior tranche bears all losses in excess of 30%.

The originate-to-distribute model played a role in the 2007–2008 crisis. Banks relaxed their mortgage lending standards so that the quality of the mortgages being originated declined. Despite this decline in quality, however, banks still managed to securitize them. In fact, they re-securitized mortgages by creating tranches from tranches. As defaults on the mortgages grew higher, losses were experienced by tranche holders. Unsurprisingly, for a period after the crisis, the originate-to-distribute model could not be used because it was not trusted by investors.

SUMMARY

Banks can engage in many different types of activities. They can take deposits, make loans, trade securities, arrange security issues, provide advice to corporations, provide services to retail investors, etc. Since 1988, global rules for regulating banks have been determined by the Basel Committee. Bank regulatory capital is currently the sum of amounts for market risk, credit risk, and operational risk.

To maintain confidence in the banking system, many countries have developed deposit insurance procedures. These protect depositors against losses arising from a bank failure. There is usually a maximum amount that can be claimed by any one depositor from the insurance fund.

There are potential conflicts of interest in a bank's activities and banks have developed internal rules to avoid them. Violations of the conflict of interest rules are likely to be costly both in terms of fines as well as loss of reputation.

Banks sometimes sell loans they have originated to investors. This is known as the originate-to-distribute model. This model was heavily used for subprime mortgages in the period immediately before the 2007–2008 financial crisis and was partly responsible for the crisis.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

1.1 What is the difference between commercial banking and investment banking?

1.2 What are the three main risks in banking mentioned in the chapter?

1.3 What is the difference between regulatory capital and economic capital?

1.4 What are the two liquidity ratios introduced as part of Basel III?

1.5 Explain the difference between a private placement and a public offering.

1.6 Explain the difference between best efforts and firm commitment when an equity issue is made.

1.7 What is an IPO?

1.8 What is a broker's discretionary account?

1.9 What is the originate-to-distribute model?

1.10 What are the two types of offers that can be made by the acquiring company in a takeover?

Practice Questions

1.11 What was the effect of the Glass–Steagall Act in the U.S.? When was it repealed? What changes along the lines of Glass–Steagall have been made since the 2007–2008 crisis?

1.12 What is the difference between a bank's trading book and its banking book? Why is the distinction between the two important from a regulatory perspective?

1.13 What is meant by moral hazard? Explain why deposit insurance can give rise to moral hazard.

1.14 What is a poison pill? Give three examples.

1.15 Give two examples of conflicts of interest that could arise from communication between investment bankers and commercial bankers. Why do the senior management of banks want to avoid conflicts of interest?

1.16 What is the difference between standardized models and internal models in regulation? What is the formula that will be used from 2027 going forward to determine regulatory capital using calculations from both models?

1.17 Explain the risks in funding long-term needs with short-term instruments. What steps have been taken to reduce these risks in Basel III?

1.18 The bids and bidders in a Dutch auction to sell 20,000 shares are as follows:

Bidder	Number of Shares	Price (USD)
A	3,000	80
B	2,000	73
C	5,000	90
D	6,000	85
E	9,000	70
F	4,000	84
G	8,000	86

At what price are the shares sold? How many shares does each bidder get?

1.19 An investment bank has been asked to underwrite an issue of 20 million shares. The share price is currently USD 24. The bank is trying to decide between a firm commitment at USD 20 versus a best efforts where it will charge 40 cents for each share sold regardless of price. Discuss the pros and cons of the two alternatives.

The following questions are intended to help candidates understand the material. They are *not* actual FRM exam questions.

ANSWERS

Short Concept Questions

1.1 Commercial banking involves the traditional banking activities of taking deposits and making loans. Investment banking involves underwriting new issues of securities, advising companies on major financial decisions, and trading activities.

1.2 Market risks, credit risks, and operational risks

1.3 Regulatory capital is capital determined by regulators. Economic capital is a bank's own estimate of the capital it needs.

1.4 Liquidity coverage ratio (to ensure that the bank can survive a 30-day period of extreme stress) and the net stable funding ratio (to determine whether the maturities of assets and liabilities are reasonably well matched).

1.5 A private placement is the placement of an issue of securities with institutional investors without offering the securities to the general public. A public offering is an offering of securities to the general public.

1.6 In a best efforts underwriting, the investment bank does its best to sell securities at the agreed price but there is no guarantee. In a firm commitment, the bank guarantees that the issue can be sold at a certain price.

1.7 An IPO (initial public offering) is an issue of a company's equity to the market for the first time.

1.8 A discretionary account is an account where the broker can trade without consulting the investor.

1.9 The originate-to-distribute model describes the situation where a bank originates loans and then sells them to investors. Often the cash flows from a portfolio of loans are channelled to different tranches.

1.10 A cash offer and a share-for-share exchange

Solved Problems

1.11 The Glass–Steagall Act established a separation between commercial banking and investment banking in the U.S. It prevented commercial banks from engaging in some investment banking activities, and vice versa. It was repealed in 1999. Since the 2007–2008 crisis, rules have been implemented in the U.S. to prevent commercial banks from engaging in proprietary trading.

1.12 A bank's trading book includes instruments that the bank expects to trade. The banking book includes instruments the bank expects to keep until maturity. Items in the trading book face market risk capital. Items in the banking book face credit risk capital.

1.13 Moral hazard is the risk that the existence of an insurance contract will change the behavior of the insured party thereby making the insurance contract riskier. Deposit insurance could encourage a bank to take more risks because it is less likely to lose depositors when they become aware of the risks being taken.

1.14 A poison pill is a change made by a company to make it more difficult for another company to acquire it. Examples of poison pills are (a) granting key employees stock options that can be exercised in the event of a takeover, (b) making it impossible to fire directors for a period of time after a takeover, (c) issuing preferred shares that automatically become common shares in the event of a takeover, (d) allowing existing shareholders to buy new shares at a bargain price in the event of a takeover, (e) changing the voting structure to give management more control, and (f) allowing remaining shareholder to sell shares to the new owners in the event of a takeover at a premium price.

1.15 One possibility: when advising about a possible acquisition, an investment banker might get confidential information about the target company from a commercial banking arm of the bank if the target happens to be a client of the bank. Another possibility: a commercial banker might use the investment banking operation to replace a loan that contains too much credit risk with a bond issue. Violations of conflicts of interest can give rise to big fines and reputational losses that far exceed the extra profits earned from the violations.

1.16 Standardized models are those determined by the Basel Committee. Internal models are those developed by the bank. By 2027 the required capital will be the maximum of (a) that given by approved internal models, and (b) 72.5% of that given by standardized models.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

1.17 Short-term funding must be rolled over frequently. If the market rightly or wrongly loses confidence in the bank this could become impossible—even if the assets being financed are sound. The Net Stable Funding Ratio requirement in Basel III is designed to ensure that banks do not make excessive use of short-term funding.

1.18 19,000 shares can be sold for 85 or more. 23,000 shares can be sold for 84 or more. Shares are therefore sold for 84. Bidders C, D, and G get the shares they have bid for. Bidder F gets 1,000 of the 4,000 shares it has bid for.

Bids sorted from highest to lowest:

Bidder	Number of Shares Requested	Cumulative Number of Shares Requested	Price Bid (USD)
C	5,000	5,000	90
G	8,000	13,000	86
D	6,000	19,000	85
F	4,000	23,000	84
A	3,000	26,000	80
B	2,000	28,000	73
E	9,000	37,000	70

1.19 Assuming that the shares can be sold for the same price, the best efforts alternative will lead to a gross profit for the bank of USD 8 million. The bank's profit from the firm commitment could be much higher. For example, if shares can be sold at the current price of USD 24, the bank will make USD 80 million. However, there is a chance that the price will collapse. For example, if the post-issue price is only USD 17, the bank will lose USD 60 million.

2

Insurance Companies and Pension Plans

Learning Objectives

After completing this reading, you should be able to:

- Describe the key features of the various categories of insurance companies and identify the risks facing insurance companies.
- Describe the use of mortality tables and calculate the premium payments for a policy holder.
- Distinguish between mortality risk and longevity risk and describe how to hedge these risks.
- Describe defined benefit plans and defined contribution plans and explain the differences between them.
- Compare the various types of life insurance policies.
- Calculate and interpret loss ratio, expense ratio, combined ratio, and operating ratio for a property-casualty insurance company.
- Describe moral hazard and adverse selection risks facing insurance companies, provide examples of each, and describe how to overcome these risks.
- Evaluate the capital requirements for life insurance and property-casualty insurance companies.
- Compare the guaranty system and the regulatory requirements for insurance companies with those for banks.

Insurance provides protection against specific adverse events. The company or individual obtaining protection is known as the *policyholder*. Typically, the policyholder must make regular payments known as *premiums*. The most likely outcome in a year is that the adverse events do not occur and so there is no cost to the insurance company. If one of the specified adverse events does occur, however, there is usually a relatively large payment from the insurance company to the policyholder that covers all or part of the losses experienced.

Most insurance contracts can be categorized as either *life insurance* or *property and casualty insurance*.[1] The simplest type of life insurance contract requires regular monthly or annual premiums for the life of the policyholder. When the policyholder dies, the premiums stop, and his or her beneficiaries receive a lump sum payment. A property and casualty insurance policy typically lasts one year and provides the policyholder with compensation for losses arising from accidents, fires, thefts, and similarly insured adverse events. The policy is usually renewed annually. While the premiums in a life insurance policy do not usually change from year to year, those in a property and casualty policy generally do.

There are similarities between pension plans and the contracts offered by life insurance companies. In an employer-sponsored pension plan, it is typically the case that both the employee and the employer make regular contributions to the plan. The contributions are used to fund a lifetime pension for the employee following the employee's retirement. We describe pension plans after presenting the activities of life insurance companies.

2.1 MORTALITY TABLES

Mortality tables are used extensively by actuaries for setting life insurance contract premiums and assessing pension plan obligations. We review the information in these tables before discussing life insurance contracts and pension plans.

Table 2.1 shows an extract from mortality tables produced by the U.S. Social Security Administration. The table is based on the death rates observed for men and women of different ages in 2016.

The Probability of Death within 1 Year column is the probability that an individual will die during the following year (providing that he or she is alive at the beginning of the year). The table shows that a man aged 70 has a 2.3122% probability of dying within the next year, whereas for a woman this probability is a more favorable 1.5413%.

The survival probability for Year n is the cumulative probability that an individual will live to Year n. The survival probability for Year zero is 1. The survival probability for Year $n + 1$ can be calculated from the survival probability for Year n and the probability of death within one year for Year n.

For example, the probability that a man will survive until age 71 is the probability that he survives until age 70 and does not die within the next year. This is

$$0.72843 \times (1 - 0.023122) = 0.71159$$

which agrees with the survival probability for age 71 in the table.[2] The probability that a man aged 71 will die in the following year is 0.025265. Thus, the probability that a man aged 70 will die between the first and second year is[3]

$$(1 - 0.023122) \times 0.025265 = 0.024681$$

Similarly, the probability that he will die between the second and third year is

$$(1 - 0.023122) \times (1 - 0.025265) \times 0.027585 = 0.026266$$

and so on. Assuming that (on average) a person dies in the middle of a year, the life expectancy of a man aged 70 (shown in the third column) can be calculated as:

$$0.023122 \times 0.5 + 0.024681 \times 1.5 + 0.026266 \times 2.5 + \cdots$$

2.2 LIFE INSURANCE

There are many different types of life insurance policies. Here we describe some of the more common ones.

Whole Life Insurance

As its name implies, whole life insurance provides insurance for the whole life of the policyholder. The policyholder makes regular monthly or annual payments until he or she dies. At that time, the face value of the policy is paid to the designated beneficiary. This means that there is certain to be a payment by the insurance company[4] and the only uncertainty for the insurance company is when the payment will occur. Usually, both the

[1] Property and casualty insurance is sometimes referred to as *non-life insurance*. Health insurance is sometimes considered to be a third category.

[2] The last decimal place is 9 rather than 8 because of the effects of rounding.

[3] This means that individual will survive the next year (while he is 70) and die during the second year (when he is 71).

[4] Life insurance is sometimes referred to as *life assurance* when a payout is certain.

Table 2.1 **Extract from Mortality Tables Published by U.S. Social Security Administration Based on 2016 Data. (See www.ssa.gov/OACT/STATS/table4c6.html)**

	Males				Females		
Age (Years)	**Probability of Death within 1 Year**	**Survival Probability**	**Life Expectancy**		**Probability of Death within 1 Year**	**Survival Probability**	**Life Expectancy**
30	0.001794	0.97238	47.72		0.000803	0.98530	52.01
31	0.001835	0.97063	46.80		0.000853	0.98451	51.05
32	0.001880	0.96885	45.89		0.000905	0.98367	50.09
33	0.001930	0.96703	44.97		0.000956	0.98278	49.14
50	0.005007	0.92209	29.69		0.003193	0.95536	33.26
51	0.005493	0.91747	28.84		0.003492	0.95231	32.36
52	0.006016	0.91243	27.99		0.003803	0.94899	31.48
53	0.006575	0.90694	27.16		0.004126	0.94538	30.59
70	0.023122	0.72843	14.40		0.015413	0.82573	16.57
71	0.025265	0.71158	13.73		0.017089	0.81301	15.82
72	0.027585	0.69360	13.07		0.018861	0.79911	15.09
73	0.030070	0.67447	12.43		0.020705	0.78404	14.37
90	0.163689	0.18303	4.08		0.129706	0.29685	4.85
91	0.181104	0.15307	3.78		0.144636	0.25835	4.50
92	0.199810	0.12535	3.50		0.160741	0.22098	4.18
93	0.219765	0.10030	3.25		0.177971	0.18546	3.88

Source: The United States Social Security Administration.

payments and the face value of the policy remain constant through time.

Consider a USD 1 million whole life policy where the policyholder is a 30-year old male. The insurance company can calculate its expected payout each year. From Table 2.1, we see that the probability of a male who is 30 years old dying within one year is 0.001794. The insurance company's expected payout (in USD) during the first year is therefore:

$$1{,}000{,}000 \times 0.001794 = 1{,}794$$

The probability that the policyholder will die during the second year is the probability that death does not occur during the first year multiplied by the probability of dying during the second year. From Table 2.1 this is

$$(1 - 0.001794) \times 0.001835 = 0.001832$$

so that the expected payout during the second year is USD 1,832. Similarly, the probability that the policyholder will die during the third year is

$$(1 - 0.001794) \times (1 - 0.001835) \times 0.001880 = 0.001873$$

so that the expected payout during the third year is USD 1,873.

The expected payout increases year by year throughout the life of the policy. From Table 2.2, we can calculate the probability that a policyholder aged 30 will survive until 70 as:

$$\frac{\text{Probability of Survival from Birth Until Age 70}}{\text{Probability of Survival from Birth Until Age 30}} = \frac{0.72843}{0.97238} = 0.74912$$

The probability that the policyholder (currently 30 years old) will die between age 70 and 71 is therefore:

$$0.74912 \times 0.023122 = 0.017321$$

Table 2.2 Calculation of Present Value of Expected Payout per Dollar of Face Value for a Female Aged 50 in Three-Year Term Life Insurance

Time (Years)	Expected Payout	Discount Factor[6]	PV of Expected Payout
0.5	0.003193	0.975900	0.003116
1.5	0.003481	0.929429	0.003235
2.5	0.003778	0.885170	0.003344
Total			0.009695

The expected payout from the policy during its forty-first year is therefore USD 17,321.

Suppose that the premium charged is USD 15,000 per year. It is clear from these calculations that the insurance company has an expected surplus during the early years and an expected deficit in later years.

On any one policy, the results are uncertain. If the insurance company sells many similar policies to men aged 30, however, the investment of the early premiums is going to be an important part of how the insurer finances later payouts. Of the USD 15,000 per USD 1 million face value received in the first year, the expected USD amount available for investment is

$$15{,}000 - 1{,}794 = 13{,}206$$

Similarly, the expected USD amount of the premium in the second year that is available for investment is

$$15{,}000 - 1{,}832 = 13{,}168$$

By contrast, in the forty-first year the insurance company has a negative expected USD cash flow of:

$$17{,}321 - 15{,}000 = 2{,}321$$

Figure 2.1 shows how a typical whole life insurance policy generates expected surpluses and expected deficits during its life. From the insurance company's perspective, the investment income from the expected surpluses must be sufficient to finance the expected deficits. Figure 2.1 is illustrative of the general pattern of expected surpluses and deficits in a whole life policy. It does not include investment income.

In many jurisdictions, there are tax benefits associated with whole life insurance. When funds are invested by the policyholder, tax is paid on the investment income each year. But when the funds are invested by the insurance company, no tax is payable until there is a payout from the insurance policy. In some jurisdictions, taxes are not levied on the payout.[5]

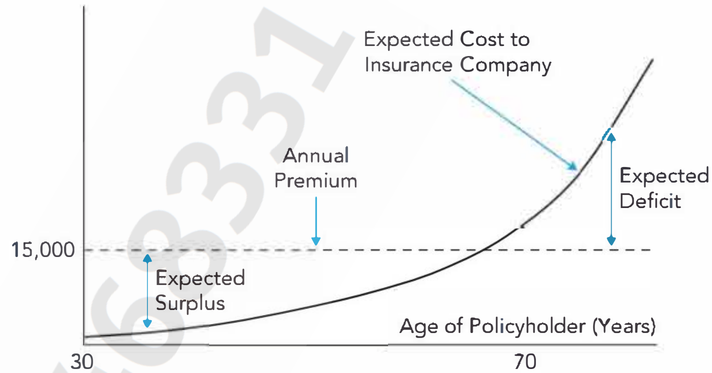


Figure 2.1 Shown is a pattern of expected surpluses and expected deficits on USD 1 million whole life insurance policy. The age of the policyholder is 30 and the premium is USD 15,000 per year.

Given that a whole life insurance policy involves funds being invested for the policyholder, a natural development is to allow the policyholder to specify how the funds are to be invested. *Variable life insurance* is the name used to describe policies where the policyholder can do this. Typically, there is a minimum guaranteed payout on the policyholder's death. However, the payout can be greater than this if the investments do well.

Normally, if the policyholder stops making premium payments, the policy no longer provides coverage. The policy is then referred to as *lapsing*. Another variation on the standard whole life policy is where the policy holder can reduce the premium down to a specified minimum. This is called *universal life insurance*. While reducing the premium reduces the benefits, it does not result in the policy lapsing. *Variable-universal* life insurance incorporates the features of both variable and universal life insurance.

[5] Note that taxes are usually payable on the proceeds if the policy is sold to a third party before the policyholder's death.

[6] As will be explained in later chapters, the discount factor for time t is the value of one dollar received at time t. In this case it is $1/1.05^t$.

Term Life Insurance

Term life insurance lasts a specified number of years. If the policyholder dies during the life of the policy, there is a payout equal to the face value of the policy. Otherwise there is no payout. Mortgages are a common reason for purchasing term life insurance. For example, a 30-year old person with a 15-year mortgage might choose to buy a 15-year term life insurance policy with a declining face amount equal to the amount outstanding on the mortgage. If the policyholder dies, the insurance will provide dependents with the funds necessary to pay off the mortgage.

Mortality tables can be used to determine breakeven premiums on term life insurance. As a simple example, consider a three-year term life insurance contract offered to a woman aged 50. The probability that she will die in the first year is given by Table 2.1 as 0.003193. The probability that she will die during the second year is

$$(1 - 0.003193) \times 0.003492 = 0.003481$$

And the probability that she will die during the third year is

$$(1 - 0.003193) \times (1 - 0.003492) \times 0.003803 = 0.003778$$

We assume that death always occurs halfway through a year and that the discount rate is 5% (with annual compounding). This leads to the present value of the expected payout per dollar of face value being calculated as indicated in Table 2.2.

Now consider the premium paid by the policyholder. Suppose this is X and that (as is the case with most insurance contracts) it is paid in advance. The insurance company is certain to receive the first premium at time zero. It will receive the second premium after one year if the policyholder is still alive. The probability of this is 0.996807 ($= 1 - 0.003193$). It will receive the third premium after two years if the policyholder is still alive at that time. The probability of this is

$$(1 - 0.003193) \times (1 - 0.003492) = 0.993326$$

The present value of the expected premiums received by the insurance company can therefore be calculated as indicated in Table 2.3.

The breakeven premium is calculated by equating the present value of expected premiums with the present value of the expected payout:

$$2.850316X = 0.009695$$

This gives $X = 0.003401$. The breakeven cost of three-year term insurance for a 50-year old female is therefore 0.003401 per dollar of face value. For a policy with a face value of USD 1 million, the breakeven premium is therefore USD 3,401 per year. (The insurance company will of course charge more than this to cover administrative costs and earn a profit.)

Endowment Life Insurance

Endowment life insurance is a type of term life insurance where there is always a payout at a pre-specified contract maturity. If the policyholder dies during the life of the policy, the payout occurs at the time of the policyholder's death. Otherwise, it occurs at the end of the life of the policy. Sometimes the payout is also made when the policyholder has a critical illness. In the case of a *with profits endowment* life insurance policy, the insurance company declares bonuses that depend on the performance of its investments. These bonuses increase the final payout (assuming that the policyholder lives until the end of the life of the policy). In a *unit-linked endowment* policy, the policyholder chooses a fund, and the payout depends on the performance of that fund.

Group Life Insurance

Group life insurance is typically purchased by companies for their employees. The premiums may be paid entirely by the company or shared between the company and its employees. Note that while individuals are normally required to undergo medical tests when applying for life insurance, these tests are usually waived in group life insurance. This is because the insurance company knows it will be taking some better-than-average and worse-than-average risks.

Table 2.3 **Calculation of the Present Value of Expected Premiums for a Female Aged 50 in Three-Year Term Life Insurance**

Time (Years)	Probability of Receiving Premium	Discount Factor	PV of Expected Premiums
1	1.000000	1.000000	X
2	0.996807	0.952381	$0.949340X$
3	0.993326	0.907029	$0.900976X$
Total			$2.850316X$

Annuity Contracts

Most life insurance companies offer annuity contracts in addition to life insurance contracts. While life insurance converts regular payments into a lump sum, annuity contracts do the reverse (i.e., they convert a lump sum into regular payments). Typically, the payments in an annuity contract last for the rest of the policyholder's life. In some cases, the annuity starts as soon as a lump sum is deposited with the insurance company. In the case of *deferred annuities*, it starts several years later.

As with life insurance, taxes can be a consideration when an annuity is purchased. Because the insurance company is investing funds on behalf of the policyholder, tax is normally payable only when the annuity is received. If the policyholder were investing the funds themselves, however, tax would be payable on the investment income each year. There are therefore two advantages to this arrangement: Taxes are deferred so that investments grow tax-free, and many policyholders have relatively low marginal tax rates when the annuity is received.

The amount to which the policyholder's funds grow in an annuity contract is referred to as the *accumulation value*. Funds can usually be withdrawn early, but there may be penalties. Some annuity contracts have embedded options designed to ensure that the accumulation value never declines. For example, the contract may be structured so that the accumulation value tracks the S&P 500 with a lower limit on the return of zero and an upper limit of 7%.[7]

In the United Kingdom, deferred annuity contracts sometimes guarantee a minimum level for the future annuity payments. For example, an insurance company might offer an annuity that will start in ten years and pay at least 8% of the accumulation value at that time.[8] If interest rates fall and life expectancies increase, however, the guarantee can prove to be very expensive. Equitable Life, a huge U.K. life insurance company founded in 1762 that had 1.5 million policyholders at its peak, was an example of an insurance company that failed because of the generous guarantees it offered.

[7] We will discuss options in a later chapter. By creating an upper and lower limit on the return each year, the insurance company has sold a put option to the policyholder and bought a call option from the policyholder.

[8] The specification of a minimum level for the annuity was regarded as a necessary part of marketing the annuity and its cost was often not calculated by the insurance company.

Longevity and Mortality Risk

Longevity risk is the risk that people will live for longer than mortality tables indicate. Recall that Table 2.1 is based on the percentages of people of different ages that died in the year 2016. It estimates that a female aged 30 has a life expectancy of 52.01 years (i.e., she will, on average, die at age 82.01). However, advances in medical science, improved nutrition, and other factors may lead to this being an underestimate. Over the last century, the life expectancy of people born in the U.S. have steadily increased to the point where someone born today can expect to live about 20 years longer than a similar person born 100 years ago.[9]

On the other hand, mortality risk is the opposite of longevity risk. It is the risk that wars, epidemics, and other factors will lead to people dying sooner than expected.

The life insurance business of an insurance company should welcome the possibility that people will live longer than expected. This will lead to premiums on whole life policies being paid for longer periods of time and payouts being (on average) later than expected. Mortality risk is more of a concern, however, because it could lead to earlier payouts without sufficient collected premiums to cover them.

The annuity business of an insurance company has the opposite exposure compared to that of the life insurance business. If people live longer than expected, they receive the annuity to which they are entitled for longer and thus make the contract more expensive for the insurance company. If they die sooner than expected, the contract will prove to be less expensive for the insurance company.

The longevity/mortality exposures on an insurance company's whole life business should (to some extent) offset the exposure on its annuity business. However, there is unlikely to be a perfect offset and the residual exposure must be managed. One approach is the use of longevity derivatives. These instruments have been structured in a number of different ways so that the payoff depends on the difference between a pre-specified (expected) mortality rate for individuals in a certain age group and the actual mortality rate. A simple payoff could be defined as:

$$\text{Payoff} = (\text{Pre-Specified Fixed Mortality Rate} - \text{Realized Mortality Rate}) \times \text{Principal} \tag{2.1}$$

This is similar to a forward contract on the mortality rate. (Forward contracts will be discussed in later chapters.) Other longevity derivatives involve bonds where either the principal

[9] See for example http://www.demog.berkeley.edu/~andrew/1918/figure2.html for life expectancy estimates by year since 1900.

or the interest rate depends on the difference between a prespecified mortality rate and an actual mortality rate.

Longevity derivatives are potentially of interest to defined benefit pension plans (which will be discussed later in this chapter) as well as insurance companies because these plans also have an exposure to longevity risk. The mortality rate used to define an instrument with the payoff in Equation (2.1) could be the mortality rate of all members of a pension plan that are currently receiving pensions.

Investments

Life insurance policies and annuity contracts generate funds for investment. The investment strategy of a life insurance company is therefore very important.

Many of these investments consist of long-term corporate bonds.[10] Life insurance companies can try to match the maturities of the bonds with the maturities of their obligations. However, corporate bonds give rise to both market risk and credit risk. The market risk relates to interest rates (as rates increase, bond prices decline). The credit risk relates to the possibility that bond issuers will default. Life insurance companies could avoid credit risk by investing in government bonds. Over the long term, however, the extra return earned on corporate bonds (over risk-free government bonds) more than compensates for losses arising from defaults.

Life insurance companies have risks on both sides of their balance sheets: market and credit risks on the asset side along with longevity and mortality risks on the liability side. In addition, they are subject to operational risks that are like those of banks, which were discussed in the previous chapter. Regulators take all these risks into account in determining minimum capital requirements.

2.3 PENSION PLANS

Pension plans are like annuity contracts in that they are designed to produce income for an individual for the remainder of his or her life following retirement. Typically, contributions to the plan are made by the individual and the individual's employer while the person is employed. These contributions are deductible for tax purposes.

There may be some indexation of the pension. For example, it might be agreed that the annuity growth rate will reflect 75% of the inflation rate. The terms of the pension plan state that the employee's spouse (and possibly other dependents) will receive a (usually smaller) pension if they are still alive when the employee dies.

In a *defined contribution plan*, the funds are invested by the employer. When the employee retires, he or she can begin withdrawing the funds.[11] Sometimes the employee can choose how the funds are invested, and sometimes he or she can opt for a lump sum payment at retirement.

In a *defined benefit plan*, funds are also usually contributed by the employer and employee. In this case, however, the contributions are pooled, and a formula is used to determine the pension received by the employee on retirement. For example, the pension plan might state that the pension is the employee's average annual income during the last three years of employment multiplied by the number of years of employment multiplied by 2%.[12]

A defined benefit plan is much riskier for an employer than a defined contribution plan. In a defined contribution plan, the company is merely acting as an agent investing the pension plan contributions on behalf of its employees. The pension obtained by an employee depends on the extent to which the employee's funds have grown.[13] In a defined benefit plan, the company has guaranteed to its employees that pensions will be calculated in a certain way.

Each year, actuaries assess the present value of a defined benefit pension plan's obligations and compare it with the plan's assets. The company is required to provide contributions to the plan to make up for any shortfall. And while it may be able to spread these contributions out over several years, the shortfall will remain a liability that reduces shareholders' equity on its balance sheet. The risk for the company and its shareholders in defined benefit plans is the reason why defined benefit plans are not initiated today. Most companies that have defined benefit plans initiated them many years ago, and many firms have switched from defined benefit plans to defined contribution plans (at least for new employees) to avoid these risks.

An important issue in estimating defined benefit plan obligations is the discount rate used. Plan outflows often stretch for many years into the future. For example, the present value of USD 1,000 paid out in 40 years with a discount rate of 2% is about USD 453. If the discount rate is 5%, however, the present

[10] Insurance companies also make equity investments. As discussed, a particular policy could require equity or other investments by the insurance company.

[11] In the U.S. a 401(k) is a type of defined contribution plan. Some other countries have similar plans.

[12] Government pension plans such as Social Security in the U.S. are a type of defined benefit plan.

[13] The only real risk for the company is a small operational risk that it will manage the funds badly and be sued by employees.

value is less than one third of this. Accounting standards now require private sector pension plan obligations to be discounted at the yield on AA-rated bonds.

We mentioned earlier that life insurance companies manage risks by investing in long-term corporate bonds. Specifically, they attempt to match the maturities of the bonds with the maturity of their life insurance and annuity obligations. It might be thought that defined benefit pension plans would follow the same investment strategy because their obligations are also fairly predictable. In practice, however, the returns provided by bonds are not necessarily sufficient for pension plans to meet their obligations. As a result, pension plans typically put much of their assets into equities.[14] If equity markets do well (as they have done in many parts of the world since 1960), these plans should be able to meet their obligations. If there is a prolonged decline in equity prices (particularly if it is combined with an increase in life expectancy), however, these plans are likely to chalk up huge deficits.

Primarily, shareholders bear the cost of deficits in defined benefit plans. In cases of bankruptcy, the cost may be borne by government-sponsored entities.[15] In either case, there is a transfer of wealth to retirees from the next generation. It has been argued that the terms of defined benefit plans must be altered so that there is some risk sharing between generations. If equity markets do well, some of the benefits can be passed on to the next generation. If they do badly, some of the costs should be borne by retirees.[16]

2.4 PROPERTY AND CASUALTY INSURANCE

Property and casualty insurance contracts are quite different from life insurance contracts. Property insurance provides protection against damage to property from fire, theft, flooding, and other loss-generating events. Casualty insurance provides coverage for liabilities arising from injuries and damages sustained by others due to the insured party's actions. Often, both property and casualty insurance are provided in a single policy. For example, homeowners' insurance typically provides insurance for losses from fire and theft as well as for liabilities if others are injured on the property. Meanwhile, car insurance typically provides coverage for theft as well as claims by others for damages caused.

[14] A common portfolio mix is 60% equity, 40% debt.

[15] For example, in the U.S., the Pension Benefit Guarantee Corporation insures private defined benefit plans.

[16] See for example K. P. Ambachtsheer, *Pension Revolution: A Solution to the Pension Crisis.* Hoboken, NJ: John Wiley & Sons, 2007.

Unlike life insurance, property and casualty insurance is typically renewed from year to year. If the insurer feels that the risks have increased, the premium may be raised. For example, car insurance premiums are likely to increase if a driver has been convicted of speeding.

The risks to a property and casualty insurance company can be divided into:

- Risks where the average payout can be predicted reasonably well from historical data because the yearly payout arises from many independent (or almost independent) claims by policyholders, and
- Risks where a single event (such as a hurricane or an earthquake) can lead to many claims at the same time.

Car insurance is in the first category. For example, assume a company has insured 100,000 car owners in a certain risk category. Furthermore, it knows from experience that 10% of drivers will make a claim in a year, with an average claim being USD 3,000. As a result, it can expect its total annual cost in meeting claims to be about USD 30 million ($= 3{,}000 \times 0.1 \times 100{,}000$). While the total annual cost might vary slightly from year to year, large differences are extremely unlikely statistically. The key point is that claims by different drivers are independent (or almost independent) of each other. Independence implies that if Driver A has an accident, this does not increase the probability of Driver B having an accident. Of course, the insurance company also monitors trends in the number of accidents per year, trends in automobile repair costs, and trends in the damages awarded to accident victims.

Risks in the second category are referred to as *catastrophe risks*. For example, a company that has insured 100,000 homes against hurricane damage in South Florida will have claims that are not independent. Either a hurricane happens (and most policyholders file a claim) or it doesn't (and no policyholders do so). Catastrophe risks are therefore a much greater threat for insurance companies. As a result, insurance companies used models produced by specialists to predict the probability of different events that they are insuring against. However, this does not alter the fact that they are (by nature) all-or-nothing risks.[17]

CAT Bonds

When a company does not want to keep catastrophe risks, it can pay a reinsurance company to take them on. It can also use derivatives known as CAT (catastrophe) bonds.

[17] Analysts often set premiums so that their coverage is three times the largest cost given by simulations.

A CAT bond is a bond issued by an insurance company that pays a higher than normal rate of interest. If payouts by the insurance company for a specified risk are in a certain range, the bond's interest (and in some cases the principal) are used to provide the payouts.

For example, suppose an insurance company has USD 100 million of exposure to Florida hurricanes and wants to reduce this exposure to USD 40 million. It could issue three bonds.

1. Bond A has a principal balance of USD 20 million and covers claims in the range USD 40 to 60 million.
2. Bond B has a principal balance of USD 20 million and covers claims in the range USD 60 to 80 million.
3. Bond C has a principal balance of USD 20 million and covers claims in the range USD 80 to 100 million.

Bond A is riskier than Bond B, which is in turn riskier than Bond C. As a result, the interest rate offered on Bond A will be greater than that on Bond B, which in turn will be greater than that on Bond C.

With these bonds, the principal is at risk in the event of a hurricane. As an alternative, the insurance company could issue bonds with much higher principal balances so that the dollar amount of promised interest is much higher and can be used to cover claims without putting the principal at risk.

It is natural to ask why anyone would buy a CAT bond. The main reason is that the risk is likely to be uncorrelated (or almost uncorrelated) with other risks in an investor's portfolio. CAT bonds therefore provide diversification benefits for investors. Capital market theory would suggest that the expected return on a security should be the risk-free rate when the security's return is uncorrelated with the return provided by the market. CAT bonds are attractive because they can offer a higher return than that suggested by capital market theory.

Loss Ratios

A key statistic for a property–casualty insurance company is its *loss ratio*. This is the ratio of payouts to premiums. A loss ratio of 70% would mean that for every USD 100 of premiums received, the insurance company pays out USD 70 in claims. The remaining 30% is to cover expenses and (hopefully) make a profit. Two major expenses are

1. Selling expenses, and
2. Expenses related to determining the validity of a claim (referred to a loss adjustment expenses).

The *expense ratio* is total expenses divided by premiums received. Table 2.4 shows an income statement for a property and casualty insurance company. The loss ratio is 70%. The expense ratio is 26%. The *combined ratio* (which is the sum of the loss ratio and the expense ratio) is 96%. Sometimes a small dividend is paid to policyholders. Table 2.4 assumes that this is 1%. The *combined ratio after dividends* is therefore 97%.

Table 2.4 Calculation of Operating Ratio for a Property–Casualty Insurance Company

Loss Ratio	70%
Expense Ratio	26%
Combined Ratio	96%
Dividends	1%
Combined Ratio After Dividends	97%
Investment Income	(2%)
Operating Ratio	95%

The insurance company's *operating ratio* is a gross profitability measure. From the loss ratio, expense ratio, and dividends it appears that the operating ratio is 97%. However, this does not consider the interest earned. Premiums are typically paid at the beginning of a year and payouts happen later. The policyholder's funds can therefore be invested in short- and medium-term bonds to earn interest. Table 2.4 assumes the investment income earned during this period is 2% of premiums received. This changes the operating ratio to 95%. Also note that Table 2.4 shows investment income as a negative number because it offsets expenses. If the company had lost money on its investments during the year, that would *increase* the operating ratio.

2.5 HEALTH INSURANCE

A final category of insurance is health insurance. In many countries, health care is provided almost entirely by the government. In other countries, public and private health care systems run side by side. In the U.S., private health insurance has historically been a necessary expenditure for many people.

One difference between health insurance and other types of insurance concerns the circumstances when premiums increase. In whole life insurance, premiums typically remain constant. Even if it becomes known that a policy holder has a very expensive medical condition, premiums cannot be increased. In property–casualty insurance, risks are reassessed every year, and premiums may increase or decrease. In health insurance, premiums may increase because the overall costs of health care have increased. However, they cannot do so because the policyholder develops unexpected health problems that were unknown at the time the policy was initiated.

2.6 MORAL HAZARD AND ADVERSE SELECTION

In this section we discuss two important risks facing insurance companies: moral hazard and adverse selection.

Moral Hazard

Moral hazard is the risk that the behavior of the policyholder will change as a result of the insurance. We mentioned moral hazard regarding deposit insurance in Chapter 1. We noted that banks might be tempted to follow riskier strategies when depositors have government insurance because the insurance makes it less likely that depositors will transfer their funds elsewhere. Other examples of moral hazard include the following.

- If your house insurance fully insures you against burglaries, you may be less likely to install alarm systems and cameras.
- As a result of buying health insurance, you may visit the doctor more frequently.

Moral hazard is not a big problem in life insurance. It is difficult to imagine that an individual would take up a risky pursuit (such as sky diving) because he or she has bought life insurance.

Insurance companies attempt to reduce moral hazard in a number of ways. The following are examples.

- Most insurance policies have a deductible so that the policyholder bears the first part of any loss.
- Sometimes there is co-insurance where the insurance company pays only a certain percentage of any loss.
- There is nearly always a limit on the amount that can be claimed.

Adverse Selection

Adverse selection is the risk that insurance will be purchased only by high-risk policy holders. If an insurance company offers the same car insurance rates to everyone, for example, the rates will seem more attractive to those who have bad driving records. It is therefore important that insurance companies know as much as possible about the risks they are taking on before quoting a premium. The risk assessments should be updated as more information is obtained.

2.7 REGULATION

In Chapter 1, we explained that the Basel Committee determines global regulatory requirements for banks. In contrast, there are no similar global regulatory requirements for insurance companies. In the European Union, however, a regulatory framework known as Solvency II was implemented in 2016 and applies to all insurance companies.

Solvency II specifies a minimum capital requirement (MCR) and a solvency capital requirement (SCR). If capital falls below the SCR, an insurance company is required to formulate a plan to bring it back up above the SCR level. If it falls below the MCR level, the insurance company may be prevented from taking new business, and existing policies might be transferred to another insurance company. The MCR is typically between 25% and 45% of the SCR.

As in the case of the Basel Committee rules, there are both standardized approaches and internal model-based approaches to determining capital requirements. There are capital charges for:

- Investment risk (this relates to the asset side of the balance sheet),
- Underwriting risk (this relates to the liabilities side of the balance sheet), and
- Operational risk.

Investment risk is divided into credit risk and market risk. Meanwhile, the capital requirements for property–casualty underwriting tend to be higher than those for life insurance underwriting. This is because the catastrophe risks associated with the former are greater than the longevity/mortality risks associated with the latter.

In the U.S., insurance is regulated at the state level rather than at the federal level. The National Association of Insurance Commissioners is a national forum for insurance regulators to exchange ideas. For example, it provides statistics on loss ratios for insurance companies across the country.[18] However, there are differences from state to state in the way insurance companies are regulated, and a large U.S. insurance company may have to deal with 50 different regulators.

The guaranty system for policyholders in the U.S. is different from the deposit insurance system for bank depositors. As mentioned in Chapter 1, premiums paid by banks create a fund administered by the Federal Deposit Insurance Corporation (FDIC) that is used to compensate depositors for losses. The federal government has added to the fund when necessary.

In contrast, there is no permanent fund to provide protection for policyholders. And insolvencies are handled on a state-by-state basis. When one insurance company fails, others are required to make contributions to a fund, and there are limits on what can be claimed as well as delays in settlement.

[18] For more information see http://www.naic.org/

SUMMARY

The two main categories of insurance are life insurance and property–casualty insurance. Life insurance companies offer a number of products whose costs are a function of how long the policyholder lives. Whole life insurance has mortality risk in that it becomes more expensive if the policyholder dies at a young age. Annuities have longevity risk in that they become more expensive the longer the policyholder lives.

Pension plans have much in common with life insurance companies. They offer annuity-type contracts whose costs depend on how long pension plan members live. There are two types of pension plans: defined contribution and defined benefit. Offering employees a defined contribution plan poses very little risk for companies because they merely invest the pension plan contributions made on behalf of their employees (who are then able to withdraw the funds when they reach the retirement age). A defined benefit plan is much riskier because the funds are pooled, and it is promised that each pension will be calculated (and paid) in a certain way. Actuaries assess defined benefit plans annually and determine whether they are overfunded or underfunded.

Property–casualty insurance companies are concerned with providing protection for loss or damage to property and for liabilities arising from injuries and damages resulting from the insured party's actions. For some risks (such as those associated with car insurance), the average payout per policy on a large number of policies can be predicted with reasonable accuracy. For other risks (such as those associated with natural disasters), the insurance company is in the position where either its underwriting business will prove to be very expensive or it will cost nothing at all.

Health insurance has some of the features of life insurance and some of the features of property–casualty insurance. Health insurance premiums can increase because the cost of providing health care increases (just as property–casualty insurance premiums can increase), but they cannot increase because a policyholder's health deteriorates. In the latter respect, health insurance is like whole life insurance.

Insurance companies must consider moral hazard and adverse selection when assessing risks and setting premiums. Moral hazard is the risk that the policyholder will take more risks because of the existence of insurance and thereby make payouts more likely. Adverse selection is the risk that individuals who are most likely to suffer losses will tend to buy the insurance.

Insurance companies are regulated to ensure that they keep enough capital for the risks they are taking. A property–casualty company must typically keep more capital for the risks it underwrites than a life insurance company. In the European Union, capital for all insurance companies is determined by Solvency II rules. In the U.S., regulation is a responsibility of states, not the federal government.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

2.1 Which of these is the biggest risk for whole life insurance?

- **A.** Longevity risk
- **B.** Mortality risk
- **C.** Policies being sold to a third party
- **D.** Inflation

2.2 Which of these is the biggest risk for property–casualty insurance?

- **A.** Longevity risk
- **B.** Mortality risk
- **C.** Natural disasters
- **D.** Inflation

2.3 Is the following statement true or false? "For a woman aged 70, five-year term insurance will have lower premiums than whole life insurance."

2.4 What two activities of life insurance companies have opposite exposures to longevity risk?

- **A.** Term life insurance and whole life insurance
- **B.** Term life insurance and endowment life insurance
- **C.** Whole life insurance and annuities
- **D.** None of the above

2.5 What are the most popular investments for life insurance companies?

- **A.** Equities
- **B.** Treasury bonds
- **C.** Commercial paper
- **D.** Long-term corporate bonds

2.6 What is variable life insurance?

- **A.** Whole life insurance where the policyholder chooses how funds will be invested
- **B.** Whole life insurance where the policyholder can vary premiums from year to year
- **C.** Term life insurance that can be renewed
- **D.** Life insurance where the policyholder can ask for an early payout in certain circumstances

2.7 What is universal life insurance?

- **A.** Whole life insurance where the policyholder chooses how funds will be invested
- **B.** Whole life insurance where the policyholder can vary premiums from year to year
- **C.** Term life insurance that can be renewed
- **D.** Life insurance where the policyholder can ask for an early payout in certain circumstances

2.8 Is the following statement true or false? "Defined contribution pension plans are more risky for a company than defined benefit plans."

2.9 How is the loss ratio of a property–casualty insurance company defined?

- **A.** Ratio of payouts to premiums
- **B.** Ratio of payouts plus selling expenses to premiums
- **C.** Ratio of all costs to premiums
- **D.** Ratio of all costs to premiums plus interest income

2.10 Which of the following is true?

- **A.** There is a single regulator for all insurance companies globally.
- **B.** The National Association of Insurance Commissioners regulates all insurance companies in the U.S.
- **C.** Insurance companies operating in the state of Maine are regulated by the state of Maine.
- **D.** German insurance companies and Italian insurance companies are subject to different regulations.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

Practice Questions

2.11 Using Table 2.1, what is the probability that a woman aged 70 will live to age 90:

A. 0.29685
B. 0.35950
C. 0.31166
D. It is impossible to tell from the information in the table.

2.12 What is the minimum USD annual premium that an insurance company should charge for a two-year term life insurance policy with face value of USD 1 million when the policyholder is a woman aged 71? (Use Table 2.1 and assume an interest rate of 3% compounded annually.)

A. 18,153
B. 17,691
C. 17,996
D. 17,767

2.13 An insurance company offers a policy that pays out if a worker becomes unemployed. Which of the following risks are applicable?

A. Moral hazard, but not adverse selection
B. Adverse selection, but not moral hazard
C. Moral hazard and adverse selection
D. Neither moral hazard nor adverse selection

2.14 A company's defined benefit plan invests primarily in equities. Which of the following creates risks for the company? (Assume that equity markets are unaffected by A, B, C, and D.)

A. High interest rates
B. Low interest rates
C. Employees working past retirement age
D. An epidemic that shortens life expectancy

2.15 Is the following statement true or false? "Defined benefit plans have similar exposures to life insurance companies. Like life insurance companies, they invest primarily in bonds matching the maturities of assets and liabilities."

2.16 In health care insurance, is it usually the case that:

A. Premiums cannot increase from year to year
B. The insurance company can change premiums as it reassesses risks
C. Premiums increase with the age of the policyholder
D. Premiums can increase as the overall cost of health care increases

2.17 Suppose that in a certain defined benefit plan the following simple situation exists. Employees work for 40 years with a salary that increases exactly in line with inflation. The pension is 60% of the final salary and increases exactly in line with inflation. Employees always live for 25 years after retirement. The funds in the pension plan are invested in bonds that earn the inflation rate. Which of the following is the best estimate of the percentage of the employee's salary that must be contributed to the pension plan? (Hint: You should do all calculations in real rather than nominal terms so that salaries and pensions are constant and the interest earned is zero.)

A. 37.5%
B. 43.75%
C. 22.5%
D. 27.5%

2.18 A life insurance company issues whole life insurance policies where the annual premiums are USD 1 million. A property–casualty insurance company issues fire, theft, and flood house insurance policies with annual premiums of USD 1 million. Is the following statement true or false? "It is likely that the life insurance company will have to keep more capital for the risks it faces."

2.19 Consider two bonds. One is a CAT bond where there is no default risk. The other is a regular corporate bond. An analysis has shown that the expected loss from default risks on the corporate bond is the same as the expected loss from insurance claims on the CAT bond. The bonds have the same coupon and the same price. Which bond would be most attractive to a fund manager with an existing portfolio of corporate bonds?

A. The bonds are likely to be equally attractive
B. The CAT bond is likely to be more attractive
C. The corporate bond is likely to be more attractive
D. Any of A, B, and C could be true.

2.20 In Solvency II, which of the following is true?

A. If capital falls below the solvency capital requirement, an insurance company is not allowed to take on further business.
B. If capital falls below the solvency capital requirement, an insurance company's policies may be transferred to another insurance company.
C. Insurance companies are only required to formulate a solvency capital plan if capital falls below the minimum capital requirement.
D. None of the above

ANSWERS

Short Concept Questions

2.1 **B.** The biggest risk for whole life insurance is that the policyholder dies earlier than expected. This is referred to as mortality risk. Longevity risk is the risk that the policy holder will live longer than expected. Sale of the policy to a third party should make no difference and policies are not inflation-linked.

2.2 **C.** Natural disasters such as hurricanes and earthquakes can lead to a high volume of claims in a year. A, B, and D are not relevant to property–casualty insurance.

2.3 **True.** Term insurance provides a payout if she dies in the next five years. Whole life insurance provides a payout whenever she dies.

2.4 **C.** Most forms of life insurance have mortality risk (i.e., the risk that policyholders will die earlier than expected). Annuities have longevity risk (i.e., the risk that policyholders will live longer than expected).

2.5 **D.** Life insurance companies tend to invest in long-term corporate bonds. This allows them to match assets and liabilities. Equities and commercial paper do not allow them to do this. They prefer long-term corporate bonds to long-term Treasury bonds because on an actuarial basis they get well compensated for taking credit risk.

2.6 **A.** In variable life insurance, the policyholder chooses how surplus funds are invested, and the final payout may increase if investments do well.

2.7 **B.** In universal life insurance, premiums can be varied from year to year.

2.8 **False.** A defined benefit plan is riskier because the company is responsible for any shortfall assessed by actuaries.

2.9 **A.** Loss ratio is the ratio of payouts to premiums.

2.10 **C.** Insurance companies are different from banks in that there are no global regulations. Regulations are enforced by the individual states in the U.S. The National Association of Insurance Commissioners is a forum for exchanging ideas but does not impose the regulations. In the European Union there is a single set of regulations, so German and Italian insurance companies should be subject to the same regulations.

Solved Problems

2.11 **B.** $$\frac{\text{Probability of Survival from birth to 90}}{\text{Probability of survival from birth to 70}} = \frac{0.29685}{0.82573} = 0.35950$$

2.12 **B.** The probability of a payout in the first year (time 0.5 years) is 0.017089. The probability of a payout in the second year (time 1.5 years) is

$$(1 - 0.017089) \times 0.018861 = 0.018539$$

The PV of the expected cost of the policy is therefore:

$$\frac{17{,}089}{1.03^{0.5}} + \frac{18{,}539}{1.03^{1.5}} = 34{,}573$$

The first premium is at time zero. The second premium, at time one year, has a probability of $1 - 0.017089 = 0.982911$ of being made. If the premium is X, the expected present value is

$$X + \frac{0.982911X}{1.03} = 1.954283X$$

The minimum premium is given by solving:

$$1.954283X = 34{,}573$$

It is 17,691.

2.13 **C.** There is both moral hazard and adverse selection. The behavior of a policyholder with a job might change (he or she will not be as concerned about being fired as he or she would be if there were no policy) and an employee without a job will not look as hard to find one. This is moral hazard. Also, the policy is likely to attract people with less secure jobs. This is adverse selection.

2.14 **B.** Low interest rates will reduce the discount rate used to assess liabilities and therefore increase the present value of the liabilities. High interest rates, employees working past retirement age, and a shortening of life expectancy all lessen risks. (The question states that the impact of A, B, C, and D on equity markets should not be considered. In practice a reduction in interest rates tends to lead to an increase in equity prices.)

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

2.15 False. Defined benefit pension plans invest in equities as well as bonds. Bonds do not provide a large enough return to meet their obligations and so they have no choice but to hope that equity markets will continue to perform well in the long term.

2.16 D. Health insurance premiums can increase from year to year but the increase should only reflect overall increases in health care costs.

2.17 A. The employee's salary is constant in real (inflation-adjusted) terms. Suppose it is X. The pension is $0.6X$. The real return earned is zero. The pension plan contributions therefore grow to $40XR$ where R is the (employer + employee) contribution rate as a percentage of the employee's salary. The present value of the benefits is $25 \times 0.6X = 15X$. We therefore require $40XR = 15X$ so that $R = 15/40 = 0.375$. Contributions equal to 37.5% of salary are therefore necessary. This simple example illustrates the problems with defined benefit plans. Total contributions of employer and employee are, in practice, 15% or less, but much higher contributions are necessary to pay promised pensions with certainty.

2.18 False. The property–casualty insurance company's risks are likely to be higher because the policies include some catastrophe risks.

2.19 B. The CAT bond is likely to be more attractive because it offers better diversification benefits. The corporate bond's return will be somewhat correlated with the return on market indices. The CAT bond has the advantage that its return is almost entirely uncorrelated with the return on market indices.

2.20 D. If capital falls below the *solvency* capital requirement, a plan to bring it back up must be formulated. If it falls below the minimum capital requirement, the insurance company may not be allowed to take on more business and its policies may be transferred.

3 Fund Management

Learning Objectives

After completing this reading, you should be able to:

- Differentiate among open-end mutual funds, closed-end mutual funds, and exchange-traded funds (ETFs).
- Identify and describe potential undesirable trading behaviors at mutual funds.
- Explain the concept of net asset value (NAV) of an open-end mutual fund and how it relates to share price.
- Explain the key differences between hedge funds and mutual funds.
- Calculate the return on a hedge fund investment and explain the incentive fee structure of a hedge fund, including the terms hurdle rate, high-water mark, and clawback.
- Describe various hedge fund strategies including long-short equity, dedicated short, distressed securities, merger arbitrage, convertible arbitrage, fixed income arbitrage, emerging markets, global macro, and managed futures, and identify the risks faced by hedge funds.
- Describe characteristics of mutual fund and hedge fund performance and explain the effect of measurement biases on performance measurement.

Fund managers invest money on behalf of individuals and companies. Funds from different clients are pooled, and the fund managers choose investments in accordance with stated investment goals and risk appetites. There are several advantages to this approach.

- Fund managers may have more investment expertise than their clients.
- Transaction costs (as a percentage of the amount traded) are usually lower for large trades than for small trades.
- It is difficult for a small investor to be well diversified, but a large fund with billions of dollars should not have any difficulty in achieving diversification.

In this chapter, we consider mutual funds, exchange-traded funds (ETFs), and hedge funds. Mutual funds and ETFs cater to individual investors, whereas hedge funds generally have high minimum investment thresholds that limit participation to wealthy individuals and institutions. Hedge funds are also subject to less regulation than mutual funds or ETFs and are free to follow a wide range of trading strategies. Furthermore, they are subject to fewer disclosure requirements than mutual funds or ETFs. However, they are subject to some additional restrictions on how they can solicit funds from investors.

3.1 MUTUAL FUNDS

Mutual funds (which are called unit trusts in some countries) have been a popular investment vehicle for small investors for many years. There has been considerable growth in the assets managed by mutual funds. In the U.S. for example, the assets of mutual funds (including ETFs) have grown from USD 0.5 billion in 1940 to more than USD 21 trillion in 2018.[1]

There are two types of mutual funds: open-end and closed-end. Open-end funds are by far the most popular and account for over 98% of mutual fund assets in the U.S.

Open-End Funds

The key feature of open-end funds is that the number of shares (and the size of the fund) expand and contract as investors choose to buy and sell shares. If more investors decide to buy shares rather than sell shares, the number of shares in the fund increases; if the reverse happens, and more investors sell, the number of shares decrease. Some open-end funds are offered by asset managers that are not banks (e.g., Fidelity), while others are offered by the asset management divisions within banks (e.g., JP Morgan Asset Management).

The assets of an open-end mutual fund are valued at 4 p.m. each day. The *net asset value* (NAV) is the value of the assets of the fund divided by the number of shares in the fund. All sales and purchases of shares take place at the 4 p.m. NAV after the order is placed. An investor might issue instructions to a broker to buy or sell shares at 10 a.m. on a business day, but the instructions will not be carried out until 4 p.m. that day.

When discussing insurance contracts in Chapter 2, we noted several tax advantages. For example, tax is deferred in many jurisdictions if funds are invested to provide an annuity or a pension. However, there are no tax advantages associated with investments in mutual funds. The investor pays taxes as though he or she owns the investments of the fund. For example, if shares are bought by the fund at USD 70 and sold at USD 90, the investor has a USD 20 capital gain that is subject to taxation.

Open-end funds can be categorized as follows:

- Money market funds,
- Bond funds, and
- Equity funds.

Funds that invest in more than one type of security are referred to as *hybrid funds* (or *multi-asset funds*).

Money market funds invest in fixed-income instruments that have a life of less than one year (e.g., commercial paper). Bond mutual funds invest in fixed-income securities that last more than one year. For many investors, money market funds are an alternative to a savings account at a bank. In fact, sometimes funds allow clients to write checks on their funds. The return on money market funds is usually higher than on bank deposits. However, there is no deposit insurance.

Most money market funds in the U.S. keep constant NAV of USD 1, with each day's gains being returned to the investor.[2] A negative return is therefore referred to as *breaking the buck* (because it causes the NAV to fall below USD 1). Breaking the buck is very unusual, but it did happen to the Reserve Primary Fund (a large money market fund) in September 2008. This happened because the fund held commercial paper issued by Lehman Brothers, which declared bankruptcy in September 2008. To avoid a panic, the government stepped in with a guaranty.

[1] *The Investment Company Factbook 2019*, The Investment Company Institute, 2019 https://www.ici.org/pdf/2019_factbook.pdf

[2] Regulations by the Securities and Exchange Commission have imposed a floating NAV requirement for some money market funds.

Equity funds are the most popular type of mutual funds. They can be subdivided into:

- Actively managed funds, and
- Index funds.

In the case of an actively managed fund, the fund manager uses his or her skill to achieve the fund's objectives. One objective might be to invest in stocks that provide a high dividend income. The fund would then trade in a manner consistent with this objective. Stocks in the portfolio that reduce their dividend would be sold, while those that increase their dividend will be purchased.

Index funds, on the other hand, attempt to track a specific index such as the S&P 500 or the FTSE 100. A simple way of doing this would be to buy all the shares in the index in amounts that reflect their weight in the index. Sometimes, a smaller representative set of shares are bought instead.

Tracking error measures how well a fund tracks its intended index. A popular tracking error is the square root of the average squared difference between the fund's return and the index's return (referred to as the *root-mean-square error*).[3] For example, consider a mutual fund that is designed to track the S&P 500. Suppose that the returns in successive years on the S&P 500 are 4.0%, 12.0%, 13.0%, −6.0%, and 2.0% and that the returns on the fund are 3.3%, 11.1%, 13.2%, −7.0%, and 2.0% (respectively). The differences between the fund's returns and the S&P 500 return are −0.7%, −0.9%, +0.2%, −1.0%, and 0.0% (respectively). The average difference is −0.48% (i.e., on average the fund has underperformed the S&P 500 by 0.48%).[4] The root mean squared error is 0.684%, or 68.4 basis points:

$$0.684 = \sqrt{\frac{[(-0.7)^2 + (-0.9)^2 + (0.2)^2 + (-1.0)^2 + 0.0^2]}{5}}$$

The *expense ratio* is the management fee charged on a yearly basis as a percentage of the value of the assets being managed. In addition, investors sometimes must pay a *front-end load* (a fee when they buy) and/or a *back-end load* (a fee when they sell). The fees charged vary from country to country. They are relatively low in the U.S. and Australia and relatively high in Canada and most European countries.[5] Fees are also much higher for actively managed funds than for index funds. Some index funds in the U.S. charge no fees,[6] whereas actively managed funds tend to charge 1% to 2% per year.

In 2008, the assets in index equity mutual funds were estimated to comprise 14% of the assets in all equity mutual funds. Since then, index mutual funds have become progressively more popular. By 2018, the assets in index mutual funds had risen to 29% of the assets in all mutual funds.[7] The popularity of index funds can be attributed to research (to be discussed later in this chapter) showing that actively managed funds do not, on average, outperform the market.

Closed-End Funds

Closed-end funds are funds where the number of shares remains constant through time. In essence, a closed-end fund is a regular company whose business is to invest in other companies. Thus, buying a share in a closed-end fund is like buying a share in any other company. Price changes balance supply (investors wanting to sell shares) with demand (investors wanting to buy shares).

Whereas open-end funds are bought and sold at their NAV, share prices for closed-end funds are typically lower than their NAV. For example, the price of a closed-end fund's shares might be USD 30 even though the NAV is USD 32 per share. This means that an arbitrageur could (in theory) profit by buying all the shares of the closed-end fund at the current market price and then selling the assets of the fund. In practice, this arbitrage opportunity would drive up the fund's share price and eliminate the profit.

Researchers have investigated the reason for the discrepancy between the NAV and the share price for closed-end funds. Stephen Ross has argued that management fees explain the discount.[8]

We can illustrate the argument with a simple example. Assume that an investor plans to keep a closed-end fund for five years and then sell it. Assume further that the fund pays no dividends and that the expected return from the fund's assets is 10%.

An investment of USD 100 in the assets of the fund (not the fund itself) has an expected value of USD 161.05 ($= 100 \times 1.1^5$) in five years. The correct risk-adjusted discount rate to apply to this expected value is 10% because this gives a present value of USD 100 and is consistent with the current USD 100 market value of the assets.

[3] Sometimes the tracking error is defined as the standard deviation of the error. Note that the root mean squared error is arguably a better measure than the standard deviation of the error. For example, if the fund always underperformed the index by 3%, the standard deviation of the error would be zero but the root mean squared error would be 3%.

[4] The underperformance may be a result of management fees, to be discussed shortly.

[5] See A. Khorana, H. Servaes, and P. Tufano, "Mutual fund fees around the world," *Review of Financial Studies*, 22 (March 2009): 1279–1310.

[6] Fidelity Investments began offering no-fee funds in 2018 https://fundresearch.fidelity.com/mutual-funds/summary/31635T708

[7] See Investment Company Institute www.ici.org/pdf/2019_factbook.pdf, page 73.

[8] See, for example, S. Ross, "Neoclassical finance, alternative finance, and the closed-end puzzle," *European Financial Management*, 8 (2002): 129–137.

Now consider what happens when the investment is in a closed-end fund and the fund invests in the USD 100 of assets. We suppose that the management fee is 1% of the fund's value each year. This means that the value of the fund grows at 9% rather than 10%. The expected value of the fund after five years is thus USD 153.8 (= 100 × 1.09^5). The correct discount rate is still 10% and therefore the value of the investment in the fund is

$$\frac{\text{USD } 100 \times 1.09^5}{1.10^5} = \text{USD } 95.54$$

We therefore expect the market value of the fund to be 4.46% less than the value of the underlying assets (USD 95.54 instead of USD 100.00).

As already mentioned, closed-end funds are much less popular than open-end funds. However, they do have certain advantages. Unlike open-end fund shares, shares of closed-end funds can be bought and sold at any time of day; they can even be shorted.[9] Also, unlike open-end funds, closed-end funds do not need to keep enough liquid assets to handle possible redemptions. This is because closed-end fund investors trade with each other, whereas open-end fund investors trade with the fund itself.

3.2 EXCHANGE-TRADED FUNDS

Exchange-traded funds (ETFs) combine features of open-end mutual funds with features of closed-end mutual funds. In a 2008 survey of investment professionals, ETFs were voted the most innovative investment product of the previous two decades. They have been trading in the U.S. since 1993 and in Europe since 1999. A well-known ETF is the SPDR S&P 500 ETF, which tracks the S&P 500 and trades under the symbol SPY.

An ETF is created when an institutional investor deposits a block of shares with the ETF and in exchange receives shares in the ETF. The shares of the ETF are then traded on an exchange just like the shares of any other company.

Institutional investors have the right to give up shares in the ETF and in exchange receive their share of the ETF's underlying assets. They also have the right to do the reverse (i.e., obtain additional shares in the ETF by adding assets). The assets added must have the same composition as the ETF's current assets. These features act as a control mechanism preventing the NAV and the share price from diverging.

[9] Shorting is a procedure where an investor borrows shares from another investor and sells them in the market, hoping to buy them back more cheaply later.

For example, consider an ETF where there are 10,000 shares worth USD 25 each. The value of the fund's assets is USD 260,000 and therefore the NAV is USD 26 (= 260,000/10,000). An institutional investor owning 1,000 shares could choose to exchange the shares for one tenth of the fund's portfolio. This would reduce the number of shares of the ETF to 9,000 and provide the institutional investor with an immediate gain of USD 1 per share.

Now suppose that the shares are worth USD 26 each and the value of the assets of the fund is USD 250,000. In this case, an institutional investor can acquire (from elsewhere) a portfolio that is the same as one tenth of the fund's portfolio and exchange that portfolio for 1,000 new shares in the fund. The number of shares in the fund would then increase to 11,000, and the institutional investor should gain USD 1 for each new share.

The fact that institutional investors have these opportunities means that in practice the share price and the net asset value do not diverge (at least not significantly or for long periods of time). While many ETFs are designed to track an index, they can be actively managed as well. However, to ensure that the mechanism for equalizing the NAV to the share price works, they must disclose their assets twice a day. Mutual funds disclose their assets much less frequently.

ETFs are gaining in popularity. By 2018, their total net assets in the U.S. was about USD 3.4 trillion.[10] As mentioned earlier, they combine features of open- and closed-end funds. Like closed-end funds they can be traded at any time, they can be shorted, and they do not have to keep liquid assets to meet redemptions. But unlike closed-end funds, ETFs trade with no discount to the NAV.

3.3 UNDESIRABLE TRADING BEHAVIOR

In the U.S., mutual funds and ETFs are heavily regulated by the Securities and Exchange Commission (SEC).[11] Complete and accurate financial information must be provided to prospective investors. There are also rules to prevent conflicts of interest and fraud. Despite these safeguards, there have been instances of undesirable behavior.

a) *Late Trading.* As mentioned, all trades to buy and sell shares in an open-end mutual fund are at 4 p.m., and thus

[10] See Investment Company Institute www.ici.org/pdf/2019_factbook.pdf, page 83.

[11] For Security and Exchange Regulations, see https://www.sec.gov/fast-answers/answersmutfundhtm.html

trade instructions should reach a broker before 4 p.m. For administrative reasons, however, those trades may not be passed to the mutual fund until well after 4 p.m. If there are market developments soon after 4 p.m., a trader might call his or her broker to cancel a trade due to be carried out at the 4 p.m. price or to put through a new trade at that price. Some brokers in the U.S. have been known to (dishonestly) accept orders after 4 p.m. This activity is known as *late trading* and is not permitted by the SEC. In fact, there have been several prosecutions leading to large fines and the involved employees losing their jobs.

b) *Market Timing*. Not all the assets of an open-end mutual fund trade actively. This may lead to the prices used to calculate NAV being stale (i.e., not reflecting recent information). The prices of securities that trade on overseas markets may also be stale because of time zone differences. For example, suppose that it is now 3:45 p.m. and prices in markets have been rising during the last few hours. The existence of stale prices means that shares in the mutual fund are probably worth slightly more than the NAV, and therefore buying at the 4 p.m. NAV is attractive. Similarly, if prices in markets have been falling, selling at the 4 p.m. NAV is attractive. Market timing trades of this sort are not illegal, but they must be quite large to be worthwhile. If the mutual fund allows the trades, the size of the fund will whipsaw up and down. This could lead to costs for all investors as the fund may have to keep additional cash to accommodate redemptions. Regulators are likely to be concerned if special trading privileges are offered to market timers.

c) *Front Running*. If a trader working for a mutual fund (or any other type of fund) knows that the fund will execute a big trade that is likely to move the market, it is tempting for the trader to trade on his or her own account immediately before putting through the fund's trade. For example, if a fund is going to buy 1 million shares of a certain stock, the trader might buy 10,000 for his or her own account first. The trader could also inform favored customers or other fund employees about what is about to happen and allow them to trade ahead of a predicted price increase or decrease. Front running is illegal in fund management.

d) *Directed Brokerage*. This involves an informal arrangement between a mutual fund and a brokerage house. The unwritten agreement is that the mutual fund will use the brokerage house for its trades if the brokerage house recommends the mutual fund to its clients. The practice is frowned upon by regulators.

3.4 HEDGE FUNDS

Hedge funds (a form of *alternative investments*) are subject to less regulation than mutual funds and ETFs. While mutual funds and ETFs cater to the needs of small investors, hedge funds usually accept only large investments from wealthy private individuals or institutions. There are several other differences. The following are examples.

- A mutual fund or ETF allows investors to redeem their shares on any day. A hedge fund may have a *lock-up period* during which time funds cannot be withdrawn. Lock-up periods of one year are common.
- The NAV of a mutual fund or ETF must be calculated and reported at least once a day. In contrast, hedge funds have no such requirements, and their NAVs are reported much less frequently.
- Mutual funds and ETFs must disclose their investment strategies. Hedge funds generally follow proprietary strategies that they see as fundamental to their competitiveness and/or value proposition. They give prospective clients some information to explain their value proposition, but do not disclose everything. Furthermore, they are not obligated to stick to one strategy.
- Mutual funds and ETFs may be restricted in their use of leverage. A hedge fund is only restricted by the amount banks are willing to lend to it.
- Hedge funds charge an incentive fee as well as a management fee. A typical hedge fund fee is *2 plus 20%*. This means that the investors are charged 2% of the value of their investment per year along with 20% of the profits (if these net profits are positive).[12]

The descriptor *hedge fund* arises from the long-short strategies followed by many hedge funds.[13] This type of strategy involves taking long positions in stocks expected to provide good returns and short positions in stocks expected to provide poor returns. However, there are also hedge fund strategies that involve little to no hedging whatsoever.

Some hedge funds have done very well. Jim Simons is a former math professor who founded hedge fund Renaissance Technologies in 1982. Its flagship Medallion fund has had a truly amazing record—returning an average of 35% per year over

[12] Fund of funds have been developed to help investors choose a portfolio of hedge funds. This creates an extra layer of fees for the investor. Fund of funds used to be able to charge as much as 1 plus 10%, but their fees are now usually much less than this.

[13] It was coined by Carol Loomis in 1966 in an article about the first hedge fund, A.W. Jones & Co.

a 20-year period.[14] It uses huge volumes of data and complex mathematical models to determine patterns and devise trading strategies, many of which are automated. As previously mentioned, a common hedge fund fee schedule is 2 plus 20%. It is reported that Renaissance has charged as much as 5 plus 44%. In 2019, Jim Simons' wealth was estimated to be USD 21.5 billion.[15]

Other well-known hedge funds are Bridgewater (founded by Ray Dalio), Soros (founded by George Soros), and Citadel (founded by Ken Griffin). All three founders have become very rich from the incentive fees charged to investors.

In this chapter, we will assume that the management fee is calculated on the assets at the beginning of the year and that the incentive fee is calculated after subtracting management fees. But it should be noted that some hedge funds do try to use a more aggressive fee structure where the management fee is calculated based on the end-of-year asset value (making it greater if the fund's value has increased) and the incentive fee is calculated before subtracting management fees (which is also more valuable to the hedge fund managers if the value of fund has increased).

The incentive fee can be thought of as a call option on the net profit produced by the hedge fund for an investor in a given year. Consider the 2 plus 20% fee schedule. Using the assumptions mentioned above, we can calculate the incentive fee as:

$$0.2 \times \max(R \times A - 0.02 \times A, 0)$$

where A is the assets under management at the beginning of the year and R is the return on the assets during the year. This is the payoff from a call option on the dollar return with a strike price equal to 2% of the assets under management. This creates a situation where a hedge fund has an upside, but no downside.

To illustrate this point, suppose that a hedge fund has USD 100 million of investors' funds and its fees are 2 plus 20%. Consider the following three strategies.

1. *Strategy A1:* Choose a safe investment that will produce a profit of USD 3 million with certainty. The expected return is 3% (= 3/100).
2. *Strategy B1:* Choose a riskier strategy that has a 50% chance of producing a profit of USD 5 million and a 50% chance of producing a profit of zero. The expected return is 2.5% (= 0.5 × 5% + 0.5 × 0%).
3. *Strategy C1:* Choose a highly risky strategy that has a 50% chance of producing a profit of USD 10 million and a 50% chance of a loss of USD 10 million. The expected return is 0% [= 0.5 × 10% + 0.5 × (−10%)].

[14] This includes a 98.2% return in 2008, a year when the S&P 500 lost 38.5%.

[15] According to the Forbes Billionaires List.

The expected returns for strategies A1, B1, and C1 are 3%, 2.5%, and 0% (respectively). Strategies B1 and C1 make no sense because the risk-return trade-off is negative (i.e., the expected return decreases as risk increases). Nevertheless, strategies B1 and C1 could be attractive to hedge fund managers (as we will now show).

The fees to the hedge fund for each of the strategies are calculated in Table 3.1. The management fee for all three strategies is USD 2 million (i.e., 2% of USD 100 million). For the first strategy, the profit after the management fee is USD 1 million and the incentive fee is thus USD 0.2 million (i.e., 20% of this profit). For the second strategy, there is a 50% chance that the incentive fee will be zero, and a 50% chance that it will be 20% of USD 3 million. For the third strategy, there is a 50% chance that the incentive fee will be zero, and a 50% chance that it will be 20% of USD 8 million.

The table shows that the expected incentive fee increases as the strategy becomes riskier even though the expected gross return declines as this occurs.

If the hedge fund manager does better as more risks are taken, the investor is likely to do worse. This is illustrated in Table 3.2.

This simple example illustrates that a hedge fund has an incentive to take risks even if the risks are not justified by a higher expected return.

It is also the case that hedge funds can benefit from risks even when there is an acceptable risk-return trade-off (i.e., the hedge fund earns a higher expected return from taking more risks).

Table 3.1 Hedge Fund Fees for Strategies A1, B1, and C1 (USD Million)

Strategy	Management Fee	Expected Incentive Fee	Total Expected Fee
A1	2	0.2	2.2
B1	2	0.3	2.3
C1	2	0.8	2.8

Table 3.2 Returns to Hedge Fund and Investors from Strategies A1, B1, and C1

Strategy	Expected Return to Hedge Fund	Expected Return to Investor	Total Expected Return
A	2.2%	0.8%	3.0%
B	2.3%	0.2%	2.5%
C	2.8%	−2.8%	0.0%

For example, suppose that we change the above example so that the three strategies become the following.

1. *Strategy A2:* Choose a safe investment that will produce a profit of USD 3 million with certainty.
2. *Strategy B2:* Choose a riskier strategy that has a 50% chance of producing a profit of USD 10 million and a 50% chance of producing a profit of zero. The expected return is 5% (= 0.5 × 10% + 0.5 × 0%).
3. *Strategy C2:* Choose a highly risky strategy that has a 50% chance of producing a profit of USD 30 million and a 50% chance of a loss of USD 12 million. The expected return is 9% (= 0.5 × 30% + 0.5 × (−12%)).

In this case, there is a positive risk-return trade-off because the expected total returns from strategies A2, B2, and C2 are 3%, 5%, and 9% (respectively). The expected fees for the hedge fund are now calculated as shown in Table 3.3.

Strategy A2 is the same as strategy A1 so that the investors earn 0.8%. The investor's expected returns after fees from strategy B2 is 2.2% (= 5% − 2.8%), and the investor's expected return after fees from strategy C2 is 4.2% (= 9% − 4.8%). The hedge fund's expected return is thus higher than the investor's expected return. Furthermore, it is never negative. These results are summarized in Table 3.4.

It would be wrong to give the impression that hedge fund managers always try to maximize the value of the embedded call option or that they always earn more than the investors. However, abuses do occur. A hedge fund called Amaranth tried to increase the value of the call option embedded in its incentive fees by taking huge risks is Amaranth. Brian Hunter, a star trader at the firm, took large highly leveraged positions in natural gas futures in 2006. He was betting that prices in the winter would rise relative to those in the summer. This type of strategy worked well the previous year because hurricanes had adversely affected natural gas supplies. Hunter's bonus from his trading in that year is reported to have been about USD 100 million. In 2006, however, Hunter's strategy lost about two thirds of the USD 9 billion that had been invested with Amaranth. But while these losses caused the fund to be wound down, Hunter was able to keep the bonus he had received the previous year.

Table 3.3 Hedge Fund Fees for Strategies A2, B2, and C2 (USD Million)

Strategy	Management Fee	Expected Incentive Fee	Total Expected Fee
A2	2	0.2	2.2
B2	2	0.8	2.8
C2	2	2.8	4.8

Table 3.4 Returns to Hedge Fund and Investors from Strategies A2, B2, and C2

Strategy	Expected Return to Hedge Fund	Expected Return to Investor	Total Expected Return
A2	2.2%	0.8%	3.0%
B2	2.8%	2.2%	5.0%
C2	4.8%	4.2%	9.0%

Investors are naturally wary of the one-sided nature of incentive fees and hedge funds often adjust the terms they offer to reflect this. For example:

- Sometimes incentive fees are payable only on returns above a certain level. This level is referred to as a *hurdle rate*.
- Sometimes the agreement between hedge funds and investors states that incentive fees only apply when cumulative profits for an investor are positive. This is known as *a high-water mark* clause. If USD 100 million is invested, and the hedge fund loses USD 10 million, for example, it must make USD 10 million before incentive fees kick in. There may also be a *proportional adjustment clause* stating that the high-water mark only applies to funds that are not withdrawn. If the investor in our example withdraws half of his or her remaining funds after the loss, the hedge fund only needs to make USD 5 million for the investor before incentive fees apply.
- There is sometimes a *clawback clause* where incentive fees paid by the investor can be used to offset future losses.

When hedge fund managers incur substantial losses, however, they have an incentive to close a fund and start a new one to avoid high-water marks and clawbacks.

Prime Brokers

A hedge fund's prime broker is the bank that handles its trades and lends it money. Many hedge funds take short positions, and the prime broker will handle these for them as well. The bank may provide risk management and hedging services as well. Furthermore, the prime broker can carry out stress tests on the hedge fund's portfolio to decide how much it is prepared to lend. The hedge fund can then post its securities with the bank as collateral.

As mentioned, hedge funds are subject to very little regulation. However, their activities may be constrained by their prime broker. There is always a danger that in some adverse economic environments (e.g., those experienced during the 2007–2008 crisis), the prime broker will reduce the borrowing limit of the hedge fund and force it to close out positions.

Some hedge fund strategies can be certain to make money in the long term while risking short term losses. If these losses occur, the bank could require additional collateral. Hedge funds must therefore consider the extent to which their prime brokers are prepared to fund short-term losses. Long Term Capital Management is an example of a hedge fund that took positions that could reasonably be expected to be profitable if held for several years. However, there were huge short-term losses in 1998 because of the impact of Russia's default on its debt. The increased collateral requirements that followed led to the fund's failure.

Large hedge funds may use more than one prime broker. This gives hedge funds additional flexibility and ensures that no one bank is able to see every trade. Prior to 2008, it was considered that nearly all the risks in the prime broker–hedge fund relationship were borne by the prime brokers. When Lehman defaulted, however, many hedge funds that used Lehman Brothers as their prime broker found that they could not access the securities that they had posted as collateral. This made the market realize that both sides were subject to risks.

3.5 TYPES OF HEDGE FUNDS

Hedge funds follow many different trading strategies. Here we discuss a few of the more common ones.[16]

Long-Short Equity

The original hedge funds were long-short funds that purchased stocks that were considered underpriced and shorted those considered to be overpriced. A long-short fund's performance should depend entirely on the fund manager's ability to pick winners and losers (and not on what happens to the market as a whole) so long as:

- The value of the shares shorted equals the value of those bought, and
- Both the long and short portfolios have the same sensitivity to market movements.

For example, consider a situation where General Motors and Ford are considered to have the same sensitivity to the S&P 500. However, Ford is considered to be undervalued, while General Motors is considered to be overvalued. A long-short strategy could therefore involve buying USD 100,000 of Ford stock and selling USD 100,000 of General Motors stock.

There are many variations on the traditional long-short strategy. If the hedge fund manager thinks that the market is more likely to go up than down, the value of the long position may be greater than the value of the short position. If the reverse is true, the manager may choose to have a larger short position.

Dedicated Short

At any given time, it is reasonable to suppose that there are as many overvalued shares as undervalued shares. A dedicated short fund devotes its attention to picking overvalued stocks. Hedge funds using dedicated short strategies look for companies that are experiencing difficulties not recognized by the market. However, they are not hedged against the overall performance of the market. It is therefore not surprising that dedicated short strategies perform badly during bull markets.

Distressed Debt

Some hedge funds specialize in trading distressed debt. They use their understanding of the bankruptcy process to find situations where they can take a big position in the debt and benefit from reorganization proposals.

Merger Arbitrage

When a company announces that it is prepared to buy another company, there is usually some uncertainty about whether the acquisition will proceed. The share price of the target company usually increases, but not to the price being offered. A merger arbitrage hedge fund might consider that there is an 80% chance that the acquisition will be successful and that the acquisition price will be higher than the current price. Buying the target company's shares could then be a good trade. If the offer on the table is a share-for-share exchange and the hedge fund expects an improvement in the terms before a deal is finally announced, it could buy the shares in the target company and short shares in the acquiring company in a ratio that reflects the current offer.

It should be emphasized that merger arbitrage is not about trading on inside (non-public) information (which is illegal).[17] It is about assessing the probability of a merger being successful and the likely final price (or exchange ratio in the case of a share-for-share exchange) at the time of the merger announcement.

[16] Renaissance Technologies does not fit into any of the standard hedge fund categories. It could be categorized as a *Quant Fund*.

[17] Ivan Boesky was sentenced to three years in prison for insider trading.

Convertible Arbitrage

Convertible bonds are bonds issued by a company that can be converted into a predetermined number of the company's shares at a future time. Convertible arbitrage hedge funds use sophisticated models to value convertible bonds. They hedge the risks associated with the company's share price, credit spreads, and interest rates. If the market price is currently different from the hedge fund's model price, the strategy can be profitable if the market price converges to the model price.

Fixed-Income Arbitrage

At any given time, some traded bonds are likely to be relatively expensive compared with other similar bonds, while others are relatively cheap. In a fixed-income arbitrage strategy, the hedge fund manager buys bonds that seem relatively cheap and shorts the ones that are relatively expensive. Typically, they use a lot of leverage to make the strategy worthwhile.

Emerging Markets

Hedge funds that specialize in emerging markets attempt to gather information about little-known equity securities in developing countries. When they consider a security to be undervalued (overvalued), they buy (short) it in the local markets. An alternative is to use American Depositary Receipts (ADRs), which are certificates backed by shares of a foreign company and traded on an exchange in the U.S. Any discrepancies between ADR prices and local prices can give rise to arbitrage opportunities.

Hedge funds can also invest in emerging market sovereign debt. However, this is fraught with risks. Countries such as Russia, Argentina, Brazil, and Venezuela have defaulted multiple times (as will be discussed in Chapter 5 of *Valuation and Risk Models*).

Global Macro

Global macro hedge funds use macroeconomic analysis to determine their trades. Specifically, they look for situations where markets are not in equilibrium using models based on factors such as exchange rates, interest rates, balance of payments, inflation rates, etc. Sometimes the results are spectacular: The Quantum Fund managed by George Soros made a profit of USD 1 billion in 1992 by betting that the British pound was overvalued. However, not all global macro trades are that successful, and economies can remain in disequilibrium for long periods of time.

Managed Futures

Managed futures strategies attempt to predict future commodity prices and take positions that will be profitable if the predictions are correct. Several different models are used and trading rules are usually *back-tested* by seeing how well they would have performed if they had been used in the past. However, there is a danger in this. Back-testing does not differentiate between strategies with a fundamental understanding of the markets and strategies that were just lucky (and thus not necessarily bound to be successful in the future). It is also important to test a trading strategy out-of-sample. This means that the historical data used to test a strategy should be separate from the historical data used to develop the strategy.

3.6 RESEARCH ON RETURNS

The key question for an investor is: "Is it worth paying a professional to invest my money for me?" It is not clear that the answer to this question is yes. Some mutual funds and hedge funds have produced excellent returns for investors, but there is always the all-too-true small print in any solicitation: "Past performance is no guarantee of future results."

Mutual Fund Research

Questions a prospective mutual fund investor might reasonably ask are as follows.

- On average, do actively managed mutual funds outperform the market?
- Do actively managed funds that outperform the market in one year have a high probability of doing so in the next year?

Michael Jensen investigated these questions in the 1960s, and the answer to both appears to be no.[18] Over the past half century, his results have been confirmed by many other researchers using more recent data.

Jensen found that actively managed mutual funds (on average) do not beat the market after expenses. This is not surprising. On average, the returns to all investors (before expenses) is the market's return. Because mutual funds and ETFs hold over 30% of all U.S. corporate equity, any outperformance by these securities would imply systematic underperformance by other investors.

To answer the second question, Jensen calculated the probability that a mutual fund that has beaten the market for one or

[18] See M. C. Jensen, "Risk, the pricing of capital assets, and the evaluation of investment portfolios," *Journal of Business*, 42 (April 1969): 167–247.

more years will do so again the following year. This is referred to as testing for *persistence*. He found that only 50% of mutual funds that beat the market one year did so again the following year. Mutual funds that beat the market two years in a row also had a roughly 50% probability of beating the market in the following year. Similar results were obtained for mutual funds that had beaten market for three, four, five, and six years. While there may be some mutual funds that can consistently beat the market, Jensen's research (which has been confirmed by other researchers using more recent data) indicates that there cannot be very many of them.

This research has led many investors to invest in index funds rather than actively managed funds. Index funds are termed *passive investments*. They charge lower fees than actively managed funds and on average provide investors with a better return.

These comments may seem to contradict the impressive returns advertised by many mutual funds. However, the fund featured in a mutual fund advertisement may be one out of many different funds offered by a fund manager. Because Jensen's research would suggest that a fund has a 1/2 probability of beating the market in one year, the probability that it will beat the market every year for four years is therefore $1/16$ ($= (1/2)^4$). If a company has 16 different funds, there is a good chance that one will beat the market every year for the last four years. That is the fund that will be advertised.

Hedge Fund Research

It is not as easy to assess hedge fund performance as it is to assess mutual fund performance. There are organizations that collect data on returns and provide return statistics for different types of hedge funds, but participation by hedge funds is (to some extent) voluntary and not all of them report their results.[19] In fact, one can speculate that funds that incur losses (and funds that are closed) will be less inclined to report returns. This tendency would in turn create an upward bias in average reported returns.

The strategies followed by hedge funds are more sophisticated than those followed by mutual funds, and a few have created a lot of wealth for investors (we mentioned the outstanding success of Jim Simons and Renaissance Technologies earlier). But others have failed badly. Sometimes hedge funds report good returns for a few years and then lose a large percentage of funds under management (e.g., Long Term Capital Management and Amaranth).

[19] These problems lead to a situation where the statistics produced by one organization are not always in agreement with those produced by another organization.

Table 3.5 Performance of Hedge Funds and the S&P 500

Year	BarclayHedge Index Net Return (%)	S&P 500 Return Including Dividends (%)
2008	−21.63	−37.00
2009	23.74	26.46
2010	10.88	15.06
2011	−5.48	2.11
2012	8.25	16.00
2013	11.12	32.39
2014	2.88	13.38
2015	0.04	1.38
2016	6.10	11.96
2017	10.36	21.83
2018	−5.08	−4.38

Prior to 2008, hedge funds performed quite well compared with the S&P 500. However, since 2008 the performance of hedge funds has been worse than that of the S&P 500. This is illustrated in Table 3.5, which uses the BarclayHedge index to compare returns of all hedge funds with returns on the S&P 500 from 2008 to 2018. In 2008, which was a watershed year for equity markets, hedge funds on average lost money but outperformed the S&P 500. From 2009 to 2018, hedge funds on average underperformed the S&P 500.

A possible reason for these results is that hedge funds tend to underperform in bull markets and outperform in bear markets. As is apparent from the descriptions of different strategies in Section 3.5, most strategies are not designed to follow market trends. For example, the long-short strategy is designed to either attenuate or eliminate the effects of market moves.

In view of the statistics in Table 3.5, a surprising fact is that hedge funds have been quite successful in attracting investors. Hedge fund assets under management were estimated to be a record USD 3.245 trillion in mid-2019 by the HFR Global Hedge Fund Industry Report.

SUMMARY

An advantage of mutual funds and ETFs for a small investor is that they provide an easy way for the investor to obtain a well-diversified portfolio. Open-end mutual funds are by far the most popular type of mutual fund. In these funds, the number of shares increases as new investors are attracted to the fund and decreases as investors redeem their shares. All trades happen at the net asset value calculated each trading day at 4 p.m. A closed-end fund has a fixed number of shares that trade throughout the day.

Exchange-traded funds are proving to be popular alternatives to mutual funds. They trade on an exchange throughout the day. They are different from closed-end funds in that there is a mechanism whereby the share price matches the net asset value of the fund's investments.

Hedge funds are used mostly by wealthy individuals and institutional investors. They follow innovative strategies and charge an incentive (performance) fee, which is a percentage of profits, as well as a management fee. However, they can be criticized because the incentive fee is (in essence) an option on the performance of the fund and encourages the fund to take risks that are not in the best interests of investors.

Research shows that, on average, actively managed mutual funds do not outperform the market. Furthermore, there is very little persistence: fund managers who have performed well for several years have probably done so by chance, and it would be a mistake to assume that they have a better than even chance of doing so in the future. These observations have increased the attractiveness of funds whose objective is to mimic an index of equity returns such as the S&P 500.

The data on the performance of hedge funds is not as reliable as that on the performance of mutual funds. Some hedge funds have performed extremely well and have shown persistence, but the statistics indicate that on average hedge funds have underperformed the S&P 500 in every year between 2009 and 2018.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

3.1 What is the difference between the share structures of open-end and closed-end mutual funds?

3.2 How is NAV defined?

3.3 Explain how trades by investors to buy or sell shares in the fund are handled by an open-end mutual fund.

3.4 What is the difference between an ETF and a closed-end mutual fund?

3.5 Give three examples of restrictions placed on mutual funds that do not apply to hedge funds.

3.6 Explain the meaning of late trading and front running.

3.7 Explain the meaning of market timing and directed brokerage.

3.8 What are (a) hurdle rates, (b) high-water marks, and (c) clawbacks in the contract between hedge funds and investors?

3.9 What does a hedge fund's prime broker do?

3.10 What is meant by persistence in mutual fund returns?

Practice Questions

3.11 The fees of a hedge fund are 2% plus 20%. What is the investor's return as an algebraic function of the hedge fund's return? Consider all possible values of the hedge fund's return. Assume that the incentive fee is applied after the management fee has been subtracted and that the management fee is applied to the beginning-of-year assets under management.

3.12 Repeat Question 3.11 assuming that the hedge fund follows the more aggressive strategy mentioned in the chapter where the incentive fee is applied before the management fee has been subtracted, and the management fee is applied to the end-of-year assets under management.

3.13 Consider how Renaissance Technologies might justify a 5 plus 44% fee schedule. Assume that the incentive fee is applied after the management fee has been subtracted, that the management fee is applied to beginning-of-year assets, and that Renaissance has averaged a 35% return on assets under management in recent years.

3.14 An investor buys 100 shares in an open-end mutual fund on January 1 of a year for USD 30. The fund earns dividends of USD 1 per share in the first year and USD 2 per share in the second year. These dividends are reinvested by the fund. The capital gains in the first year are USD 3, and the capital gains in the second year are USD 4. The investor sells the shares for USD 43 in the third year. Explain how the investor is taxed.

3.15 A mutual fund investor tells you, "I only invest in funds that have beaten the market over the last three years." How would you respond?

3.16 What are index funds? Why has their popularity increased in recent years?

3.17 Why does an ETF not have to worry about the liquidity of its fund, whereas an open-end mutual fund does?

3.18 What is the mechanism that leads to the share price of an ETF being close to its NAV?

3.19 Why does the incentive component of the fee of a hedge fund involve an option? How can the hedge fund increase the value of the option?

3.20 Why do hedge funds tend to beat a bear market and perform less well than a bull market?

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

3.1 An open-end fund has a number of shares that increase and decrease as investors buy new shares or redeem existing shares. A closed-end fund has a fixed number of shares.

3.2 NAV is the value of the assets under management divided by the number of shares of the mutual fund.

3.3 All trades are carried out at a price equal to the NAV calculated at 4 p.m.

3.4 In an ETF there is a mechanism to ensure that the share price equals the NAV.

3.5 Mutual funds must be redeemable on any day. They must calculate net asset value at least once a day. They must disclose their investment policies. They may be restricted in the leverage they can take on.

3.6 Late trading is the illegal practice of finding a way of trading an open-end mutual fund at the 4 p.m. price after 4 p.m. Front running is the illegal practice of trading to profit from the fact that a large trade by a fund is expected to move the market.

3.7 Market timing involves an open-end mutual fund investor embarking on large trades motivated by the fact that some of the security prices used to calculate an NAV are stale. Directed brokerage is an arrangement between a fund and a broker where the broker recommends the fund and the fund uses the broker for trading.

3.8 A hurdle rate is the return that must be exceeded for an incentive fee to apply. A high-water mark is a loss that must be recouped before incentive fees apply. A clawback clause is a clause allowing investors to apply previously paid incentive fees to losses.

3.9 The hedge fund's prime broker is a bank that carries out trades for it and lends it money. Typically, the securities of the hedge fund reside with the prime broker as collateral for loans and trades carried out.

3.10 Persistence measures the extent to which a fund manager who has performed well in the past will continue to do so.

Solved Problems

3.11 The investor's return as a function of the hedge fund's return, R_H, is

$$0.8(R_H - 0.02) \text{ if } R_H > 0.02$$

$$R_H - 0.02 \quad \text{if } R_H \le 0.02$$

3.12 The investor's return as a function of the hedge fund's return, R_H, is

$$0.8R_H - 0.02(1 + R_H) \text{ if } R_H > 0$$

$$R_H - 0.02(1 + R_H) \text{ if } R_H \le 0$$

3.13 Renaissance Technologies has performed spectacularly well, with returns averaging 35% per year. If it continues to earn 35%, investors will earn 0.56 × 30% or 16.8% per annum.

3.14 The investor will pay tax on the investment income from dividends of USD 100 in the first year and on USD 200 in the second year. The investor will also pay tax on capital gains of USD 300 in the first year and USD 400 in the second year. The investor's basis at the start of the third year will be USD 40 per share (= 30 + 1 + 2 + 3 + 4) or USD 4,000 in total. The shares are sold for USD 4,300. Tax is therefore payable on USD 300 in the third year.

3.15 Research shows that mutual funds that have beaten the market for three years only have a probability of about 50% of beating the market in the fourth year (i.e., there is very little persistence in returns).

3.16 Index funds are funds that attempt to mirror the performance of a stock index such as the S&P 500. Their popularity has increased because investors have realized that, on average, actively managed funds do not consistently beat the market and there is very little persistence in their returns.

3.17 When an investor redeems shares in an open-end mutual fund, he or she trades with the fund. If the fund does not have liquid assets, it must sell part of its portfolio to provide the investor with cash for the shares. When an investor redeems shares in an ETF, he or she trades with another investor, so the fund does not have this problem.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

3.18 Institutional investors have the right to exchange shares in the fund for their share of the assets of the fund. They can also obtain new shares in the fund by contributing to the fund a portfolio that reflects the fund's current asset mix.

3.19 The fund receives an incentive fee if its return is positive (or above a hurdle rate). The incentive fee is never negative. The fund's fees are therefore a call option on the return of the fund. The value of call options increases as volatility increases. This means that the incentive fee actually incentivizes the fund to take as much risk as possible.

3.20 Many hedge funds follow strategies that have little to do with the return from the market. If the market does very well, this may not be reflected in the hedge fund's returns. If the market does badly, this also may not be reflected in the hedge fund's returns.

4

Introduction to Derivatives

Learning Objectives

After completing this reading, you should be able to:

- Define derivatives, describe the features and uses of derivatives, and compare linear and non-linear derivatives.
- Describe the specifics of exchange-traded and over-the-counter markets, and evaluate the advantages and disadvantages of each.
- Differentiate between options, forwards, and futures contracts.
- Identify and calculate option and forward contract payoffs.
- Differentiate among the broad categories of traders: hedgers, speculators, and arbitrageurs.
- Calculate and compare the payoffs from hedging strategies involving forward contracts and options.
- Calculate and compare the payoffs from speculative strategies involving futures and options.
- Describe arbitrageurs' strategy and calculate an arbitrage payoff.
- Describe some of the risks that can arise from the use of derivatives.

The next three chapters introduce derivatives and the ways they are traded. Derivatives are contracts whose values depend on (or derive from) the values of one or more financial variables (e.g., equity prices, exchange rates, and interest rates). These variables are referred to as *underlyings*.

Derivatives can be categorized into linear and non-linear products. Linear derivatives provide a payoff that is linearly related to the value of the underlying asset(s). Forward contracts are an example of linear derivatives. Specifically, they are agreements between two parties to buy or sell an asset at a specified price at a future time. As we will see in later chapters, the value of a forward contract prior to maturity (as well as its payoff at maturity) is linearly dependent on the value of the underlying asset.

Options, on the other hand, are non-linear derivatives (i.e., their payoff is a non-linear function of the value of their underlying assets). They are contracts where the holder has the right (but not the obligation) to buy or sell an asset for a specified price at a future time. The payoff from an option is a non-linear function of the value of its underlying(s).

The value of the underlying(s) is central to the valuation of both linear and non-linear derivatives. That being said, other variables (e.g., interest rates and volatilities) can also play an important role.

Derivatives have existed for many years. Before the development of money, for example, goods could be sometimes be exchanged for crops that would be harvested in the future. In the nineteenth century, the Chicago Board of Trade was set up to facilitate agreements to exchange commodities in the future.

Derivatives now trade on exchanges as well as in over-the-counter markets. Both types of markets have experienced unprecedented growth as derivatives have become widely used.

Examples of the applications of derivatives include:

- Derivatives are used by companies to manage interest rate risk, foreign exchange risk, the risk arising from commodity price changes, and many other risks.
- Derivatives sometimes form part of a corporate bond issuance. They can give bond issuers the right to repay bonds early or give bond holders the right to demand early repayment. Sometimes, bond holders have the right to convert a bond into equity of the issuing company.
- Employee compensation plans sometimes give employees options to buy shares from the company at future times for a predetermined price.
- Capital investment opportunities often have embedded options. For example, a company embarking on an investment may be able to abandon it if things go badly or expand if things go well. These are referred to as "real options" because they involve physical assets (rather than financial assets). It is now a common practice to consider real options when valuing capital investment opportunities.
- Homeowners sometimes have derivatives embedded in their mortgages. For example, a homeowner can have the right to repay the mortgage early and refinance at a lower rate. (We will discuss this option in some detail in Chapter 18.)

The underlyings in derivatives are often financial variables (e.g., interest rates, exchange rates, and stock prices). However, almost every observable variable can be an underlying for a derivative. Examples include

- The price of hogs,
- The price of electricity in a particular region,
- The amount of snow falling at a certain ski resort,
- The temperature at a weather station,
- Earthquake damage claims made by an insurance company's policyholders.

Derivatives can be used for either hedging or speculation. If a trader has a particular exposure, a derivatives trade can reduce that exposure. If the trader has no exposure, however, that same trade is speculative. As we will see in this chapter, speculation can be extremely risky.

Derivatives have been criticized for their role in the 2007–2008 credit crisis. In the years leading up to the crisis, banks in the U.S. relaxed their lending standards on mortgages. This created a huge number of *subprime mortgages* (i.e., mortgages to less creditworthy borrowers). Furthermore, banks did not simply keep these mortgages on their balance sheets. Instead, they bundled them into portfolios and created complex derivatives whose values depended on the losses from defaults on these mortgages.

The increased availability of mortgages led to a significant increase in demand for housing that in turn led to a sharp increase in house prices. Meanwhile, defaults and foreclosures began to rise as many subprime borrowers realized they could not afford their mortgages. This led to an increase in the supply of houses for sale and a reduction in housing prices. This in turn led to negative equity positions (i.e., situations where the amount owed on a mortgage was greater than the value of the house).

Some homeowners with negative equity defaulted even though they could afford to service their mortgages, creating more foreclosures and depressing house prices even further. As a result, there were losses on many of the derivatives created from mortgages, and investors throughout the world (temporarily) lost their appetite for risky debt of any sort. As a result, the world was plunged into the worst recession in 75 years.

Of course, derivatives have many attractive features for society. They allow risks to be transferred from one party to another in ways that benefit both sides. The following are examples.

- Corporate treasurers can manage exchange rate risk, interest rate risk, and commodity price risk in ways that would otherwise not be possible.
- Fund managers can diversify their exposures using derivatives.
- Ski slope operators can avoid being forced out of business due to a single unseasonably warm winter.

The challenge for regulators is to find ways to benefit from derivatives while discouraging situations where huge risks are taken or unnecessarily complex instruments are created.

In this chapter, we take a first look at derivatives markets by examining futures, forwards, and options. We also describe the types of trades used by hedgers, speculators, and arbitrageurs. Later chapters will go into more detail on the markets and the ways these financial instruments are traded.

4.1 THE MARKETS

Derivatives trade on exchanges as well as in over-the-counter markets. An exchange is a market where investors trade standardized contracts that have been defined by the exchange. Over-the-counter markets are markets where participants contact each other directly (or possibly by using a broker as an intermediary) to trade. Chapters 5 and 6 describe the operation of the two markets in some detail. Here we provide a brief overview.

Exchange-Traded Markets

Derivatives exchanges have existed for many years. For example, the Chicago Board of Trade (CBOT) was established in 1848 to allow farmers to trade with merchants. Within a few years, a forerunner of futures contracts known as "to-arrive" contracts began to be traded. There are now futures exchanges in many parts of the world.

In 1973, the CBOT launched the Chicago Board Options Exchange (CBOE).[1] The CBOE established well-defined contracts and a mechanism to minimize the probability of losses from defaults on the contracts. Today, the CBOE is one of many exchanges around the world trading options on stocks and stock indices.

[1] At first, the exchange only featured call options and it did not trade put options until 1977. Call and put options will be formally defined later in this chapter.

The year 1973 also saw the publication of the famous Black-Scholes-Merton options valuation model.[2] Within a short period of time, special-purpose calculators were developed and put into use by options traders.

Traditionally, derivatives exchanges have used what is referred to as the *open-outcry* system. This involves traders meeting on the floor of the exchange and indicating their proposed trades with hand signals and shouting. (Tall traders may have had an advantage because it was easier for them to attract the attention of other traders.) Most trading is now done electronically, however, with computers being used to match buyers and sellers. Sometimes, electronic trading is initiated by computer algorithms without any human intervention at all.[3]

Over-the-Counter Markets

An advantage of over-the-counter (OTC) markets is that the contracts traded do not have to be the standard contracts defined by exchanges: market participants can agree to any contract they like.

OTC market participants can be categorized as either end users or dealers. End users are corporations, fund managers, and other financial institutions who use derivatives to manage their risks or to acquire specific exposures. Dealers are large financial institutions that provide both bid and ask quotes for commonly traded derivatives. They are also prepared to make one-sided quotes for highly structured derivatives when requested. Dealers typically offset the risks from their trades with end users by trading with other dealers in what is referred to as the *interdealer market*.

While end users typically contact dealers directly, dealers often use *interdealer brokers* when trading with other dealers. The advantage of an interdealer broker is that a dealer does not have to indicate a desired trade to other dealers.[4] When finding a dealer to be a counterparty to its client, the broker does not reveal its client's name until the trade has been finalized.

Before the 2007–2008 credit crisis, the over-the-counter market was largely unregulated. Since the crisis there have been several new regulations. The following are examples.

[2] See F. Black and M. Scholes, "The pricing of options and corporate liabilities," Journal of Political Economy, 81 (May/June 1973): 637–659; R. C. Merton, "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4 (Spring 1973): 141–183.

[3] This has not been without its problems. For example, the flash crash in 2010, where markets declined by about 7% in 15 minutes before rebounding, was partly the result of the unintended consequences of a trading algorithm.

[4] This can be important. If it is known that a dealer must get rid of a large exposure, other traders may trade ahead of the dealer because they know that the dealer's trade will potentially move the market price. A good interdealer broker will try to feed the dealer's trades into the market anonymously.

- In the U.S., standardized OTC derivatives traded between dealers must (whenever possible) be traded on platforms known as swap *execution facilities*. These are like exchanges and feature market participants posting bid and ask prices.
- A central counterparty (CCP) must be used for standardized transactions between dealers. (CCPs are discussed in the next two chapters.)
- All trades must be reported to a central registry. (Previously, trades in the OTC market were considered private transactions, and there were no reporting requirements.)

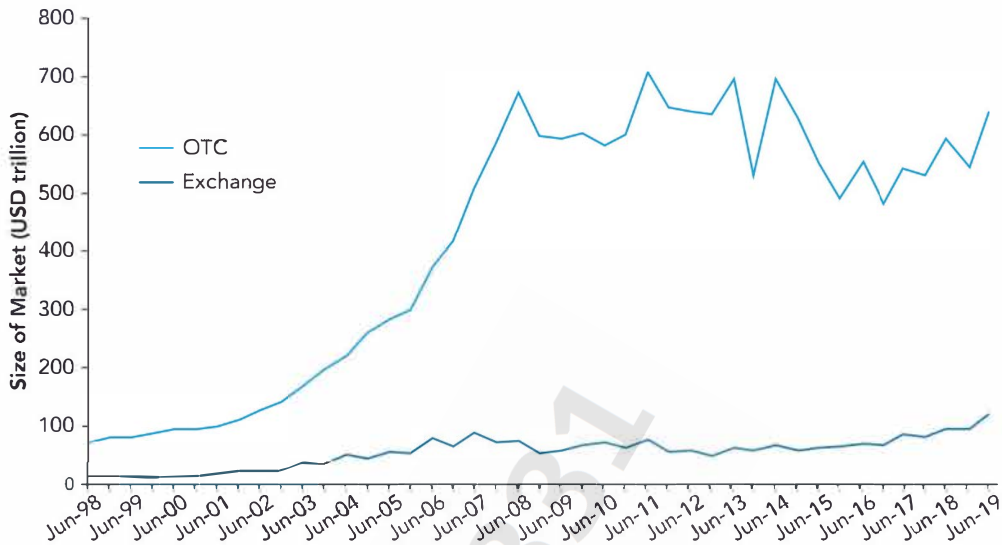


Figure 4.1 **Size of Exchange-traded and OTC market measured in terms of the value of underlying assets between 1998 to 2019.**

Source: https://stats.bis.org/statx/toc/DER.html

Market Size

The Bank for International Settlements began collecting data on derivatives markets in 1998. Note that statistics for the exchange-traded market show the value of the assets underlying outstanding exchange-traded contracts, while statistics for the OTC market show the total principal underlying outstanding transactions. Table 4.1 shows how the two markets have grown between June 1998 and June 2019.

The statistics indicate that both markets grew by a factor of about 9 between 1998 and 2019. To put the numbers in Table 4.1 in perspective, note that the world's gross domestic product (GDP) in 2018 was about USD 85 trillion. This means that the value of the assets underlying outstanding derivatives contracts (exchange-traded plus OTC) is about nine times the world's GDP.

Figure 4.1 illustrates how the sizes of the two markets have changed between 1998 and 2019. In interpreting Figure 4.1 and the other statistics presented, it is important to emphasize that the size statistics do not provide the value of the transactions. In the case of the exchange-traded market, they measure the value of the underlying assets. If an option gives the holder the right to purchase 100 shares worth USD 40 per share for USD 45 per share, for example, the size of the contract would be recorded as USD 4,000 (= 100 × 40). In the case of the OTC market, the statistics measure the principal underlying outstanding transactions. This means that a forward contract to buy 1 million British pounds at an exchange rate of USD 1.2500 in the future would be recorded at the current value of 1 million British pounds and not at the value of the forward contract (which might be only a few thousand dollars).

The decline in the size of the OTC market between 2013 and 2015 can be attributed to the widespread use of *compression*. This is a procedure where two or more market participants restructure transactions with the result being that their underlying principal (and therefore the amount of capital they are required to keep) is reduced. Compression will be explained in the next two chapters.

Table 4.1 **Value of Underlying Assets in Derivatives Markets in June 1998 and June 2019 (Trillions of USD)**

	Exchange-Traded Market	Over-the-Counter (OTC) Market
June 1998	13.3	72.1
June 2019	120.3	640.4

Source: https://stats.bis.org/statx/toc/DER.html

4.2 FORWARD CONTRACTS

A forward contract is an over-the-counter contract where one party agrees to buy an asset for a predetermined price at a future time and the other party agrees to sell the asset for the predetermined price at the future time. Forward contracts can be contrasted with spot contracts, which are agreements to buy or sell an asset almost immediately.[5]

[5] Settlement may take a day or two in a spot contract.

The party that has agreed to buy has a *long forward position* while the party that has agreed to sell has a *short forward position*. The specified asset price in a forward contract is referred to as the *forward price*. The determination of forward prices will be discussed in later chapters.

Forward contracts on foreign currency are very popular. If Company A knows it will receive a certain amount of a foreign currency at a certain future time, it can use a forward contract to lock in the exchange rate by selling the foreign currency at the forward exchange rate. Similarly, if Company B knows that it will pay a certain amount of a foreign currency at a certain future time, it can enter into a long forward contract to buy the foreign currency at the forward exchange rate.

Table 4.2 shows quotes for spot and forward contracts on the USD–euro exchange rate (as they might have been made by a derivatives dealer) on December 23, 2019. The bid price is the price at which the dealer is prepared to buy and the ask price is the price at which the dealer is prepared to sell. The quotes show that the dealer is prepared to pay USD 1.1090 per euro and sell at a price of USD 1.1093 per euro in the spot market. The second row shows that the dealer is prepared to commit to pay USD 1.1117 per euro in one month or sell at a price of USD 1.1121 per euro in one month. The other rows provide quotes for buying and selling in three and six months. Note that forward prices are not equal to spot prices. We will explain how they are calculated in a later chapter.

Earlier, we mentioned the distinction between linear and nonlinear derivatives. Forward contracts are linear derivatives because their payoff is linearly related to the value of the underlying asset at maturity.

Suppose that Company A enters into a long forward contract to buy 1 million euros in six months. From Table 4.2, the exchange rate will be USD 1.1129. Suppose further that Company B enters into a short forward contract to sell 1 million euros in six months. Table 4.2 shows that the exchange rate applicable to this trade will be USD 1.1124. The payoff from the contracts depends on the actual exchange rate in six months.

Table 4.2 Spot and Forward Quotes on the USD–Euro Exchange Rate

	Bid	Ask
Spot	1.1090	1.1093
One-Month Forward	1.1117	1.1121
Three-Month Forward	1.1159	1.1164
Six-Month Forward	1.1124	1.1129

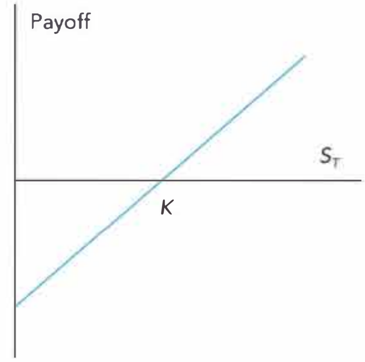


Figure 4.2 Payoff from a long position in a forward contract. The delivery price is K and the value of the asset at maturity is S_T.

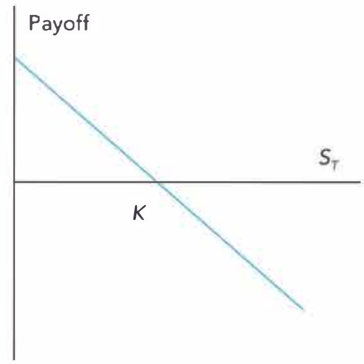


Figure 4.3 Payoffs from a short position in a forward contract. The delivery price is K and the value of the asset at maturity is S_T.

If the (mid between bid and ask) exchange rate is USD 1.1500, for example, Company A buys 1 million euros for USD 1.1129 when one euro is worth USD 1.1500. This leads to a gain of:

$$1{,}000{,}000 \times (1.1500 - 1.1129) = 37{,}100$$

Meanwhile, Company B sells 1 million euros for USD 1.1124 when they are worth USD 1.1500. This leads to a payoff of:

$$1{,}000{,}000 \times (1.1124 - 1.1500) = -37{,}600$$

In general, suppose that S_T is the asset price at the maturity of a forward contract and K is the delivery price (i.e., the forward price when the contract was initiated).[6] The payoff from a long forward contract on one unit of the asset is

$$S_T - K$$

This is shown in Figure 4.2. The payoff from a short forward contract on one unit of the asset is:

$$K - S_T$$

This is shown in Figure 4.3.

[6] K is the forward price at the time the contract is entered but is not necessarily the forward price for the contract at future times.

4.3 FUTURES CONTRACTS

As mentioned, forward contracts are traded in the over-the-counter market. A futures contract provides a similar payoff to a forward contract, but it trades on an exchange. The exchange defines the asset and specifies the maturity dates that can be traded. As we will see in later chapters, the exchange organizes trading so that there is very little credit risk (i.e., risk that the agreement will not be honored) even though the two parties to a trade may not know each other.

Forward contracts are most popular on exchange rates and interest rates, whereas futures trade on a much wider range of underlyings than forwards. These underlyings include

- The prices of agricultural products such as corn, wheat, and live cattle;
- The prices of metals such as gold, silver, copper, and platinum;
- Equity indices such as the S&P 500 and the NASDAQ 100;
- The prices of energy products such as oil, natural gas, and electricity;
- Real estate indices;
- Temperatures in particular cities; and
- Cryptocurrencies such as bitcoin.

Futures contracts are covered further in later chapters.

4.4 OPTIONS

Options are derivatives that give the holder the right (but not the obligation) to buy or sell an asset at a predetermined price in the future. They trade on exchanges as well as in the over-the-counter market.

To see how options work, note that there are two sides to every options contract. In the case of a call option, the party with a long position has the right (but not the obligation) to buy an asset from the party with a short position for a certain price (known as the *strike price* or *exercise price*) at one or more future times. If the party exercises this right, the party with the short position must sell the asset for the strike price.

A put option is a contract where the party with a long position has the right (but not the obligation) to sell an asset to the party with a short position for a certain price (the strike price) in the future.

The date specified in an options contract is known as the expiration date (or maturity date). A *European option* can only be exercised at expiration. An *American option* can be exercised at any time up until expiration.

While it costs nothing to enter into a forward contract, an option has a price (known as the premium) to be paid at the outset. Table 4.3 shows price quotes provided by the CBOE for call options on IBM on December 23, 2019. Table 4.4 does the same for put options on IBM. The options on stocks traded by exchanges are American. Options trade with several different future expiration dates.[7] The price of IBM stock at the time of the quotes was bid USD 135.66 and ask USD 135.68.

We see from the tables that call option prices decrease as the strike price increases, whereas put option prices increase as the strike price increases. The tables also show that an option price increases as the time to maturity increases. Bid-ask spreads (particularly if they are expressed as a proportion of the price) are much higher for options on stocks than they are for the stocks themselves.

Consider a European call option that can be exercised at time T. Suppose that K is the strike price and S_T is the option price at time T. Consider first the position of the trader that has bought the option (i.e., has a long position in the call option). If $S_T > K$, the trader exercises the option. This means that he or she pays K for an asset that can be immediately sold for S_T. The payoff to the trader is therefore $S_T - K$. If $S_T < K$, the option is not exercised and the payoff to the trader is zero. Putting the $S_T > K$ and $S_T < K$ outcomes together, we see that the option payoff is

$$\max(S_T - K, 0)$$

This payoff is illustrated in Figure 4.4.

To the trader with a short position (i.e., the trader who has sold the option), the payoff is

$$-\max(S_T - K, 0)$$

If $S_T > K$, this trader must sell an asset worth S_T for K. If $S_T < K$, the trader is not required to make the sale. This payoff is shown in Figure 4.5.

Next consider a European put option with strike price of K and asset price at maturity of S_T. In this case, the trader who has bought the option has the right to sell the asset for K. If $S_T < K$, the trader will exercise this right, and an asset worth S_T is then sold for K. This creates a payoff of $K - S_T$. If $S_T > K$, the option is not exercised, and there is no payoff to the option owner. The payoff to the option owner is therefore

$$\max(K - S_T, 0)$$

[7] The expiration dates in Tables 4.3 and 4.4 are the third Friday of the month. The exchange has introduced *weeklys* and *quarterlys* that expire at other times. These will be discussed in later chapters.

Table 4.3 Call Option Premiums on IBM on December 23, 2019, with Different Expiration Dates

	March 20, 2020		June 19,2020		December 18, 2020	
Strike Price	**Bid**	**Ask**	**Bid**	**Ask**	**Bid**	**Ask**
125	11.95	12.30	13.15	13.80	14.90	16.05
130	8.20	8.40	9.85	10.10	11.65	13.25
135	5.05	5.15	6.95	7.20	9.35	10.20
140	2.70	2.78	4.60	4.80	7.55	7.90
145	1.26	1.32	2.91	3.05	5.45	6.35

Table 4.4 Put Option Premiums on IBM on December 23, 2019, with Different Expiration Dates

	March 20, 2020		June 19,2020		December 18, 2020	
Strike Price	**Bid**	**Ask**	**Bid**	**Ask**	**Bid**	**Ask**
125	1.86	1.91	4.05	4.20	7.95	8.90
130	3.10	3.20	5.70	5.85	10.10	10.70
135	5.05	5.20	7.90	8.10	12.30	13.35
140	7.85	8.00	10.65	10.90	15.10	16.45
145	11.50	11.70	13.95	14.20	18.20	19.65

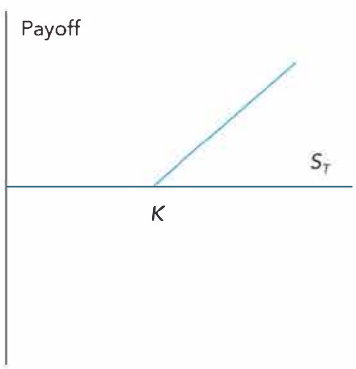


Figure 4.4 Payoff to trader who has a bought a European call option. K is the strike price and S_T is the asset price on the expiration date.

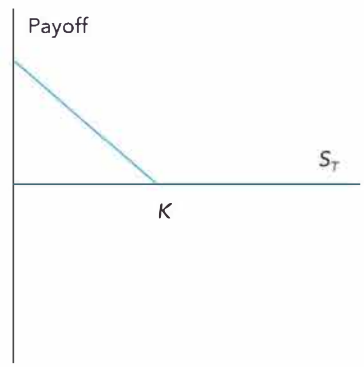


Figure 4.6 Payoff to trader who has a bought a European put option. K is the strike price and S_T is the asset price on the expiration date.

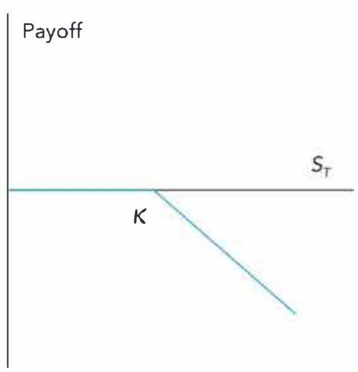


Figure 4.5 Payoff to trader who has sold a European call option. K is the strike price and S_T is the asset price on the expiration date.

This payoff is shown in Figure 4.6.

To the trader with a short position (i.e., the trader who has sold the option), the payoff is

$$-\max(K - S_T, 0)$$

If $S_T < K$, the trader (who is short the put option) must buy an asset worth S_T for K. If $S_T > K$, and the trader is not required to buy the asset. This payoff is shown in Figure 4.7.

Options are more complex derivatives than forward or futures contracts. As Figures 4.4 to 4.7 show, their payoffs are non-linear functions of the underlying asset price. This has two consequences.

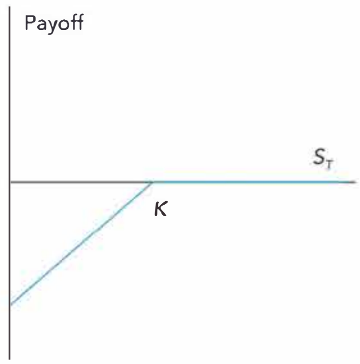


Figure 4.7 **Payoff to trader who has sold a European put option. K is the strike price and S_T is the asset price on the expiration date.**

1. The value of an option is a non-linear function of the value of the underlying.
2. The value of an option is dependent on the volatility of the underlying.

We will discuss the properties of options in more detail later in this book. The *Valuation and Risk Models* book covers the way options are valued and hedged.

4.5 MARKET PARTICIPANTS

There are three main categories of traders in derivatives markets.

1. Hedgers,
2. Speculators, and
3. Arbitrageurs.

Hedgers

Hedgers use derivatives to reduce or eliminate risk exposure. We have already seen how traders can use forward contracts on a foreign currency to manage foreign exchange risk. Options can also be used to hedge currency risk. Unlike forward contracts, options allow traders get downside risk protection while preserving some upside potential.

As an example, consider a U.S. corporate treasurer due to pay 1 million euros in six months. The treasurer could buy a European call option to buy 1 million euros in six months with a strike price of USD 1.1150. If the exchange rate in six months is greater than 1.1150, the treasurer will exercise the option and acquire 1 million euros at an exchange rate of USD 1.1150. If the exchange rate is less than USD 1.1150, the treasurer does not exercise the option and is instead able to buy the euros at a more favorable exchange rate than 1.1150.

Next, consider the treasurer who is due to receive euros in six months. This treasurer could buy a European put option to sell 1 million euros at an exchange rate of USD 1.1150. If the exchange rate in six months' time is less than this, the treasurer exercises the option and sells the euros that are received for USD 1.115 million. If the exchange rate in six months is greater than this, the option is not exercised. However, the treasurer can benefit from a better exchange rate than 1.1150 when the 1 million euros are sold.

These examples show that while a forward contract locks in the price applicable to a future transaction, an option provides protection against adverse price movements.

It is important to note, however, that it does not cost anything (except for the bid-ask spread) to lock in the forward price. By contrast, an option requires that the buyer pay a premium. In our example, the call (put) option for 1 million euros with a strike price of USD 1.1150 might cost USD 40,000. While the treasurer could reduce the cost by increasing (reducing) the strike price, this would lower the protection as well.

Speculators

Derivatives also allow risks to be taken with a relatively small upfront payment. In this sense, they have much the same effect as leverage and can be attractive to speculators.

To illustrate this point, suppose a stock is currently worth USD 40 and a trader is convinced that the price will increase in the next three months. Furthermore, a call option on the stock with a strike price of USD 42 has a price of USD 2. One strategy would be to buy 100 shares of the stock for USD 4,000. An alternative strategy would be to use the USD 4,000 to buy 2,000 options.

Both strategies involve an initial investment of USD 4,000. As indicated in Table 4.5, however, they provide very different outcomes.

Consider the outcomes where the share price is USD 45 at the end of the three months. The results from using the two speculative strategies would be as follows.

1. *Share purchase strategy*: There is a USD 5 profit on each share purchased. The total profit is USD 500 $= ((45 - 40) \times 100)$.
2. *Option strategy*: The options strategy allows 2,000 shares worth USD 45 to be purchased for USD 42. This generates a profit of USD 6000 ($= 2000 \times (45 - 42)$). When the initial cost of the options is accounted for, the profit is USD 2,000. (We ignore the impact of discounting.)

Table 4.5 Profits from Two Alternative Strategies that are Speculating that the Price of an Asset will Increase

Asset Price in Three Months (USD)	Profit from Asset Purchase (USD)	Profit from Options (USD)
30	−1,000	−4,000
35	−500	−4,000
40	0	−4,000
45	500	+2,000
50	1,000	+12,000

Table 4.5 indicates that the option strategy is four times as profitable as the share purchase strategy if the share price proves to be USD 45 and is twelve times as profitable if the share price ends up being USD 50. However, the option strategy leads to a loss of USD 4,000 if the share price proves to be less or equal to than USD 42. In these scenarios, the option is not exercised, and there is no payoff to offset the USD 4,000 option premium. In contrast, the share purchase strategy involves no option premium and thus produces a profit (or loss) that is solely dependent on the difference between the share price at purchase and the share price in three months.

Arbitrageurs

Arbitrage involves taking advantage of inconsistent pricing across two or more markets. For example, suppose that an asset that provides no income is priced at USD 50, the borrowing rate is 3%, and the asset's one-year forward price is USD 52. An investor can

- Borrow USD 50 million for one year at 3% to buy 1 million units of the asset for USD 50 per unit,
- Enter into a forward contract to sell 1 million units of the asset for USD 52 per unit in one year, and
- Repay the loan in one year at a cost of USD 51,500,000 (= 50,000,000 × 1.03) and sell the 1 million units of the asset for USD 52,000,000.

This is arbitrage leads to a profit of USD 500,000.

4.6 DERIVATIVES RISKS

We have just shown that derivatives markets attract many kinds of traders. This is one of their strengths and a major reason for their success. However, the leverage that speculators can obtain means that it is very easy for traders to take significant risks.

A key problem is that the rewards from successful speculation are very high and many traders are tempted to speculate even when their mandate is to hedge risks or to search for arbitrage opportunities. If controls are not in place, these traders may start speculating without the knowledge of others in their organization.

A speculating trader who loses money may seek to offset losses by taking increasingly large and risky positions. If the offsetting gains do not materialize, the trader may increase the risks taken again and again until the losses reach catastrophic levels. Examples of situations where this has occurred include:

- John Rusnak at Allied Irish Bank lost USD 700 million trading in foreign currencies and managed to conceal his losses by creating fictitious option trades;
- Nick Leeson at Barings Bank lost about USD 1 billion by making unauthorized large bets on the future direction of the Nikkei index and managed to hide his losses from his superiors for some time;
- Jérôme Kerviel at Société Générale lost about USD 7 billion by speculating on equity indices while giving the appearance of being an arbitrageur;
- Kweku Adoboli at UBS lost about USD 2.3 billion taking unauthorized speculative positions in stock market indices.

SUMMARY

Derivatives trading takes place in both exchange-traded markets and over-the-counter markets. These markets have been introduced in this chapter and are discussed in more detail in the next two chapters.

Forward contracts are over-the-counter agreements where two parties agree to trade a certain asset at a certain price on a given future date. These contracts can be used to lock in the price of an asset that will be bought or sold on the future date. Futures contracts are similar to forward contracts but are traded on exchanges.

Options trade on both exchanges and in over-the-counter markets. While forwards and futures commit a trader to buying or selling an asset for a certain price in the future, options give the holder the right (but not the obligation) to buy or sell at a certain price in the future. Whereas forwards and futures can be used to lock in the price for a future transaction, options can be used to buy protection against unfavorable price movements while allowing the holder to benefit from favorable movements.

Derivatives are versatile instruments that can be used for hedging, speculation, or arbitrage. This is one of their strengths, but it also creates risks. For many traders, speculation is more fun (and potentially more rewarding) than hedging. Unless there are good controls in place, there is the danger of traders engaging in unauthorized speculation (with disastrous results).

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

4.1 What is meant by a linear and a non-linear derivative? Give an example of each.

4.2 What is the difference between a long forward position and a short forward position in an asset?

4.3 What are the two main types of markets for trading derivatives?

4.4 Which of the following is closest to the ratio of the size of the OTC market to size of the exchange traded market in 2019: (a) 2, (b) 4, (c) 5, or (d) 16?

4.5 Which of the following are traded on exchanges: (a) forwards, (b) futures, (c) options?

4.6 What is the payoff from a long position in a call option in terms of the asset price, S_T, and the strike price, K?

4.7 What is the payoff from a short position in a put option in terms of the asset price, S_T, and the strike price, K?

4.8 What are the three types of traders in derivatives markets?

4.9 Traders such as Jérôme Kerviel and Nick Leeson had a mandate to carry out hedging or arbitrage trades. How did they lose billions of dollars?

4.10 "Options provide leverage." Explain this statement.

Practice Questions

4.11 A trader enters into a short futures contract to sell 100 units of an asset for USD 50. What is the trader's gain or loss if the price of the asset at maturity is (a) USD 55 and (b) USD 48?

4.12 For a corporate treasurer wanting to hedge exchange rate risk, what is an advantage of the OTC market over the exchange-traded market?

4.13 What is the difference between selling a call option and buying a put option?

4.14 A trader buys a call option on a stock with a strike price of USD 50 when the stock price is USD 49. The cost of the option is USD 2. Under what circumstances does the trader make a profit? (Ignore the impact of discounting.)

4.15 A trader sells a put option on a stock with a strike price of USD 50 when the stock price is USD 51. The price of the option is USD 3. Under what circumstances does the trader make a profit? (Ignore the impact of discounting.)

4.16 Why does compression reduce the size of the OTC markets as measured by the Bank for International Settlements?

4.17 What is the payoff from a portfolio consisting of a short forward contract with maturity T and a long call option with maturity T? Assume that the strike price for the option is the forward price.

4.18 What is the payoff from a portfolio consisting of a short forward contract with maturity T and a short put option with maturity T? Assume that the strike price for the option is the forward price.

4.19 A trader thinks that the price of a stock, which is currently USD 70, will increase. The trader is trying to choose between buying 100 shares and buying European call options on 1,000 shares. The options cost USD 7 per option and have a strike price of USD 70 with a maturity of six months. Explain the difference between the two trading strategies.

4.20 In Question 4.19, what is the breakeven stock price for which the two strategies give the same result? (Ignore the impact of discounting.)

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

4.1 A linear derivative is a derivative where the payoff at a future time is linearly dependent on the value of the underlying at that time. A non-linear derivative is a derivative where the payoff is a non-linear function of the value of the underlying. A forward contract is a linear derivative. An option is a non-linear derivative.

4.2 A long forward position is a position where a trader has agreed to buy an asset at a certain price at a certain future time. A short forward position is a position where a trader has agreed to sell an asset at a certain price at a certain future time.

4.3 Exchanges and the over-the-counter markets

4.4 (c). The OTC market was about five times as big as the exchange-traded market in June 2019.

4.5 Futures and options (Options are also traded in the OTC market.)

4.6 $\max(S_T - K, 0)$

4.7 $-\max(K - S_T, 0)$

4.8 Hedgers, speculators, and arbitrageurs

4.9 They switched from being hedgers/arbitrageurs to being speculators without anyone realizing that this had happened.

4.10 Buying a call option on an asset is like borrowing money to buy the asset, in that it allows big risks to be taken with a small initial investment. The gains and losses are accentuated.

Solved Problems

4.11 (a) Trader loses (in USD) $100 \times (55 - 50) = 500$.

(b) Trader gains (in USD) $100 \times (50 - 48) = 200$.

4.12 The transaction does not need to have the standard features determined by the exchange. The maturity date and (in the case of options) the strike price can be negotiated to meet the treasurer's precise needs.

4.13 When a trader sells a call option, the trader must sell the asset when the asset's price is greater than the strike price. When a trader buys a put option, the trader has the option to sell the asset when the asset price is less than the strike price.

4.14 The payoff is $\max(S_T - 50{,}0)$. The trader will make profit when:

$$S_T - 50 > 2 \text{ or } S_T > 52$$

4.15 The payoff is $-\max(50 - S_T{,}0)$. The trader will make profit when:

$$50 - S_T < 3 \text{ or } S_T > 47$$

4.16 Compression reorganizes the trades between several market participants, with the result being that the underlying principal is reduced. The BIS measures the size of the market as the total outstanding underlying principal.

4.17 With our usual notation, the payoff is

$$K - S_T + \max(S_T - K{,}0) = \max(0{,}K - S_T)$$

To make sure you understand why this is true, you should consider the $S_T > K$ and the $S_T < K$ cases separately. The payoff is the payoff from a long put option with a strike price equal to the forward price.

	$S_T > K$	$S_T < K$
(a) Long Call	$S_T - K$	0
(b) Short Forward	$K - S_T$	$K - S_T$
Net (a) + (b)	0	$K - S_T$

4.18 With our usual notation, the payoff is

$$K - S_T - \max(K - S_T{,}0) = -\max(0{,}S_T - K)$$

To make sure you understand why this is true, you should consider the $S_T < K$ and the $S_T > K$ cases separately. The payoff is the payoff from a short call option with a strike price equal to the forward price.

	$S_T > K$	$S_T < K$
(c) Short Put	0	$-(K - S_T)$
(d) Short Forward	$K - S_T$	$K - S_T$
Net (a) + (b)	$K - S_T$	0

Note that in this case, $S_T > K$ implies that the short forward is losing money. Thus, the solution is $-\max(S_T - K{,}0)$ that is the same as $\min(K - S_T{,}0)$, because $K - S_T < 0$.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

4.19 The option trading strategy is considerably riskier as indicated by the following table:

Stock Price (USD)	Profit from Buying Stock (USD)	Profit from Buying Stock Options (USD)
50	−2,000	−7,000
55	−1,500	−7,000
60	−1,000	−7,000
65	−500	−7,000
70	0	−7,000
75	+500	−2,000
80	+1,000	+3,000
85	+1,500	+8,000
90	+2,000	13,000

4.20 We require

$$100(S_T - 70) = 1{,}000(S_T - 70) - 7{,}000$$

or

$$900S_T = 70{,}000$$

The breakeven stock price is therefore 70,000/900 or 77.78.

5

Exchanges and OTC Markets

Learning Objectives

After completing this reading, you should be able to:

- Describe how exchanges can be used to alleviate counterparty risk.
- Explain the developments in clearing that reduce risk.
- Define netting and describe a netting process.
- Describe the implementation of a margining process, explain the determinants of and calculate initial and variation margin requirements.
- Describe the process of buying stock on margin without using CCP and calculate margin requirements.
- Compare exchange-traded and OTC markets and describe their uses.
- Identify risks associated with OTC markets and explain how these risks can be mitigated.
- Describe the role of collateralization in the OTC market and compare it to the margining system.
- Explain the use of special purpose vehicles (SPVs) in the OTC derivatives market.

This chapter examines how derivatives trading is organized. As mentioned in the previous chapter, derivatives trading takes place on exchanges and in the over-the-counter (OTC) markets. This chapter is devoted to explaining how both these markets have operated in the past and how they operate today.

This chapter also introduces central counterparties (CCPs) in the context of exchange-traded markets. In the next chapter, we explain how they operate in the OTC market and discuss the potential risks associated with them.

5.1 EXCHANGES

Exchanges have existed for many years. The contracts traded when exchanges were first set up were precursors of the futures contracts trading today.

An exchange is an organization with members who trade with each other. Originally, the main role of an exchange was to simply provide a forum where members could meet and agree to trades. These early exchanges also defined standard contracts, resolved disputes between members, and expelled members who reneged on agreed transactions. However, they provided few other services. For example, there were no mechanisms in place to protect members from losses associated with counterparty defaults.

A natural development was for members to protect themselves by requiring margin. *Margin* refers to the assets transferred from one trader to another for protection against counterparty default.

Consider Traders A and B, who have entered a trade where Trader A agrees to buy 10,000 bushels of corn from Trader B in September at a price of 400 cents per bushel in September. Suppose further that the price of corn then declines by 5 cents soon after the agreement is made.[1] If there is a margin agreement in place, Trader B can request USD 500 (= 10,000 × 5 cents) from Trader A. If the price subsequently declined by a further 10 cents per bushel, a further USD 1,000 would be requested, and so on. If the price of corn moved in the other direction, however, then Trader B would be required to provide margin to Trader A. For example, a 20-cent increase in the price of corn would lead to a cumulative transfer of USD 2,000 from Trader B to Trader A.

In both cases, margin transfers get rid of the incentive for a party to back out of the transaction to take advantage of more favorable market prices in the market.

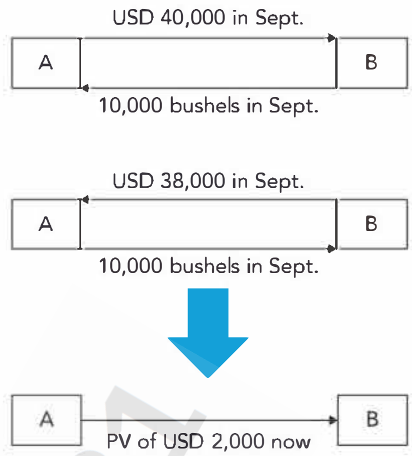


Figure 5.1 Trader A and Trader B net two offsetting contracts for the future delivery of corn.

Another development to protect members from losses was netting.[2] Netting is a procedure where short positions and long positions in a particular contract offset each other. For example, if Trader A (from the prior example) subsequently agrees to sell 10,000 bushels of corn for September delivery to Trader B for 380 cents per bushel, the contracts could be netted by Trader A paying USD 2,000 to Trader B in September. Alternatively, the present value of USD 2,000 could be paid at the starting date of the second contract. The key point is that once the two traders have entered into offsetting contracts, there is no need for corn to be exchanged in September. This is illustrated in Figure 5.1.

Other netting arrangements are more complicated. For example, suppose that:

- Trader A agrees to buy 10,000 bushels of corn from Trader B for 400 cents per bushel, and
- Trader B agrees to buy 10,000 bushels of corn from Trader C for 400 cents per bushel.

Suppose further that both contracts having the same delivery date. The contracts can then be collapsed into a single contract where Trader A agrees to buy 10,000 bushels of corn from Trader C for 400 cents per bushel. This is illustrated in Figure 5.2.

If the contract between Trader B and Trader C is for 380 cents per bushel, rather than 400 cents per bushel, we can use a combination of Figure 5.1 and Figure 5.2 to collapse the two contracts into a single contract in one of two ways.

[1] As indicated later, traders should base the calculation of margin on the price for future delivery in September, rather than at the current (spot) price.

[2] The standardization of contracts in terms of size, quality of the commodity, and delivery dates facilitated netting.

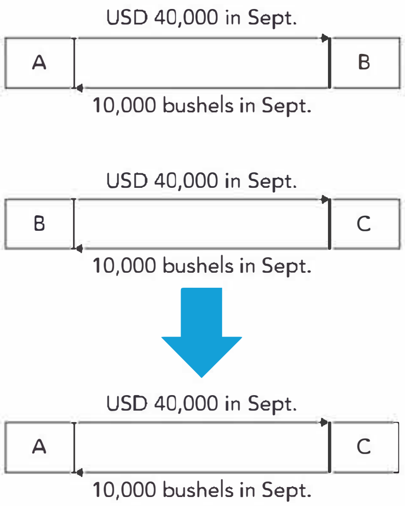


Figure 5.2 **Netting of contracts when more than two traders are involved.**

1. Trader A agrees to buy 10,000 bushels of corn from Trader C at 380 cents per bushel and agrees to make a payment equal to the present value of USD 2,000 to Trader B.
2. Trader A agrees to buy 10,000 bushels of corn from Trader C at 400 cents per bushel, and Trader C agrees to make a payment of the present value of USD 2,000 to Trader B.

One issue arising from netting arrangements involving more than two market participants (such as that in Figure 5.2) is that the parties involved may have different credit qualities. For example, Trader C may be wary about changing a contract with Trader B for one with Trader A if Trader A is more likely to default. This is where the margin arrangements mentioned earlier are relevant. If Trader A agrees to post margin in the event that the price of corn declines, Trader C's credit exposure to Trader A would be reduced. Note that margin already posted by Trader A with Trader B would have to be transferred to Trader C at the time of the netting.

This can get quite complicated, and a natural market development was for exchanges to handle margin arrangements so that traders did not need to worry about the credit quality of other traders.

Central Counterparties

Exchanges today are more heavily involved in organizing trading than they were in the past. Specifically, they operate clearing houses known as central counterparties (CCPs) that clear all transactions between members.

For example, assume that Member A agrees to buy 5,000 bushels of corn (which is defined by the Chicago Mercantile Exchange as one contract) from Member B for delivery in September at 400 cents per bushel (USD 20,000 in total). The exchange (through its CCP) then becomes the counterparty to both members. What this means is that CCP agrees to buy 5,000 bushels of corn from Member B at 400 cents per bushel while Member A agrees to buy 5,000 bushels of corn from the CCP at 400 cents per bushel.

Thus, when Member A and Member B agree on a certain transaction, the exchange stands between them as illustrated in Figure 5.3.

The key point is that Member A no longer needs to worry about the creditworthiness of Member B (and vice versa). Indeed, the two members might agree on a trade (either on the floor of the exchange or electronically) without even knowing each other. The CCP becomes the counterparty to both and is a clearing house for all transactions.

Another advantage of CCPs is that it is much easier for exchange members to close out positions. To see how this works, note that Member A has a long position in 5,000 bushels (one contract) of September corn in Figure 5.3. (*September corn* means corn that will be delivered in September.) If Member A decides to close out this position, he or she could agree to sell 5,000 bushels of September corn to any other member of the exchange. That trade will also be transferred to the CCP and Member A would then have two trades with the CCP that offset each other. Without a CCP, Member A would either have to approach Member B to close out the position (and there is no guarantee that Member B will be interested in doing this) or short 5,000 bushels of corn with another member. In the latter case, Member A would have to worry about the creditworthiness of the two members of the exchange that he or she has traded with.

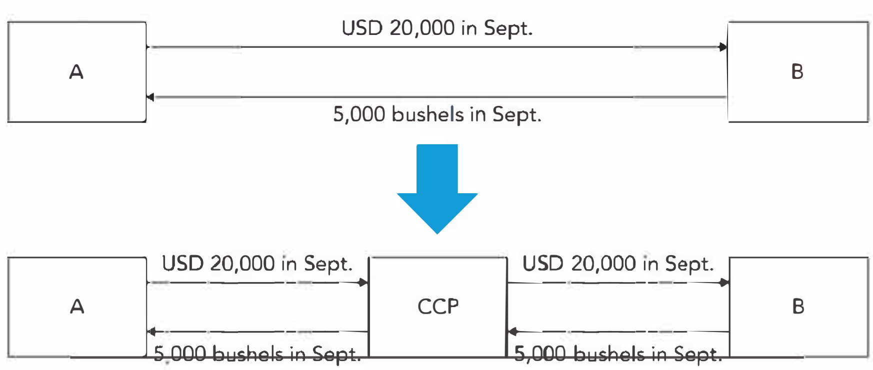


Figure 5.3 **The exchange clearing house positions itself between two of its members and becomes the counterparty to each member.**

5.2 HOW CCPs HANDLE CREDIT RISK

Once an exchange has decided to establish a CCP, it must find a way of managing the associated credit risk. It can do this with a combination of the following:

- Netting,
- Variation margin and daily settlement,
- Initial margin, and/or
- Default fund contributions.

Netting

As mentioned earlier, netting means that long and short positions are combined to determine a CCP's net exposure to a member. For example, suppose that Member X shorts one September corn contract. If it enters into a trade to buy four September corn contracts, this will also become a trade between Member X and the CCP. These two trades would be collapsed to a net long position of three September corn contracts (i.e., contracts to buy 15,000 bushels of corn in September).

Variation Margin and Daily Settlement

A futures contract is not settled at maturity. Rather, it is settled day-by-day during the time to maturity. Consider Trader X from the previous example (who is long three September corn contracts) and suppose that the September futures price is 400 cents per bushel at the close of trading on Day 1 and 395 cents per bushel at the close of Day 2. Trader X has lost

$$15{,}000 \times 5 \text{ cents}$$

or USD 750. This is because September corn is now worth five cents less per bushel than it was worth at the close of trading on Day 1. The trader is thus required to pay USD 750 to the CCP. If September corn is 405 cents per bushel at the close of Day 3, Trader X has gained

$$15{,}000 \times 10 \text{ cents}$$

or USD 1,500. In this case, the exchange pays the trader USD 1,500.

Each day, members who have lost money pay an amount equal to their loss to the exchange CCP, while members who have gained receive an amount equal to their gain from the exchange CCP. These payments are known as *variation margin*, and they typically occur once per day. However, variation margin may be exchanged more often than once a day when markets are highly volatile.

When a futures price for a contract increases from the close of trading on one day to the close of trading on the next day, funds flow through the CCP from members who have net short positions to members who have net long positions. This is indicated in Figure 5.4.

When a futures price for a contract decreases from the close of trading one day to the close of trading the next day, funds flow through the CCP from members who have net long positions to members who have net short positions. This is indicated in Figure 5.5.

The number of long positions always equals the number of short positions. This means that while funds are flowing between the members of CCP, there is no net cash inflow or outflow to the CCP.

Daily settlement has another important advantage: It makes closing out futures contracts much simpler. A member does not have to worry about when a contract was entered or what the futures price was at that time.

For example, suppose that a long contract is closed out at 11 a.m. on a particular day by trading a short contract at 375 cents per bushel. Daily settlement means that the member's gain or loss up to the close of trading on the previous day has already been recognized. If the futures price was 372 cents per bushel at close of trading the previous day, only a further gain of three cents per bushel needs to be added to the member's account as a result of the market movement.

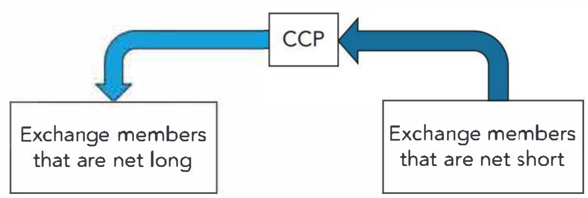


Figure 5.4 **Flow of variation margin for a futures contract when the futures price increases from the close of trading on one day to the close of trading the next day.**

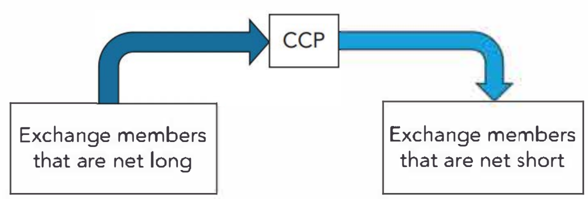


Figure 5.5 **Flow of variation margin for a futures contract when the futures price decreases from the close of trading one day to the close of trading the next day.**

Initial Margin

In addition to variation margin, the CCP requires *initial margin*. These are funds or marketable securities that must be deposited with the CCP in addition to variation margin.

To understand the role of initial margin, suppose that a member who has a net long position in 20 September corn contracts is required to make a variation margin payment of USD 7,000 and fails to do so. The CCP then must close out the member's position by selling 20 September corn contracts. Without initial margin, the CCP would be looking at an immediate loss of USD 7,000 because the variation margin paid to members with short positions would be USD 7,000 higher than that received from members with long positions. Furthermore, the price of September corn could decline by one cent per bushel before the CCP is able to close out the member's position. This would lead to a further loss for the CCP of USD 1,000 (= 0.01 × 5,000 × 20) and increase the total loss to USD 8,000.

Initial margin is designed to prevent these type of losses. In the situation described, the CCP might set the initial margin equal to USD 700 per contract. (The initial margin for a contract is the same regardless of whether the position is short or long). In this situation, the CCP would have required an initial margin of USD 14,000 (= 20 × USD 700). This would be more than enough to cover the USD 8,000 loss.

Note that the initial margin for a futures contract is set by the exchange and reflects the volatility of the futures price. The exchange reserves the right to change the initial margin at any time if market conditions change.

CCPs do not pay interest on variation margin payments because futures contracts are settled daily (and not at maturity). However, CCPs do pay interest on initial margin because it belongs to that member that contributed it.

If the interest rate paid by the CCP is considered unsatisfactory, a member may be able to post securities such as Treasury bills instead of cash. In that case, the CCP would reduce the value of the securities by a certain percentage in determining their cash margin equivalence. This reduction is referred to as a *haircut*. A haircut for a particular asset is usually increased if the price volatility of the asset increases.

Multiple contracts on the same asset can affect variation and initial margin requirements. For example, suppose that a member is long one September corn contract (to buy corn for delivery in September) and short one December corn contract (to sell corn for delivery in December). When the member is required to pay variation margin for the September contract, the member will probably also receive variation margin for the December contract (and vice versa). Therefore, there is an automatic netting of variation margin across contracts. Exchanges also typically have rules that reduce initial margin requirements so that the total initial margin for the long September and short December contract is less than the sum of the initial margins for the two contracts considered separately.

Default Fund Contributions

As a final safety net for the CCP, members are required to make default fund contributions. If the initial margin is not enough to cover losses during a member default, the default fund contributions of that member are used to make up the difference. If those funds are still insufficient, the default fund contributions of other members are used. In the unlikely event that the losses are greater than the sum of the defaulting member's initial margin and the default fund contributions from all members, the equity of the CCP becomes at risk.

5.3 USE OF MARGIN ACCOUNTS IN OTHER SITUATIONS

The margin accounts we have talked about so far are those between CCPs and their members. If a retail trader contacts a broker to do a futures trade, the trader will be required to post margin with the broker. And if the broker is not a member of the CCP, the broker will have to pass the trade to a member, and there will be a margin account kept between the broker and that member.

The margin accounts between retail traders and brokers are somewhat different from the those between CCPs and their members. For instance, there is an initial margin as well as a maintenance margin.[3]

To see how maintenance margin works, note that contracts are settled daily (as it is for members) with gains (losses) being added to (subtracted from) margin accounts. Funds in a margin account in excess of the initial margin requirement can be withdrawn. If the balance in the margin account falls below the maintenance margin level, the trader is required to supply additional margin to bring the account back up to the initial margin level. If the trader does not supply the additional margin, the broker closes out the retail trader's position by entering an offsetting trade on behalf of the trader. The maintenance margin is typically 75% of the initial margin.

[3] The exchange specifies minimum levels for the initial margin and the maintenance margin. The initial margin for a trade by a retail trader is higher than the initial margin would be for the same trade by a CCP member.

Margin accounts are used to mitigate credit risk in many other situations. We will now review several examples. In the next section and the next chapter, we will explain how margin accounts are used by CCPs in the OTC market.

Options on Stocks

A trader with a net long position in a particular exchange-traded stock option has no potential future liability. The options are usually paid for upfront, and they may or may not be exercised (as explained in Chapter 4). Therefore, there is no reason for the exchange to require margin from a trader with a net long position in a call or put option.

However, a trader with a net short position in an option contract (i.e., a trader who has sold call or put options) does have potential future liability. If the options are exercised, the trader must sell or buy the underlying stock at an unfavorable price. Traders with short positions are therefore required to post margin with the CCP.

The Chicago Board Options Exchange calculates margin that must be maintained each day as follows.

For a short call option, the margin requirement is the greater of:

- 100% of the value of the option plus 20% of the underlying stock price less the amount (if any) that the option is out-of-the-money, or
- 100% of the value of the option plus 10% of the underlying stock price.

For a short put option, the margin requirement is the greater of:

- 100% of the value of the option plus 20% of the underlying stock price less the amount (if any) that the option is out-of-the-money, or
- 100% of the value of the option plus 10% of the strike price.

Consider the situation where 100 call options on a stock are sold for USD 5 per option when the stock price is USD 47. If the strike price is USD 50, the option is USD 3 out-of-the-money and the margin required per option (USD) is

$$max(5 + 0.2 \times 47 - 3,5 + 0.1 \times 47) = 11.4$$

The total margin requirement is therefore USD 1,140. As the stock price and option price change, the margin requirement changes and a trader may be requested to contribute more funds to the margin account. If the trader does not post additional margin when required, the trader's position is closed out. Options (unlike futures) are usually settled at maturity. The margin posted by a trader therefore belongs to the trader and interest is paid by the CCP on a cash margin balance.

Short Sales

Shorting a stock involves borrowing shares and selling them in the usual manner. At some later date, the shares are repurchased and returned to the account from which they were borrowed. A retail trader who chooses to short a stock in the U.S. is typically required to post margin equal to 150% of the stock price at the time the short position is initiated. The proceeds of the sale account for two-thirds (100%/150%) of the margin. The trader must therefore contribute a further 50% of the stock price.

The account is adjusted for changes in the stock price. When the stock price declines, the balance in the margin account increases; when the stock price increases, the balance in the margin account decreases. A maintenance margin is typically set at 125% of the stock price. If the margin account balance falls below the maintenance margin, additional margin is required to bring it up to the maintenance margin level.[4]

For example, suppose a trader shorts 100 shares when the stock price is USD 30. The proceeds of the sale (USD 3,000) belong to the trader. The margin that must be initially posted, however, is USD 4,500 (i.e., 150% of USD 3,000). The trader must therefore post margin equal to the proceeds of the sale plus an additional USD 1,500.

Suppose further that the share price rises to USD 35. This is bad news for the trader because a short position is designed to do well (poorly) when the price decreases (increases). The value of the shares that have been shorted is USD 3,500, and the maintenance margin is now USD, 4,375 (= 1.25 × USD 3,500). The USD 4,500 initial margin covers this. If the share price rises again to USD 40, however, the maintenance margin becomes USD 5,000 (= 1.25 × USD 4,000) and there is a USD 500 margin call. If this margin call is not met, the position is closed out. The margin account balance belongs to the trader, and interest should be paid on the balance to the trader.

Buying on Margin

Buying on margin refers to the practice of borrowing funds from a broker to buy shares or other assets. As an example, consider a situation where a retail trader buys 1,000 shares for USD 60 per share on margin. The trader's broker states that the initial margin is 50% and the maintenance margin is 25%. The initial margin is the minimum percentage of the trade cost that must

[4] Note that the short seller is required to bring the balance up to the maintenance margin level. This can be contrasted with a retail trader in futures who is required to bring the balance up to the initial margin level when it drops below the maintenance margin level.

be provided by the trader at the time of the trade. In this case, the trader must therefore deposit at least USD 30,000 in cash or marginable securities with the broker.

Suppose that the trader deposits USD 30,000 in cash. The remaining USD 30,000 that is required to buy the shares is borrowed from the broker, who keeps the shares as collateral. The balance in the margin account is calculated as the value of the shares minus the amount owed to the broker. Initially, the balance in the margin account is USD 30,000 (= USD 60,000 − USD 30,000). Gains and losses on the shares (as well as interest charged by the broker) are reflected in the margin account balance (which can also be viewed as the trader's equity in the position). The maintenance margin (25% in this case) is the minimum margin balance as a percentage of the value of the shares purchased. If the margin balance drops below this minimum, there is a margin call requiring the trader to provide additional margin in order to bring the balance up to the maintenance margin level.

Now suppose that the price of the security declines by USD 5. The value of the shares purchased falls to USD 55,000 and the balance in the margin account falls to USD 25,000. The margin as a percentage of the value of the shares purchased is around 45% (= 25,000/55,000). This is more than 25%, so there is no margin call. (The calculations here ignore the interest that would be charged by the broker.)

If the price of the security falls further to USD 39, the cumulative loss on the position is USD 21,000 (= USD 60,000 − USD 39,000). The balance in the margin account therefore falls to USD 9,000 (= USD 30,000 − USD 21,000), and the value of the shares is now USD 39,000.

The balance in the margin account has now fallen to 23.1% (= 9,000/39,000) of the value of the shares. Because this is less than 25%, there is a margin call. This requires the trader to bring the margin balance up to 25% of the value of the shares. To do this, the trader must add USD 750 (= 0.25 × USD 39,000 − USD 9,000) to the margin balance. If it is not provided, the broker sells the shares. If it is provided, the position is maintained, and the amount borrowed from the broker falls to USD 29,250.

5.4 OVER-THE-COUNTER MARKETS

Over-the-counter derivatives markets have existed for many years. Traditionally, they have involved two market participants interacting directly and agreeing to a transaction. As mentioned in Chapter 4, participants in the market can be categorized as end users and dealers. Dealers satisfy the needs of end users by entering derivative transactions and sometimes take positions in derivatives themselves. Dealers typically hedge their risks by trading with other dealers.[5]

The attractiveness of the OTC market is that derivatives are not standardized by an exchange and they can therefore be tailored to meet the needs of end users.

OTC derivatives have traditionally been cleared bilaterally. This involves the two parties to a transaction agreeing how it will be cleared, what netting arrangements will apply, and what collateral (if any) will be posted. However, CCPs have existed for some time in the OTC markets, and they have been increasingly used following the 2007–2008 global financial crisis. (The reason for this is discussed in more detail in the next chapter.)

Table 5.1 provides some statistics on the OTC market in June 2019. As Table 4.1 shows, the OTC market was more than five times the size of the exchange-traded market at that time (when measured in terms of the notional values of underlying assets). Table 5.1 shows that the value of all transactions in the OTC market was less than 2% of the value of the underlying assets.[6]

Interest rate derivatives are by far the most popular type of derivative in the OTC markets. They account for about 82% the value of the market (measured in terms of underlying assets) and about 73% of the value of outstanding derivatives. Most interest rate derivative transactions are interest rate swaps. These are agreements to exchange a fixed interest rate on a certain notional principal for a floating interest rate on the same notional principal. They will be discussed in more detail in Chapter 20. For now, note that the principal is *notional* because it is not exchanged.

It should be emphasized that value of the underlying assets in Table 5.1 is the notional principal of outstanding transactions, which is much greater than the value of the transactions.

A major disadvantage of the OTC markets has traditionally been related to credit risk. As we have explained, the exchange-traded markets use margin accounts and CCPs to mitigate credit risk. In the early days of the OTC markets, however, transactions were generally cleared bilaterally, and measures to alleviate credit risks were relatively rare. Instead, two market participants would simply agree to certain contingent future cash flows. If one side experienced financial difficulties and

[5] As mentioned in Chapter 4, interdealer brokers are often used for OTC transactions between dealers.

[6] If a transaction has a positive value of X to one side and a value of −X to the other side, the transaction would contribute X to the statistics on market value.

Table 5.1 Statistics Produced for the OTC Market by the Bank for International Settlements in June 2019 (See www.bis.org)

	Value of Underlying Assets (USD Billions)	Value of Transactions (USD Billions)	Ratio of Value of Transactions to Underlying Assets
Foreign Exchange	98,651	2,229	2.26%
Interest Rate	523,960	8,806	1.68%
Equity	7,046	579	8.22%
Commodity	2,114	198	9.37%
Credit Default Swaps	8,418	235	2.79%
Other	253	14	5.53%
Total	640,442	12,061	1.88%

was unable to meet its obligations, the other side was likely to experience a loss.

It should be emphasized that the credit exposure on a derivative (such as an interest rate swap with a certain notional principal) is much less than that on a bond or a loan with the same principal. If no collateral is posted, the credit exposure to a trader on a derivative is max(V, 0), where V is the value of the derivative to the trader. If the derivative has a negative value there is no exposure. If the derivative has a positive value, however, the potential loss equals that positive value. The Value of Transactions column in Figure 5.1 is therefore a better indication of aggregate potential credit exposures than the underlying principal.[7]

The total expected cost of defaults on a derivatives portfolio with a counterparty depends on the lives of the derivatives. This is because of the interplay between the following factors.

- There is a greater probability of the counterparty experiencing financial difficulties during the life of a derivative as the life of the derivative increases.
- The market variables that determine the value of derivatives are likely to move more during the life of a derivative as the life of the derivative increases.

Bilateral Netting in OTC Markets

Earlier in this chapter, we discussed the use of netting in exchange-traded contracts. Netting was adopted fairly early in the development of the bilaterally cleared OTC markets. Two market participants would enter into a master agreement that would apply to all the derivatives they traded. In the event of a default by one side, all the outstanding derivatives transactions between the two participants would be considered as a single transaction.

For example, suppose that A and B are two companies trading derivatives with each other in the OTC market and that at a point in time there are the four outstanding transactions between them as listed in Table 5.2.

Suppose that Company B gets into financial difficulties and declares bankruptcy. Without netting, Company B will default on Transactions 1 and 3, but keep Transactions 2 and 4. (The liquidators of Company B might keep Transactions 2 and 4 or sell them to a third party.) The potential loss to Company A is then USD 60 million. With netting, all transactions will be considered as a single transaction worth –USD 20 million to Company B. Company B's default then leads to a potential loss for Company A of only USD 20 million (instead of USD 60 million).

If Company A gets into financial difficulties and declares bankruptcy, there is a gain from netting to Company B. Without netting, Company B has a potential loss of USD 40 million (on Transactions 2 and 4). With netting, there is no potential loss (In fact, Company B will have to pay the liquidators of Company A USD 20 million to settle outstanding contracts).[8]

The Bank for International Settlements estimates that, when enforceable netting agreements are considered, the total exposure of participants in derivatives markets was about USD 2.7 trillion in June 2019. This is less than 25% of the market value of transactions, indicating that the beneficial effect of netting agreements for credit exposures in derivatives markets is considerable.

[7] In practice, netting reduces aggregate exposure below this as is discussed in the next section.

[8] The rules are a little more complicated than this. If one side defaults, the other side must replace outstanding transactions and is allowed to value transactions at the bid or the ask, whichever is more favorable.

Table 5.2 Outstanding Transactions between Company A and Company B

Transaction	Value to Company A (USD Million)	Value to Company B (USD Million)
1	+40	−40
2	−30	+30
3	+20	−20
4	−10	+10

Collateral

In recent years, the posting of collateral in OTC derivatives markets has become increasingly common. The credit support annex (CSA) of a master agreement between two parties specifies how the required collateral is to be calculated and what securities can be posted. Typically, outstanding derivatives are valued every day, and the net value is used to determine the extra collateral that must be posted. The terminology of the exchange-traded markets is sometimes used (with collateral being referred to as margin).

For example, suppose that Companies A and B are trading derivatives. On a given day, the net value of outstanding transactions increases by USD 1 million to Company B (and therefore decreases by USD 1 million to Company A). Under the terms of the CSA, Company A might be required to post collateral of USD 1 million to Company B.[9]

As already mentioned, CCPs have become an important feature of OTC markets and will be discussed in more detail in the next chapter. At this point, we discuss other ways in which OTC market participants have attempted to handle credit risk.

Special Purpose Vehicles

Special Purpose Vehicles (SPV), also called Special Purpose Entities (SPE), are companies created by another company in such a way that the credit risks are kept legally separate. SPVs and SPEs are sometimes created to manage a large project without the organization setting it up being put at risk.

For example, suppose that Company Y creates SPV/SPE Company X. Typically, Y transfers assets to X and may not control those assets. (In some jurisdictions, Y is not even allowed to own X.) If Y goes bankrupt, X should be able to continue to fulfil its obligations (and vice versa). Company X will typically have a AAA credit rating, but only after rating agencies have carefully examined the legal arrangements and mechanics of how the company operates.

SPVs and SPEs are frequently used to create derivatives from portfolios of assets such as mortgages or other types of loans. The company setting up the SPV/SPE is not responsible for the payoffs on the derivatives, and anyone who purchases the derivatives has payoffs that may be affected by defaults on the underlying loan portfolio. Because of the SPV/SPE's high credit rating, however, the payoffs will not be affected by the creditworthiness of the SPV/SPE.

Derivative Product Companies

One historic approach for handling credit risk in derivatives markets involved the use of derivative product companies (DPCs). These were well-capitalized subsidiaries of dealers designed to receive AAA credit ratings. When the dealer traded with Company X, the DPC (rather than the dealer) became Company X's counterparty. This means that while a dealer may have had a poor credit rating, a well-capitalized DPC would have a AAA credit rating, and therefore counterparties would be comfortable trading with it.

DPCs were set up so that they took on virtually no market risk. When they traded with a counterparty, they normally entered an offsetting trade with the parent so that the parent is responsible for managing the risk.

One point to bear in mind is that credit risk is not eliminated by a DPC. The DPC itself may be virtually riskless, but the parent company is exposed to risk, and a default by the parent company would have had consequences for a DPC's counterparties. Typically, it would have led to the DPC being sold to another entity or all transactions being closed out at mid-market prices. DPCs tried to alleviate the concerns of counterparties by documenting exactly what would happen if the parent experienced financial difficulties.

DPCs have become virtually nonexistent since the credit crisis of 2007–2008. However, they were made redundant well before then by the increasing use of collateral in the OTC market.

Credit Default Swaps

One way of managing credit risk is to use credit default swaps. These are insurance-like contracts between a protection buyer and a protection seller. The buyer pays a regular premium to the seller, and if there is a default by a specified entity (not the

[9] The CSA may require collateral to be posted only when the value of outstanding transactions to one side exceeds a certain threshold level. Also, to avoid the administrative costs of small transfers, a minimum transfer amount is usually specified.

buyer or seller), the seller makes a payment to the buyer. The credit default swap market grew rapidly between 2000 and 2007, but it has declined since then.

Some companies have specialized in selling credit protection using credit default swaps. Monolines are companies with good credit ratings that do this as their main activity. Also, some insurance companies have sold protection as an extension of their other insurance activities. The most well-known of these insurance companies is AIG, which (through its subsidiary AIG Financial Products) sold a huge amount of protection on products created from mortgage portfolios.

The 2007–2008 crisis led to many failures among monolines. Also, AIG suffered severe losses arising from credit default swaps and required a USD 180 billion bailout from the U.S. government (the funds have since been repaid). Banks such as Citigroup and Merrill Lynch that bought protection from monolines also lost several billion dollars.

SUMMARY

Exchange-traded markets have developed procedures to significantly reduce credit risk and make it easy for market participants to close out positions before maturity. These features include netting, central clearing, margin requirements, and default fund contributions. Netting allows traders to offset long and short positions. Central clearing means that an exchange's CCP is always a member trader's counterparty. By requiring initial margin and variation margin, exchanges ensure that they are unlikely to lose money from a member's default. Default fund contributions are an added source of protection for exchanges.

Over-the-counter markets have copied many of the tools first introduced in exchange-traded markets. For example, netting has been a feature of bilaterally cleared OTC markets almost from the beginning. Margin has become a progressively more popular feature of bilaterally cleared OTC markets. As we will discuss in the next chapter, CCPs have become more widely used in the OTC markets following the 2007–2008 crisis.

OTC markets have experimented with several other ways of managing credit risks. Special purpose vehicles and special purpose entities are useful in some situations. Derivative product companies were once seen as a way for dealers to convince their counterparties that there was no risk in trading with them. Credit derivatives have also been used. However, none of these alternatives have stood the test of time. Increasingly, risks in the OTC market are being managed in much the same way as they are in the exchange-traded market.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

QUESTIONS

Short Concept Questions

5.1 What is the main role of an exchange CCP?

5.2 Explain why CCPs make it easy for traders to close out positions in exchange-traded markets.

5.3 What is margin in an exchange-traded futures contract? Explain the difference between variation margin and initial margin.

5.4 Explain how netting works in futures markets.

5.5 What is the role of default fund contributions in CCPs?

5.6 What does *buying on margin* mean?

5.7 In the futures market, how is the margin account between a retail trader and his/her broker managed differently from that between a CCP member and the CCP?

5.8 What are the most actively traded OTC instruments?

5.9 How does netting reduce credit risk in the bilaterally cleared OTC market?

5.10 What is the difference between collateral and margin?

Practice Questions

5.11 Suppose that A agreed to buy a certain amount of soybeans for future delivery from B before the existence of CCPs. How could A have exited from this trade. What problems might A have encountered?

5.12 Before the existence of central clearing, Trader A agreed to buy 10,000 bushels of wheat for delivery in March from Trader B for 500 cents per bushel. A few weeks later, Trader A sold 5,000 bushels of wheat for March delivery to Trader C for 520 cents per bushel and 5,000 bushels to Trader D for March delivery for 530 cents per bushel. What is Trader A's net profit or loss?

5.13 What factors are likely to influence a CCP's determination of initial margin for a futures contract?

5.14 What funds are available to a CCP to fund losses arising from a default by a member? In what order are they used?

5.15 A trader contacts a broker to enter into a futures contract to sell 5,000 bushels of wheat for 600 cents per bushel. The initial margin is USD 3,000, and the maintenance margin is USD 2,000. Under what circum-stance is the trader required to provide more margin? How much margin is required? Under what circumstances can USD 500 be withdrawn from the margin account?

5.16 A U.S. trader sells 500 put options. The option price is USD 3, the strike price is USD 31, and the stock price is USD 30. What is the margin requirement?

5.17 A trader shorts 100 shares when the price is USD 50. The initial margin and maintenance margin are 150% and 125%. What is the initial margin required? How high can the share price go before further margin is required? (Ignore interest payments.)

5.18 The counterparty to a dealer in the OTC market has agreed to post margin equal to max(V, 0) where V is the value of outstanding transactions to the dealer. What credit risk is the dealer taking?

5.19 Explain the operation of a derivative product company.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

5.1 The CCP stands between the two sides when a derivative is traded on an exchange. Suppose A trades a derivative contract with B. The transaction is moved to the CCP so that A is buying from the CCP and B is selling to the CCP.

5.2 A CCP becomes the counterparty to all traders. A long position entered into with one counterparty can therefore be closed out by a short position in the same contract entered into with another counterparty.

5.3 Margin is the word used to describe cash or marketable securities that must be provided by a trader. Variation margin is an amount paid or received to settle a futures position daily. Initial margin is cash or marketable securities provided by a trader to a CCP to reduce the possibility of losses from a default.

5.4 Long positions to buy an asset at a future time are netted with short positions entered into by the same trader to sell the asset at the same time. For example, if a trader has five long contracts (to buy) and then enters into three short contracts (to sell), the trader's position going forward is two long contracts.

5.5 Default fund contributions are used to fund losses when the initial margin of the defaulting member of the CCP proves inadequate. First the default fund contributions of the defaulting party are used. Then the default fund contributions of other members are used.

5.6 Buying on margin means borrowing part of the funds necessary to buy an asset. The asset is held as collateral. If its value declines sufficiently far, the borrower will have to either provide further funds or be closed out.

5.7 In the case of a retail trader, there is an initial margin and a maintenance margin. The margin account balance is adjusted each day to reflect gains and losses. If the balance in the margin account falls below the maintenance margin level, the trader is asked for funds to bring the balance up to the initial margin level. For the member of an exchange, the maintenance margin and the initial margin are effectively the same.

5.8 Interest rate derivatives, in particular, interest rate swaps.

5.9 In the event of a default by a counterparty, positive-valued transactions are netted against negative-valued transactions, and the resulting portfolio is considered as a single transaction. The counterparty cannot therefore default on transactions with negative value to itself and keep transactions with a positive value to itself.

5.10 Margin is the word used for collateral in the exchange-traded market. Increasingly, it is also used in the OTC market.

Solved Problems

5.11 A could approach B to agree on a payment that will close out the transaction. If no agreement is reached, A can enter into an offsetting transaction with a third party, C. However, A is then taking on the credit risk of both B and C.

5.12 The trader's net profit on the delivery date will be

$$5{,}000 \times 20\ cents + 5{,}000 \times 30\ cents = \text{USD } 2{,}500$$

Since the introduction of daily settlement in futures markets, the net profits such as this have been realized day-by-day.

5.13 Initial margin will depend on the volatility of the futures price and how long it will take the exchange to close out the member if the member defaults.

5.14 Funds are used in the following order:

1. Initial margin provided by the member,
2. Default fund contribution of the member,
3. Default fund contributions of other members, and
4. Equity capital of the exchange.

5.15 The trader is required to provide more margin if more than USD 1,000 is lost from the margin account. This will happen if the price of wheat falls by more than 20 cents because variation margin will then have reduced the margin balance by more than $5{,}000 \times 20$ cents = USD 1,000. The trader is required to bring the balance in the margin account up to the initial margin level of USD 3,000. USD 500 can be withdrawn from the margin account if the price rises by 10 cents or more because the balance in the margin account will have risen by at least USD 500 (5,000 × 10 cents = USD 500) over the required initial margin of USD 3,000.

5.16 The margin per option in USD is

$$\max(3 + 0.2 \times 30,\ 3 + 0.1 \times 31) = 9$$

The total margin requirement is therefore USD 4,500.

5.17 The trader is initially required to contribute USD 2,500 in addition to the USD 5,000 obtained from selling the shares, creating a margin balance of USD 7,500. If the USD share price rises to X, the maintenance margin becomes $100 \times 1.25X = 125X$. This is greater than the 7,500 margin balance when $X > 7{,}500/125$, that is, when $X > 60$.

5.18 The risk to the dealer corresponds to a possible increase in the value of its transactions with the counterparty around the time that the counterparty defaults. In particular, it has exposure to increases in value between the last time that the counterparty posts collateral and the time when the dealer is able to close out its position with the counterparty.

5.19 A derivative product company is a wholly owned subsidiary of a dealer and the company through which it conducts its derivatives transactions. It is given enough capital to be rated AAA and takes no market risk. When it enters into a transaction with a counterparty, it also enters into an offsetting transaction with the parent company.

Central Clearing

6

Learning Objectives

After completing this reading, you should be able to:

- Provide examples of the mechanics of a central counterparty (CCP).
- Describe the role of CCPs and distinguish between bilateral and centralized clearing.
- Describe advantages and disadvantages of central clearing of OTC derivatives.
- Explain regulatory initiatives for the OTC derivatives market and their impact on central clearing.
- Compare margin requirements in centrally cleared and bilateral markets and explain how margin can mitigate risk.
- Compare netting in bilateral markets vs centrally cleared markets.
- Assess the impact of central clearing on the broader financial markets.
- Identify and explain the types of risks faced by CCPs.
- Identify and distinguish between the risks to clearing members and to non-members.

Central counterparties (CCPs) have been used for trading derivatives in exchange-traded markets for many years. As explained in Chapter 5, a derivatives exchange is organized so that its CCP (also known as its clearing house) is the counterparty to all its member traders (whether they are buyers or sellers).

The rules developed by CCPs for members posting margin allow the exchanges to handle credit risks efficiently.[1] As a result, failures of CCPs handling exchange-traded products have been rare.

One example of such a failure is that of French clearing house Caisse de Liquidation in 1974. Following a steep decline in sugar futures prices, the exchange was required to make large variation margin payments to its members with short positions. However, some members with long positions failed to make their variation margin payments. As a result, their initial margin balances were insufficient to make the variation margin payments even when combined with other funds the CCP could access.

A similar failure is that of the Kuala Lumpur Commodity Clearing House. The Malaysian exchange was forced to close in 1983 after a steep decline in palm oil futures prices left the CCP unable to pay margin owed to members with short positions. As with Caisse, this inability to pay was the result of members with long futures positions failing to post variation margin when required.[2]

It should be emphasized that failures such as these are rare and futures markets have almost always been able to survive periods of market volatility. For example, futures markets in the U.S. were tested on October 19, 1987, when the S&P 500 index declined by over 20% in one day. This created a stressful situation for exchanges trading S&P 500 futures, such as the Chicago Mercantile Exchange (CME). Despite the dramatic fall in prices, however, the CME was able to pay in full all members with short futures positions.[3]

Exchanges have learned from the failures and near-failures of CCPs in the past and are now considered to be extremely safe. For example, initial margin requirements are now adjusted more frequently to reflect changing market conditions. Furthermore, initial and variation margin payments may be required from members during the day (as well as at the end of a day) when asset prices are especially volatile. This is important given that a member typically has only one or two hours to meet a margin call. If a margin call is not met, the member's position is closed out. The positions of members are monitored carefully by the CCP and members may be required to reduce their exposures by the CCP in some circumstances.

This chapter focuses on CCPs in the over-the-counter (OTC) market.[4] These CCPs operate similarly to the CCPs used by exchanges. As with exchange CCPs, members of OTC CCPs are required to post initial and variation margin as well as make contributions to the default fund. However, products being cleared in the OTC market differ from those being cleared by exchanges. For example, most exchange-traded contracts last only a few months, and very few last more than three years. In contrast, OTC contracts can last ten years or longer. The average futures trade on an exchange is also much smaller than the average trade in the OTC market.

Another key difference is that while exchange-traded futures contracts trade continuously, OTC contracts trade only intermittently. As a result, OTC contracts are less liquid than exchange-traded contracts. The differences in trading frequencies between the two types of markets also affects variation margin calculations. While exchange-traded variation margin requirements can be determined directly from market prices, it is often necessary to use a model when determining margin requirements in the OTC markets.

Three large CCPs for clearing OTC transactions are

- SwapClear (part of LCH Clearnet in London),
- ClearPort (part of the CME Group in Chicago), and
- ICE Clear Credit (part of the Intercontinental Exchange).

These large CCPs are critical for the smooth functioning of the global financial system and are regarded as "too big to fail" by financial regulators. This means that in the event of financial difficulties, they would almost certainly receive some type of bailout. It is therefore important to examine their risks.

There are also several other smaller, more localized OTC CCPs that might need to be bailed out in the event of financial difficulties. This is because some authorities regard it as important to have local OTC CCPs to service financial institutions in their region and clear transactions denominated in local currency.

While cooperation between CCPs is limited, there are significant economies of scale involved with running a CCP. Thus, we

[1] As explained in Chapter 5, margin is a form of collateral collected by the CCP to cover the potential losses associated with a counterparty default.

[2] For more details on CCP failures, see J. Gregory, *Central Counterparties: Mandatory Clearing and Bilateral Margin Requirements for OTC Derivatives*, Wiley 2014, Chapter 14.

[3] The Hong Kong futures exchange was not so lucky at the time of the 1987 crash and had to be bailed out by the Hong Kong government. It is also worth noting that some brokers within the U.S. went bankrupt because their clients with long S&P futures positions failed to provide variation margin when requested. As a result they were unable to meet their obligations on the long futures positions they had entered into on behalf of their clients.

[4] Chapter 5 discusses CCPs in the exchange-traded market.

should expect mergers between CCPs as well as takeovers of smaller CCPs by larger ones. Such consolidation has already begun with exchanges. For example, the Chicago Board of Trade and the Chicago Mercantile Group have merged to form the CME Group, which includes the New York Mercantile Exchange and the Kansas City Board of Trade. Another example is the merger between Euronext and the New York Stock Exchange that created NYSE Euronext (which itself was acquired by the Intercontinental Exchange).

6.1 THE OPERATION OF CCPS

CCPs clearing trades in the OTC markets operate in much the same way as CCPs clearing trades on exchanges. Members are required to post initial margin and variation margin as well as make contributions to the default fund. The variation margin is paid or received daily (or even more frequently) by members and reflects the change in the value of each member's portfolio of transactions with the CCP. In absence of a member default, the variation margin received by the CCP should always equal the variation margin paid by the CCP. This is because the CCP has a matched book and therefore it takes no market risk. When there is a default, there is market risk as the CCP closes out the positions of the defaulting member.

The initial margin required from each member is calculated using historical data. The key question for a CCP is "How much could be lost if the member defaults and market price movements reduce the value of the member's position as it is being closed out?"

Typically, the CCP sets the initial margin so that if it takes five days to close out the member's position, the CCP is 99% certain that the initial margin will cover the losses. If the defaulting member's initial margin proves insufficient to cover losses, however, the default fund contributions from the defaulting member are used. If this is still not enough, the contributions from other members are used. Only when this amount is insufficient does the CCP's equity become at risk.

When a member defaults, the exchange typically holds an auction inviting other members to bid for transactions that offset the defaulting member's transactions. Members have an incentive to cooperate; if a CCP can quickly close out a defaulting members positions, the remaining members' default fund contributions will be safe, and they can continue to clear transactions through the CCP. If the auction fails, however, the CCP may have the right to allocate losses to members who have made recent gains. Additionally, the CCP may choose to *tear up* transactions. This procedure involves closing out transactions between members and the defaulting party at prices that leave the non-defaulting members with some loss.[5]

CCPs cover their costs by charging fees per trade. They may also earn interest on initial margin in excess of that paid to members. For CCPs owned by their members (or a subset of their members), excess profits are distributed to member-owners. Other CCPs are owned by outside investors and are under pressure to generate profits. Competition between CCPs can benefit users by providing choices and encouraging CCPs to improve their systems. However, there is a danger that CCPs will try to compete with each other by reducing initial margin requirements and default fund contributions. That in turn would increase the risks that CCPs pose within the financial system.

OTC CCPs are subject to a great deal of regulation. For example, the Financial Conduct Authority (FCA) in the United Kingdom closely monitors risks taken by LCH Clearnet.

It is important that OTC CCPs not be allowed to take risks unrelated to their main activity of clearing OTC transactions. For example, it would be inappropriate for them to engage in speculative trading activities. In this respect, an OTC CCP should behave like a public utility.

6.2 REGULATION OF OTC DERIVATIVES MARKET

Regulations introduced since the 2007–2008 global financial crisis have led to an increase in the use of CCPs in the OTC derivatives market. These regulations were prompted by the belief that complex OTC-traded derivatives, specifically those created from portfolios of subprime (i.e., riskier than average) mortgages, played a role in causing the crisis.

During the time these mortgage derivatives were being traded, OTC markets were largely unregulated. Market participants could execute and clear trades in any way they chose without reporting their trades to a central authority.

So, when the G-20 leaders met in Pittsburgh in September 2009, they were anxious to rein in the OTC market.[6] They were particularly concerned about *systemic risk*. This is the risk that a default by one derivatives dealer could lead to losses being incurred by other derivatives dealers on their transactions with

[5] For a discussion of this, see ISDA, "CCP loss allocation at the end of the waterfall" https://www.isda.org/a/jTDDE/ccp-loss-allocation-waterfall-0807.pdf

[6] The G-20 (or Group of Twenty) is an international forum for the governments and central bank governors from 20 countries.

the defaulting dealer. This in turn could result in further defaults and further losses by other dealers. In the worst-case scenario, this interconnectedness of derivatives dealers would lead to a collapse of the global financial system.

The statement issued by the G-20 leaders after the Pittsburgh summit included the following paragraph.[7]

> *All standardized OTC derivative contracts should be traded on exchanges or electronic trading platforms, where appropriate, and cleared through central counterparties by end-2012 at the latest. OTC derivative contracts should be reported to trade repositories. Noncentrally cleared contracts should be subject to higher capital requirements. We ask the FSB and its relevant members to assess regularly implementation and whether it is sufficient to improve transparency in the derivatives markets, mitigate systemic risk, and protect against market abuse.*[8]

The G-20 Pittsburgh meeting resulted in three major regulations affecting OTC derivatives.

1. A requirement that all standardized OTC derivatives be cleared through CCPs. Standardized derivatives include plain vanilla interest rate swaps (which account for most traded OTC derivatives) and credit default swaps on indices. The purpose of this requirement is to create an environment where dealers have less credit exposure to each other, reducing interconnectedness and systemic risk.
2. A requirement that standardized OTC derivatives be traded on electronic platforms to improve price transparency. If there is an electronic platform for matching buyers and sellers, the prices at which products trade should be readily available to all market participants. These platforms are called *swap execution facilities* (SEFs) in the U.S. and *organized trading facilities* (OTFs) in Europe. In practice, standardized products are passed automatically to a CCP once they have been traded on these platforms.
3. A requirement that all trades in the OTC market be reported to a central trade repository. This requirement provides regulators with important information on the risks being taken by participants in the OTC market.

[7] The New Rules for OTC Derivatives- https://fincad.com/blog/new-rules-otc-derivatives

[8] The Financial Stability Board (FSB) is an international institution that works with the G-20 to promote global financial stability by coordinating the development of regulatory, supervisory, and other financial sector policies. It monitors and makes recommendations about the health of the global financial system.

The first two of these regulations apply only to transactions between two financial institutions (or between a financial institution and a non-financial company deemed systemically important due to the volume of its OTC derivatives trading). This means that derivatives dealers do not have to use electronic platforms and CCPs when trading standardized contracts with most of their non-financial end users. The requirement that CCPs be used for standard interdealer transactions (e.g., interest rate swaps) has led to a huge growth in the volume of OTC transactions being cleared through CCPs.

6.3 STANDARD AND NON-STANDARD TRANSACTIONS

The meaning of the term *standard transaction* is obviously important to the application of the rules we have just mentioned. Standard transactions are transactions that CCPs are prepared to clear. For a CCP to clear a product, several conditions must be satisfied.

- The legal and economic terms of the product must be standard within the market.
- There must be generally accepted models for valuing the product (because the CCP needs to determine variation margin at least once a day).
- The product must trade actively. If this is not the case, it may be difficult to unwind a member's position if the member fails to produce margin when required. It may also be difficult to obtain up-to-date valuations for non-actively traded products. A related point is that CCPs will not consider it worthwhile to develop the systems to support the clearing of a product if their members do not trade it frequently.
- Extensive historical data on the price of the product should be made available to enable initial margin requirements to be determined.

The main product categories currently classified as standard are interest rate swaps and credit default swaps on indices. Other products, such as options on interest rate swaps and single-name credit default swaps, may be added at some stage. However, it seems likely that most exotic derivative products will be classified as non-standard for the foreseeable future.

Transactions that are cleared bilaterally (rather than through a CCP) are referred to as *uncleared*.

Regulators recognized that derivatives dealers could avoid the intent of the regulations described above by adding features to standard transactions that would make them slightly non-standard. A 2011 G-20 meeting resulted in uncleared derivatives joining their CCP-cleared counterparts in being subject to initial and variation margin requirements.

The new rules, which were phased in between 2016 and 2020, require margin to be posted for uncleared derivatives traded between two financial institutions (or between a financial institution and a systemically important non-financial institution engaging in a large volume of transactions). Under the new rules, initial margin and variation margin must be posted. This is a significant change because initial margin between OTC market dealers in the pre-crisis period was rarely used (unlike variation margin, which was common in the bilaterally cleared OTC markets).[9]

Variation margin on uncleared trades is usually transmitted from one counterparty to the other directly. However, initial margin cannot be handled in this way. To see why, suppose counterparty A transferred USD 1 million of initial margin to counterparty B and B transferred USD 1 million of initial margin to A. Since these transfers would cancel each other, the regulations therefore require initial margin on uncleared transactions to be transmitted to a third party to be held in trust.

The initial margin that A has to post should cover the greatest decrease in the value of the contracts to itself (and therefore the greatest increase in their value to B) estimated to occur over a ten-day period with 99% confidence in stressed market conditions.[10] This requirement recognizes that if A defaults, it could take B up to ten days to close out (or replace) its positions with A. This means that the close out could be quite expensive if the value of these positions to B increase during that time.

6.4 THE MOVE TO CENTRAL CLEARING

As previously mentioned, the new regulations have led to a huge increase in the volume of OTC derivatives cleared though CCPs. Statistics produced by the Bank for International Settlements shows that 78% of interest rate derivatives and 54% of credit default swaps were cleared through CCPs in June 2019.[11]

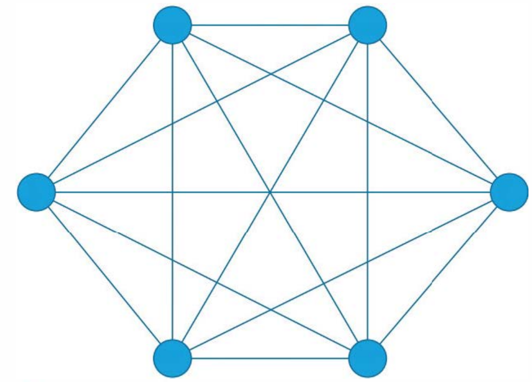


Figure 6.1 Six market participants trade derivatives with each other bilaterally. Each line represents a master agreement between a pair of dealers.

To examine the impact of CCPs on OTC markets, Figure 6.1 shows a simplified situation where there are six participants in an OTC derivatives market trading with each other. It assumes that all contracts are cleared bilaterally and (as explained in Chapter 5) there are master agreements between every pair of participants covering all contracts between them. The agreements outline netting arrangements, collateral arrangements, and what happens in the event of a default by one side. Standard documentation for master agreements has been provided by the International Swaps and Derivatives Association.

Figure 6.2 shows the situation when all transactions between the six market participants are centrally cleared through the same CCP. In this case, we assume that the six market participants are members of the CCP. The positions of market participants are transferred to the CCP, and each market participant agrees to adhere to the terms set by the CCP.[12] Specifically, each participant agrees to post initial margin and variation margin. The participants also agree to make contributions to the default fund as required.

In practice, the current environment for trading derivatives is a mixture of what is shown in Figure 6.1 and what is shown in Figure 6.2. Some transactions (i.e., non-standard transactions between financial institutions and a subset of transactions with non-financial end users) are cleared bilaterally as indicated in Figure 6.1. Standard transactions between financial institutions (and some standard transactions with non-financial end users)

[9] When entering into a transaction with a much less creditworthy counterparty, a derivatives dealer might insist on the counterparty posting initial margin. Posting of initial margin by both sides was almost unheard of in the bilaterally cleared market pre-crisis.

[10] Market participants use a model referred to as SIMM (Standard Initial Margin Model) to calculate initial margin for uncleared trades. This model was developed by the International Swaps and Derivatives Association in conjunction with market participants. We mentioned earlier that the initial margin for cleared transactions is based on a five-day close out period. One reason for the difference is that standard transactions are more liquid and can therefore be expected to be closed out more quickly.

[11] See www.bis.org

[12] The word used to describe the transfer of a contract from one party to another is *novation*.

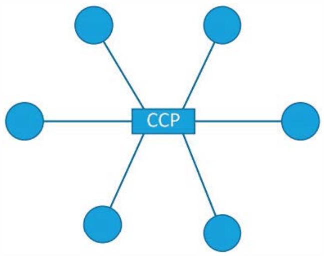


Figure 6.2 Six members of the CCP trade derivatives with each other and clear their transactions through a single CCP. Each line represents an agreement between a member and the CCP.

are cleared though CCPs as indicated in Figure 6.2. A further complication is that there is more than one CCP. This means that even if all trades by the six market participants are cleared centrally, they might be cleared though different CCPs.

Members of a CCP clear the trades of non-members bilaterally. These non-members are small financial institutions and non-financial companies. (Retail investors do not generally trade in the OTC market.) The non-members must post initial and variation margin with the member who is clearing their trades. This is similar to the manner in which the non-members' trades are cleared by exchange CCPs.

OTC transactions are different from futures in that they are not settled daily. Cash flows settling the contracts occur periodically and sometimes (e.g., in the case of European options) all cash flows are settled at the end of the contract's life. However, CCPs value OTC contracts at least once a day and transfer the required variation margin reflecting the change in net value of outstanding contracts. This means, for example, that the variation margin transferred from Party A to Party B belongs to Party A until the contractual cash flows occur. As a result, Party B must pay interest on the variation margin it receives from A. If clearing is through a CCP, the interest is paid by Party B to the CCP and by the CCP to Party A. In addition to this, interest on initial margin balances is paid by the CCP.

Even though interest is paid on margin transfers as appropriate, financial institutions still regard them as a cost. Typically, institutions compare the interest rate paid on their margin with their average cost of borrowed funds and calculate a cost relating to the difference.[13]

Financial institutions understand that the margin they post could be put to a better use. However, the effect of the margin requirements within the central clearing ecosystem is that these firms are much less likely to lose money because of a default by another financial institution. Furthermore, these benefits can only be obtained if all financial institutions are part of the CCP ecosystem.

CCPs may give rise to an increase in netting.[14] To see how this can be the case, consider the situation in Figure 6.3 where three market participants trade bilaterally. The arrows in this figure indicate that Party A has transactions with Party C that are worth +USD 8 million to Party A and −USD 8 million to Party C. Party A also has transactions with Party B that are worth +USD 5 million to Party B and −USD 5 million to Party A. Finally, Party B has transactions with Party C that are worth +USD 2 million to Party C and −USD 2 million to Party B. Positive values lead to potential credit exposures, while negative values do not. In the absence of any credit mitigation procedures:

- Party A has a credit exposure of USD 8 million to Party C and none to Party B,
- Party B has a credit exposure of USD 5 million to Party A and none to Party C, and
- Party C has a credit exposure of USD 2 million to Party B and none to Party A.

The total credit exposure of all three parties is USD 15 million (= 8 million + 5 million + 2 million).

Next, suppose that all transactions are cleared through a single CCP. Figure 6.4 indicates how this would work. In this example, a transaction between A and B would be converted into a transaction between A and the CCP as well a transaction between the CCP and B.

The CCP would then net the transactions in the manner shown in Figure 6.5.

This demonstrates that the CCP has increased netting.

- Party A is able to net its transactions with Party B against its transactions with Party C.
- Party B is able to net its transactions with Party A against its transactions with Party C.
- Party C is able to net its transactions with Party A against its transactions with Party B.

Note that the overall credit exposure in the system is lower. Parties A and B each have a credit exposure to the CCP of 3 million USD and Party C has no credit exposure at all. If the

[13] In the case of variation margin, the cost (which can be negative) is referred to as the funding value adjustment (FVA). In the case of initial margin, the cost is referred to as margin value adjustment (MVA). Whether the average cost of borrowed funds should be used in the calculation is debatable. This is because margin is a relatively safe use of the financial institution's funds (somewhat like buying Treasury bills).

[14] See Sections 5.1 and 5.4 of Chapter 5 for a discussion of netting in exchange-traded and OTC markets.

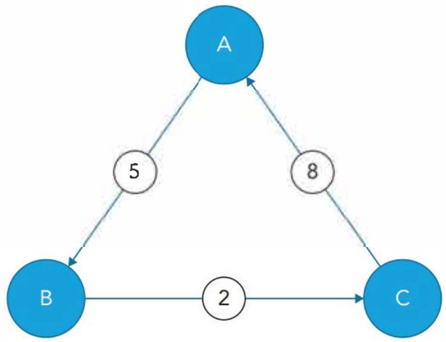


Figure 6.3 Three market participants clearing bilaterally.

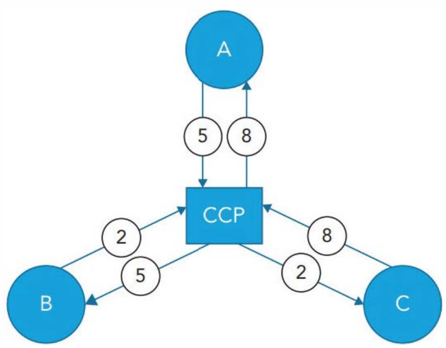


Figure 6.4 The transactions in Figure 6.3 are presented to a CCP.

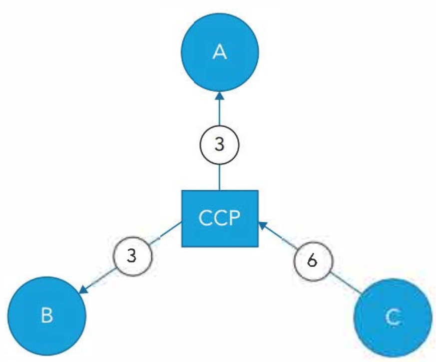


Figure 6.5 Situation after CCP has netted the transactions shown above in Figure 6.4.

credit exposures were handled with variation margin and initial margin, there would be a saving of initial margin.

Clearing through CCPs usually increases netting, but this is not always the case. For example, the transactions with one CCP cannot usually be netted against transactions with another CCP.

Furthermore, some transactions (e.g., non-standard transactions) cannot be cleared through CCPs. And while standard transactions and non-standard transactions between two parties can be netted when cleared bilaterally, this is not the case if the standard transactions go through a CCP (while the non-standard transactions continue to be cleared bilaterally). This means that the loss of netting could outweigh the extra netting provided by diverting transactions through CCPs even when the same CCP is used for all transactions.

6.5 ADVANTAGES AND DISADVANTAGES OF CCPS

A key advantage of the central clearing model is that it is much easier for market participants to exit a CCP transaction than a bilaterally cleared transaction.

When transactions are cleared bilaterally, a market participant can only exit a trade by approaching the original counterparty and trying to negotiate a close-out (which usually would involve a payment from one side to the other). If the original counterparty does not offer reasonable close-out terms, the market participant would need to enter into an offsetting transaction with another counterparty. This arrangement creates credit risk because either the original counterparty or the new counterparty may default. However, the credit risk is eliminated if both the original trade and the offsetting trade are cleared with the same CCP.

If the original transaction and the new transaction are cleared with different CCPs, the market participant will have to post initial margin with each one. In the future, it may be possible to arrange for trades with different CCPs to be netted so that the initial margin is avoided. This arrangement is known as *interoperability* and is not (as of yet) common practice.

In bilateral clearing, the risk of a default by a market participant is borne entirely by its counterparties. If trades are cleared through a CCP, however, the risks are shared by all members of the CCP (some of which may never have traded with the defaulting counterparty). This sharing of credit risk is referred to as *loss mutualization* and is attractive to regulators because it has the effect of reducing systemic risk. It does this by dispersing the impact of a default by a market participant throughout the market.

Another advantage of CCPs is that they manage the margining, netting, settlement, and default resolution that would typically be handled by each market participant in the case of bilateral clearing. In fact, with the vast resources devoted to these functions and intense regulatory oversight, it is expected that a CCP can handle these functions better than any single market participant in the bilaterally cleared market. However, any operational problems experienced by the CCP could have widespread consequences because it is responsible for a far larger number of transactions than any single market participant.

CCPs may also improve liquidity in the OTC market by making it much easier for market participants to net and exit from transactions. Another positive aspect of central clearing is that it encourages the development of standard documentation for OTC derivative transactions.[15]

Chapter 2 explained two key risks faced by insurance companies: moral hazard and adverse selection. Recall that moral hazard is the risk that insurance encourages the insured party to take on more risk than they would otherwise, while adverse selection is the tendency for an insurance company to attract clients with more risks than the general population. A disadvantage for CCPs is that they entail both moral hazard and adverse selection risks.

- Moral hazard exists because market participants have less incentive to concern themselves with the riskiness of the companies they trade with when much of the risk will be passed on to the CCP.
- Adverse selection exists when a dealer has a choice between clearing a transaction through a CCP or clearing it bilaterally (e.g., when it is trading a standard derivative with a non-financial end user). If the dealer considers the credit risk of the counterparty to be high, it might persuade the counterparty to clear through a CCP. If the counterparty is financially strong, however, the dealer may be comfortable clearing the transaction bilaterally.

Another disadvantage of CCPs is that they tend to increase the severity of adverse economic events (i.e., they are pro-cyclical). When markets are highly volatile or there is a financial crisis, for example, many financial institutions are likely to have liquidity shortages. At the same time, CCPs are likely to increase initial margin requirements and default fund contributions. These actions would thereby exacerbate the liquidity shortages faced by financial institutions.

It is also difficult for CCP members to evaluate the credit risk they are taking. Note that a CCP member's default fund contributions (and even some of its variation margin gains) may be at risk if another member defaults. However, a CCP member may have very little information about trades done by other members. This can be contrasted with bilateral clearing, where exposures are more concentrated but better understood.

Conversely, a CCP has a great deal of information about the portfolios of its members and can act to address excessive risk taking. If a member has a particularly large exposure, for example, the CCP may limit trading or increase the required initial margin. In the bilaterally cleared market, however, participants typically see a much smaller percentage of their trading partners' portfolios.

[15] CCPs do not usually accept transactions with non-standard documentation.

6.6 CCP RISKS

It can be argued that the new derivatives regulations do nothing more than replace too-big-to-fail banks with too-big-to-fail CCPs. It certainly would be a disaster for the financial system if a major CCP were to fail.

However, a critical feature of CCPs is that they are much simpler organizations than banks and are therefore much easier to regulate.

The key activities of CCPs are

- Admitting members,
- Valuing transactions,
- Determining initial margin and default fund contributions, and
- Managing systems for netting, transferring variation margin, and so on.

In contrast, large banks engage in many more complex and less transparent activities. For example, the many different types of loan agreements that banks enter into give rise to credit risk, while bank trading activities lead to market risk. Derivatives trading leads to both market risks and credit risks (which can be quite difficult to evaluate).

Additionally, a bank's funding strategy can give rise to liquidity risks. Banks also face many operational risks arising from cybersecurity, anti-money laundering legislation, internal fraud, external fraud, and so on.

Thus, the rise of CCPs has seen risk transferred from companies that are very difficult to regulate to companies whose activities are more amenable to oversight.

It should also be noted that OTC CCPs have (up to now) functioned well. When Lehman Brothers declared bankruptcy in September 2008, for example, it was the largest bankruptcy in U.S history. However, CCPs (both those clearing exchange-traded products and those clearing OTC transactions) managed to close out Lehman's positions within a matter of days.[16] By contrast, disputes concerning Lehman Brothers' bilaterally cleared OTC transactions has dragged on for many years at great cost to everyone involved.

[16] See an account of this by the Global Association of Central Counterparties in "Central Counterparty Default Management and the Collapse of Lehman Brothers" http://ccp12.org/wp-content/uploads/2017/05/CCPDefaultManagementandtheCollapseofLehmanBrothers.pdf

Despite these points, it would be a mistake to believe that OTC CCPs pose no risks. CCPs clearing OTC transactions are critically important to financial markets, and some of them are very large. For example, at the beginning of 2020, LCH Clearnet's SwapClear cleared interest rate swaps with a total notional principal of over USD 300 trillion.[17] If LCH Clearnet failed, it would have to be bailed out (likely by the British government).[18]

One significant problem with CCPs is that there is a positive correlation among member defaults. If one member defaults because of difficult economic conditions, others are likely to do so as well. Recognizing this, regulators ask CCPs to consider scenarios where multiple members default at the same time. Regulators also require CCPs to conduct stress tests involving imaginary adverse events to determine whether they would survive and conduct close-outs efficiently.

Note that CCPs tend to treat all members in the same manner when calculating initial margin and default fund contributions. A consequence of this is that CCPs do not take the credit quality of its counterparties into account in the same way that a dealer does in the bilaterally cleared market. Once a dealer has been admitted as a member, it is usually treated in the same way as all other members.[19]

A risk for CCPs is that the auction processes for closing out defaulting members could fail in the turbulent markets. It may then be compelled to force other members to share in the losses and thereby cause more defaults. It might also lead to resignations among members unwilling to stay in the central clearing model.[20] This in turn could lead to a loss of reputation for the CCP and further resignations.

Other risks faced by a CCP are

- Fraud,
- Computer systems failure/hacking,
- Litigation costs, and
- Losses on investments of the initial margin and variation margin.

There may be correlations between losses arising from defaults and these types of losses. In the stressed market conditions that can lead to defaults, investments may perform poorly and litigation costs may increase.

Model Risk

As mentioned previously, OTC CCP transactions are different than exchange CCP transactions. OTC transactions last longer, are less standard, have less price transparency, and trade less frequently.

As a result, OTC CCPs are more reliant on valuation models in determining transaction values and clearing variation margin transfers. If these models function poorly, the operation of the CCP may be compromised and member disputes may follow.

OTC CCPs also rely on models to determine initial margin requirements. Whereas exchange CCPs tend to have standard rules for determining initial margin, an OTC CCP must run models to determine how much initial margin is appropriate for each member's position. A lesson from the failures of exchange CCPs mentioned in the introduction is that initial margin requirements should be updated regularly as volatilities change. Members should be required to post initial and variation margin almost immediately upon request.

Liquidity Risk

Like all other businesses, CCPs are subject to liquidity risk. A large CCP holding tens of billions of dollars of initial margin is faced with a trade-off between the return it gets by investing this cash and the liquidity constraints of its investments. Liquid investments (e.g., Treasury bills) tend to provide lower returns than less liquid investments (e.g., corporate bonds). At the same time, CCPs need some of their investments to be readily convertible into cash in the event one or more members default. Furthermore, it is important that the liquidity of an investment be assessed assuming stressed market conditions because member defaults are likely to be accompanied by turbulent market conditions (which typically reduce investment liquidity).

SUMMARY

As a result of regulations introduced since the 2007–2008 crisis, the volume of OTC transactions cleared through CCPs has risen dramatically. As of mid-2019, 78% of interest rate derivatives and 54% of credit default swaps were cleared through CCPs.

OTC CCPs operate much like the exchange CCPs that were discussed in Chapter 5. Arguably, OTC CCPs are more difficult to manage than exchange CCPs because their instruments are less standard, last longer, and only trade intermittently.

[17] LCH. Monthly Volumes—SwapClear Global. Retrieved from https://www.lch.com/services/swapclear/volumes/monthly-volumes-swapclear-global

[18] As mentioned earlier, it is subject to a great deal of oversight from the Financial Conduct Authority in the U.K.

[19] Regulators may be comfortable with this approach as assessing the quality of members would involve some subjectivity on the part of CCPs.

[20] Typically, a member would have to eliminate its CCP exposures, and give one month's notice, in order to resign.

CCPs require members to post initial margin, variation margin, and to make contributions to a default fund. While this model increases the funds that must be provided to support trading, it reduces counterparty credit risk. CCPs also make it easier for market participants to close out positions and tend to increase the amount of netting that is available to market participants.

Some OTC CCPs are extremely large and play a key role in the global financial system. If one of them were to fail, there could be disastrous consequences. While OTC CCPs have been tested by difficult markets (e.g., disruption caused by the default of Lehman Brothers), the central clearing model still contains risks. Market turbulence could lead to defaults by several members simultaneously and a loss of confidence in OTC central clearing. Like many other businesses, CCPs also face operational and liquidity risks. They are also reliant on models for determining variation margin transfers and initial margin requests.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

6.1 What types of transactions are required to be cleared through CCPs?

6.2 Give two examples of transactions referred to as "standard" in the regulations governing OTC derivative transactions.

6.3 What are "uncleared" transactions?

6.4 Why does initial margin for uncleared derivatives have to be posted with third parties?

6.5 Has the use of bilateral clearing for OTC derivatives increased or decreased since the 2007–2008 crisis?

6.6 Give two reasons why netting might not increase as a result of the new rules for central clearing.

6.7 What are *tear ups*?

6.8 How do CCPs give rise to moral hazard?

6.9 How do CCPs give rise to adverse selection?

6.10 What are pro-cyclical actions? Why might the actions of a CCP be pro-cyclical?

Practice Questions

6.11 Give three differences between the properties of OTC derivatives and exchange-traded derivatives that make central clearing more difficult.

6.12 Describe the OTC transactions that continue to be cleared bilaterally.

6.13 Give three reasons why a CCP requires the OTC derivatives it clears to be traded actively.

6.14 X, Y, and Z have entered into many derivative transactions. When transactions between X and Y are netted, the net value to X is 60. When transactions between Y and Z are netted, the net value to Y is 70. When transactions between Z and X are netted, the net value to Z is 80. Suppose that all transactions are cleared through a CCP rather than bilaterally. What is the net position of X, Y, and Z?

6.15 Explain why interest is paid on variation margin in OTC markets but not in exchange-traded futures markets.

6.16 What are the pros and cons regarding competition between the CCPs that clear OTC products?

6.17 What is the difference between the criteria for setting initial margin in the cleared market versus the uncleared market?

6.18 Explain what is meant by (a) loss mutualization and (b) systemic risk.

6.19 Why are CCPs easier to regulate than banks?

6.20 Why is an OTC CCP more dependent on valuation models than an exchange CCP?

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

6.1 All standard transactions between financial institutions must be cleared through CCPs. If a non-financial institution is engaging in a large volume of transactions, its standard transactions must also be cleared through CCPs. Standard transactions include interest rate swaps and credit default swaps on credit indices.

6.2 Interest rate swaps and credit default swaps on credit indices.

6.3 Uncleared transactions are transactions not cleared though a CCP.

6.4 If equal amounts of initial margin were posted from Party A to Party B and from Party B to Party A, the net initial margin posted by each side would be zero, and thus the initial margin would not be available to cover default losses.

6.5 It has decreased for interest rate derivatives and credit default swaps because of regulatory requirements. Both interest rate derivatives and credit default swaps are popular categories of OTC derivatives (see Table 5.1).

6.6 Transactions might be cleared through different CCPs. Netting between standard and non-standard transactions is no longer possible.

6.7 Tear ups involve closing out transactions of members that were originally entered into with the defaulting party at prices that leave the members with some loss.

6.8 A dealer may not monitor the credit quality of counterparties carefully.

6.9 When trading standard transactions with an end user, a dealer is more likely to insist that the transactions are cleared through a CCP when the credit quality of the end user is low.

6.10 Actions that reinforce the severity of economic cycles are referred to as procyclical. During stressed periods, a CCP is likely to ask for more initial margin from its members. This could exacerbate liquidity problems already being experienced by members and therefore be procyclical.

Solved Problems

6.11 OTC derivatives last longer and may be much larger than exchange-traded derivatives. They are less standard. They trade less frequently, have lower volume and are less liquid. They may have less historical data.

6.12 Non-standard transactions and some standard transactions between financial and non-financial companies continue being cleared bilaterally.

6.13 It needs to value transactions at least once a day. It needs to unwind a member's position expeditiously if the member defaults. It is not worth developing the systems to clear a particular transaction type if members do not trade it very often. The CCP uses trading data to set initial margin.

6.14 When they are cleared centrally, they net to −20 for X, +10 to Y, and +10 to Z.

6.15 Futures contracts are settled daily, so when variation margin is paid from A to B, it belongs to B. OTC derivatives are not usually settled daily, so variation margin from A to B belongs to A until contractual payments on the derivatives are required.

6.16 Competition gives market participants choice and may provide incentives for CCPs to become more efficient and improve their systems. However, it may also lead to initial margin and default fund contributions being reduced to attract more clients and, thereby, increase the probability of a CCP failure.

6.17 In the cleared market, CCPs want initial margin to cover five-day price changes with 99% certainty. In the uncleared market, it is anticipated that close-outs will take longer and initial margin covers ten-day price changes with 99% certainty.

6.18 Loss mutualization involves spreading losses over many different entities. Systemic risk is the risk that the interconnectedness of financial institutions will lead to a crisis whereby a failure of one financial institution leads to losses by other financial institutions—leading to even more failures.

6.19 CCPs are engaged in activities that are simpler than those of banks. Thus, they are easier to be supervised and regulated.

6.20 An OTC CCP needs models to value the portfolios of its members and to determine the initial margin that should be posted by its members. Exchange CCPs can usually observe prices from contracts being traded in the market and have developed standard rules for determining initial margin.

7

Futures Markets

Learning Objectives

After completing this reading, you should be able to:

- Define and describe the key features and specifications of a futures contract, including the underlying asset, the contract price and size, trading volume, open interest, delivery, and limits.
- Explain the convergence of futures and spot prices.
- Describe the role of an exchange in futures transactions.
- Explain the differences between a normal and an inverted futures market.
- Describe the mechanics of the delivery process and contrast it with cash settlement.
- Describe and compare different trading order types.
- Describe the application of marking to market and hedge accounting for futures.
- Compare and contrast forward and futures contracts.

Futures contracts are popular exchange-traded products. They are agreements to buy or sell an asset at a future time for a certain price. We have already discussed how variation margin and initial margin are used to handle credit risks in futures markets. In this chapter, we further explain how contracts are specified and how futures markets operate.

7.1 EXCHANGES

Futures contracts are actively traded all over the world. The largest futures exchange in the world is the CME Group, which formed the result of a merger of the Chicago Board of Trade (CBOT) and the Chicago Mercantile Exchange (CME) along with subsequent acquisitions of the New York Mercantile Exchange (NYMEX) and the Commodity Exchange, Inc. (COMEX).

Other large futures exchanges are the National Stock Exchange of India, the Intercontinental Exchange, B3 (formed in 2017 by the merger of Brazilian exchanges BM&FBOVESPA and CETIP), Eurex (based in Germany), the Shanghai Futures Exchange, and the Dalian Commodity Exchange. Table 7.1 shows the number of options and futures contracts traded in 2017 by the ten largest exchanges (as reported by the Futures Industry Association).[1]

Table 7.1 Futures and Options Contracts Traded by the Ten Largest Exchanges in 2017

Exchange	Futures and Options Contracts Traded (Million)
CME Group	4,089
National Stock Exchange of India	2,465
Intercontinental Exchange	2,125
CBOE Holdings	1,810
B3	1,809
NASDAQ	1,677
Eurex	1,676
Moscow Exchange	1,585
Shanghai Futures Exchange	1,364
Dalian Commodity Exchange	1,101

Source: www.fia.org

7.2 OPERATION OF EXCHANGES

The following list reviews what we have learned in earlier chapters about the operation of exchanges.

- Standard contracts are defined by an exchange.
- An exchange's CCP becomes the counterparty to all members of the exchange when they trade.
- Trades are settled daily by variation margin flowing from one side to the other.
- Positions can be closed out by entering into offsetting positions (e.g., one long[2] contract to buy an asset in September can be closed out by entering into a short contract to sell the asset in September).
- Members post initial margin and contribute to a CCP's default fund to protect the CCP and its members against losses.
- Members clear the trades of non-members and maintain margin accounts with them to provide protection for themselves against non-member defaults.

Because it is so easy to close out futures contracts, most are closed out before delivery. The number of contracts in existence at any time is referred to as the *open interest*. This is the number of net long contracts held by members and it is equivalent to the number of net short contracts held by members (because the number of long positions must always equal the number of short positions). When trading in a certain contract starts, the open interest is zero. As members trade with each other, the open interest increases. Typically, the open interest peaks shortly before the delivery period specified in the contract. Members then start to close out their positions, and the open interest plummets.

A trade of a futures contract between two members has one of the following effects on the open interest of that contract.

- When both members are taking new positions, the open interest increases by one.
- When one member is closing out a position while the other member is taking a new position, the open interest remains the same.
- When both members are closing out their respective positions, the open interest decreases by one.

[1] It is not always easy to distinguish the volume of trading in futures from that in options from published statistics. Options on futures for example are not always reported separately from futures. However, other statistics published by the Futures Industry Association show that more than half the futures and options contracts traded worldwide are futures. It is estimated that in total 14.84 billion futures contracts and 10.36 billion option contracts were traded in 2017.

[2] Recall that taking a long position is referred to as buying a futures contract, while taking a short position is referred to as selling a futures contract.

The number of contracts traded in a day is referred to as the *trading volume*. The trading volume can be greater than the open interest at the end of the day if many traders are closing out their positions (which, as mentioned, tends to occur towards the end of the life of a contract). It can also happen if there is a large amount of *day trading*, which is when trades are entered into and closed out on the same day.

7.3 SPECIFICATION OF CONTRACTS

Exchanges must define what is being traded in detail. Whenever there is a choice about what is to be delivered, where it is to be delivered, and when it is to be delivered, it is nearly always the case that the party with the short position (the party making the delivery) has the right to choose.[3]

The Underlying Asset

In the case of financial assets (e.g., foreign currencies or stock indices), the definition of the underlying asset is usually straightforward. For example, the asset underlying one contract on the euro traded by the CME Group is 125,000 euros. The asset underlying one contract traded by the CME Group on the S&P 500 Index is USD 250 multiplied by the index.[4]

When the underlying asset is a commodity, the grade (in terms of quality) of the commodity that could be delivered must be specified. For example, the Intercontinental Exchange has specified the asset in its orange juice futures contract as "frozen concentrates that are U.S. Grade A with a Brix value of not less than 62.5 degrees."

In some instances, different grades can be delivered with a corresponding price adjustment. In the CME Group's corn futures, for example, the grade is specified as "No. 2 Yellow." However, "No. 1 Yellow" is deliverable for a 1.5 cents higher price per bushel than "No. 2 Yellow," while "No. 3 Yellow" is deliverable for 1.5 cents less per bushel than "No. 2 Yellow."[5] Failure by the exchange to adequately distinguish grades of the underlying asset could cause the contract to fail. This is because the party with a short position will always choose to deliver the lowest quality product possible, and this is likely to be considered unacceptable by the party with a long position.

Contract Size

Each exchange is responsible for determining the size of its contracts. Exchanges want to attract both retail investors (who usually want to do small trades) and large corporations (who may have large positions to hedge). Typically, the value of what is delivered for a contract on a financial asset tends to be much greater compared to that of an agricultural product. However, exchanges have also developed smaller contracts that are meant to appeal to retail investors seeking to hedge smaller exposures or take smaller speculative positions.

For example, the CME Group offers both a regular NASDAQ contract (which is USD 100 multiplied by the NASDAQ 100 index) and a mini NASDAQ contract (which is USD 20 multiplied by the index). Interestingly, the mini NASDAQ contracts trade more actively than the regular NASDAQ contracts.

Delivery Location

In the case of commodities, transportation costs make the specification of the delivery location very important. For example, the CME Group's crude oil futures contract specifies that "delivery shall be made free-on-board ("F.O.B.") at any pipeline or storage facility in Cushing, Oklahoma with pipeline access to Enterprise, Cushing storage or Enbridge." For some contracts, the delivery location may factor into the price of the underlying asset.

Delivery Time

Futures contracts are referred to by their delivery months. The precise dates during the delivery month when a delivery can be made varies from contract to contract. The party with the short position can choose among the delivery dates specified by the exchange. Some contracts (e.g., the CME's crude oil futures) allow for delivery on any day during a given month. Other contracts specify a more restricted set of delivery dates.

Exchanges determine the delivery months for which a contract trades, the time when a contract starts trading, and the time when it finishes trading. For example, consider the corn futures contracts traded by the CME Group. These contracts have delivery months in March, May, July, September, and

---

[3] A rare exception is in the CME Group live cattle futures contract where the contract was changed in 1995 so that the buyer was given some delivery options.

[4] The CME Group also trades an E-mini S&P futures contract on USD 50 times the index to appeal to smaller investors.

[5] From 2019 the discount applicable to No. 3 Yellow will be "between 2 and 4 cents depending on broken corn and foreign material and damage grade factors."

December. Trading in the December 2020 contract began on October 09, 2017, and continues until December 14, 2020.[6]

A day's *settlement price* is the futures price at the close of trading and it is used for determining daily settlement (i.e., variation margin transfers). If the settlement price increases from close of one trading day to close of the next trading day, funds are taken out of the accounts of traders with short positions and added to the accounts of traders with long positions. If the settlement price decreases, funds are taken out of the accounts of traders with long positions and added to the accounts of traders with short positions.

The futures price converges to the spot price as the delivery period approaches. Figure 7.1 shows the situation where the futures price starts above the spot price, while Figure 7.2 shows the reverse. If the futures price is above the spot price during the delivery period, traders would have a clear arbitrage opportunity that can be implemented by:

- Shorting futures,
- Buying the asset, and
- Making the delivery.

Arbitrage opportunities such as this will not last long as traders take advantage of them. Also, if the futures price is below the spot price during the delivery period, those wanting to acquire the underlying asset will find it profitable to take a long futures position and wait for delivery to be made. As they do this, the futures price will rise towards the spot price.

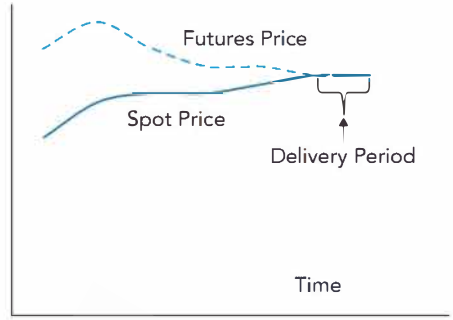


Figure 7.1 **Futures price converges to spot price from above.**

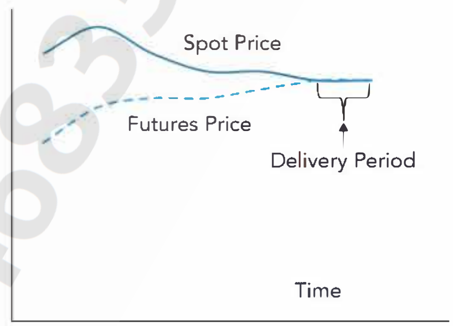


Figure 7.2 **Futures price converges to spot price from below.**

Price Quotes

Traders need to know how prices of contracts will be quoted and how accurately they will know the value of their positions. An exchange specifies the quotation convention it will use as well as the minimum price movements for each contract. In the case of the CME Group's corn futures, prices are quoted as cents per bushel, and the minimum price movement is 0.25 cents per bushel. Because one contract is on 5,000 bushels, the minimum price movement per contract is USD 12.5.

Price Limit

For most contracts, exchanges set limits on how much a futures price can move in one day. However, these price limits may be changed from time to time. If the price during a day moves up or down by the price limit (referred to as a *limit move*), trading is normally halted for the day. If the price limit is reached with a price increase, the contract is referred to as *limit up*. If it is reached with a price decrease, the contract is referred to as *limit down*. A typical limit move for the CME Group's corn futures would be 50 cents (i.e., 200 times the minimum price movement). However, the exchange does have the authority to step in and increase the limit.

The purpose of price limits is to prevent large price movements resulting from speculation. However, these price limits can also hinder the determination of true market prices if limit moves arise from new information reaching the market.

Position Limits

A position limit (as the name implies) is a limit on the size of a position that a speculator can hold. Its purpose is to prevent speculators from exercising an unreasonable influence on the market. Position limits are often in the tens of thousands of contracts and do not affect most traders.

[6] In the case of corn futures, trading finishes on the business day prior to the fifteenth calendar day of the contract month. Other rules apply to other contracts.

7.4 DELIVERY MECHANICS

As mentioned before, very few futures contracts lead to the delivery of an underlying asset because traders usually prefer to close out contracts before the delivery period. If traders wish to buy or sell the underlying asset, they can then do so in the spot market. However, it is important to note that the futures price is tied to the spot price by the *potential* for final delivery of the specified asset. Therefore, the mechanics of the delivery process are an important aspect of futures markets.

The time when a delivery can be made varies from contract to contract. The delivery process begins when a member with a short position issues a notice of intention to deliver to the exchange CCP. This notice indicates how many contracts will be delivered. For commodity contracts, the notice states where delivery will be made and (if applicable) what grade will be delivered. The exchange must then choose one or more parties with long positions to accept deliveries. Standard procedure is for exchanges to allocate delivery notices to members who have had net long positions for the longest period of time. However, sometimes delivery notices are allocated at random. Members cannot refuse to accept delivery notices, but sometimes they are given a short period of time to transfer their contracts to other members.

Delivery of commodities can involve warehousing costs. In the case of livestock, there may be costs associated with feeding and looking after the animals. There are stories of futures traders who inadvertently held long livestock positions during the delivery period and received notices of delivery from the exchange. They then ended up owning livestock in a warehouse several thousand miles away.[7]

The price to be paid for the asset is the most recent settlement price. In some cases, this is adjusted for grade, delivery location, (etc).

The *first notice day* is the first day when a notice to deliver can be submitted to the CCP. The *last notice day* is the last day when this can happen. The *last trading day* is generally a few days before the last notice day. For the December 2020 corn futures contract we considered earlier, delivery notices can be issued any time between November 30, 2020, and December 15, 2020, (with delivery happening one business day later). The last trading day (as mentioned earlier) is December 14, 2020.

[7] The delivery of financial assets can usually be handled electronically and does not give rise to these types of problems.

Cash Settlement

In theory, futures contracts could be designed so that they are settled in cash. Traders would then avoid the (sometimes inconvenient) delivery process. However, regulators do not like cash settlement because it makes futures contracts seem like gambling. As such, they prefer physical settlements (i.e., the delivery of the asset) to take place whenever possible.

However, the CME Group's futures contracts on the S&P 500 are settled in cash. This is because physical delivery would involve the delivery of a portfolio containing the 500 stocks underlying the S&P 500; this would be an expensive and inconvenient undertaking. Instead, all contracts are settled on the third Friday of the delivery month by a final exchange of variation margin. The final settlement price is the opening value of the S&P 500 on that day.

Examples of other cash settled contracts include those dependent on weather and real estate prices. The CME Group's popular Eurodollar futures contract (which will be discussed in a later chapter) is also settled in cash.

7.5 PATTERNS OF FUTURES PRICES

Table 7.2 shows settlement futures prices for gold on June 25, 2018. This is referred to as a *normal market* because the futures price increases as time to maturity increases. Specifically, the futures price increases from USD 1,265.6 (which is close to the spot price) for June 2018 delivery to USD 1,480.1 for delivery in December 2023. In the case of gold futures, delivery can take place on any day during the delivery month (with notice of intention to deliver being made one day earlier).

The CME Group futures for crude oil had a very different pattern on June 25, 2018. As shown by Table 7.3, the future prices declined from USD 68.08 to USD 53.57 as maturity increased. This is referred to as an *inverted market*.

Some assets have patterns that are partly normal and partly inverted. For example, soybean futures prices on June 25, 2018, increased for maturities out to July 2019 and then decreased. Furthermore, futures prices do not always have the same pattern. Oil futures, for example, sometimes exhibit a normal market rather than the inverted market as presented in Table 7.3. The determination of futures prices and whether a normal market or an inverted market is likely to be observed will be discussed in a later chapter.

Table 7.2 Settlement Prices for Gold Futures on June 25, 2018

Maturity Month	Settlement Price (USD per Ounce)
June 2018	1,256.6
July 2018	1,267.2
August 2018	1,268.9
October 2018	1,274.7
December 2018	1,280.9
February 2019	1,287.1
April 2019	1,293.2
June 2019	1,299.6
August 2019	1,306.0
October 2019	1,312.6
December 2019	1,319.2
February 2020	1,325.8
April 2020	1,332.2
June 2020	1,339.2
December 2020	1,359.6
June 2021	1,379.8
December 2021	1,399.9
June 2022	1,420.3
December 2022	1,440.7
June 2023	1,460.4
December 2023	1,480.1

Source: www.cmegroup.com

Table 7.3 Settlement Prices for Selected Crude Oil Futures Contracts on June 25, 2018

Maturity Month	Settlement Price (USD per Barrel)
August 2018	68.08
October 2018	66.18
December 2018	65.47
June 2019	63.36
December 2019	61.64
June 2020	60.16
December 2020	59.01
June 2021	57.91
December 2021	56.97
June 2022	56.15
December 2022	55.53
June 2023	54.91
December 2023	54.52
June 2024	54.08
December 2024	53.82
June 2025	53.60
December 2025	53.57

Source: www.cmegroup.com

7.6 MARKET PARTICIPANTS

Futures market participants can be classified as (a) those that solicit or accept orders to trade from retail and other clients, and (b) those that trade for their own account. The former are termed *futures commission merchants* or *introducing brokers*. Under the U.S. Commodity Exchange Act, they are subject to registration and minimum capital requirements. Futures commission merchants manage customer funds (including margin requirements), whereas introducing brokers do not. Futures commission brokers may or may not be members of an exchange. Those that are not exchange members must clear their clients' trades through members.

Market participants who trade for their own accounts are referred to as *locals*. While they are not typically members of an exchange, they do tend to have a close relationship with a member who clears their trades. Locals and the clients of futures commission brokers can be classified as *scalpers*, *day traders*, or *position traders*. Scalpers generally hold futures positions for a very short period (perhaps only a few minutes) before closing them out.[8] Day traders enter positions with the intention of closing them out during the same trading day and (like scalpers) hope to make a profit from short-term intra-day price movements. However, they are prepared to wait longer than scalpers before closing out. Finally, position traders have views on how the market will move over much longer periods of time and take positions to reflect those views.

7.7 PLACING ORDERS

Traders of futures and other securities can place many different types of orders.

[8] Some exchanges use market makers. These add liquidity to the market by always being prepared to quote a bid and an ask price. The trading activity of market makers is similar to that of scalpers.

Market Orders

The simplest order is a *market order*, which is a request to buy or sell futures (i.e., take a long or short position) as quickly as possible at the best available price. The disadvantage of a market order is that a trader may end up buying (selling) at a much higher (lower) price than expected.

Limit Orders

The main alternative to a market order is a *limit order*, where the trader specifies a price limit. A limit order can only be executed at the specified price or at a price more favorable to the trader. In the case of a buy order, the limit price is the maximum price at which the trader is prepared to buy; in the case of a sell order, the limit price is the minimum price at which the trader is prepared to sell. Limit orders usually remain in effect for one day, but traders can specify a longer or shorter period.

To see how limit orders work, suppose that a trader sees that the current futures price is USD 32.5. He or she could then put in a limit order to buy at USD 32.3, which would be executed only if the price declines slightly. If the limit is USD 32.6, however, the order is much more likely to be executed if the exchange operates on a price-time priority basis (i.e., it executes higher priced orders first).

Stop-Loss Order

In a *stop-loss order* (sometimes referred to as a *stop order*), the order becomes a market order once the asset reaches a specified or a less favorable price. Stop-loss orders (as the name implies) are orders that are designed to limit a trader's loss on a certain position.

For example, suppose that the price is currently USD 50 and that the trader has a short futures position. The trader could issue a stop-loss to buy at USD 52. This becomes a market order to buy as soon as the future price reaches USD 52. If the price continues to rise, the order might be executed at USD 53. Alternatively, if the price reaches USD 52 and then falls, the order could be executed at USD 51.

Stop-Limit Orders

Whereas a stop-loss order becomes a market order when the stop price is reached, a stop-limit order becomes a limit order. For a stop-limit order, two prices must therefore be specified: the stop price and the limit price. In the example just given, suppose that the stop price is USD 52 and that the limit price is USD 52.5. Once the price equals USD 52, the order becomes a limit order that is executed at a price of USD 52.5 or lower. Sometimes the stop price and the limit price are the same, in which case the order is called a *stop-and-limit order*.

Market-if-Touched Orders

A *market-if-touched* (MIT) order is an order that becomes a market order if a trade occurs at the specified price or a more favorable price. It is also known as a *board order* and is a way in which a trader can ensure that profits are taken if there are sufficiently favorable price movements. Consider again a trader with a short futures position when the price is USD 50. If the price specified in an MIT order is USD 45, the trader is indicating that he or she wants to take profits by exiting from the position as soon as the price reaches USD 45.

Discretionary Orders

A *discretionary order*, also called a *market-not-held order*, is an order that the broker can delay filling in hopes of getting a better price.

Duration of Orders

As mentioned, the default arrangement is that an order continues to exist for one trading day. If it is not filled by the end of the trading day, it is cancelled. Traders can specify other periods of time during which an order is active. At one extreme, a *fill-or-kill order* is an order that is automatically cancelled if it is not fully executed immediately (i.e., within a few seconds). At the other extreme, an *open order* or a *good-till-cancelled order* remains open for the remaining life of the futures contract unless it is cancelled by the trader.

7.8 REGULATION OF FUTURES MARKETS

Futures markets are regulated in different ways in different countries. In the U.S., futures markets are regulated by the Commodity Futures Trading Commission (CFTC). This government agency aims to ensure that futures markets are open, transparent, competitive, and financially sound. The CFTC is responsible for licensing individuals who offer their services to the public and handling complaints that are raised by futures markets participants. It also oversees the setting of position limits.

Some of the responsibilities of the CFTC have been delegated to the National Futures Association (NFA). This is a self-regulatory organization consisting of members who participate in futures markets. Its aim is to protect investors and to ensure that members meet their regulatory responsibilities. It monitors trading, resolves disputes, and takes disciplinary action when appropriate.

In earlier chapters, we have mentioned the regulatory changes in the OTC market introduced since the 2007–2008 crisis. These have led to increased responsibilities for the CFTC, which now ensures that standard OTC derivatives (e.g., swaps) are traded and cleared in accordance with the new rules.

7.9 ACCOUNTING

Normal accounting rules call for gains and losses from futures to be accounted for as they occur. For example, consider a gold mining company with a fiscal year ending in December. Suppose that it sells 200 two-year futures contracts on gold in June when the futures price is USD 1,300 per ounce. Each contract is for the sale of 100 ounces. Suppose further the following scenarios.

- In December of the first calendar year, the futures price is USD 1,240 per ounce.
- In December of the second calendar year, the futures price is USD 1,160 per ounce.
- The contract is closed out at USD 1,190 per ounce in June of the third calendar year.

The USD profit is then reported as follows.

First fiscal year: $(1{,}300 - 1{,}240) \times 100 \times 200 =$	1,200,000
Second fiscal year: $(1{,}240 - 1{,}160) \times 100 \times 200 =$	1,600,000
Third fiscal year: $(1{,}160 - 1{,}190) \times 100 \times 200 =$	−600,000

Futures are settled daily so that the cash corresponding to the profits is realized in the years in which the profits are accounted for. The valuation process is referred to as *marking to market*.

Realizing and accounting for gains and losses year-by-year when hedging could lead to an increase in reported earnings volatility, rather than a reduction in volatility as would be expected when performing hedging activities.

If the gold company is hedging gold that it expects to produce in two years, however, the contracts it has entered may qualify for hedge accounting. This provides an exception to the general rule we have just mentioned and allows the gain (or loss) from hedging transactions to be recognized at the same time as the loss (or gain) on the items being hedged. The Financial Accounting Standard Board (FASB) has issued FAS 133 and ASC 815 to explain when hedge accounting can and cannot be used by companies in the U.S. The International Accounting Standards Board (IASB) has similarly issued IAS 39 and IFRS 9.

If the gold producer we have considered qualifies for hedge accounting, the entire gain of USD 2,200,000 (= $(1{,}300 - 1{,}190) \times 100 \times 200$) will be realized in the third year. This is likely to be attractive to the company as it looks to reduce its earnings volatility.

The rules concerning the use of hedge accounting are strict. The hedge must be fully documented, for example, with the hedged item and the hedging instrument being clearly identified. The hedge must also be classified as *effective*, which means that an economic relationship that is not dominated by the effect of credit risk must exist between the hedging instrument and the hedged item.[9] The effectiveness of a hedge must be tested periodically.

Taxation raises similar issues for futures trading. In many jurisdictions, futures contracts are normally treated for tax purposes as though they are closed out at the end of the tax year. While hedging transactions in the U.S. are exempt from this rule, the definition of a hedging transaction for tax purposes is different from the definition used for accounting purposes. For tax purposes, a transaction is a hedging transaction if it is entered in the normal course of business primarily to reduce risk exposures. Given the different criteria, it is possible for a transaction to qualify as a hedge for tax purposes but not for accounting purposes.

7.10 FORWARDS COMPARED WITH FUTURES

Forward and futures contracts are similar in that both are agreements to buy or sell an asset in the future. However, there are some key differences. For example, most forward contracts are on foreign exchange or interest rates. In contrast, futures contracts are on a wide range of financial and non-financial assets. We gave examples of how forward contracts on foreign exchange can be used for hedging in Chapter 4. We will describe forward contracts on interest rates later in the book.

Whereas a futures contract is traded on an exchange, a forward contract is an over-the-counter product and is (in many situations) subject to more credit risk. There are several other differences. A futures contract is settled daily, while a forward contract is settled at the end of its life. Delivery of the underlying asset is relatively rare for futures contracts because traders usually close out their positions before the delivery period specified in the contract. In the case of a forward contract, delivery is

[9] IFRS 39 has been replaced by IFRS 9, which eases some the hedge accounting requirements. Similarly, ASC 815 eases some of the rules in FAS 133.

usually made. Closing out is not as easy as it is for a futures contract because a company with a forward contract must approach its counterparty and negotiate a close out.

Forward contracts usually specify a single delivery date. In contrast, futures contracts specify a period (sometimes a whole month) during which time delivery can be made. Because futures contracts are traded on an exchange, they are standardized financial products. Forward contracts have the advantage in that the delivery date can be chosen to meet the precise needs of the client.

SUMMARY

Futures allow parties to contract for the future delivery of a wide range of commodities and financial assets. They are designed so that there is very little credit risk and it is very easy for market participants to close out their positions. Because of these qualities, futures are very successful financial products and trade on exchanges throughout the world.

The specification of contracts is an important activity for a futures exchange. It is necessary for an exchange to define what can be delivered, where it can be delivered, and when it can be delivered. When there are choices, it is the party with the short position that chooses. If a contract is inadequately specified, it is likely to fail.

In general, the futures price of an asset is not equal to its spot price and instead converges to the spot price as the specified delivery period approaches. With normal markets, the futures price increases with the maturity of the contract. With inverted markets, the futures price decreases with the maturity of the futures contract. Some futures prices show patterns that are normal for some periods and inverted for others.

Many different types of orders can be placed in futures markets. The simplest is a market order, which indicates that the trade is to be executed as soon as possible at the best available price. A limit order specifies how high the price can be in the case of a buy order and how low it can be in the case of a sell order. Other orders indicate the circumstances in which a trader would want to cut losses after adverse price movements or take profits because of favorable movements.

The over-the-counter alternative to a futures contract is a forward contract. Unlike futures contracts, forward contracts have a single delivery date and are not settled daily. Forward contracts are not standardized and, unlike futures contracts, usually lead to delivery of the underlying assets.

We will discuss how futures are used for hedging in the next chapter. Later chapters will discuss the nature and pricing of commodity, foreign exchange, and interest rate futures.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

7.1 Name three of the largest ten futures and options exchanges in the world.

7.2 What is meant by the open interest of a futures contract?

7.3 How does the open interest change following a trade between two members who are both closing out positions?

7.4 What is the difference between a market order and a limit order?

7.5 Several different grades of corn can be delivered in a futures contract. Can the party with a long position who has entered into a particular contract specify which grade it wants?

7.6 Why do exchanges trade "mini" contracts?

7.7 List three types of futures contracts that are cash settled.

7.8 What is meant by (a) a normal market and (b) an inverted market?

7.9 What is the difference between a local and a futures commission merchant?

7.10 What is the difference between a scalper and a day trader?

Practice Questions

7.11 A futures contract is still trading as the delivery period is reached. What trades would you make if the settlement futures price is above the spot price?

7.12 Futures contracts are referred to by their delivery month. When during the delivery month can the actual delivery be made?

7.13 What would happen to a contract if its underlying product had an inadequately specified quality?

7.14 Explain the difference between a market-if-touched order and a stop-loss order.

7.15 If the options (concerning delivery times and what can be delivered) available to the party with a short position are increased, do you think the futures price increases or decreases? [*Hint*: An option always has a positive value to the holder of the option.]

7.16 Suppose that both the one-year forward and one-year futures price of the British pound is USD 1.3000. During the year the futures price decreases to USD 1.2000 and then rises to USD 1.4000 at the end of the year. The forward price is also 1.4000 at the end of the year. How are the timing of cash flows different for (a) a trader entering into a forward contract to buy 1 million British pounds in one year to (b) a trader entering into a futures contract to buy 1 million British pounds in one year.

7.17 Suppose that on a particular day there are 3,000 trades in a futures contract. Of the buyers in those 3,000 trades, 1,800 were closing out positions and 1,200 were taking new positions. Of the sellers in those 3,000 trades, 1,400 were closing out positions and 1,600 were taking new positions. What is the change in open interest during the day? Does it increase or decrease?

7.18 A company enters into a futures contract to buy 10,000 units of an asset for USD 50 per unit in two years. At the end of the first year, the futures price is USD 45. At the end of the second year, the futures price is USD 52. The contract is closed out during the third year when the futures price is USD 54. What is the accounting for the profit or losses from the trade each year if the company (a) uses hedge accounting and (b) does not use hedge accounting?

7.19 Explain the relationship between the Commodity Futures Trading Commission and the National Futures Association.

7.20 List six differences between futures and forwards.

ANSWERS

Short Concept Questions

7.1 Exchanges (as listed in Table 7.1) are the CME group, the National Stock Exchange of India, CBOE Holdings, B3, NASDAQ, Eurex, Moscow Exchange, Shanghai Futures Exchange, and Dalian Commodity Exchange.

7.2 The open interest is the number of contracts outstanding. It equals the number of long positions or equivalently the number of short positions.

7.3 The open interest goes down by one.

7.4 A market order is to be filled as quickly as possible at the best available price. A limit order specifies a limit on how high the price can be when buying or how low it can be when selling.

7.5 No. Traders with short positions choose what will be delivered. When issuing a notice of intention, a trader with a short position specifies which grade will be delivered.

7.6 Mini contracts are designed to attract retail traders who want to take relatively small positions.

7.7 The CME Group contracts on indices such as the S&P 500 and the NASDAQ 100 are cash settled, as is the Eurodollar futures contract. Contracts on real estate and the weather are also cash settled.

7.8 In a normal market, the futures price is an increasing function of maturity, whereas in an inverted market it is a decreasing function of maturity.

7.9 A local is a speculator trading for himself or herself. A futures commission merchant processes trades for clients.

7.10 Both take positions and close them out during the same day. A scalper is looking for very short-term trends in futures prices and often keeps a position open for only a few minutes. A day trader keeps the position open longer.

Solved Problems

7.11 Short futures, buy the asset, and make delivery. Delivery will be at the most recent futures settlement price.

7.12 This varies from contract to contract. Sometimes the delivery period is a single day, sometimes it is the whole month, and sometimes it is part of the month.

7.13 The contract would fail. When the delivery period is reached, shorts would deliver the cheapest version (lowest quality) of the product they could, and longs would be dissatisfied.

7.14 A stop-loss order becomes a market order when there is a bid or offer at a specified price or at a less favorable price. It is designed to limit losses. A market-if-touched order becomes a market order if there is a trade at the specified price or a more favorable price. It is designed to enable a trader to take profits in the event of sufficiently favorable price movements.

7.15 The more options that the party with a short position has the more attractive the contract is to that party. They are therefore prepared to agree to a lower delivery price. Options available to the party with a short position therefore tend to lower the delivery price.

7.16 Both traders will make a profit of USD 100,000. The forward trader will make all the profit at the end of the year.

The futures trader will have negative cash flows followed by positive cash flows because of daily settlement. On a present value basis, the forward trader will do better. Note that if the price had increased to USD 1.5000 and then reduced to USD 1.4000, the futures trader would do better on a present value basis.

7.17 The open interest goes down by 200. For purposes of calculating the change in open interest, the 1,400 sales that close out positions can be matched with 1,400 of the buys that also close out positions, which reduces open interest by 1,400. The 1,200 new long positions can be matched with 1,200 of the new short positions, which increases open interest by 1,200. The remaining 400 buys that closed positions and 400 sales that established new short positions offset each other in terms of open interest. Therefore, $-1{,}400 + 1{,}200 + 0 = -200$.

7.18 If the company is able to use hedge accounting, all the total profit of USD 40,000 ($(54 - 50) \times 10{,}000$) is realized in the third year. If the company is not able to use hedge accounting, there is a USD 50,000 loss in the first year ($(45 - 50) \times 10{,}000$), a USD 70,000 profit in the second year ($(52 - 45) \times 10{,}000$), and a USD 20,000 profit in the third year ($(54 - 52) \times 10{,}000$).

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

7.19 THE CFTC is tasked with regulating the futures market in the U.S., but some of its responsibilities have been delegated to the National Futures Association, which is a self-regulating organization of futures market participants.

7.20 Futures are available on a much wider range of assets than forwards. Futures are traded on an exchange, whereas forwards are traded OTC. Futures are standardized by the exchange, while forwards can be non-standard. Forwards have a single delivery date, whereas futures usually have a range of delivery dates. Futures are settled daily, whereas forwards are settled at the end of their life. Forwards are usually held to maturity, whereas futures are usually closed out before maturity. In the case of futures, credit risk is very low because the exchange requires initial and variation margin. Credit risk may be higher in the case of forwards.

8

Using Futures for Hedging

Learning Objectives

After completing this reading, you should be able to:

- Define and differentiate between short and long hedges and identify their appropriate uses.
- Describe the arguments for and against hedging and the potential impact of hedging on firm profitability.
- Define and calculate the basis, discuss various sources of basis risk, and explain how basis risks arise when hedging with futures.
- Define cross hedging and compute and interpret the hedge ratio and hedge effectiveness.
- Calculate the profit and loss on a short or a long hedge.
- Compute the optimal number of futures contracts needed to hedge an exposure and explain and calculate the "tailing the hedge" adjustment.
- Explain how to use stock index futures contracts to change a stock portfolio's beta.
- Explain how to create a long-term hedge using a stack-and-roll strategy and describe some of the risks that arise from this strategy.

Like other derivatives, futures can be used for either speculation or hedging. If a trader has exposures to variables such as exchange rates, interest rates, equity indices, and commodity prices, a position in futures contracts can reduce that exposure. But if the trader has no exposure, that same position is speculative. In this chapter, we focus on the use of futures for hedging.

Eliminating all risk exposures using futures is usually impossible. It is therefore important to develop a way of calculating an optimal hedge (i.e., a hedge that reduces risk as much as possible). Initially, we will treat futures as forwards and assume that the hedge position is left untouched for the duration of the exposure. Later, we discuss the impact of daily settlement on the construction of a hedge and discuss how long-term hedges can be created from positions in a series of short-term futures contracts.

8.1 LONG AND SHORT HEDGES

We start by providing simple examples of situations where short and long futures positions can be used as hedges.

Short Hedge

A short futures position is appropriate in the following situations.

- A company owns a certain quantity of an asset and knows that it will sell it at a certain time in the future.
- A company knows that it will receive a certain quantity of an asset in the future and plans to sell it.

For example, consider a company that will receive 2 million barrels of crude oil in three months. Assume that it plans to sell the oil as soon as it is received and knows that it will lose (gain) USD 20,000 (= 2,000,000 × USD 0.01) from each one-cent decrease (increase) in the price of oil. The company regards this risk as unacceptable.

Oil futures contracts traded by the CME Group are for the purchase or sale of 1,000 barrels of oil. The company can therefore hedge its position by taking a short futures position in 2,000 (= 2,000,000/1,000) three-month oil futures contracts. A hedge that involves taking a short futures position is known as a *short hedge*.

We assume that the oil is received during the delivery period of the futures contracts. As explained in the previous chapter, the futures price should equal (or be very close to) the spot price during this period. The company can either deliver the oil under the terms of the futures contract or close out the futures contracts and sell the oil in the usual manner.

This strategy locks in the price received for the oil at the current three-month futures price. To illustrate this, suppose that the current spot price of oil is USD 59.50 per barrel and the three-month futures price is USD 60.00 per barrel. Consider two situations.

1. The spot price of oil at the time it is delivered equals USD 50.
2. The spot price of oil at the time it is delivered equals USD 68.

In the first situation, the price received for the oil in the market in USD is

$$2{,}000{,}000 \times 50 = 100{,}000{,}000$$

while the gain (USD) on the 2,000 short futures contracts is

$$2{,}000 \times (60 - 50) \times 1{,}000 = 20{,}000{,}000$$

Combining the price received for the oil in the market with the gain on the 2,000 futures contracts, the net price received is USD 120 million (= 100,000,000 + 20,000,000). This corresponds to USD 60 (= 120,000,000/2,000,000) per barrel.

In the second case, where the spot price in three months equals USD 68 per barrel, the price received in the market for the oil in USD is

$$2{,}000{,}000 \times 68 = 136{,}000{,}000$$

However, there is a USD loss on the short futures position of:

$$2{,}000 \times (68 - 60) \times 1{,}000 = 16{,}000{,}000$$

When this loss is considered, the net price received for the oil is again USD 120 million (= 136,000,000 − 16,000,000), or USD 60 per barrel. Table 8.1 shows the result of the strategy we have considered for a range of oil prices. Note that the hedger locks in the initial futures price (USD 60) rather than the initial spot price (USD 59.50).[1]

By assuming that the hedger locks in the futures price, we assumed that the company can either:

- Deliver the oil under the terms of the futures contract as soon as it receives it, or
- Close out the futures contract at the spot price when the oil is received with the oil then being sold in the usual manner.

In practice, futures hedges rarely work as they do in this idealized scenario (as we will discuss in later sections).

Long Hedges

A *long hedge* is the opposite of a short hedge and it can be used when a company knows it will have to buy a certain quantity of an asset in the future.

[1] Because there is no upper bound to the price of oil in three months, the potential loss on the short futures contract is unlimited.

Table 8.1 The Impact of Hedging the Sale of 2 Million Barrels of Oil: In all Cases the Price Received for the Oil, when Adjusted for the Gain or Loss on the Futures Position, Is USD 120 Million

Price of Oil in Three Months (USD per Barrel)	Price Received for Oil (USD Millions)	Gain on Short Futures Position (USD Millions)
20	40	80
30	60	60
40	80	40
50	100	20
60	120	0
70	140	−20
80	160	−40
90	180	−60
100	200	−80

Consider a company that will need to purchase 150,000 pounds of copper in two months and wants to hedge its price risk. The spot price of copper is USD 3.00 per pound and the two-month futures price is USD 2.90 per pound. One futures contract traded by the CME Group is on 25,000 pounds of copper. The company can therefore hedge its risk by taking a long position in 6 (= 150,000/25,000) two-month copper futures contracts. Consider two situations.

1. The spot price of copper in two months equals USD 3.30 per pound.
2. The spot price of copper in two months equals USD 2.70 per pound.

If the spot price of copper becomes USD 3.30, the 150,000 pounds of copper can be purchased in the market for:

$$150{,}000 \times \text{USD } 3.30 = \text{USD } 495{,}000$$

If we assume that the copper is purchased when the futures price equals the spot price, the futures price can be closed out at the time of purchase. There will then be a gain on the six copper futures contracts of:

$$6 \times 25{,}000 \times (\text{USD } 3.30 - \text{USD } 2.90) = \text{USD } 60{,}000$$

This reduces the net cost of the copper purchased to USD 435,000 (= 495,000 − 60,000), or USD 2.90 per pound.

Consider next the situation where the spot price of copper in two months is USD 2.70. The cost of the copper purchased in the market is

$$150{,}000 \times \text{USD } 2.70 = \text{USD } 405{,}000$$

In this case there is a loss on the futures contracts of:

$$6 \times 25{,}000 \times (\text{USD } 2.90 - \text{USD } 2.70) = \text{USD } 30{,}000$$

This increases the cost of the copper purchase to USD 435,000 (= 405,000 + 30,000), or USD 2.90 per pound.
As in the case of the short hedge, we see that the long hedge locks in a price equal to the current futures price of USD 2.90 per pound. Table 8.2 shows the results for other copper prices.

Table 8.2 The Impact of Hedging the Purchase of 150,000 Pounds of Copper: In all Cases, when it is Adjusted for the Gain or Loss on Futures Contracts, the Price Paid is USD 435,000

Price of Copper in Two Months (USD per Pound)	Price Paid for Copper (USD)	Gain on Long Futures Position (USD)
2.00	300,000	−135,000
2.20	330,000	−105,000
2.40	360,000	−75,000
2.60	390,000	−45,000
2.80	420,000	−15,000
3.00	450,000	15,000
3.20	480,000	45,000
3.40	510,000	75,000
3.60	540,000	105,000
3.80	570,000	135,000

One alternative to using futures is to buy the copper in the spot market today. There are two reasons why the hedger would probably find this less attractive.

1. In this example, copper is more expensive in the spot market than in the futures market.[2]
2. The hedger would bear the cost of financing the purchase and storing the copper for two months.

If the futures price were higher than the spot price, buying the copper today could be more attractive to the hedger. In that case, however, the financing and storage costs would be at least as great as the difference between the cost of the asset in the futures market and the cost of the asset in the spot market. Otherwise, any trader could execute a simple arbitrage strategy:

- Buy the asset in the spot market and store it, and
- Sell the asset in the futures market.

8.2 PROS AND CONS OF HEDGING

As we have just shown, hedging can reduce the risk arising from changes in asset prices. Hedging can help firms reduce the volatility of their earnings and potentially make themselves more attractive to investors.

In practice, however, many companies do not hedge. Here we consider some of the reasons.

Shareholders May Prefer No Hedging

It is sometimes argued that there is no reason for a company to hedge its risks because shareholders can do their own hedging. By not hedging, a company allows its shareholders to decide whether they want to take on a particular risk.

An argument in favor of leaving things to the shareholders is that they typically invest across multiple companies. Shareholders can diversify risks by choosing a portfolio of companies in different industries and operating in different geographical regions. This diversification can substantially reduce risks that might otherwise be hedged. For example, while one company in an investor's portfolio might suffer from a decrease in oil prices, another company in the same investor's portfolio might gain from such a price change.

A related point is that firms can be tempted to diversify by entering new lines of business or buying other companies. However, a company should always consider whether its shareholders can do this type of diversification more easily than the company can. Arguably, a company should diversify outside its traditional areas of expertise only when there is synergy (i.e., when it is combining two or more different business activities such that the value of the whole is greater than that of the sum of the parts).

Two arguments against leaving hedging to shareholders are the following.

1. The shareholders of a public company are less informed than management about the risks being taken by the company. Therefore, the company's managers are in the best position to assess risks and hedge them.
2. Shareholders that do understand the risks faced by the company may find it difficult to hedge because their positions may require only a small fraction of one futures contract.

Whether a firm chooses to hedge or not, its hedging strategy should be set by its board and clearly communicated to its shareholders so that they know and understand the risks they are taking.

For example, consider gold mining companies. It can take several years to develop a mine for extraction. Given that the price of gold can move adversely during that period, some gold mining companies choose to hedge the price that will apply to their future production. On the other hand, some firms choose not to hedge. In either case, most gold mining companies are careful to explain their hedging strategies to their shareholders. If investors want an exposure to the price of gold, they will buy the shares of gold producers that do not hedge. If investors do not want this exposure, they will buy the shares of companies that choose to hedge.

There May Be Little or No Exposure

In determining the size of an exposure, it is important to consider a company's complete risk profile.

As an example, consider a jewelry manufacturer that produces 24 carat gold jewelry and buys 100 ounces of gold every two months. Because the firm is exposed to the price of gold, it chooses to lock in its purchases for the next two years with a series of long futures contracts.

However, suppose that an analysis of available data shows that economic pressures cause the wholesale price of jewelry to reflect the price of the gold it contains. When the price of gold increases (decreases), there is a corresponding increase (decrease) in the price of gold jewelry. This would mean that if the price of gold declines during the two-year period, the manufacturer would lose money on the hedges and the expected improvement to the manufacturer's gross margin would not materialize. Similarly, if the price of gold increases during the two-year period, neither the manufacturer's gains on the hedges nor the expected worsening of the manufacturer's gross margin would materialize.

[2] This is not always the case.

The result of the "hedging" strategy followed by the jewelry manufacturer is to increase, rather than reduce, risk.

Of course, it can be argued that the demand for jewelry may be affected by the price of gold and that this might justify a long gold hedge. However, the size of the hedge that is necessary for that purpose will be quite different from the size necessary to hedge the jewelry manufacturer's gold purchases. If the firm overestimates its exposure and takes a long position in ten gold futures contracts when its true exposure requires only one contract, the extra nine contracts are unintended speculation. If the price of gold declines, the hedger will incur a loss.

Hedging May Lose Money

It is often assumed that the purpose of hedging is to increase profits. This is incorrect. The purpose of hedging is to reduce the *variability* of profits. The outcome with hedging will sometimes be worse than the outcome without hedging, whereas other times it will be better. However, the outcome should always be more certain as a result of hedging.

The fact that hedging can sometimes lead to losses can make some corporate treasurers reluctant to hedge. For example, suppose that a treasurer working for an oil producer sells oil in the futures market to lock in the price received for the company's future production. If the spot price of oil decreases, the hedging will result in increased profits. However, we can only expect this to happen in some scenarios. Under other scenarios, the price of oil will increase and the hedging will result in decreased profits. The treasurer may then get criticized (or worse) because of these losses.

The reluctance to make a decision that could adversely affect profits has led some treasurers to prefer the use of options for hedging. Instead of locking in a price, options provide insurance. If the treasurer bought put options on the future price of oil, the company would be protected in the event of fall in price and yet still be able to benefit from a sharp increase in price. Of course, this result is not achieved without a cost. As explained in Chapter 4, options require a premium to be paid upfront by the purchaser.

8.3 BASIS RISK

In the examples we looked at in Section 8.1, we made several simplifying assumptions. In particular:

- We assumed that the asset underlying the futures contract is the same as the asset price to which the hedger is exposed.
- We assumed that the futures price and the spot price are equal at the time the hedge is closed out.

While the first assumption may be reasonable in some hedging situations, the second assumption is rarely true.

As mentioned in the previous chapter, most futures positions are closed out prior to the delivery period specified in the contract and the market for futures contracts can be unreliable during the delivery month. With these facts in mind, a sensible rule of thumb for hedgers is that the maturity for a futures contract should be the earliest possible month after the maturity of the desired hedge.

For example, consider a futures contract with maturity months in March, May, July, September, and December. The March contract would be used for December, January, and February exposures; the May contract would be used for March and April exposures; and so on.

However, closing out a futures contract before maturity exposes the hedger to what is called *basis risk*.

The basis for a futures contract at a given time is defined by:[3]

$$\text{Basis} = \text{Spot Price} - \text{Futures Price}$$

If the asset being hedged is not the same as the asset underlying the futures contract, we can extend this definition to:

$$\text{Basis} = \text{Spot Price of Hedged Asset} - \text{Futures Price of Asset Underlying Futures Contract}$$

Basis risk is the risk associated with the basis at the time a hedge is closed.

Suppose a hedger is due to sell an asset in the future and therefore implements a short hedge (similar to that of the oil producer in Section 8.1). Defining terms:

- F_0: Futures price at the time the hedge is initiated,
- F_t: Futures price at the time the hedge is closed,
- S_t: Spot price of asset being hedged at the time the hedge is closed, and
- b_t Basis at time t $(= S_t - F_t)$.

At time t, the price received for the asset is S_t and the gain on the short position is $F_0 - F_t$. It follows that:

$$\text{Net price received when a short hedge is used} = S_t + (F_0 - F_t)$$

This can be written as:

$$\begin{aligned}\text{Net price received when short hedge is used} &= F_0 + (S_t - F_t)\\ &= F_0 + b_t\end{aligned}$$

In the oil example in Section 8.1, we assumed that the spot price was the same as the futures price at the time the hedge

[3] Note that sometimes the alternative definition: Basis = *Futures Price* − *Spot Price* is used, particularly when the futures contract is on a financial asset. However, we will use the Basis = Spot Price − Futures Price definition throughout.

was closed out. This meant that the basis was zero at the time of the close out and the net price received for the oil was always the futures price at hedge initiation (i.e., F_0). When the hedge is closed before the delivery period and/or the crude oil underlying the futures is different from the crude oil that will be delivered, then b_t is uncertain and the hedger is subject to basis risk.

Consider next a long hedge used when an asset is going to be purchased in the future (e.g., position of the manufacturer in Section 8.1 that knows it will need to purchase copper in the future). At time t, the price paid for the asset is S_t and the gain on the long futures is $F_t - F_0$. This means that:

Net cost of asset when long hedge is used

$$= S_t - (F_t - F_0)$$

This can be written as:

Net cost of asset when long hedge is used

$$= F_0 + (S_t - F_t)$$
$$= F_0 + b_t$$

This shows that the net price received when a short hedge is implemented for the future sale of an asset is the same as the net price paid when a long hedge is implemented for the future purchase of an asset. Both are equal to $F_0 + b_t$.

The futures price (F_0) is known at the time the hedge is initiated. The uncertainty about the performance of the hedge is therefore due entirely to the uncertainty about the future basis (b_t). This uncertainty tends to be greater for futures on commodities than for futures on financial instruments.

As a first example of basis risk, consider a U.S. company that is planning to hedge an impending purchase of 250,000 British pounds (GBP) in February by using the CME Group's March futures contract. Given that each contract is on GBP 62,500, the company needs a long position in four contracts.

Suppose that the futures price at the time the hedge is initiated is 1.25 USD per GBP and the futures price at the time of the close out in February is 1.30 USD per GBP. However, the spot price in February is 1.31 USD per GBP. After the gain from hedging has been considered, the cost (in USD) of the British pounds is

$$250{,}000 \times 1.31 - (1.30 - 1.25) \times 4 \times 62{,}500 = 315{,}000$$

This indicates that the hedger's net exchange rate is USD 1.26 (= 315,000/250,000) per GBP. The situation is as follows:

- The exchange rate increases and so hedging improves (reduces) the price paid (compared to the spot price) by 0.05 USD per GBP.
- However, the basis is USD 0.01 (= 1.31 − 1.30). This means that while the hedger can cash out the futures contracts at USD 1.30, it must buy the underlying asset at USD 1.31. With no basis, the hedger would pay the initial futures price of USD 1.25 (= 1.30 − 0.05). Instead, the basis increases this to USD 1.26 (= 1.31 − 0.05).

As a second example, suppose that a hedger plans to sell 50,000 bushels of corn in June and uses the July futures contract for hedging. Each contract is on 5,000 bushels of corn and therefore a total of ten contracts are shorted. Suppose that the futures price at the time the hedge is initiated is 300 cents per bushel and that the futures price is 320 cents per bushel when the hedge is closed out in June. The spot price for the corn being sold in June (which may be a different type of corn from that underlying the futures contract) is 325 cents per bushel. The USD price received for the corn when the loss from hedging is considered is

$$50{,}000 \times 3.25 + 10 \times 5{,}000 \times (3.00 - 3.20) = 152{,}500$$

This shows the net price received is 305 (= 152,500/50,000) cents per bushel. The situation is as follows:

- The price of corn increases, so the corn producer loses 20 (= 320 − 300) cents per bushel when it closes the futures contracts.
- However, the basis is 5 (= 325 − 320) cents per bushel. When the hedger sells the corn, it does so at the spot price of 325 cents per bushel. Combined with the 20 cents per bushel loss on the hedge, the net price received is 305 cents per bushel. This more than the initial futures price of 300 cents per bushel.

These examples emphasize that the purpose of hedging is to make the outcome more certain. However, basis risk means that there are still some uncertainties associated with hedge.

8.4 OPTIMAL HEDGE RATIOS

In the examples presented so far, we have assumed that futures positions equal the exposures being hedged. Thus, when 50,000 bushels of corn were being sold in an earlier example, we assumed that the hedger would enter into futures contracts for delivery of a total of 50,000 bushels. In this section, we explore situations where the asset underlying the futures contract is different from the asset being hedged.

The *hedge ratio* is the ratio of the position in the futures contract to the position in the underlying asset.

Hedging an exposure to the price of one asset with a futures position in another asset is referred to as *cross hedging*. Cross hedging can be analyzed by considering the relationship between the following:

ΔS: Change in spot price during a period equal to the life of the hedge, and

ΔF: Change in futures price during a period equal to the life of the hedge.

Suppose that historical data, used in conjunction with linear regression, shows that the best fit linear relationship between ΔS and ΔF is

$$\Delta S = a + b\Delta F + \varepsilon$$

where a and b are constants and ε is an error term. Suppose further that h is the hedge ratio. The change in the value of the position per unit of the asset being hedged is then:

$$\Delta S - h\Delta F = a + (b - h)\Delta F + \varepsilon$$

The variance of the right-hand side is minimized by setting $h = b$ (so that the second term is zero).

Defining terms:

ρ: Coefficient of correlation between ΔS and ΔF,

σ_S: Standard deviation of ΔS,

σ_F: Standard deviation of ΔF, and

h^*: Optimal hedge ratio.

An expression for b is[4]

$$b = \rho \frac{\sigma_S}{\sigma_F}$$

It follows that this is also an expression for the optimal hedge ratio h^*:

$$h^* = \rho \frac{\sigma_S}{\sigma_F} \tag{8.1}$$

The *hedge effectiveness* is the proportion of the variance in ΔS that is eliminated by the hedging. This is usually referred to as the R^2 of the regression and equals ρ^2 in the case of linear models with a single variable.

If we have perfect correlation ($\rho = 1$) and $\sigma_S = \sigma_F$, then $h^* = 1$. If we have perfect correlation but σ_S is 20% higher than σ_F, then $h^* = 1.2$. This is also as expected because spot price changes are always 20% greater than futures price changes.

The parameters, ρ, σ_S, and σ_F, are estimated from historical data on ΔS and ΔF. Ideally, the time period over which changes are measured should be equal to the life of the hedge. In practice, shorter time periods are often used to increase the number of observations that can be used in generating estimates.

Defining terms:

Q_A: Size of the position being hedged (units),

Q_F: Number of units of the asset underlying one futures contract, and

N^*: Optimal number of futures contracts for hedging.

It follows that:

$$N^* = \frac{h^* Q_A}{Q_F} \tag{8.2}$$

A common example of cross hedging is related to the use of heating oil futures by airlines to hedge jet fuel. Note that while jet fuel futures do exist, they do not trade as actively as heating oil futures. As a result, airlines often hedge using heating oil futures instead.

Suppose that an airline estimates that the correlation between monthly changes in heating oil futures prices and jet fuel prices is 0.9. The standard deviation of monthly changes in heating oil futures price per gallon is USD 0.03, while the price of jet fuel per gallon is USD 0.025. The optimal hedge ratio from Equation (8.1) is

$$h^* = 0.9 \times \frac{0.025}{0.03} = 0.75$$

This indicates that the optimal hedge ratio is 75%.

Suppose that the airline wishes to hedge the purchase of 1 million gallons of jet fuel in one month. Each heating oil futures contract is on 42,000 gallons of heating oil. The number of long contracts required for hedging can be obtained using Equation (8.2):

$$N^* = \frac{0.75 \times 1{,}000{,}000}{42{,}000} = 17.86$$

or 18, when rounded to the nearest whole number.

Tailing the Hedge

The analysis presented so far is correct when hedging is carried out with forward contracts. A small adjustment to the analysis known as *tailing the hedge* is (in theory) necessary when futures contracts are used.[5] This recognizes that futures are settled daily so that the hedger is effectively implementing a series of one-day hedges. The hedge ratio for a particular day should be given by Equation (8.1), where the correlation ρ is the correlation between one-day changes while σ_S and σ_F are the standard deviations of one-day changes.

---

[4] Recall that the slope coefficient in the linear model is given by $b = \frac{cov(\Delta S, \Delta F)}{\sigma^2_{\Delta F}} = \frac{\rho \sigma_{\Delta S}\sigma_{\Delta F}}{\sigma^2_{\Delta F}} \rightarrow b = \rho \frac{\sigma_{\Delta S}}{\sigma_{\Delta F}}$

[5] The term, "tailing the hedge," comes from the fact that, when it is recognized that a hedge using futures is a series of one-day hedges, the hedge should in theory be adjusted (tailed) with the passage of time.

A related point is that analysts often work with the standard deviation of one-day returns (i.e., daily volatilities) rather than the standard deviation of price changes. This is because the former can be expected to be less variable. Define:

$\hat{\sigma}_S$: Standard deviation of the one-day return in spot price (i.e., the percentage change in spot price);

$\hat{\sigma}_F$: Standard deviation of the one-day return provided by futures price (i.e., the percentage change in futures price);

$\hat{\rho}$: Correlation between the one-day spot return and the futures return;

S: Spot price;

F: Futures price;

V_A: Value of position being hedged, which equals SQ_A where (as above) Q_A is the number of units of the assets being hedged; and

V_F: Value of one futures contract, which is defined as FQ_F, where Q_F is the number of units of the asset underlying one futures contract.

The standard deviation of the daily spot price change is $\hat{\sigma}_S S$, and the standard deviation of the daily futures price change is $\hat{\sigma}_F F$ (both in absolute terms). This means that the optimal hedge ratio given by Equation (8.1) can be shown as:

$$h^* = \hat{\rho}\frac{\hat{\sigma}_S S}{\hat{\sigma}_F F}$$

Meanwhile, applying Equation (8.2) for a one-day hedge gives the optimal number of contracts:

$$N^* = \hat{\rho}\frac{\hat{\sigma}_S S}{\hat{\sigma}_F F}\frac{Q_A}{Q_F} = \hat{\rho}\frac{\hat{\sigma}_S}{\hat{\sigma}_F}\frac{V_A}{V_F}$$

The hedge ratio is sometimes also defined as:

$$\hat{h} = \hat{\rho}\frac{\hat{\sigma}_S}{\hat{\sigma}_F} \tag{8.3}$$

This leads to:

$$N^* = \hat{h}\frac{V_A}{V_F} \tag{8.4}$$

Note the differences between Equations (8.1) and (8.3). In Equation (8.1), σ_S and σ_F are the standard deviations of the change in S and F (respectively) over the life of the hedge, while ρ is the correlation between these changes. In Equation (8.3) $\hat{\sigma}_S$ and $\hat{\sigma}_F$ are the standard deviation of daily returns in S and F (respectively), while $\hat{\rho}$ is the correlation between these daily returns.

In theory, the following are true.

- The values of V_A and V_F are appropriate for successive one-day hedges. They change as the price of the asset and the futures price of the contract change.
- It is necessary to allow for the time difference between the settlement of each one-day hedge and the maturity of the futures contract. This means a discount factor must be applied to the hedge ratio calculated for each day.

When applying Equation (8.4), V_A/V_F is usually assumed to be constant at its initial value and discount factor adjustments are not usually made.

As an example, suppose that the standard deviations of the daily returns in the futures price and the spot price are 1% and 1.2% (respectively) while the correlation between the two is 0.88. The value of the assets being hedged is 1 million USD, while the value of one futures contract is USD 20,000. In this case, $V_A = 1{,}000{,}000$ and $V_F = 20{,}000$. From Equation (8.3):

$$\hat{h} = 0.88 \times \frac{0.012}{0.01} = 1.056$$

The optimal number of contracts is from Equation (8.4):

$$1.056 \times \frac{1{,}000{,}000}{20{,}000} = 52.8$$

or 53, when rounded to the nearest whole number.

8.5 HEDGING EQUITY POSITIONS

Stock index futures are popular products that can be used to change an investor's exposure to the stock market. For example, an investor who believes that the market will be particularly volatile during the next three months might want to temporarily decrease or eliminate exposure to the market. On the other hand, an investor who is bullish about the market's prospects over the next three months might want to use futures markets to temporarily increase exposure.

Stock index futures can be a way of increasing or reducing exposure to the market without incurring significant transaction costs. Take the example of an investor who is good at picking stocks but has no views on the future direction of the market. The investor might want to form a portfolio and then hedge the return from the market so that he or she is left with the excess return of the chosen stocks over the return from the market.

To see how this could work, suppose an investor has a portfolio of USD 1 million that is well diversified and closely tracks the S&P 500. The hedger wants to have no exposure to the market for the next two months and decides to use the mini CME futures contract on the S&P 500. This is on USD 50 multiplied by the index.[6]

[6] The regular contract is on 250 times the index, but the investor might choose the mini contract because it trades more actively and allows positions to be hedged more precisely.

In this case, the standard deviation of the daily return on the futures price and the standard deviation of the daily return of the asset being hedged can be assumed to be the same. Also, the correlation between the daily returns can be assumed to be 1.0. Equations (8.3) and (8.4) therefore give the optimal hedge ratio as:

$$N^* = \frac{V_A}{V_F}$$

If the futures price is USD 2,500, the number of contracts that should be shorted is therefore:

$$N^* = \frac{1{,}000{,}000}{50 \times 2{,}500} = 8$$

Managing Beta

The beta (β) of a portfolio is the sensitivity of its return to the return of the market portfolio. If a portfolio has a beta of 1.0, it mirrors what the market does. If the portfolio has a beta of 0.5, it is half as volatile as the market. When the beta is 2.0, it is twice as volatile as the market. For an investor in U.S. stocks, we can define the market portfolio as the S&P 500 Index.

The capital asset pricing model (CAPM) relates the expected return on a portfolio to its beta. The model states that:[7]

$$E(R_P) - R_F = \beta(R_M - R_F) \tag{8.5}$$

where $E(R_P)$ is the expected return on a portfolio, R_F is the risk-free rate, R_M is the return on the market portfolio, and β is the beta of the portfolio. The parameter beta magnifies the excess return over the risk-free rate. Therefore, a portfolio with a beta of 1.5 has a 50% higher excess return over the risk-free rate than the market.

Suppose that the risk-free rate is 3%. If the market return is 7%, the excess return over the risk-free rate is 4%. For a portfolio with a beta of 1.5, the expected excess return over the risk-free rate is therefore 6% (= 1.5 × 4%), and the total expected return of the portfolio is 9%. If the return on the market is −3%, however, then the difference from the risk-free rate is −6% (= −3% − 3%), and therefore the expected portfolio return relative to the risk-free rate is −9% (= −6% × 1.5). This means that the expected return on the portfolio is −6% (= −9% + 3%).

Suppose we wish to use S&P 500 futures to hedge a well-diversified portfolio with a beta of β. In this case, the standard deviation of the daily return for the asset being hedged can be assumed to be β multiplied by the standard deviation of the return provided by the futures price. The correlation between the daily returns on the asset and the futures is approximately 1.0. Therefore, equations (8.3) and (8.4) give the optimal hedge ratio as:

$$N^* = \beta \frac{V_A}{V_F}$$

Consider a portfolio worth USD 1 million when the index futures price for a six-month contract is USD 2,500. If the beta of the portfolio is 1.25, the number of contracts required to hedge the portfolio is

$$N^* = 1.25 \times \frac{1{,}000{,}000}{50 \times 2{,}500} = 10$$

We will explain in a later chapter that the futures price of an index should equal its spot price compounded forward at the excess of the risk-free rate over the dividend yield for the life of the futures contract. Suppose that the hedge lasts for six months, the risk-free rate is 4% per year and the dividend yield on the index is 2% per year. In this case, the initial futures price is the initial spot price compounded forward for six months at 2% (= 4% − 2%) per year (i.e., it is about 1% (= 0.5 × 2%) higher than the spot price). Because the futures price of the index is USD 2,500, the spot price must be about USD 2,475 (= 2,500/1.01).

Consider the situation where the index level at the end of six months is USD 2,300. The percentage return on the market consists of a capital loss of −7.07% (= −175/2,475) and a dividend gain of 1% (half of the annual dividend yield). The total return on the index is therefore −6.07%. From the capital asset pricing model in Equation (8.4), the expected return of the portfolio during the six-month period is

$$0.02 + 1.25 \times (-0.0607 - 0.02) = -0.0809$$

or −8.09%.[8] The value of the portfolio is therefore expected to decrease from USD 1,000,000 to USD 919,100 (= 1,000,000 × (1 − 0.0809)). The USD gain on the futures position should be

$$10 \times 50 \times (2{,}500 - 2{,}300) = 100{,}000$$

This would bring the value of the portfolio up to USD 1,019,100.

Table 8.3 summarizes this calculation and shows similar calculations for other values of the S&P 500 in six months. It can be seen that the hedge works very well. The total value of the hedger's position is close to 1,020,000 in six months in all scenarios. Note that this is the value of the portfolio if it had been invested in a risk-free asset.[9]

[7] For an explanation of the arguments leading to CAPM see J. Hull, "Risk Management and Financial Institutions," 5th edition, 2018.

[8] Note that the risk-free rate for six months is half the risk-free rate per year.

[9] The table ignores the impact of daily settlement.

Table 8.3 Performance of Hedge of a Portfolio with a Beta of 1.25

Index Level in Six Months	2,100	2,300	2,500	2,700	2,900
Capital Gain	−15.15%	−7.07%	1.01%	9.09%	17.17%
Dividend	1.00%	1.00%	1.00%	1.00%	1.00%
Total Index Return	−14.15%	−6.07%	2.01%	10.09%	18.17%
Portfolio Return	−18.19%	−8.09%	2.01%	12.11%	22.21%
Value of Portfolio in Six Months	818,100	919,100	1,020,100	1,121,100	1,222,100
Gain on Futures	200,000	100,000	0	−100,000	−200,000
Final Value	1,018,100	1,019,100	1,020,100	1,021,100	1,022,100

Our example assumes that the investor wants to totally hedge the risks in a portfolio with a beta of 1.25 (i.e., the hedger wants to reduce the portfolio beta to zero). Stock index futures can also be used to modify beta.

Suppose that the current beta is β and the desired beta is β^*. When $\beta > \beta^*$, the number of futures contracts the investor should short is

$$(\beta - \beta^*)\frac{V_A}{V_F}$$

If the hedger we have just considered wants to reduce beta from 1.25 to 0.5, for example, then 6 (= 1.25 − 0.5 × [1,000,000/(50 × 2,500)]) contracts should be shorted.

When $\beta < \beta^*$, a long position of

$$(\beta^* - \beta)\frac{V_A}{V_F}$$

contracts is required. For example, a long position in six contracts would increase beta to 2.0 (because 6 = (2 − 1.25) × [1,000,000/(50 × 2,500)]).

8.6 CREATING LONG-TERM HEDGES

Sometimes hedgers are faced with a lack of liquid (i.e., actively traded) futures contracts for the required hedge maturities. Because the most liquid futures contracts are those with relatively short maturities, a hedger can work around this issue by following what is termed a *stack and roll* strategy. This involves

- Implementing a short-maturity futures hedge,
- Closing the hedge out just prior to the delivery period and replacing it with another short-maturity futures hedge, and
- Closing the new hedge out just prior to the delivery period and replacing it with yet another short-maturity futures hedge, and so on.

For example, suppose that it is currently January of Year 1 and a company knows that it will have to sell 5,000 ounces of gold in 18 months. However, it finds that the only contracts with sufficient liquidity are those with maturities of seven months or fewer. The CME Group's gold futures contracts trade with maturities in the months of February, April, June, August, October, and December. Because one contract is on 100 ounces of gold, a possible hedging strategy is as follows.

January, Year 1:	Sell 50 futures contracts maturing in August of Year 1.
July, Year 1:	Close out the futures maturing in August of Year 1 and sell 50 futures contracts maturing in February of Year 2.
January, Year 2:	Close out the futures maturing in February of Year 2 and sell 50 futures contracts maturing in August of Year 2.
July, Year 2:	Close out contracts maturing in August of Year 2.

Suppose the following.

- The August of Year 1 contracts are sold at USD 1,500 and closed out at USD 1,450 in July of Year 1.
- The February of Year 2 contracts are sold for USD 1,470 and closed out at USD 1,410 in January of Year 2.
- The August of Year 2 contracts are sold for USD 1,430 and closed out in July of Year 2 at USD 1,440, when the spot price is USD 1,435.

The gain per contract on the short futures positions in USD is

$$100 \times (1{,}500 - 1{,}450) + 100 \times (1{,}470 - 1{,}410) + 100 \times (1{,}430 - 1{,}440) = 10{,}000$$

so that the total gain from the short futures positions is USD 500,000 (= 50 × 10,000). The futures hedge is worth USD 100 (= 10,000/100) per ounce so that the price realized for the gold is USD 1,535 (= USD 1,435 + USD 100) per ounce. This is a good result and probably close to the result that would have been achieved from shorting futures contracts maturing in August of Year 2 (if that had been possible). However, the hedger is subject to multiple risks. This is because there is some uncertainty as to the difference between the futures prices every time an old contract is closed out and a new one is entered.

In practice, companies usually have exposure to the price of an asset every month, rather than in just one future month. The hedger must then enter into enough short-maturity contracts to hedge each future maturity month and roll them forward in the way we have described (thus the term *stack and roll*).

8.7 CASH FLOW CONSIDERATIONS

Because futures are settled daily, there is a mismatch between the cash flows from a futures contract used for hedging and the cash flows from the exposure being hedged. This difference can be especially important in the case of long-term hedges (such as the ones just considered). A company should therefore ensure that the losses on its futures contracts can be financed without difficulty until the corresponding gains are made on the position being hedged.

In the early 1990s, German company Metallgesellschaft sold a huge volume of 5- to 10-year heating oil and gasoline fixed-price contracts to its customers. It hedged its exposure by rolling over long positions in short-maturity futures contracts in the way described in the previous section. The price of heating oil and gasoline then began to fall, so the company expected to eventually benefit from its fixed-price contracts. However, the immediate losses on the futures contracts led to huge cash outflows that could not be financed. As a result, the fixed-price contracts were abandoned at a cost to the company of over USD 1 billion.

SUMMARY

A long futures position can be used to hedge the price of an asset that will be purchased in the future, whereas a short futures position can be used to hedge the price of an asset that will be sold in the future. If the asset underlying the futures contract and the asset being hedged are the same, then the number of futures contracts traded should be calculated by dividing the size of the exposure by the quantity of underlying assets per contract.

However, there are several reasons why many exposures are left unhedged. Sometimes it may simply not be in shareholders' best interests to hedge a particular risk. In other cases, the market price of a product may reflect the cost of its inputs, and therefore the manufacturer of the product has little exposure to the cost of the inputs. Finally, hedging using futures may be unattractive to some treasurers because of the possibility that the hedge results in a loss. Hedging, it should be remembered, is about making an outcome more certain—not about increasing profits.

Basis risk arises from the difference between the spot price of the hedged asset and the futures price for the contract used for hedging at the time the hedge is closed out. This difference can either improve or worsen the position of a hedger.

When there are differences between the asset underlying the futures contract and the asset whose price is being hedged, an optimal hedge ratio with minimum variance can be calculated. This ratio depends on the correlation between the change in the futures price and the change in the price of the asset being hedged, as well as on the standard deviation of each. Adjustments to the minimum variance hedge ratio can be made to allow for daily settlement and to use percentages, rather than absolute changes in the prices.

Hedging using stock index futures is popular. Stock index futures are a way of reducing or increasing an investor's exposure to the market for a period of time. This is likely to be attractive to an investor who has no views on the future direction of the market but believes that it is possible to pick stocks that will outperform the market.

Stack and roll is a procedure where sequences of short-maturity futures contracts are used to hedge relatively long maturity exposures. Under this approach, each contract is periodically closed out and replaced with a new one for as long as the hedge is needed.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

8.1 When should (a) a short hedge and (b) a long hedge be used?

8.2 Is hedging always profitable? Explain.

8.3 What is basis risk?

8.4 Consider the situation where the cost of an asset to be purchased in the future is being hedged. What is the impact of a positive basis when the hedge is closed out?

8.5 What should the hedge ratio be when the asset price and the futures price have of a correlation of 0?

8.6 How is the number of contracts that should be used calculated from the optimal hedge ratio? Assume that the optimal hedge ratio is calculated from actual, rather than proportional, changes in the asset price and the futures price.

8.7 Explain why hedging using futures is a series of one-day hedges.

8.8 Give two examples of situations where an investor might want to use stock index futures.

8.9 Explain what the beta (β) of a portfolio measures.

8.10 How could a hedging strategy using futures lead to cash flow problems?

Practice Questions

8.11 Explain why hedging is sometimes in the best interests of shareholders and sometimes it is not.

8.12 It is now February. A company knows that in May it will have to sell 10,000 barrels of crude oil. It uses the CME Group June futures contract for hedging. Each contract is on 1,000 barrels of light sweet crude. What position should it take? What are the price risks that it is exposed to after taking the position?

8.13 20 futures contracts are used to hedge an exposure to the price of soybeans. Each futures contract is on 5,000 bushels. At the time the hedge is closed out, the basis is 20 cents per bushel. What is the effect of the basis on the hedger if (a) the purchase of soybeans is being hedged and (b) the sale of soybeans is being hedged?

8.14 The standard deviation of quarterly (three-month) changes in the price of a commodity is 80 cents, and the standard deviation of quarterly changes in the futures price of a related commodity is 90 cents. The correlation between the two changes is 0.81. What is the optimal hedge ratio for a three-month hedge? How should it be interpreted?

8.15 In Question 8.14 the amount of the commodity being hedged is 200,000 units, and one futures contract is on 5,000 units of the commodity. How many contracts should be used in hedging? (Round to the nearest whole number.)

8.16 "In theory tailing the hedge involves adjusting the number of contracts used with the passage of time." Explain this statement.

8.17 Two alternative formulas for the hedge ratio are

$$N^* = h^* \frac{Q_A}{Q_F}$$

$$N^* = \hat{h} \frac{V_A}{V_F}$$

Explain the meaning of Q_A, Q_F, V_A and V_F. How are h^* and $\hat{h}$ defined?

8.18 A company has a portfolio of stocks worth 1 million dollars with a beta of 1.5. An index futures price is currently at 3,000, and each contract is for delivery of 50 times the index. How many contracts are necessary to hedge the market risk of the portfolio? Should long or short contracts be used?

8.19 In Question 8.18, how can beta be reduced to 0.9? How can it be increased to 1.8?

8.20 On January 15 of Year 1, a company decides to hedge the purchase of 100,000 bushels of corn on February 15 of Year 2. The following table gives futures prices (cents per bushel) of three selected contracts on four different dates. Explain how the company can use the contracts to create the required hedge. What is the net (after hedging) price paid for the corn as a function of the spot price on February 15 of Year 2? Each corn contract is on 5,000 bushels.

	January 15, Year 1	April 15, Year 1	August 15, Year 1	February 15, Year 2
May, Year 1 Futures Price	300	320		
Sep. Year 1 Futures Price		330	320	
March, Year 2 Futures Price			325	300

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

ANSWERS

Short Concept Questions

8.1 A short hedge should be used when an asset will be sold in the future. A long hedge should be used when an asset will be purchased in the future.

8.2 Hedging is designed to make outcomes less variable. Sometimes hedging leads to a profit relative to the no-hedging situation. Sometimes hedging leads to a loss relative to the no-hedging situation.

8.3 The basis is the spot price minus the futures price. Basis risk is the risk associated with the uncertainty about what the basis will be when a hedge is closed out.

8.4 The cost of the asset is the initial futures price plus the basis. A positive basis therefore worsens the hedger's situation.

8.5 The hedge ratio should be zero. This is common sense and is what is given by the Equation (8.1) when we substitute $\rho = 0$.

8.6 The number of contracts that should be used is h^*Q_A/Q_F, where h^* is the optimal hedge ratio, Q_A is the number of units of the asset being hedged, and Q_F is the number of units of the asset underlying a futures contract.

8.7 Because futures contracts are settled daily a hedge using futures contracts can be regarded as a series of one-day hedges. Note that if forward contracts are used for hedging, daily settlement is not an issue.

8.8 Possible examples: an investor wants to be out of the market for a period; an investor wants to change the beta of his or her portfolio; an investor specializes in picking stocks but does not want an exposure to the return on the market portfolio.

8.9 The beta of a portfolio measures sensitivity of the return from a portfolio to the return from the market.

8.10 Futures are settled daily. Losses on a futures hedge should eventually be matched by gains on the price of the asset being hedged, but the timing difference can give rise to cash flow problems.

Solved Problems

8.11 A company's management knows more about its risks than the shareholders and is therefore in the best position to hedge these risks. Also, the size of a futures contract may be too large to be useful to a single shareholder who wants to hedge a particular risk faced by a company. However, diversification is one way by which risks can be reduced, and it is easier for shareholders to diversify than it is for companies to do so.

8.12 The company should short 10 (= 10,000/1,000) contracts. It is exposed to basis risk. There are two components to this: the excess of the spot price of light sweet crude over the futures price when the hedge is closed out in May and the difference between the spot price of light sweet crude and the crude oil that the company is selling.

8.13 The basis increases the net price after hedging by 20 × 5,000 × USD 0.20 or USD 20,000. In (a) this is an extra cost to the hedger. In (b) it is an extra amount received from the sale of soybeans.

8.14 The optimal hedge ratio is

$$0.81 \times \frac{0.80}{0.90} = 0.72$$

This the proportion of the exposure that should be hedged.

8.15 The number of contracts is

$$0.72 \times \frac{200{,}000}{5{,}000} = 28.8$$

or 29, when rounded to the nearest whole number.

8.16 Because futures contracts are settled daily, a hedge using futures contracts is actually a series of one-day hedges. In theory, the hedge ratio changes because the asset price and the futures price change, and because a discount factor is necessary to account for the time difference between each one-day hedge and the end of the life of the hedge. In practice, it is often the case that the hedge ratio is calculated throughout the life of the hedge using the initial values of the asset price and the futures price, and the effects of the discount factors are ignored.

8.17 Q_A is the quantity of the assets to be hedged, V_A is the value of the assets to be hedged, Q_F is the quantity of the assets underlying one futures contract, and V_F is the value of the assets underlying one futures contract. The hedge ratio, h^*, is calculated from the standard deviation of actual changes in the asset price and the futures price during the life of the hedge and the hedge ratio, $\hat{h}$, is calculated from the standard deviation of percentage changes (returns) in the asset price and the futures price over one day.

8.18 The number of contracts that should be shorted is

$$1.5 \times \frac{1{,}000{,}000}{50 \times 3{,}000} = 10$$

8.19 To reduce beta to 0.9 the number of contracts that should be shorted is

$$(1.5 - 0.9) \times \frac{1{,}000{,}000}{50 \times 3{,}000} = 4$$

To increase beta to 1.8 the number long contracts required is

$$(1.8 - 1.5) \times \frac{1{,}000{,}000}{50 \times 3{,}000} = 2$$

8.20 The company should go long 20 May contracts on January 15 of Year 1 and close them out by selling 20 May contracts on April 15 of Year 1. It should go long 20 September contracts on April 15 of Year 1 and close them out by selling 20 September contracts on August 15 of Year 1. It should go long 20 March contracts on August 15 of Year 1 and close them out on February 15 of Year 2. The result of the short positions in cents per bushel is:

$(300 - 320) + (330 - 320) + (325 - 300) = -15$, or a loss of 15 cents per bushel.

The price paid is therefore $S + 15$ cents per bushel, where S is the spot price on February 15 of Year 2. In total, the cost in USD is $1{,}000 \times (S + 15)$.

9

Foreign Exchange Markets

Learning Objectives

After completing this reading, you should be able to:

- Explain and describe the mechanics of spot quotes, forward quotes, and futures quotes in the foreign exchange markets distinguish between bid and ask exchange rates.
- Calculate a bid-ask spread and explain why the bid-ask spread for spot quotes may be different from the bid-ask spread for forward quotes.
- Compare outright (forward) and swap transactions.
- Define, compare, and contrast transaction risk, translation risk, and economic risk.
- Describe examples of transaction, translation, and economic risks and explain how to hedge these risks.
- Describe the rationale for multi-currency hedging using options.
- Identify and explain the factors that determine exchange rates.
- Calculate and explain the effect of an appreciation/depreciation of one currency relative to another.
- Explain the purchasing power parity theorem and use this theorem to calculate the appreciation or depreciation of a foreign currency.
- Describe the relationship between nominal and real interest rates.
- Describe how a non-arbitrage assumption in the foreign exchange markets leads to the interest rate parity theorem and use this theorem to calculate forward foreign exchange rates.
- Distinguish between covered and uncovered interest rate parity conditions.

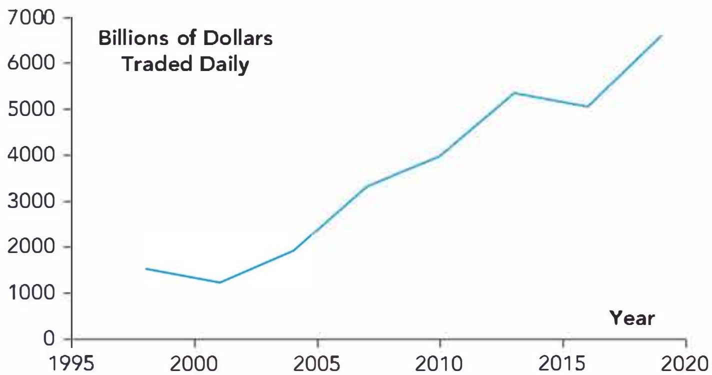


Figure 9.1 Growth of foreign exchange trading through time.

Source: BIS Triennial Central Bank Survey, September 2016.

The foreign exchange market (also referred to as the Forex, FX, or currency market) is the market where participants exchange one currency for another. As discussed in Chapter 4, we can distinguish between spot trades (where there is an agreement for the immediate or almost immediate exchange of currencies) and forward trades (where there is an agreement to exchange currencies at a future time).[1] The Forex market attracts both hedgers and speculators.

In terms of notional trading volume, the foreign exchange market is by far the largest market in the world. The 2019 Triennial Central Bank Survey of the Foreign Exchange and OTC Derivatives Markets Activity by the Bank for International Settlements (BIS) shows that trading in foreign exchange markets averaged USD 6.6 trillion per day in April 2019.[2] Figure 9.1 shows that the volume of trading has grown quite rapidly since the BIS started producing statistics in 1998. In 2019, 88% of the trading was between the U.S. dollar (USD) and another currency. The seven most popular currency pairs (listed by their global Forex market shares) were

- USD and the euro (24.0%),
- USD and the Japanese yen (13.2%),
- USD and the British pound (9.6%),
- USD and the Australian dollar (5.4%),
- USD and the Canadian dollar (4.4%),
- USD and the Chinese yuan (4.1%), and
- USD and the Swiss franc (3.5%).

[1] The standard settlement time for foreign exchange transactions is two days. This is referred to as T + 2. A spot trade is therefore actually a two-day forward trade. A few currency exchanges, such as exchanges between the U.S. dollar and the Canadian dollar, settle in one day (i.e., are T + 1).

[2] Triennial Central Bank Survey of foreign exchange and OTC derivatives markets in 2019. Retrieved from https://www.bis.org/statistics/rpfx19_fx.htm

Exchange rates can have a large effect on reported profits for firms that operate in multiple countries. Specifically, these firms are subject to foreign exchange gains and losses from both operating cash flows in foreign currencies as well as from the translation of asset and liability values in those currencies.

9.1 QUOTES

When an exchange rate is quoted, there is a base currency and a quote currency. Currency pairs are typically indicated as XXXYYY or XXX/YYY (with XXX as the base currency and YYY as the quote currency). The exchange rate shows how much of the quote currency is needed to buy one unit of the base currency. For example, a EURUSD quote of 1.2345 indicates that 1.2345 U.S. dollars are needed to buy one euro. A USDSEK quote of 8.7654 would indicate that 8.7654 Swedish kronor are needed to buy one U.S. dollar. The three-letter abbreviations for some traded currencies are shown in Table 9.1.[3]

The most common exchange rate quotes are between USD and another currency. Other quotes (e.g., between GBP and EUR) are known as *cross-currency* quotes. Currency traders have conventions about which currency is the base currency when exchange rates are quoted. In the case of the exchange rate between the U.S. dollar and the British pound, for example, the U.S. dollar is the quote currency. This is also the case when the U.S. dollar is quoted with the euro, the Australian dollar, and the New Zealand dollar. In most other cases, however, the U.S. dollar is the base currency and the other currency is the quote currency. The quote for the Canadian dollar, for example, is USDCAD, and it indicates the number of Canadian dollars that are equivalent to one U.S. dollar.[4]

Spot exchange rates are typically quoted with four to five decimal places. The bid-ask spread faced by corporations when they trade large amounts of a currency is quite small. At a particular time on January 16, 2020, for example, EURUSD was quoted as bid 1.11347 and ask 1.11354. The bid–ask spread is 0.00007 (=1.11354 − 1.11347).[5]

[3] The euro (EUR) is the official currency of the European Union. It was used by 19 of the 28 member states in 2019: Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Portugal, Slovakia, Slovenia, and Spain.

[4] News reports frequently quote the exchange rate the other way around as CADUSD.

[5] For the small currency exchanges necessary when traveling, however, bid-ask spreads are of course much larger.

Table 9.1 Currency Abbreviations

Country/Currency	Abbreviation
Argentina Peso	ARS
Australian Dollar	AUD
Brazil Real	BRL
Britain Pound	GBP
Canadian Dollar	CAD
China Yuan/Renminbi[6]	CNY
Columbia Peso	COP
Czech Koruna	CZK
Denmark Krone	DKK
Egypt Pound	EGP
Euro	EUR
Hong Kong Dollar	HKD
Iceland Krona	ISK
India Rupee	INR
Iran Rial	IRR
Iraq Dinar	IQD
Israel New Shekel	ILS
Japan Yen	JPY
Malaysia Ringgit	MYR
Mexico Peso	MXN
New Zealand Dollar	NZD
Norway Krona	NOK
Pakistan Rupee	PKR
Philippines Peso	PHP
Poland Zloty	PLN
Russia Ruble	RUB
Saudi Arabia Riyal	SAR
Singapore Dollar	SGD
South Africa Rand	ZAR
South Korea Won	KRW
Sweden Krona	SEK
Switzerland Franc	CHF
Taiwan Dollar	TWD
Thailand Baht	THB
USA Dollar	USD

[6] The renminbi is the name of the official currency of China. The yuan is the unit of account.

Table 9.2 EURUSD Forward Rates on January 16, 2020; the Spot Rate is Bid 1.11347, Ask 1.11354

Maturity	Bid	Ask
1 Week	4.6	4.7
2 Weeks	9.4	9.5
3 Weeks	14.2	14.3
1 Month	21.1	21.3
2 Months	42.6	42.8
3 Months	63.2	63.5
4 Months	84.1	84.5
5 Months	105.9	106.4
6 Months	126.1	126.7
7 Months	146.8	147.8
8 Months	167.5	168.5
9 Months	187.8	188.8
10 Months	209.1	210.1
11 Months	227.1	228.1
1 Year	251.0	252.3
2 Years	480.0	485.0
3 Years	707.0	717.0
4 Years	930.0	945.0
5 Years	1,157.0	1,172.0
6 Years	1,370.0	1,410.0
7 Years	1,593.0	1,633.0
10 Years	2,216.0	2,296.0

Forward exchange rates are quoted with the same base currency as spot exchange rates. They are usually shown as points that are multiplied by 1/10,000 and then added to or subtracted from the spot quote.[7] Table 9.2 shows the points quoted for EURUSD on January 16, 2020.

As an example, consider the EURUSD three-month forward quote with bid 63.2 and ask 63.5. Because the spot rate was bid 1.11347 and ask 1.11354, this means that the three-month forward bid quote is

$$1.11347 + 0.00632 = 1.11979$$

and the forward ask quote is

$$1.11354 + 0.00635 = 1.11989$$

[7] In the case of the Japanese yen, the points are multiplied by 1/100.

The bid-ask spread for the points is 0.3 (= 63.5 − 63.2). This increases the bid-ask spread for the three-month forwards by 0.00003 relative to the bid-ask spread for spot trades and so that the bid-ask spread for the forward quote is 0.00010 (= 0.00007 + 0.00003). The bid-ask spread tends to increase as the maturity of the forward contract increases. For example, the 10-year forward rate is bid:

$$1.11347 + 0.22160 = 1.33507$$

and ask

$$1.11354 + 0.22960 = 1.34314$$

for a bid-ask spread of 0.00807.

The quotes indicate that EUR, when purchased with USD, is more expensive in the forward market than in the spot market. We will discuss the reason for this later in this chapter in the context of covered interest parity.

For another example, consider USDCAD quotes from July 2018. At that time, the CAD interest rate was lower than the USD rate. As will be explained later, this means that the forward exchange rate (expressed as the number of CAD per USD) was lower than the spot exchange rate. When the spot exchange rate is bid 1.30815, ask 1.30825, forward rates (quoted in points) might be as shown in Table 9.3.

The six-month bid forward rate is then

$$1.30815 - 0.00436 = 1.30379$$

while the ask forward rate is

$$1.30825 - 0.00411 = 1.30414$$

Similarly, the bid and ask forward rates for one year are 1.29993 and 1.30043, respectively.

Note that sometimes the minus signs in quotes, such as those in Table 9.3, are omitted. However, traders know to subtract the points when the ask quote is less than the bid quote. This is because the bid-offer spread for forward rates is higher than that for spot rates, and this would not be the case if the points were added.

Outrights and Swaps

A forward foreign exchange transaction, where two parties agree on an exchange at some future date, is termed an *outright transaction* or a *forward outright transaction*. It can be contrasted with an *FX swap transaction*, where currency is exchanged on two different dates. Typically, an FX swap involves a foreign currency being bought (sold) in the spot market and then sold (bought) in the forward market. An FX swap is a way of funding an asset denominated in a foreign currency by paying interest in the domestic currency.

Table 9.3 **Sample Forward Quotes in 2018 for USDCAD**

Maturity	Bid	Ask
6 Months	−43.6	−41.1
1 Year	−82.2	−78.2

For example, a U.S. company can fund its European operations by borrowing in USD and buying 1 million EUR today while at the same time agreeing to sell 1 million EUR for USD in one month. This has the effect of funding the European operation in the domestic currency. We see from Table 9.2 that the bid quote for one-month forward EUR is 21.1 points. This is a measure of the amount by which EUR is more valuable in the forward market. In this case, the points reduce the net funding cost in USD because more USD are going to be received for EUR in one month compared to the amount that could have been received today.

We mentioned earlier that Forex trading was estimated to be worth USD 6.6 trillion per day in 2019. Table 9.4 shows a breakdown of this trading between spot trades, forward trades, FX swaps, currency swaps, and other products. While an FX swap involves the exchange of currency on two different dates (as has been described), a currency swap (also known as a cross-currency swap) involves the exchange of principal and a stream of interest payments in one currency for principal and a stream of interest payments in another

Table 9.4 **Daily Volume of Different Types of Forex Trading in 2019**

Type of Transaction	Daily Volume (Billions of USD)
Spot	1,987
Outright Forwards	999
FX Swaps	3,202
Currency Swaps	108
Other Products (Incl. Options)	294
Total	6,590

Source: BIS Triennial Central Bank Survey, September 2019.

currency. Currency swaps will be discussed in detail with examples in Chapter 20.

Futures Quotes

Forex futures trade actively on exchanges throughout the world. The CME Group in the U.S. trades many different futures contracts on exchange rates between the U.S. dollar and other currencies. These are always quoted with USD as the quote currency. This is because (from the perspective of the exchange) a foreign currency is treated like any other asset and is valued in U.S. dollars. For example, a six-month forward quote for the USDCAD of 1.3000 corresponds to a six-month futures quote of 0.7692 (= 1/1.3000) USD per CAD.

Popular contracts traded by the CME Group are on 100,000 AUD; 62,500 GBP; 100,000 CAD; 125,0000 EUR; 12.5 million JPY; and 125,000 CHF. The maturity months available on a given date include the following three months along with March, June, September, and December for the next 20 months. The CME Group also trades contracts on several cross rates.

9.2 ESTIMATING FX RISK

Firms need to quantify their exposures to exchange rates at different times in the future. Once this is done, they must then decide whether their exposures are acceptable or whether some hedging is necessary.

This section examines three categories of risk:

1. Transaction risk,
2. Translation risk, and
3. Economic risk.

Transaction Risk

Transaction risk is the risk related to receivables and payables. For example, consider a British company that imports goods from South Africa and pays for the goods in South African rand. It is therefore exposed to GBPZAR (i.e., the number of rand per British pound) risk. If ZAR strengthens relative to GBP, the company will find that its profits suffer when it must buy ZAR to pay its suppliers.

Suppose further that the British company sells goods in Portugal and prices its goods in euros. In this case, it is exposed to EURGBP risk. If EUR weakens relative to GBP, the company will find that its profits suffer when it exchanges its EUR revenues to GBP.

Transaction risk can be hedged with outright forward transactions. For the company in the previous example, buying ZAR forward would lock in the exchange rate paid to South African suppliers, while selling EUR forward would lock in the exchange rate applicable to EUR revenues.

An FX swap is useful when a company owns foreign currency that will be used for purchases at a future time and yet wants to earn interest in its domestic currency. The swap would enable the company to sell the foreign currency in exchange for its domestic currency in the spot market and buy it back at a future time in the forward market.

Translation Risk

Translation risk can arise from assets and liabilities denominated in a foreign currency. These must be valued in a firm's domestic currency when financial statements are produced. This can lead to foreign exchange gains or losses. It also arises from earnings in foreign subsidiaries that are not repatriated, which are typically translated at the average exchange rate over the fiscal year.

As an example of translation risk, suppose that a U.S. company has a manufacturing facility in the U.K. At the end of Year 1, the facility is valued at GBP 10 million and the GBPUSD exchange rate is 1.3500. At the end of Year 2, the value of the facility in GBP has not changed and it is still valued at GBP 10 million. However, the GBPUSD exchange rate is now 1.25. The company will record a foreign exchange loss (in USD) of:

$$(1.3500 - 1.2500) \times 10{,}000{,}000 = 1{,}000{,}000$$

Borrowings in a foreign currency can also lead to foreign exchange gains and losses. To see how this is the case, suppose that a U.S. company has a loan of 20 million euros that will be paid back in five years. Interest is paid in euros, and thus the firm is exposed to transaction risk. However, the loan principal to be paid back also gives rise to translation risk, which can be much greater.[8]

Suppose that the EURUSD exchange rate at the end of Year 1 is 1.2000 and that at the end of Year 2 it is 1.1500. Assuming the loan is valued at par, loan will be valued in USD at the end of Year 1 as:

$$20{,}000{,}000 \times 1.2000 = 24{,}000{,}000$$

[8] The translation risk becomes a transaction risk when the loan has to be repaid.

Assuming the loan is valued at par, value of the loan at the end of Year 2 is

$$20{,}000{,}000 \times 1.1500 = 23{,}000{,}000$$

The company has a foreign exchange gain of USD 1 million because the euro has weakened. If the euro had strengthened during the year, the company would have incurred a foreign exchange loss.

Translation risk is fundamentally different from transaction risk. Whereas transaction risk directly affects a company's cash flows, translation risk does not. However, it can have a big effect on its reported earnings.

Note that it only makes sense to hedge translation exposure on one future date. For example, it would be over-hedging to hedge the FX exposure to the value of the assets in one and in two years because the price increase (or decrease) over the first year is then considered twice.

Hedging translation risk with forward contracts makes accounting profits less volatile on that reporting date. However, it is questionable whether this is a good idea unless there is a plan to sell foreign currency assets or retire foreign currency liabilities or repatriate foreign income at a particular time in the future. This is because hedging replaces accounting risk with cash flow risk (because forward contracts do affect future cash flows). While translation risk is reduced, the transaction risk relating to the cash flows from the forward contracts is increased.

A better way of avoiding translation risk can be to finance the assets in a country with borrowings in that country.[9] In that case, gains (losses) on assets and income from them are offset by losses (gains) on liabilities.

Consider again the U.S. company with a GBP 10 million manufacturing facility in the U.K. If the translation risk is considered unacceptable, the facility can be financed by GBP 10 million of borrowings. There will then be no net translation gain or loss.[10]

Economic Risk

Economic risk is the risk that a company's future cash flows will be affected by exchange rate movements. For example, a U.S. firm that sells software in Brazil and denominates the price of the software in USD has no transaction risk. However, the firm does have economic risk. If the real (BRL) declines in value relative to the USD, the company's customers in Brazil will find its software more expensive. As a result, either the demand for the software will decrease or the firm will find it necessary to reduce the USD price of the software when it is sold in Brazil.

Sometimes exchange rate movements can affect a firm's competitive position in its domestic market. Consider a U.K. firm with no production or sales overseas. Exchange rate movements might make it more profitable for a foreign competitor to increase its activities in the U.K. in a way that adversely affects the firm.

Economic risk is more difficult to quantify than transaction or translation risk, but possible exchange rate movements should be considered when key strategic decisions are being made. For example, foreign exchange considerations might play a role in a decision to move production overseas.

9.3 MULTI-CURRENCY HEDGING USING OPTIONS

Multinational companies have exposures to many different currencies. These exposures can lower FX risk because exchange rate movements across different currencies are not perfectly correlated.[11] In general, the volatility for an investment portfolio with multiple stocks is less than that of one with a single investment. In the same way, volatility arising from exposures to many different currencies is usually less than that arising from an exposure to a single currency.

As mentioned in a previous chapter, treasurers often prefer options to forward contracts when hedging. This is because options provide downside protection against adverse exchange rate movements while still allowing a firm to benefit from favorable movements. One FX hedging strategy is to buy options on individual currencies to cover each possible adverse exchange rate movement. A less expensive alternative, however, is for a firm to identify the portfolio of currencies to which it is exposed and buy an option on that portfolio in the over-the-counter market.

[9] When a debt instrument in a foreign currency matures, it can be replaced with a new debt instrument denominated in the same currency.

[10] The interest on the borrowings would give rise to transaction risk, which if desired can be hedged separately with forward contracts.

[11] A number of factors influence exchange rate movements, but even after these factors have been taken into account, there is still a huge amount of uncertainty about the relative values of different currencies in the future.

For example, a corporation could buy an option on a portfolio consisting of:

- A long position in 100,000 units of currency A,
- A long position in 200,000 units of currency B, and
- A short position in 75,000 units of currency C.

This type of option is known as a *basket option*.

Multinational firms often have exposures to an exchange rate in every month of the year. One way to limit this exposure is to trade options with monthly maturities. A less expensive alternative is to trade options on the average exchange rate during the year. These options are known as *Asian options*.[12]

9.4 DETERMINATION OF EXCHANGE RATES

Exchange rates are determined by many interrelated factors. While this section describes some of those factors, note that future exchange rates cannot be predicted with any precision. Exchange rates (like the prices of all financial assets) are ultimately determined by supply and demand, which are in turn influenced by many factors.

What follows is a discussion of the most important economic variables that influence exchange rates.

Balance of Payments and Trade Flows

The balance of payments between two countries measures the difference between the value of exports and the value of imports. For example, suppose exports from Country A to Country B increase. When exporters exchange their foreign currency-denominated revenues for their domestic currency, it will increase the demand for Country A's currency and strengthen it relative to Country B's currency. If imports from Country A to Country B increase, however, Country A's currency will weaken relative to that of Country B (because importers would have to buy Country B's currency to pay for the goods they are importing).

There are some equilibrating forces at work here. As Country A increases its exports to Country B, its currency strengthens. This causes its exports to become more expensive for customers in Country B. As a result, there is lower demand for those exported goods in Country B. Similarly, Country A imports increase as its currency weakens. As a result, goods imported from Country B become more expensive, and this in turn reduces the demand for goods imported from Country B.

An illustration of the importance of trade flows is provided by the USDCAD exchange rate. Because Canada is an oil exporting nation, the value of the Canadian dollar is influenced by the price of oil. The Canadian dollar was worth more than the U.S. dollar from 2011–2014 when the price of crude oil was high. When the price of oil declined, so did the Canadian dollar.

Inflation

If USDCAD is 1.2500, we would expect the CAD price of goods in Canada to be 25% higher than the USD price of goods in the U.S. If this is not the case, there is a theoretical arbitrage opportunity.

For example, suppose that a product costs USD 100 and CAD 130. An arbitrageur can buy the product in the U.S. and sell it in Canada for a profit of CAD 5 per unit. Similarly, if the product costs USD 100 and CAD 120, the arbitrageur can buy the product in Canada and sell it in the U.S. for a similar profit.

These arbitrage opportunities may not exist in practice due to several costs that the arbitrageur may incur (e.g., transportation costs and possibly tariffs). However, this relationship is the basis of what is known as *purchasing power parity*.

To see how purchasing power parity works, suppose that inflation is 3% per year in the U.S. and 1% per year in Switzerland. The cost of a representative basket of goods in the U.S. measured in USD increases at 3% per year, while the cost of the same basket of goods in Switzerland increases at 1% per year. Suppose further that the initial USDCHF exchange rate is 1.05. If purchasing power parity holds, the cost of a basket of goods worth 100 USD is now worth 105 CHF. After one year, the same basket of goods will be worth USD 103 (= 100 × 1.03) and CHF 106.05 (= 105 × 1.01). Purchasing power parity suggests that the exchange rate should become

$$\frac{106.05}{103} = 1.0296$$

This is an (approximate) 2% strengthening of CHF relative to USD.

The general (approximate) purchasing power parity formula is

Percent Strengthening of Domestic Spot Rate = Foreign *Inflation Rate* − *Domestic Inflation Rate*

[12] The two ideas mentioned here can be combined so that the multinational firm hedges risks with an Asian basket option.

In our example, the foreign (CHF) inflation rate minus the domestic (USD) inflation rate is −2%, and thus the domestic spot rate weakens by about 2%.

Monetary Policy

The value of a country's currency is also influenced by the monetary policy of its central bank. If Country A increases its money supply by 25% while Country B keeps its money supply unchanged, the value of Country A's currency will tend to decline by 25% relative to Country B's currency (with all else being equal). This is because 25% more of Country A's currency is being used to purchase the same amount of goods.

9.5 REAL VERSUS NOMINAL INTEREST RATES

Analysts distinguish between *nominal interest rates* and *real interest rates*. Nominal interest rates are usually quoted in the market and indicate the return that will be earned on a currency. An interest rate of 4% per year for a currency of a country indicates that 100 of that currency will grow to 104 in one year.

Real interest rates are adjusted for inflation. As an example, consider a basket of goods that costs 100 at the beginning of the year. If inflation in the country is 3%, then the basket of goods will cost 103 at the end of the year. An individual that starts with 100 can buy either buy one basket of the goods at the beginning of the year or invest at 4% and buy

$$\frac{104}{103} = 1.0097$$

baskets at the end of the year. This shows that the investor's real purchasing power has increased by only 0.97%. This is referred to as the investor's real interest rate.

In general:

$$R_{real} = \frac{1 + R_{nom}}{1 + R_{infl}} - 1$$

where R_{real} is the real interest rate, R_{nom} is the nominal interest rate, and R_{infl} is the rate of inflation.[13] This is often approximated as:

$$R_{real} = R_{nom} - R_{infl}$$

In the previous example, using this approximation yields a real rate of interest of 1% instead of 0.97%.

[13] This assumes that interest rates and inflation rates are expressed with annual compounding. See Chapter 16 for a discussion of compounding frequency issues.

The real interest rate (and sometimes even the nominal interest rate) can be negative. For example, the nominal interest rates in the Swedish krone, Japanese yen, Danish krone, euro, and Swiss franc were negative in period following the 2007–2008 financial crisis. It is estimated that there were about USD 13 trillion of government bonds worldwide offering yields below zero in 2019.[14] It is sometimes argued that interest rates cannot become negative because it is always possible to hold cash (which has an interest rate of zero). However, storage costs for large amounts of cash are not trivial, and thus negative rates do not necessarily present arbitrage opportunities.

9.6 COVERED INTEREST PARITY

Purchasing power parity, which was discussed in Section 9.4, provides results that are at best approximately true in the long-term. Over short periods of time, there can be significant deviations from purchasing power parity. We now explain a result (illustrated by Figure 9.2) relating to forward exchange rates, spot exchange rates, and interest rates that holds much more precisely because it is based on arbitrage arguments.[15]

Suppose a U.S. trader starts with 100 GBP and wants to end up with USD in T years. As indicated in Figure 9.2, there are two ways to do this.

1. The trader can invest the funds at the GBP risk-free rate (R_{GBP}) so that they grow to $100(1 + R_{GBP})^T$ at time T. At the same time, the trader can enter into a forward contract to exchange $100(1 + R_{GBP})^T$ for USD at time T. This leads to:[16]

$$100(1 + R_{GBP})^T F$$

where F is the T-year forward GBPUSD exchange rate.

2. The trader can exchange the funds immediately for USD and then invest the USD funds at the USD risk-free rate. The 100 GBP first becomes $100S$ USD, where S is the current GBPUSD spot rate. At time T, this becomes

$$100S(1 + R_{USD})^T$$

[14] See https://www.bloomberg.com/graphics/2019-negative-yield-debt/ for a discussion of this.

[15] Surprisingly there have been departures from covered interest parity since the 2007–2008 crisis. See for example C. Borio, R. N. McCauley, P. McGuire, and V. Susko, "Covered Interest Parity Lost: Understanding the Cross-Currency Basis," Bank for International Settlements, 2016: https://www.bis.org/publ/qtrpdf/r_qt1609e.htm. The authors attribute this to changes in the demand for FX hedging and banks not having the balance sheet capacity to engage in the type of arbitrage required.

[16] Interest rates are here expressed with annual compounding.

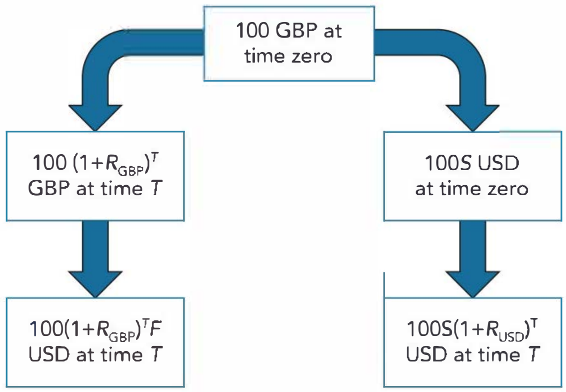


Figure 9.2 **Two ways of converting 100 GBP to USD at time *T*. *S* and *F* are the GBPUSD spot and forward exchange rates, respectively. The variables R_{USD} and R_{GBP} are the risk-free interest rates in USD and GBP.**

where R_{USD} is the annual USD risk-free rate.

There is no uncertainty about the amount of USD that will be obtained at time T for either scenario. In the absence of arbitrage opportunities, they must therefore give the same result:

$$100(1 + R_{GBP})^T F = 100S(1 + R_{USD})^T$$

so that[17]

$$F = S\frac{(1 + R_{USD})^T}{(1 + R_{GBP})^T} \tag{9.1}$$

Equation (9.1) shows what is commonly known as the *covered interest parity*. If F is less than the exchange rate given by Equation (9.1):

$$F < S\frac{(1 + R_{USD})^T}{(1 + R_{GBP})^T} \tag{9.2}$$

an arbitrageur can

- Borrow 100 GBP for T years at R_{GBP},
- Convert the funds to 100S USD,
- Invest the USD at R_{USD} for T years to obtain $100S(1 + R_{USD})^T$ USD at time T, and
- Enter into a forward contract to convert this to $100S(1 + R_{USD})^T/F$ GBP at time T.

From Inequality (9.2):

$$\frac{100S(1 + R_{USD})^T}{F} > 100(1 + R_{GBP})^T$$

Thus, the trader would have more GBP than required to repay the funds borrowed in GBP.

If F is greater than the exchange rate given by Equation (9.1):

$$F > S\frac{(1 + R_{USD})^T}{(1 + R_{GBP})^T} \tag{9.3}$$

an arbitrageur can

- Borrow 100S USD for T years at R_{USD},
- Convert the funds to 100 GBP,
- Invest at R_{GBP} for T years to obtain $100(1 + R_{GBP})^T$ GBP at time T, and
- Enter into a forward contract to convert this to $100(1 + R_{GBP})^T F$ USD at time T.

From Inequality (9.3):

$$100(1 + R_{GBP})^T F > 100S(1 + R_{USD})^T$$

Thus, the trader has more USD than required to repay the funds borrowed in USD.

This analysis assumes that the trader can borrow and lend at the same interest rate (i.e., that the borrowing and lending risk-free interest rates in the U.S. are the same and equal to R_{USD}). For a large bank, this is close to true. If we take the bank's spread between borrowing and lending rates into account, however, we would obtain a narrow range of possible values for F (instead of just one value).

In general, for an exchange rate XXXYYY, we have

$$F = S\frac{(1 + R_{YYY})^T}{(1 + R_{XXX})^T} \tag{9.4}$$

If the risk-free rate for currency XXX is higher than that for currency YYY, XXX is weaker in the forward market than in the spot market (i.e., it takes less units of YYY to buy one unit of XXX in the forward market than it does in the spot market). If the risk-free rate for currency XXX is lower than that for currency YYY, XXX is stronger in the forward market than in the spot market (i.e., it takes more units of YYY to buy one unit of XXX in the forward market than it does in the spot market).

Table 9.2 shows that the USD, when exchanged for EUR, was weaker in the forward market than in the spot market on January 16, 2020. The spot exchange rate was 1.113505 (= (1.11347 + 1.11354)/2), whereas the mid market three-month forward rate was 1.11984 (= (1.11979 + 1.11989)/2), indicating that it took more USD to buy EUR in the forward market. From Equation (9.4):

$$1.11984 = 1.113505\frac{(1 + R_{USD})^{0.25}}{(1 + R_{EUR})^{0.25}}$$

which reduces to

$$\frac{1 + R_{USD}}{1 + R_{EUR}} = 1.0229$$

---

[17] When R_{USD} and R_{GBP} are expressed with continuous compounding, equation (9.1) becomes

$$F = Se^{(R_{USD} - R_{GBP})T}$$

This indicates that the three-month USD interest rate was higher than the EUR interest rate on January 16, 2020 bid. Indeed, this was the case: The three-month interbank borrowing rate in USD was about 1.84% while the three-month interbank borrowing rate in euros was about −0.39%. These interest rates give

$$\frac{1 + R_{USD}}{1 + R_{EUR}} = \frac{1.0184}{0.9961} = 1.0224$$

which is close to the 1.0229 predicted by forward rates.

As another application of Equation (9.2), suppose that interest rates in currencies XXX and YYY are 3% and 5% per annum (respectively) and that the XXXYYY spot rate is 1.2500. The six-month XXXYYY forward rate will be

$$1.2500\frac{1.05^{0.5}}{1.03^{0.5}} = 1.2621$$

Note that it takes 1.25 units of YYY to buy XXX in the spot market and 1.2621 units of YYY to buy XXX in the forward market. XXX is therefore stronger than YYY in the forward market than it is in the spot market.

Interpretation of Points

From Equation (9.4), it is approximately true that when $T < 1$:[18]

$$\frac{F}{S} = \frac{1 + R_{YYY}T}{1 + R_{XXX}T}$$

This can be written

$$\frac{F}{S} - 1 = \frac{1 + R_{YYY}T}{1 + R_{XXX}T} - 1$$

or

$$\frac{F - S}{S} = \frac{1 + R_{YYY}T - 1 - R_{XXX}T}{1 + R_{XXX}T}$$

so that

$$\frac{F - S}{S} = \frac{R_{YYY}T - R_{XXX}T}{1 + R_{XXX}}$$

or approximately

$$\frac{F - S}{S} = (R_{YYY} - R_{XXX})T$$

This provides an interpretation of the forward rate points. The term $F - S$ is the points divided by 10,000 and (when expressed as a percentage of the spot rate) is approximately equal to the interest rate differential applied to time T.

[18] This is exactly true if the interest rates are expressed with a compound period of T.

Consider the previous example where the interest rates in currencies XXX and YYY are 3% and 5% per annum (respectively), the spot rate is 1.25, and the forward rate is 1.2621. In this case, the forward rate would be quoted as 121 points (= (1.2621 − 1.25) × 10,000) and:

$$\frac{121/10{,}000}{1.25} = 0.0097$$

This corresponds to 0.97%, which is approximately equal to the 2% per year interest rate differential applied to six months.

9.7 UNCOVERED INTEREST PARITY

Covered interest parity concerns forward exchange rates and can be expected to hold well because it depends on arbitrage arguments.

Uncovered interest parity is an argument concerned with exchange rates themselves and is just one of the many interacting factors that determine how exchange rates move. It argues that investors should earn the same interest rate in all currencies when expected exchange rate movements are considered.

Consider currencies X and Y with risk-free rates of 2% and 6% (respectively). In equilibrium, the two currencies should be equally attractive. According to uncovered interest rate parity, this means that the investor should expect the value of currency Y to weaken by about 4% relative to the value of currency X.

Arguably, there are many potential violations of uncovered interest parity. In 2020, the interest rate in USD was much higher than the interest rate in EUR. However, the U.S. economy was generally considered to be much stronger than many European economies. Furthermore, many market participants did not consider the interest rate differential to be indicative of a stronger euro in the future.

If both covered and uncovered interest rate parity held, the forward exchange rate would equal the expected future spot exchange rate. We will discuss the relationship between the forward/futures price of an asset and its spot price in Chapter 11.

SUMMARY

The foreign exchange market has developed standard ways of quoting exchange rates. For example, each currency is presented as a three-letter abbreviation. When an exchange rate is referred to as XXXYYY or XXX/YYY, it is the number of units of currency YYY (the quote currency) that is equal to one unit of currency XXX (the base currency). USD is the quote currency when traders quote exchange rates between USD and the British pound, the euro, the Australian dollar, or the New Zealand dollar. In most other cases, USD is the base currency. Forward exchange rates are usually quoted as points that are added to or subtracted from spot exchange rates.

Several different types of trades are carried out in the foreign exchange markets. A spot trade is an exchange of one currency for another that takes place immediately (or almost immediately). An outright forward is an agreement to exchange currencies that will take place at a certain time in the future. An FX swap is a transaction that involves a currency exchange at one time and the opposite exchange at a later time. A currency swap involves interest and principal in one currency being exchanged for interest and principal in another currency. The FX market is large and involves transactions involving several trillion dollars every day.

A company has three types of foreign exchange risk: transaction, translation, and economic. Transaction risk relates to the risk associated with the exchange rate that will apply when (a) a foreign currency is purchased with the domestic currency in the future to buy goods and services overseas, or (b) future revenues in foreign currencies are converted into the domestic currency. Translation risk arises from the need to value foreign assets and liabilities and the earnings of subsidiaries when financial statements are produced. Economic risk arises from longer-term changes to a company's competitive environment arising from future exchange rate movements.

Large multinationals often have exposures to many different exchange rates and have risk diversification advantages. Some use fairly complex derivatives that focus on the residual risk from multiple exchange rate exposures or average exposures over several months.

Many factors are involved in the determination of exchange rates, and there is no precise way to estimate a future exchange rate. In the long term, however, a country's balance of payments with another country will affect the exchange rate between its currency and that of the other country. Monetary policies and inflation rates are also important, but exchange rates can be immune to theoretical macroeconomic forces for periods of time.

There is a no-arbitrage relationship between forward exchange rates and spot exchange rates that involves interest rates. If the interest rate for currency A is less than that for currency B, for example, currency A will be stronger in the forward market than in the spot market. The percentage amount by which the forward exchange is better than the spot exchange is a reflection of the interest rate differential between the two currencies.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

9.1 Suppose that the quote between currency A and currency B is 1.2000 and that currency B is the base currency. How many units of currency A should be exchanged for 100 units of currency B?

9.2 Name three European countries that do not use the euro.

9.3 A three-month forward foreign exchange rate is specified as bid 38.5, ask 40.5. What does this mean?

9.4 What is an FX swap?

9.5 What is the difference between transaction and translation risk?

9.6 How can a multinational that uses options for hedging foreign exchange risk reduce the cost of its hedging?

9.7 When Country A increases its exports to Country B, what tends to happen to the exchange rate between the currencies of A and B?

9.8 When inflation increases faster in Country A than in Country B, what tends to happen to the exchange rate between the currencies of A and B?

9.9 What is the difference between real and nominal interest rates?

9.10 Under what circumstances is a currency weaker in the forward market than in the spot market?

Practice Questions

9.11 The exchange rate between USD and some foreign currencies is quoted by traders with USD as the quote currency. Give three foreign currencies for which this is the case.

9.12 In futures markets in the U.S., is the exchange rate between USD and currency XXX specified as USDXXX or XXXUSD? Explain.

9.13 The spot rate for XXXYYY is 1.4251 and the one-year forward rate is expressed as 200 basis points. How would the spot rate and the forward rate have been expressed if the currency had been YYYXXX?

9.14 Suppose there is a cash need for one year in a foreign currency. Explain how a company can borrow domestically and use an FX swap to fund the cash need. What is the difference between that and borrowing funds in the foreign currency to fund the cash need?

9.15 A company calculates its translation risk for its three-month and six-month statements and hedges each with a forward contract. Is this a sensible strategy? Discuss.

9.16 What is economic FX risk?

9.17 If the nominal interest rate is 2% and the rate of inflation is 3%, what is the real interest rate? How should it be interpreted?

9.18 Interest rates in a currency XXX increase and interest rates in currency YYY stay the same. The exchange rate is expressed as XXXYYY. Do forward rates increase or decrease?

9.19 The interest rate in XXX is 1% and in YYY 4%. The XXXYYY spot rate is 1.3000. How would three-month forward rate be quoted using points?

9.20 Until recently, Country A and Country B had similar interest rates. The central bank of Country A has just increased interest rates. A speculator thinks this will lead to international investors moving funds from Country B's currency to Country A's currency to earn the higher interest rate. This will increase the demand for currency A, and as a result, currency A will strengthen relative to currency B. What spot or forward trades should the speculator do?

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

9.1 A is the quote currency. The quote indicates that 120 units of A should be exchanged for 100 units of B.

9.2 Among the possibilities are the U.K., Sweden, Norway, Denmark, Czechoslovakia, Iceland, Poland, and Switzerland.

9.3 0.00385 should be added to the spot bid quote to get the forward bid exchange rate. 0.00405 should be added to the spot ask quote to get the forward ask exchange rate.

9.4 An FX swap is an agreement to buy (sell) a certain amount of a currency at one time and sell (buy) it at another later time.

9.5 Transaction risk exposure arises from cash inflows and outflows in a foreign currency. Translation risk exposure arises from FX gains and losses when assets and liabilities denominated in a foreign currency are converted to the domestic currency for the purposes of producing financial statements.

9.6 The multinational has an exposure each month to a basket of currencies. It can hedge by buying an option on the basket rather than on each currency. It can also buy an option on the average exposure that will apply across several months, rather than buying one option for its exposure in each month.

9.7 Country A's currency tends to strengthen relative to Country B's currency because importers in Country B will find it necessary to buy Country A's currency to pay for goods.

9.8 Purchasing power parity suggests that Country A's currency will weaken relative to Country B's currency in order to equalize the cost of a basket of goods in the two countries.

9.9 A nominal interest rate is the interest rate usually quoted. It is the rate earned in units of the currency. The real interest rate is the nominal interest rate minus the inflation rate. It allows for the fact that inflation is causing the value of the currency to decline.

9.10 Currency A relative to currency B is weaker in the forward market than in the spot market if interest rates are higher in currency A than in currency B.

Solved Problems

9.11 Possible answers are GBP, AUD, NZD, and EUR. The exchange rate for these currencies is quoted by traders as the number of units of USD per unit of the foreign currency.

9.12 It quotes as XXXUSD because the exchange rate is the number of units of USD per unit of the foreign currency.

9.13 The spot rate and forward rate are 1.4251 and $1.4251 + 0.0200 = 1.4451$. If the currency had been quoted the other way around, the spot and forward would be $1/1.4251 = 0.7017$ and $1/1.4451 = 0.6920$. The forward quote would be -97 because the forward rate is 97/10,000 less than the spot rate.

9.14 When the company borrows domestically, an FX swap can convert the borrowing to the foreign currency at the beginning of the loan and back to the domestic currency at the end of the loan. The difference between this and borrowing in the foreign currency is that interest is paid in the domestic currency instead of the foreign currency.

9.15 Whether or not balance sheet risks should be hedged with forward contracts is debatable. (Some argue that reducing the volatility of earnings is desirable; others argue that non-cash-flow risks should not be hedged.) If they are, it only makes sense to hedge the risk at one future time. Otherwise there is over-hedging. This is a key difference between transaction and translation risk.

9.16 Economic FX risk is the risk that the company's competitiveness in either domestic or foreign markets may be affected by FX movements.

9.17 The real interest rate is -1%. When an investor earns 2%, the investor's purchasing power actually decreases by 1% per year because of inflation.

9.18 Currency XXX becomes weaker in the forward market. Exchange rates are expressed as the number of units of YYY that would be exchanged for one unit of XXX. As a result of the increase in interest rates, it takes less units of YYY to buy one unit of XXX in the forward market and so the forward rate decreases.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

9.19 The forward rate is

$$1.3000 \times \frac{(1.04)^{0.25}}{(1.01)^{0.25}} = 1.3095$$

The forward rate would be quoted as 95.

9.20 The speculator thinks that currency A will strengthen. However, interest rate parity indicates that it is weaker in the forward market. If the speculator is right, he or she will make money by buying currency A with currency B in the forward market and then selling it on the delivery date.

10

Pricing Financial Forwards and Futures

Learning Objectives

After completing this reading, you should be able to:

- Define and describe financial assets.
- Define short-selling and calculate the net profit of a short sale of a dividend-paying stock.
- Describe the differences between forward and futures contracts and explain the relationship between forward and spot prices.
- Calculate the forward price given the underlying asset's spot price and describe an arbitrage argument between spot and forward prices.
- Distinguish between the forward price and the value of a forward contract.
- Calculate the value of a forward contract on a financial asset that does or does not provide income or yield.
- Explain the relationship between forward and futures prices.
- Calculate the value of a stock index futures contract and explain the concept of index arbitrage.

Chapter 9 examined the relationship between the forward price and the spot price of a foreign currency using the covered interest parity no-arbitrage argument. This chapter uses similar no-arbitrage arguments to estimate forward and futures prices for financial assets (e.g., stocks and bonds).

A *financial* asset is an asset whose value derives from a claim of some sort. An *investment* asset is an asset held by market participants for investment purposes. All financial assets (and a small number of non-financial assets) are investment assets. (Non-investment assets are sometimes referred to as *consumption* assets.)

This chapter considers three types of financial assets:

1. Assets providing no income,
2. Assets providing a known income that is a fixed amount, and
3. Assets providing a known income that is a percentage of their value.

We will confirm that the results for the above assets are consistent with those for forward contracts on currencies and use examples to show how they can be applied to forward contracts on bonds and stock indices. We argue that (in theory) futures prices and forward prices for contracts with the same maturity on the same asset should be approximately equal. This means that results produced for forward prices are approximately true for futures prices.

10.1 SHORT SELLING

Some of the no-arbitrage arguments used in this chapter involve shorting (short selling) an asset. This is the sale of an asset that is not owned with the intention of buying it back later.

Short selling is profitable if the asset price declines but incurs losses if the asset price increases. For example, suppose an investor contacts his or her broker to short 100 shares of Company X. The broker will borrow the shares from another investor and proceed to sell them in the market in the usual way. At a later time, the investor would contact the broker again to cover (i.e., close-out) the short position. The broker would then buy 100 shares of Company X and replace them in the account from which they were borrowed. If a dividend is paid while the shares were shorted, the investor would be required to pay that dividend, and the funds would be passed on to the account from which the shares were borrowed.

In the U.S., if another investor has bought shares of Company X on margin (see Section 5.3 for a description of buying on margin), the broker can borrow the shares without asking the investor's permission. Small investors may find it difficult to short if their brokers do not have access to shares that have been bought on margin by other investors. Large investors can borrow the financial assets they want to short from financial institutions (e.g., State Street Corporation) for a fee. The fee is usually quite low (less than 50 basis points per year), but it may rise if the asset is scarce.

Table 10.1 Profit from a Short Sale (USD)

May:	Shares shorted	+5,000
July:	Dividend paid	−200
Sept:	Position covered	−3,000
Total		1,800

Table 10.2 Profit from a Long Position (USD)

May:	Shares bought	−5,000
July:	Dividend paid	+200
Sept:	Shares sold	+3,000
Total		−1,800

As an example, suppose 100 shares are shorted in May when the share price is USD 50. The position is then covered in September when the share price is USD 30. Suppose further that a dividend of USD 2 per share is paid on the shares in July. The profit from the short position (ignoring brokerage fees and assuming there is no borrowing fee) is USD 1,800 (calculated as indicated in Table 10.1). Note that the cash flows for an investor with a long position are the mirror image of the cash flows of an investor with a short position. Thus, an investor with a long position in 100 shares between May and September would lose USD 1,800 as indicated in Table 10.2 (again, ignoring brokerage costs).

Short selling attracts the attention of regulators from time to time. During the 2007–2008 global financial crisis, several countries banned investors from short selling financial stocks because regulators believed such trades would exacerbate the crisis. However, many financial analysts believe short selling is an important part of the price discovery process.

10.2 THE NO INCOME CASE

This section considers the relationship between the spot price and the forward prices of a financial asset that provides no income. The asset could be a non-dividend-paying stock, a Treasury bill, or a zero-coupon bond.

For example, suppose a financial asset that will provide no income is priced at USD 70 per unit and a large bank can borrow (or lend) at 5% per year. We will consider two scenarios that give rise to an arbitrage opportunity for the bank's traders.

Scenario A: The one-year forward price of the asset is USD 80.

Scenario B: The one-year forward price of the asset is USD 65.

Consider Scenario A. A trader can buy the asset today for USD 70 and sell it in the forward market for USD 80. The cost of funding the purchase of the asset for one year is

$$0.05 \times \text{USD } 70 = \text{USD } 3.50$$

The trader therefore makes a profit of:

$$\text{USD } 80.00 - \text{USD } 70.00 - \text{USD } 3.50 = \text{USD } 6.50$$

Clearly, the forward price of USD 80 is too high in this example. In fact, any forward price above USD 73.50 would lead to the trader making a zero-risk profit. As a result, we can expect the actions of arbitrageurs to drive down the forward price to a point where the arbitrage is no longer profitable.[1] If there are to be no arbitrage opportunities, the forward price must be less than or equal to USD 73.50.

Now suppose the forward price is USD 65 (i.e., USD 5 less than the spot price). A trader can short the asset and enter into a forward contract to buy it back in one year. Assuming there is no fee for borrowing the asset, this trade leads to a profit of USD 5 (= USD 70 − USD 65). Additionally, the short position generates USD 70 in cash, which can be invested at the risk-free rate for one year. This gives a profit of:

$$0.05 \times \text{USD } 70 = \text{USD } 3.5$$

The trading strategy therefore generates a total profit of USD 8.50 (= USD 5.00 + USD 3.50).

This trading strategy is profitable for any forward price below USD 73.50. As traders take advantage of this arbitrage opportunity, however, the forward price is driven up until the strategy is no longer profitable.

Table 10.3 summarizes these scenarios. The first trading strategy generates profits when the forward price is greater than USD 73.50. The second trading strategy generates profits when the forward price is less than USD 73.50. If there are to be no arbitrage opportunities, the forward price must therefore be equal to USD 73.50.

[1] In practice, the existence of arbitrageurs means that arbitrage opportunities (if they exist at all) are very limited.

Table 10.3 Trading Possibilities for an Arbitrageur When an Asset Provides no Income and Is Priced at USD 70 with a One-Year Borrowing/Lending Rate of 5% per Year

Scenario A: One-Year Forward Price Is USD 80	Scenario B: One-Year Forward Price Is USD 65
Buy asset for USD 70 and enter into a short forward contract to sell it for USD 80 in one year.	Short asset for USD 70 and enter into a long forward contract to buy it in one year for USD 65.
Profit equals USD 6.50	Profit equals USD 8.50

We have assumed that banks can borrow and lend at the same rate. While this is not exactly true, it is close to being true. The relevant rate for a bank trader is sometimes referred to as the bank's opportunity cost of capital. In practice, it is the interbank rate (i.e., the rate at which banks borrow and lend between themselves). The spread between the borrowing rate and lending rate in the interbank market is typically around 0.1%. If the borrowing and lending rates in our example are 5.1% and 5.0%, respectively, then the no-arbitrage argument presented with Scenario A shows that the forward price must be less than:

$$70 \times 1.051 = 73.57$$

The no-arbitrage argument presented in Scenario B is unchanged and we can conclude that the forward price must lie between USD 73.50 and USD 73.57.

We will continue to ignore the distinction between borrowing and lending rates and refer to the interest rate at which a bank can borrow or lend as the risk-free rate. There is some ambiguity here (as risk-free rates are often assumed to be government borrowing rates).

The no-arbitrage arguments presented here also assume that the asset can be shorted with no borrowing-related fees. The shorting (borrowing) fee would make the Scenario B trading strategy more expensive. However, to avoid the borrowing fee we can consider an investor who already owns the asset (and there must be such investors for any financial asset). The investor in Scenario B can then sell the asset for USD 70, release the funds for investment at the risk-free rate, and buy it back in one year for USD 65 to produce USD 8.50 in profit.

BOX 10.1 JOSEPH JETT'S LOSS

In the early 1990s, Kidder Peabody trader Joseph Jett engaged in rogue trading activities related to the arbitrage strategies mentioned above. Jett would buy zero-coupon Treasury bonds (known as strips) and sell them in the forward market. Due to a flaw in Kidder Peabody's computer systems, however, he was able to immediately record the difference between the forward price and the spot price as a profit.

Consider the example in Table 10.3. If the forward price was USD 73.50 (as it should be for no-arbitrage opportunities), he would buy the asset for USD 70, sell a one-year forward for USD 73.50, and record a profit of USD 3.50.

The USD 3.50 is (of course) the cost of financing the trade and should not be recorded as a profit. In fact, the firm's systems would account for this cost only when the forward contract was a settled. Jett therefore had to continually increase the size of his position to maintain the appearance of profit.

Unaware of the fictitious nature of his profits, Kidder Peabody rewarded Jett with over USD 13 million in bonuses between 1992 and 1994. Eventually, when the nature of his trading was understood, the firm had to revise its profits downward by over USD 300 million.

Generalization

We will now generalize the relationship between the forward price and the spot price of a financial asset that provides no income. Defining terms:

F: Forward price of asset,

T: Time to maturity of the forward contract,

S: Spot price of asset, and

R: Risk-free interest rate per year for maturity T (compounded annually).

As mentioned earlier, the risk-free rate R is the relevant opportunity cost of funds. This rate is usually an interbank borrowing/lending rate.

The no-arbitrage forward price is given by:[2]

$$F = S(1 + R)^T \tag{10.1}$$

The forward price is the spot price compounded forward at the interest rate R for time T.

If $F > S(1 + R)^T$, then we are in Scenario A of Table 10.3. An arbitrageur can lock in a profit of $F - S(1 + R)^T$ by buying the asset and shorting the forward contract.

If $F < S(1 + R)^T$, then we are in Scenario B of Table 10.3. An arbitrageur can lock in a profit of $S(1 + R)^T - F$ by shorting the asset (or selling the asset if it is owned) and taking a long position in the forward contract.

Another way to understand this result is by noting that a trader can exchange a cash flow of $-S$ today for a cash flow of $+F$ at time T with 100% certainty by:

- Buying the asset today, and
- Entering a forward contract to sell it at time T.

Because the cash flows are certain, S must be the present value of F received at time T discounted at the risk-free rate, so that:

$$S = F(1 + R)^{-T}$$

or

$$F = S(1 + R)^T$$

Example:

Consider a forward contract to sell a non-dividend-paying stock in three months. The current stock price is USD 50 and the three-month risk-free rate (annually compounded) is 4% per year.[3] Equation (10.1) gives a forward price of:

$$F = 50(1 + 0.04)^{0.25} = 50.49$$

10.3 THE KNOWN INCOME CASE

We now move on to consider a forward contract on a financial asset paying a known cash income. The asset could be a bond with a known coupon or a stock providing a dividend known in advance.[4]

Reconsider the example in Section 10.2 where an asset has a price of USD 70. Now suppose the asset will provide a cash flow of USD 5 in six months. Assume that interest rates are not the same for all maturities. In particular, assume the following.

- The one-year interest rate is 5% per year annually compounded (as in our earlier example).
- The six-month interest rate is 4% per year annually compounded.

[2] When continuous (instead of annual) compounding is used for R, equation (10.1) becomes $F = Se^{RT}$. Compounding frequency issues are discussed in Chapter 16.

[3] It would be more natural to express the three-month interest rate with quarterly compounding rather than annual compounding. We use annual compounding because R is defined with annual compounding in Equation (10.1). If the rate were 4% per annum with quarterly compounding, the forward price would be

$$50 \times 1.01 = 50.50$$

We will discuss compounding frequency issues in Chapter 16.

[4] The dividend could be known because it has been declared before the ex-dividend date has been reached. Alternatively, management may have indicated that it plans to maintain the current level of dividends.

First, consider Scenario A (where the forward price is USD 80): a trader can follow the strategy described in Section 10.2 and buy the asset for USD 70 while selling a one-year forward for USD 80. In this case, the trader will receive USD 5 in six months. The present value of the income is

$$\frac{5}{(1.04)^{1/2}} = 4.903$$

The trader can therefore borrow USD 4.903 for six months using the income from the asset (equal to USD 5) to pay off this loan in the sixth month when the income is received. He or she can then borrow the remaining USD 65.097 (= 70 − 4.903) at 5% for one year. At the end of the year, the amount required to pay off this loan is

$$\text{USD } 65.097 \times 1.05 = \text{USD } 68.352$$

The profit to the trader is then:

$$\text{USD } 80 - \text{USD } 68.352 = \text{USD } 11.648$$

This trading strategy is profitable for any forward price above USD 68.352. For there to be no arbitrage, it follows that the forward price must be less than USD 68.352.

Now consider Scenario B (where the forward price is USD 65). As in Section 10.2, a trader can short the asset and enter into a forward contract to buy it back for USD 65. In this situation, the trader must pay the income of USD 5 provided by the asset to maintain the short position. The trader can do this by investing USD 4.903 from the proceeds of the short sale for six months at 4% per year and the remaining USD 65.097 for one year at 5% per year. At the end of the sixth months, the first investment generates USD 5 (= $4.903(1.04)^{1/2}$), which is paid as a dividend to the asset owner. The second investment generates USD 68.352 (= 65.097 × 1.05) in one year. The trader therefore makes a profit of:

$$\text{USD } 68.352 - \text{USD } 65 = \text{USD } 3.352$$

This strategy is profitable for any forward price below USD 68.352.

Table 10.4 summarizes these scenarios. The first trading strategy generates profits when the forward price is greater than USD 68.352, while the second trading strategy generates profits when the forward price is less than USD 68.352. For there to be no-arbitrage opportunities, the forward price must therefore be exactly USD 68.352.

Generalization

Define I as the present value of the income from the asset. The rest of the notation is the same as in the previous section (i.e., S is the spot price of the asset, F is the forward price of the asset, T is the time to maturity of the forward contract, and R is the interest rate for maturity T.) The relationship between F and S is

$$F = (S - I)(1 + R)^T \tag{10.2}$$

In the example in Table 10.4:

$$I = \frac{5}{(1.04)^{0.5}} = 4.903$$

and

$$F = (70 - 4.903) \times 1.05 = 68.352$$

If $F > (S - I)(1 + R)^T$, buying the asset and selling it in the forward market will lead to a profit.

If $F < (S - I)(1 + R)^T$, however, shorting the asset and buying it in the forward market will lead to an arbitrage profit.

For another way of deriving Equation (10.2), consider the following trading strategy.

- Buy the asset for S at time zero.
- Enter into a forward contract to sell it for F at time T.

This is certain to lead to a cash outflow of S at time zero, cash inflows with a present value of I during the life of the forward contract, and an inflow of F at time T. Setting the present value of inflows equal to the present value of the outflows yields

$$S = I + \frac{F}{(1 + R)^T}$$

Table 10.4 Trading Possibilities for an Arbitrageur. The Asset Price Is USD 70 and an Income of USD 5 Is Expected in Six Months. The Six-Month Interest Rate (Borrowing or Lending) Is 4%, Whereas the One-Year Interest Rate (Borrowing or Lending) Is 5%

Scenario A: One-Year Forward Price Is USD 80	Scenario B: One-Year Forward Price Is USD 65
Buy asset for USD 70 and enter into a short forward contract to sell it for USD 80 in one year. In six months, collect USD 5.00 of income.	Short asset for USD 70 and enter into a long forward contract to buy it in one year for USD 65. In six months, pay income on the asset of USD 5.00.
Profit equals USD 11.648	Profit equals USD 3.352

This can be written as:

$$F = (S - I)(1 + R)^T$$

Example:

Consider a ten-month forward contract on a bond paying a USD 4 coupon in three months and in nine months. We assume the risk-free rate for all maturities is 6% per year and the cash price of the bond is USD 105. The present value of the coupons in is

$$\frac{4}{1.06^{0.25}} + \frac{4}{1.06^{0.75}} = \text{USD } 7.771$$

The forward price of the bond is therefore:

$$(105 - 7.771)(1.06)^{\frac{10}{12}} = \text{USD } 102.067$$

10.4 THE KNOWN YIELD CASE

Now consider the case where a financial asset provides a known yield during the life of the forward contract. This means the known income is expressed as a percentage of the price of the asset (rather than in cash).

We assume that the yield is Q per year with annual compounding.[5] If the income is reinvested in the asset, the number of units held grows at rate Q. One unit of the asset grows to $(1 + Q)$ units of the asset in one year and to $(1 + Q)^T$ units by time T. Consider the following trading strategy.

- Buy one unit of the asset for S at time zero.
- Enter into a forward contract to sell $(1 + Q)^T$ units of the asset for price F per unit at time T.

This strategy exchanges S at time zero for $F(1 + Q)^T$ at time T. If we continue denoting the risk-free interest rate for maturity T as R (with annual compounding), this means that S is the present value of $F(1 + Q)^T$ so that:

$$S = \frac{F(1 + Q)^T}{(1 + R)^T}$$

or[6]

$$F = S\left(\frac{1 + R}{1 + Q}\right)^T \tag{10.3}$$

[5] Compounding frequencies will be discussed in Chapter 16. Just as an interest rate can be expressed as 4% with semi-annual compounding or equivalently as 4.04% (= 1.02 × 1.02 − 1) with annual compounding, a yield can be expressed with different compounding frequencies. As the compounding frequency increases, the numerical value of the yield decreases.

[6] When R and Q are expressed with continuous compounding, equation (10.3) becomes $F = Se^{(R-Q)T}$

Example:

Consider an asset expected to provide a 3% yield per year over the next three years. The risk-free rate is 4% per year and the current spot price of the asset is USD 30. The forward price (USD) is

$$F = 30\left(\frac{1.04}{1.03}\right)^3 = 30.88$$

10.5 VALUING FORWARD CONTRACTS

The value of a forward contract is quite different from the forward price. When a forward contract on a financial asset is first entered, its forward price is calculated in the manner described in Sections 10.2 to 10.4. However, the value of the forward contract itself is zero (or very close to zero). If this were not so, one party would require a payment from the other at the outset (akin to the premium paid for options). Forward contracts are normally structured so that there is no such payment. As time passes, however, the asset price changes and the value of the forward contract may become positive or negative. While the value of the contract changes, the price at which the asset will be eventually bought or sold continues to equal the original forward price.

Suppose that we are valuing a long forward contract to buy an asset for price K. We assume this is not a new contract, but rather one entered some time ago. The value of the contract depends on movements in the price of the underlying asset and can be positive or negative.

We can value the contract by comparing it with a similar contract that could be entered today. Define K as the forward price at the time the contract was originally entered, F as the current forward price for the contract, and T as the contract's current time to maturity.

The two different forward contracts we consider are

1. A forward contract to buy the asset for price K at time T (the contract we are interested in valuing), and
2. A forward contract to buy the asset for price F at time T (a contract that could be entered today).

The only difference between these two contracts lies in the price paid at time T. The value of the second contract minus the value of the first contract is the present value of $F - K$ (which can be positive or negative.) However, the second contract is worth zero because it is entered at the current forward price. Therefore, the first contract is worth the present value of $F - K$. Specifically:

$$\text{Value of Long Forward Contract} = \frac{F - K}{(1 + R)^T} \tag{10.4}$$

Similarly:

$$\text{Value of Short Forward Contract} = -\frac{F - K}{(1 + R)^T} = \frac{K - F}{(1 + R)^T} \quad (10.5)$$

These formulas are true for all forward contracts on all assets (not just financial assets).

In the case of financial assets, we can use Equations (10.1), (10.2), or (10.3) to calculate F. For example, in the case of an asset providing no income:

$$\text{Value of Long Forward Contract} = \frac{S(1 + R)^T - K}{(1 + R)^T} = S - \frac{K}{(1 + R)^T} \quad (10.6)$$

When an income with a present value of I is to be paid during the remaining life of the forward contract:

$$\text{Value of Long Forward Contract} = \frac{(S - I)(1 + R)^T - K}{(1 + R)^T} = S - I - \frac{K}{(1 + R)^T} \quad (10.7)$$

When a yield at rate Q is provided

$$\text{Value of Long Forward Contract} = \frac{S(1 + R)^T/(1 + Q)^T - K}{(1 + R)^T} = \frac{S}{(1 + Q)^T} - \frac{K}{(1 + R)^T} \quad (10.8)$$

Example:

Let us return to the forward contract we first considered (where the asset price is USD 70, there is no income, and the one-year interest rate is 5%). The current no-arbitrage forward price F is USD 73.50. Suppose some time ago a long forward contract was entered to buy the asset at USD 78. Equation (10.6) gives the value of the contract as:

$$70 - \frac{78}{1.05} = -4.286$$

Alternatively, we can use Equation (10.4) to get

$$\frac{73.5 - 78}{1.05} = -4.286$$

10.6 FORWARD VERSUS FUTURES

Recall that futures contracts are settled daily, while forward contracts are settled at maturity.

Now consider two contracts that are the same in every aspect except that one is a futures contract and the other is a forward contract. It can be shown that if interest rates are constant (or if they change in a perfectly predictable way), the theoretical no-arbitrage forward and futures prices are the same.[7]

In practice, however, interest rates do vary unpredictably and futures prices are therefore different from forward prices. This difference is due to the correlations between the returns from the underlying assets and interest rates.

As an example, suppose that the price of an asset is positively correlated with interest rates. If the asset price increases, there will be an immediate gain from a long futures contract. This gain can then be invested at a relatively high interest rate because (on average) interest rates increase when the asset price increases. If the asset price decreases, there will be an immediate loss on the long futures contract. However, the funds can be financed at a low interest rate because interest rates (on average) decline when the asset price declines. This makes the long futures contract slightly more attractive than a long forward contract and the futures price would therefore be slightly higher than the forward price. When the correlation between the return from the underlying asset and interest rate is negative, however, this argument is reversed and the theoretical futures price is slightly lower than the theoretical forward price.

The differences between futures and forward prices are small and can usually be ignored. However, an exception is the case of Eurodollar futures contracts (which we will discuss in Chapter 19).

Another issue is that while futures contracts can have a range of delivery dates, forward contracts do not. As noted in previous chapters, it is the party with the short position that chooses the delivery time.

Consider the case of financial assets. If the interest rate is greater than the income generated by a financial asset, it is optimal for the party with the short position to deliver as early as possible to avoid financing costs. If the income is greater than the interest rate, the reverse is true and the party with the short position will deliver as late as possible to earn the maximum income on the asset. This observation can be used to determine the equivalent maturity for a forward contract.

To summarize, Equations (10.1) to (10.3) can be used to determine futures prices as well as forward prices. Note that Equations (10.4) to (10.8) are irrelevant for futures prices because daily settlement ensures the value of a futures contract at the end of each day during its life is zero.

The fact that the forward price equals the futures price (to a good approximation) does not mean the profits or losses from the two contracts are the same. Consider a situation where the

[7] See J. C. Cox, J. E. Ingersoll, and S. A. Ross, "The relation between forward prices and futures prices," *Journal of Financial Economics*, 9 (December 1981): 321–346.

one-year forward and futures price for an asset is USD 2.00. Suppose that during a single day, the forward and futures prices increase to USD 2.10. Trader A, who has a long futures contract on 1,000 units of the asset, makes an immediate profit of USD 100 (because of daily settlement). Trader B, who has a long forward contract, also gains USD 100. However, this is in one year's time (because there is no daily settlement for forward contracts). The accounting systems will therefore show Trader A's position has increased by USD 100, whereas Trader B's position has increased by the present value of USD 100. This difference between forwards and futures is symmetrical. If the forward/futures price declined by USD 0.10, Trader A would lose USD 100 immediately, whereas Trader B would lose slightly less on paper (i.e., the present value of USD 100).

10.7 EXCHANGE RATES REVISITED

Chapter 9 showed that the no-arbitrage argument led to the relationship between forward rates and spot rates for the GBP/USD exchange rate being:

$$F = S\frac{(1 + R_{USD})^T}{(1 + R_{GBP})^T} \qquad (10.9)$$

Recall that GBP/USD represents the exchange rate as the number of USD (U.S. Dollar) per GBP (British Pound Sterling).

From the perspective of a U.S. investor, GBP can be treated in the same way as any other asset valued in USD. In this case, the asset provides income equal to the GBP interest rate. However, the income is received in GBP and not USD, and thus the value of the income (to a U.S. investor) is the GBP interest rate multiplied by the value of GBP in USD. In other words, the income from one unit of the asset is R_{GBP} multiplied by its price. This makes the foreign risk-free rate a yield (rather than known income), and thus Equation (10.3) applies. It is therefore not surprising that when we set R equal to R_{USD} and Q equal to R_{GBP}, Equation (10.3) becomes Equation (10.9).

10.8 STOCK INDICES

A stock index tracks the value of a hypothetical stock portfolio (i.e., if the value of the hypothetical portfolio increases by X%, the index also increases by X%). The CME Group offers exchange trading platforms for futures on several different stock indices. Among these are the S&P 500 (a portfolio of 500 stocks), the NASDAQ-100 (a portfolio of 100 stocks traded on the Nasdaq Stock Market), and the Dow Jones Index (a portfolio of 30 large stocks).

The weight of a stock in a portfolio is the percentage of the portfolio invested in the stock. In the case of the S&P 500 and the NASDAQ-100, the weight of a stock is proportional to its market capitalization (i.e., its share price multiplied by the number of shares outstanding). In the case of the Dow Jones Index, the weight is proportional to the share price.

The two S&P 500 futures contracts traded on the CME are on USD 50 and USD 250 multiplied by the index. The two CME futures contracts on the NASDAQ are on USD 20 and USD 100 multiplied by the index. The two CME futures contracts on the Dow Jones are on USD 5 and USD 10 multiplied by the index. All contracts are settled in cash (rather than by delivering the portfolio). For example, final settlement of the S&P 500 contract is the opening price of the S&P 500 on the third Friday of the delivery month.

Similar futures contracts trade actively in other countries. For example, futures on the CSI 300 Index (a market-capitalization weighted portfolio of 300 Chinese stocks) trades on the China Financial Futures Exchange (CFFEX).

A stock index can be regarded as a financial asset that pays dividends.[8] Because the asset is the portfolio of stocks underlying the index, the dividends are the dividends received by an investor holding the portfolio. It is therefore possible to go through the stocks in the portfolio and estimate the dividend on each to produce the total estimated cash income on the index. This would allow the use of Equation (10.2). In practice, it is usually assumed the dividends on an index provide a known yield so that Equation (10.3) applies.[9] The value of Q is the average dividend yield during the life of the forward/futures contract.

Example:

Consider an index that is 2,500 in a situation when the risk-free rate is 5% per year and the dividend yield is 3% per year (both are assumed to be the same for all maturities). The futures price for a contract where the final settlement will be in six months is

$$2{,}500 \times \left(\frac{1.05}{1.03}\right)^{0.5} = 2{,}524$$

Index Arbitrage

If an index futures price is greater than its theoretical value, an arbitrageur can buy the portfolio of stocks underlying the index

[8] Dividend payments are not usually considered when an index is calculated and so the index does not represent the total return that would be earned by an investor in the stocks underlying the index. An exception is a total return index, which is created by reinvesting dividends from the hypothetical portfolio in the hypothetical portfolio.

[9] Several financial institutions provide estimates of the dividend yields on the indices that are commonly referred to by investors and on which futures contracts trade.

and sell the futures. If the futures price is less than the theoretical price, the arbitrageur can short the stocks underlying the index and take a long futures position. Both of these trading strategies can be categorized as *index arbitrage*.

Typically, a computer program is used to send all the required trades (for the stocks underlying the index) to an exchange at the same time as the futures contract is traded. This is known as *program trading*. Sometimes index arbitrage is done by trading a representative subset of the stocks underlying the index.

Index arbitrage typically ensures that the theoretical relationship between the index and futures on the index holds. Occasionally, there are exceptions. On October 19, 1987, (Black Monday) the market declined by more than 20% and the number of shares traded on the New York Stock Exchange easily exceeded all previous records. This led to delays in processing orders and index arbitrage strategies could not be carried out efficiently. As a result, the futures price of the index was well below the theoretical price given by Equation (10.3).

Indices Not Representing Tradable Portfolios

For the no-arbitrage Equation (10.3) to be applicable, it must be possible to trade a portfolio whose price always equals the index. While this is usually the case, there are scenarios where the index (although well defined) does not correspond to the value of a tradable portfolio.

For example, the Chicago Mercantile Exchange offers a futures contract on the Japanese Nikkei 225 Index that is settled in dollars rather than yen. Suppose the Nikkei Index is denoted by Z. A trader can (in theory) trade a portfolio that is always worth Z yen. Due to fluctuating exchange rates, however, the trader cannot trade a portfolio that is always worth Z dollars.

SUMMARY

This chapter has considered futures prices and forward prices for financial assets. The results are summarized in Table 10.5. In most circumstances, the futures prices and forward prices for contracts on the same asset with the same maturity are (approximately) the same.

Interest rate futures are an important category of financial futures, which have not been considered in any detail so far. Although similar no-arbitrage approaches can be used to determine their prices, interest rate futures have special features and therefore deserve special attention. They are covered in greater detail in Chapter 19.

Table 10.5 Summary of Results for a Forward Contract on a Financial Asset; *S* Is Current Price of Asset, *T* Is Maturity of Contract, and *R* Is the Interest Rate

Asset	Forward or Futures Price	Value of Long Forward with Agreed Price K
Provides No Income	$S(1+R)^T$	$S - \frac{K}{(1+R)^T}$
Provides Known Income with Present Value I	$(S - I)(1+R)^T$	$S - I - \frac{K}{(1+R)^T}$
Provides a Known Yield Equal to Q	$S\left(\frac{1+R}{1+Q}\right)^T$	$\frac{S}{(1+Q)^T} - \frac{K}{(1+R)^T}$

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

10.1 Explain what happens when shares are shorted.

10.2 "If you buy a non-income-producing financial asset for 100 and enter into a forward contract to sell it for 110 in one year, you have made a profit of 10." Is this statement true? Why or why not? Consider the situations where the interest rate is 4% per year and 10% per year.

10.3 What is the formula for the futures price of a financial asset that provides no income?

10.4 What is the formula for the futures price of a financial asset that provides income with a present value of I?

10.5 What is meant by an asset that provides a constant yield?

10.6 What is the formula for the futures price of a financial asset that provides a constant yield at rate Q?

10.7 What is the difference between a forward price and the value of a forward contract?

10.8 If the return from an asset is positively correlated with interest rates, would you prefer to enter into a long forward contract or a similar long futures contract? Explain.

10.9 Explain why the interest rate earned in a foreign currency can be regarded as a yield.

10.10 Explain how index arbitrage is accomplished when the futures price is higher than its theoretical value.

Practice Questions

10.11 Suppose that you enter into a two-year forward contract on a non-dividend-paying stock when the stock price is USD 40 and the risk-free rate (annually compounded) is 10% per year. What do you expect the forward price to be?

10.12 Is the futures price of a non-dividend paying stock likely to be greater or less than the expected future stock price? Explain your argument.

10.13 The cash price of a bond is USD 90. It is expected to provide a coupon of USD 3 in six months and 12 months. The risk-free rate for all maturities is 5% per year (with annual compounding). What is the 15-month forward price of the bond?

10.14 A stock index is 3,000, the risk-free rate is 8% per year, and the dividend yield on the index is 3% per year (both expressed with annual compounding). What should the one-year futures price of the index be?

10.15 Explain the argument used for calculating the value of a forward contract.

10.16 A one-year forward contract to buy a non-dividend-paying stock is entered into when the stock price is USD 50 and the risk-free rate is 5% per year (with annual compounding). What is the (a) forward price and (b) value of the forward contract?

10.17 Six months after the forward contract in Question 10.16 was entered into, the spot price is USD 56 and the risk-free rate is still 5% per year. What is the (a) forward price and (b) value of the forward contract?

10.18 If the contract in Questions 10.16 and 10.17 had been a futures contract, how would your answers change?

10.19 If the return from an asset is negatively correlated with interest rates, would you expect the forward price to be greater than or less than the futures price? Explain.

10.20 If a stock index, interest rate, and dividend yield remain constant, derive a formula for the futures price at time t in terms of the futures price at time zero. Suppose that the risk-free rate is 5% per year and the dividend yield on an index is 3% per year. If the stock index stays constant, at what rate does the futures price grow? (All rates are expressed with annual compounding.)

[*Hint*: If the futures contract initially has a time to maturity equal to T, it has a time to maturity equal to $T - t$ at time t. Use Equation (10.3) to calculate the futures price F_0 at time zero and the future price F_t at time t.]

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

10.1 The stock is borrowed from another investor and sold in the market in the usual way. At a later time, it is purchased and returned to the account from which it was borrowed. Any income on the security has to be paid by the borrower. Sometimes there is a small fee for borrowing the stock.

10.2 The profit is less than 10 because the investment of 100 must be financed. If the interest is 4% per year, the profit is only 6. If it is 10% per year there is no profit. This was the issue that led to Joseph Jett's false recording of profit.

10.3 $F = S(1 + R)^T$

10.4 $F = (S - I)(1 + R)^T$

10.5 The income when expressed as a proportion of the asset price is constant. If a yield at rate Q is reinvested in the asset, the holding of the asset grows at rate Q.

10.6 $F = S\left(\frac{1 + R}{1 + Q}\right)^T$

10.7 The forward price is the delivery price that would be negotiated today in a forward contract. The value of a forward contract is initially zero and then moves up or down as the asset price changes.

10.8 You would prefer to own a long futures contract because daily gains will tend to be invested at a relatively high rate and daily losses will tend to be financed at a relatively low rate.

10.9 If interest is earned in the foreign currency, its value to a domestic investor is proportional to the value of the foreign currency.

10.10 A trader buys the stocks underlying the index and sells futures.

Solved Problems

10.11 The forward price is USD 40 $\times$ 1.1^2 or USD 48.4.

10.12 The futures price is the spot price compounded forward at the risk-free rate. Most stocks can be expected to provide a return greater than the risk-free rate. Hence, the expected future stock price is greater than the forward price.

10.13 The present value of the income (USD) is

$$\frac{3}{1.05^{0.5}} + \frac{3}{1.05} = 5.785$$

The forward price of the bond is

$$(90 - 5.785) \times 1.05^{1.25} = 89.51$$

10.14 The one-year futures price is

$$3{,}000 \times \frac{1.08}{1.03} = 3145.63$$

10.15 The forward contract is compared to a forward contract with the same maturity that would be entered into today. The present value of the difference between their payoffs can be calculated. We know that the value of a forward contract that would be entered into today is zero and can therefore deduce the value of the other forward contract.

10.16 The forward price is $50 \times 1.05 = 52.5$. The value of the forward contract is zero.

10.17 The forward price is $56 \times 1.05^{0.5} = 57.383$. The value of the forward contract is

$$\frac{57.383 - 52.5}{1.05^{0.5}} = 4.77$$

10.18 The value of the contract in Question 10.17 would be zero because of daily settlement. Other answers are the same.

10.19 When there is a negative correlation between return on assets and interest rates, when the asset price increases, funds tend to be invested at a relatively lower rate. On the other hand, when the asset price decreases, funds can be invested at a relatively higher rate. This makes a long futures contract less attractive than a long forward contract and the futures price will be lower than the forward price.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

10.20 The relationship between the futures price, F_t, at time t and the spot price is with the notation in the chapter:

$$F_t = S\left(\frac{1+R}{1+Q}\right)^{T-t} = S\left(\frac{1+R}{1+Q}\right)^{T}\left(\frac{1+Q}{1+R}\right)^{t} = F_0\left(\frac{1+Q}{1+R}\right)^{t}$$

where S, R, and Q are the index level, risk-free rate, and dividend yield, respectively and T is the initial time to maturity. This shows that the futures price grows at:

$$\frac{1+Q}{1+R} - 1$$

When $R = 5\%$ and $Q = 3\%$, the growth rate of the futures price per year is

$$\frac{1+Q}{1+R} - 1 = \frac{1.03}{1.05} - 1 = -0.019$$

or -1.9%.

11

Commodity Forwards and Futures

Learning Objectives

After completing this reading, you should be able to:

- Explain the key differences between commodities and financial assets.
- Define and apply commodity concepts such as storage costs, carry markets, lease rate, and convenience yield.
- Identify factors that impact prices on agricultural commodities, metals, energy, and weather derivatives.
- Explain the formula for pricing commodity forwards.
- Describe an arbitrage transaction in commodity forwards and compute the potential arbitrage profit.
- Define the lease rate and explain how it determines the no-arbitrage values for commodity forwards and futures.
- Describe the cost of carry model and determine the impact of storage costs and convenience yields on commodity forward prices and no-arbitrage bounds.
- Compute the forward price of a commodity with storage costs.
- Explain how to create a synthetic commodity position and use it to explain the relationship between the forward price and the expected future spot price.
- Explain the impact of systematic and nonsystematic risk on current futures prices and expected future spot prices.
- Define and interpret normal backwardation and contango.

This chapter considers forward and futures contracts on commodities. Chapter 10 showed how the price of a futures contract and the price of a forward contract with the same underlying asset and maturity are approximately equal in most cases. This is true whether the asset is a commodity or a financial asset. This chapter will therefore treat futures as forwards (i.e., we will ignore the daily settlement).[1]

Most commodities are consumption assets. This means that they are rarely held for purely investment reasons (metals such as gold and silver are exceptions). Commodity owners usually intend to use the commodity in some way, after which it ceases to be available for sale. The no-arbitrage arguments presented in Chapter 10 did not account for this and therefore are not fully applicable to consumption assets.

Chapter 10 showed that the no-arbitrage forward/futures price of a financial asset can be computed from risk-free interest rates and the income generated by the asset. For most commodities, however, no-arbitrage arguments can only be used to obtain an upper bound for the futures price. An additional unobserved parameter, known as the convenience yield, is required to determine the commodity futures price in the market.

11.1 WHY COMMODITIES ARE DIFFERENT

There are several important differences between commodities and financial assets. Some of these are as follows.

- The storage costs associated with financial assets (e.g., stocks and bonds) are negligible. The storage costs for commodities, however, can be quite substantial. These costs include insurance, which can vary over time. Some commodities deteriorate with time and require costly care in storage. Assets such as corn and natural gas are frequently stored for use at a particular time of the year. Other assets, such as oil and copper, are consumed throughout the year.
- Commodities can be costly to transport and thus their prices can depend on their location. By contrast, financial assets are usually transported electronically at virtually no cost.
- A commodity held for investment purposes (e.g., gold or silver) can be borrowed for shorting (to be discussed later). However, the borrower must pay what is called a lease rate. This lease rate can exceed the fees charged when financial assets are borrowed for shorting (see Section 10.1).
- A financial asset provides investors with an expected financial return that reflects its risk. Most commodities do not have this property. Indeed, it can be argued that the prices of most commodities are mean reverting. This means that although the price of a commodity can be quite volatile, it tends to get pulled back toward some central value. When the price of a commodity is relatively high, its production will become attractive and its supply will therefore tend to increase. At the same time, users of the commodity may search for less expensive alternatives. Combined, these actions will tend to reduce the price of the commodity. If the price is at a relatively low level, on the other hand, the production of the commodity will become less attractive, while its use will tend to increase. As a result, its price will tend to rise.

As a result of these distinctions, the futures prices of commodities can behave quite differently from those of financial assets.

11.2 TYPES OF COMMODITIES

This section reviews the properties of various commodity categories.

Agricultural Commodities

Agricultural commodities with futures contracts include products that are grown (e.g., corn, wheat, soybeans, cocoa, and sugar) as well as livestock (e.g., cattle and hogs).

It is expensive to store agricultural commodities; even under optimal conditions, some products can only be stored for a limited period of time. There is also an interdependence among various agricultural commodities. For example, the cost of feeding livestock can depend on the prices of grown agricultural commodities (e.g., corn).

The prices of agricultural commodities can be seasonal. For example, consider the prices of corn and soybeans. At harvesting time (October to November), their prices tend to be relatively low. Outside of this period, however, their prices may be higher as farmers and other distributors incur storage costs. This seasonality is sometimes reflected in futures prices, causing them to display a mixture of normal and inverted pricing patterns.[2]

[1] In futures contracts, there are often a range of delivery dates. If the futures price is a decreasing function of maturity, the benefits of holding the commodity in inventory exceed the financing and storage costs so that the party with the short position will tend to deliver at the latest possible time and the futures contract can be considered equivalent to a forward contract with this late maturity. If the futures price is an increasing function of maturity, the reverse is true and the futures contract can be considered equivalent to a forward contract with a maturity equal to the earliest possible delivery date.

[2] See Section 7.5 for a discussion of normal and inverted patterns for futures prices.

There are also other factors that may affect the market's view on the future price of a commodity. If there has been (or if there is expected to be) a good (or bad) harvest, market participants will expect prices to be relatively low (or relatively high). Political considerations can also affect futures prices.[3]

Weather is also an important consideration for many agricultural commodities. Bad weather in Florida, for example, can increase the futures price of frozen orange juice. Meanwhile, frosts in Brazil are liable to drastically reduce Brazilian coffee production and increase coffee futures prices.

Metals

Commodity metals include gold, silver, platinum, palladium, copper, tin, lead, zinc, nickel, and aluminum. Their properties are quite different from those of agricultural products. For example, metal prices are not affected by the weather and are not seasonal (because they are extracted from the ground). Additionally, their storage costs are typically lower than those of agricultural products.

Some metals are held purely for investment purposes. This means (as we will explain later) that their futures price can be more easily obtained from observable variables. For someone looking to hold a metal for investment, owning a futures contract can be an acceptable alternative to owning the physical asset itself.

As in the case of agricultural commodities, inventory levels are important in determining prices. Most metals are extracted in one country and consumed in another. As a result, exchange rates may affect prices. Other crucial factors include the scope of a metal's industrial application and the rate at which new sources are found. Changes in extraction methods, actions by governments (and/or cartels), and environmental regulations can also impact metal prices. Sometimes metal prices can even be affected by recycling processes. For example, a metal used in a production process one year can be recycled and re-enter the market many years later.

Energy

Energy products are another important category of commodities. There are futures contracts on crude oil and crude oil extracts (e.g., petroleum and heating oil). Futures also trade on natural gas and electricity.

The crude oil market is the largest commodity market in the world and global demand is estimated to be about 100 million barrels per day.[4] There are many grades of crude oil reflecting variations in gravity (density) and sulfur content. Two important benchmarks are Brent crude oil (which is sourced from the North Sea) and West Texas Intermediate (WTI) crude oil.

Natural gas is used for heating and generating electricity. It can be stored (either above ground or underground) for indefinite periods, but its storage costs are quite high. Natural gas is also expensive to transport, and hence the price of natural gas can vary regionally.

CME Group offers a futures contract on 10,000 million British thermal units (BTUs) of natural gas. The contract requires delivery to be made at a roughly uniform rate to a hub in Louisiana. Intercontinental Exchange (ICE) also offers a futures contract on natural gas.

The demand for natural gas is high in the winter (for heating purposes) and to a lesser extent the summer (to produce electricity for air conditioning). This creates seasonality in the futures prices as illustrated by Table 11.1 and Figure 11.1.

Electricity is an unusual commodity because it is almost impossible to store.[5] The maximum supply of electricity in a region is determined by the capacity of all electricity generating plants in that region. In the U.S., there are 140 such regions and the wholesale price of electricity is determined by the price charged in each of them.[6] The non-storability of electricity can lead to huge price swings. For example, heat waves in the summer can drive up the demand for air conditioning and have been known to increase the cost of electricity by as much as 1,000%. Once the heat wave is over, however, the price quickly returns to normal levels (an example of the mean reversion tendencies we mentioned earlier).

Futures contracts on electricity do exist, but they are less actively traded than futures contracts on natural gas and crude oil. Electricity futures contracts also trade in the over-the-counter market. A typical contract allows one side to receive a specified number of megawatt hours for a specified price at a specified location during a particular month. In a 5×8 contract, power is received during the off-peak period (11 p.m. to 7 a.m.) from Monday to Friday. In a 5×16 contract, power is received during the on-peak period (7 a.m. to 11 p.m.) from Monday to Friday for the specified month. In a 7×24 contract, power is received 24 hours a day, seven days a week for the specified month. Prices are highly seasonal, reflecting increased demand in the summer.

[3] For example, escalating trade tensions between the U.S. and China in 2018–2019 made it difficult for U.S. farmers to export commodities such as corn and soybeans.

[4] International Energy Agency. (n.d.). Oil Market Report.

[5] There some artificial ways of storing electricity. For example, electricity can be used to pump water to the top of a hydroelectric plant so that it can be released in a way that produces electricity at a later time.

[6] Whether electricity can be sold across regions depends on the existence of, and the capacity of, transmission lines.

Table 11.1 Seasonality of CME Natural Gas Futures Prices on January 24, 2020

Maturity Month	Settlement Price
February 2020	1.926
March 2020	1.904
April 2020	1.940
May 2020	1.997
June 2020	2.078
July 2020	2.141
August 2020	2.170
September 2020	2.166
October 2020	2.199
November 2020	2.307
December 2020	2.501
January 2021	2.609
February 2021	2.576
March 2021	2.478
April 2021	2.255
May 2021	2.237
June 2021	2.271
July 2021	2.306
August 2021	2.310
September 2021	2.296
October 2021	2.320
November 2021	2.384
December 2021	2.553
January 2022	2.672
February 2022	2.636
March 2022	2.507
April 2022	2.259

Source: www.cmegroup.com

Weather

Derivative contracts on weather are available in both the exchange-traded and over-the-counter markets. The most popular contracts are those with payoffs contingent on temperature (which are used by energy companies as hedges).

Two important weather derivative variables are heating degree days (HDDs) and cooling degree days (CDDs). The HDD and CDD for a day are (respectively) defined as:

$$\text{HDD} = \max(0, 65 - A)$$

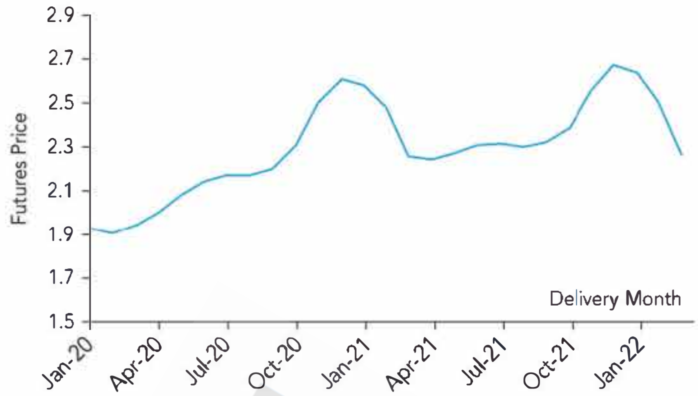


Figure 11.1 Plot of futures prices for natural gas from Table 11.1.

and

$$\text{CDD} = \max(0, A - 65)$$

where A is the average of the highest and lowest temperature during the day at a specified weather station (as measured in degrees Fahrenheit). For example, if the minimum temperature during a day (midnight to midnight) is 40 degrees Fahrenheit and the maximum is 60 degrees Fahrenheit, $A = (60 + 40)/2 = 50$. The daily CDD is therefore zero, and the daily HDD is 15. Contracts are usually defined in terms of the cumulative HDD and CDD for all the days in a given month.

11.3 COMMODITIES HELD FOR INVESTMENT

Some precious metals are held for investment purposes. Gold and silver (and to a lesser extent platinum and palladium) are in this category. While these metals have industrial applications, some purchasers hold these commodities purely for investment purposes. For these people, owning a futures or forward contract can be a practical substitute for owning the commodity itself.

The storage costs of the metals held for investment are generally low compared to metal prices and can therefore be ignored. Additionally, there is a generally small lease rate associated with metals held for investment. For example, gold is like a financial asset in that it can be lent from one entity to another to earn interest.

In analyzing the futures prices for investment commodities, we first ignore the lease rate. The commodities are then akin to financial assets providing no income and therefore Equation (10.1) of Chapter 10 should apply. This means that the relationship between the forward price and the spot price is

$$F = S(1 + R)^T$$

where T is the time to maturity of the forward (measured in years) and R is the annually compounded risk-free interest rate for maturity T.[7] The no-arbitrage arguments are as follows.

- *Arbitrage A:* If the forward price is greater than $S(1 + R)^T$, a trader can buy the investment commodity at price S (by financing the purchase at rate R) and at the same time enter into a forward contract to sell it at time T.
- *Arbitrage B:* If the forward price is less than $S(1 + R)^T$, a trader who owns the investment commodity will find it profitable to sell it at price S and at the same time enter into a forward contract to buy it back at time T.

Lease Rates

The lease rate for an investment commodity is the interest rate charged to borrow the underlying asset.

Suppose a gold producer enters into a contract with an investment bank to sell its future production forward at a predetermined price. The bank will then hedge the risk it is taking on (with respect to the price of gold) by borrowing gold for a time period equal to the life of the forward contract and selling it in the spot market. This synthetically creates a short forward contract that offsets the contract it has with the gold producer. The main lenders of gold are central banks.[8]

Let ℓ be the lease rate. The relationship between the forward price and the spot price is determined from Equation (10.3) as:[9]

$$F = S\left(\frac{1 + R}{1 + \ell}\right)^T$$

This equation can be used to calculate an implied lease rate:

$$\ell = \left(\frac{S}{F}\right)^{\frac{1}{T}}(1 + R) - 1 \tag{11.1}$$

Example:

Assume that the spot price of gold is USD 1,240, the six-month futures price is USD 1,250, and the six-month risk-free rate is 4% (with annual compounding). The implied lease rate is given by:

$$\left(\frac{1{,}240}{1{,}250}\right)^{\frac{1}{0.5}} \times 1.04 - 1 = 0.023$$

or 2.3%.

The lease rate of gold varies with the supply of gold that can be borrowed along with the demand to borrow gold. Recall that when gold producers hedge future production, the banks on the other side of the transaction borrow gold. As gold producers hedge more (less), the demand for borrowed gold will increase (decrease) and the lease rate will rise (fall). Similarly, as asset owners become more (less) willing to lend gold, the lease rate will tend to fall (rise). Occasionally the lease rate is negative and may therefore allow arbitrageurs to buy the metal and sell it forward for a profit.

11.4 CONVENIENCE YIELDS

Now consider commodities that are held almost always for consumption purposes (e.g., crude oil, copper, and corn). Assume that there are no storage costs. If $F > S(1 + R)^T$, traders can undertake Arbitrage A in Section 11.3. Even though the commodity is not an investment asset, traders can still buy the asset by financing the purchase at rate R and entering a forward contract to sell it at price F at time T.

Therefore, the upper bound to the forward price is the spot price compounded forward at the risk-free rate:

$$F < S(1 + R)^T$$

When the storage costs with present value U are considered, the arbitrageur has extra costs to finance, and there is an extra repayment of $U(1 + R)^T$ required at time T. This means that the upper bound is given by:

$$F < (S + U)(1 + R)^T \tag{11.2}$$

A storage cost can be regarded as negative income. Chapter 10 showed that when there is income with present value I, the no-arbitrage condition is

$$F < (S - I)(1 + R)^T$$

Equation (11.2) can be obtained by setting $I = -U$.

Now consider Arbitrage B in Section 11.3. This arbitrage involves entities that own the asset choosing to sell it and enter into a forward contract to buy it back at time T. In the case of a consumption asset, individuals and companies that own the asset must plan to use it in some way. If they sell the asset, they can no longer do that. Therefore, asset owners are usually not prepared to take this arbitrage opportunity.

To summarize, all that can be deduced for a consumption asset is that Equation (11.2) holds, giving an upper bound to forward/futures prices.

Despite this limitation, however, we can measure the extent to which the owners of a consumption asset value having it

[7] When R is expressed with continuous compounding, this equation becomes $F = Se^{RT}$.

[8] The Federal Reserve Board does not lend gold, but other central governments do.

[9] This assumes that the interest rate and the lease rate are expressed with annual compounding. When they are expressed with continuous compounding, this equation becomes $F = Se^{(R-\ell)T}$.

readily available in their inventory by using what is called the *convenience yield*.

The convenience yield essentially addresses the following question: If the underlying commodity were an investment asset, what yield Y would be necessary to explain the observed futures price? In other words, Y (measured with annual compounding) satisfies

$$F = (S + U)\left(\frac{1 + R}{1 + Y}\right)^T$$

so that

$$Y = \left(\frac{S + U}{F}\right)^{\frac{1}{T}}(1 + R) - 1 \qquad (11.3)$$

The convenience yield measures the benefits to the asset holder of having it in their inventory as a protection against future shortages or delivery delays.

Suppose there is expected to be a plentiful supply of the asset during the life of a futures contract. This means it can be ordered at any time for almost immediate delivery. In this case, the convenience yield is likely to be close to zero and $F = (S + U)(1 + R)^T$ might be a reasonable approximation. If inventory levels of the asset are low and there are concerns about shortages, however, the convenience yield will likely be quite high. In that case, F is likely to be less than $(S + U)(1 + R)^T$.

Example:

Suppose the spot price of oil is USD 69 per barrel, and the six-month futures price is USD 65 per barrel. The cost of storing oil for six months has a present value of USD 1 per barrel, and the risk-free rate is 2% per year. The convenience yield satisfies

$$65 = (69 + 1)\left(\frac{1.02}{1 + Y}\right)^{1/2}$$

so that

$$Y = 1.02\left(\frac{70}{65}\right)^2 - 1 = 0.183$$

That is, the convenience yield is 18.3%.

One way of interpreting this number is to view it as the cost of borrowing oil (if that were possible). Because physical oil provides benefits to the holder at the rate of 18.3% per year, this is arguably the rate that would be charged to borrow it.[10]

As we will explain later, the futures price of an asset can be related to the expected future spot price. In our example, it appears that the market was expecting the price of oil to decline over the six-month period. Under such circumstances, holding oil as an investment makes no sense. (Indeed, if oil were an investment asset, its price would reflect market sentiment.) The only reason someone would hold oil is to use it. An expectation that the price of an asset will decline sharply is therefore consistent with the physical asset having a high convenience yield.

11.5 COST OF CARRY

The cost of carry for an asset reflects the impact of:

- Storage costs,
- Financing costs, and
- Income earned on the asset.

In the case of a financial asset, there are no storage costs. If the financing cost is R and the yield is Q (both expressed with annual compounding), the cost of carry per year is

$$\frac{1 + R}{1 + Q} - 1 \qquad (11.4)$$

which is often simply approximated as $R - Q$.

Example:

Consider a foreign currency. If the domestic interest rate is 4% and the foreign interest rate is 3%, the cost of carry is approximately 1%. If the foreign interest rate were 6%, the cost of carry would be around −2% (this is referred to as a *negative carry*).

In the case of a commodity, there is usually no income.[11] The cost of carry therefore consists of the interest rate and the storage cost. In the example at the end of Section 11.4, the cost of carrying oil for six months is given by:

- A 2% per annum interest cost, and
- A storage cost of about 1/69 = 1.45% per six months (or about 2.9% per year).[12]

The total cost of carry is therefore around 4.9% (= 2% + 2.9%) per year.

[10] In this context, note that Equation (11.3), when there are no storage costs ($U = 0$), is the same as Equation (11.1).

[11] The lease rate on investment commodities is an exception.

[12] This is an approximation because the USD 1 of storage costs is a present value and not the amount paid in arrears for storage. Storage costs would be 2.9% per year with annual compounding if they equaled 2.9% of the value at the beginning of the year and were paid at the end of each year. Storage costs would be 2.9% per year with semiannual compounding if the storage costs were paid at the end of each six-month period and equaled 1.45% of the value at the beginning of the six-month period.

A Note on Compounding Frequencies for Interest Rates

Chapter 16 discusses compounding frequencies in detail and introduces continuously compounded interest rates (which are widely used in the valuation of options and other derivatives). A continuously compounded interest rate can be thought of as an interest rate where compounding takes place very frequently (to the point where the balance grows by a small amount every instant). A continuously compounded rate is equivalent to an annually compounded rate if both provide the same result at the end of a year.

When R and Q are expressed with continuous compounding, Equation (11.4) becomes

$$R - Q$$

without any approximation. Meanwhile, the forward price formula in Equation (10.3) becomes

$$F = Se^{(R-Q)T}$$

where e is the mathematical constant that is approximately equal to 2.71828. When the storage cost (as a fraction of the asset price) is expressed with continuous compounding, the cost of carry for a consumption commodity is the interest rate plus the storage cost. The relationship between the futures price and the spot price is

$$F = Se^{(C-Y)T} \tag{11.5}$$

where C is the cost of carry and Y is the convenience yield (both expressed with continuous compounding).

When the annually compounded rate is low, it more closely resembles the equivalent continuously compounded rate and thus it is often reasonable to ignore compounding frequency issues. As the rate increases in magnitude, there is a greater approximation error in treating these two rates as equivalent.

We can illustrate this with the oil example. The interest rate was 2% per annum and the storage cost was about 2.9% per annum. As an approximation, we treat these as continuously compounded and set $C = 4.9\%$ in Equation (11.5). However, as shown in Chapter 16, the annually compounded convenience yield on oil of 18.3% (which we calculated earlier) is equivalent to a continuously compounded rate of 16.8%.[13] To see that Equation (11.5) is correct, substitute for the continuously compounded estimates of C and Y to get

$$F = 69e^{(0.049-0.168)\times 0.5}$$

which is close to the assumed futures price of USD 65.

[13] This is because $\ln(1.183) = 0.168$.

11.6 EXPECTED FUTURE SPOT PRICES

The expected future spot price of an asset is the market's average opinion about what the spot price will be in the future. Natural questions are as follows.

- Does the futures price of an asset equal the expected future spot price?
- Is the futures price a good forecast of the future spot price?

The futures price converges to the spot price at maturity of the contract. If an investor thinks the spot price at maturity will be greater than the current futures price, the investor can take a long futures position. Similarly, if an investor thinks the spot price at the maturity will be less than the current futures price, the investor can take a short futures position. In either case, an investor that is correct will be able to close out the futures contract near the maturity for a profit.

These trading strategies do not involve storing the commodity or investing in a physical asset. They only involve trading futures contracts. In the example where the six-month futures price of oil is USD 65, investors who think this is too low will take long futures positions, and investors who think it is too high will take short futures positions.

Early Work

Economist John Maynard Keynes was one of the first researchers to consider the relationship between futures prices and expected future spot prices.[14] He argued that speculators require compensation for the risks they are bearing, whereas hedgers derive benefits from futures and thus do not require any such compensation. Indeed, the hedgers might be prepared to lose money (on average) because their overall market risks are reduced by hedging.

If hedgers tend to hold short positions, and speculators tend to hold long positions, Keynes' argument suggests the longs will tend to make money on average and the shorts will tend to lose money on average. The futures price should therefore be less than the expected future spot price. If the reverse is true (i.e., hedgers tend to hold long positions and the speculators hold short positions), the shorts will make money on average, and the futures price will be greater than the expected future spot price.

Modern Theory

More recent work on this subject has involved the capital asset pricing model (CAPM). Using CAPM, it is argued that

[14] See J. M. Keynes, *A Treatise on Money*, London Macmillan, 1930.

an investor should earn a return greater than the risk-free rate when the systematic risk of his or her portfolio is positive. Systematic risk is the risk related to the performance of the market as a whole and cannot be diversified away. Non-systematic risk, by contrast, can be virtually eliminated by investing in a well-diversified portfolio.

In this context, the "market" is generally considered to be a well-diversified stock portfolio (such as that underlying the S&P 500 Index). CAPM implies that if the return from an investment is positively correlated with that of the S&P 500, it should earn an expected return greater than the risk-free rate. If the return from the investment is uncorrelated with the S&P 500, the risk is non-systematic, and the expected investment return should equal the risk-free rate. Moreover, if the return from the investment is negatively correlated with the S&P 500 (so that there is negative systematic risk), the expected investment return should be less than the risk-free rate.

Consider our prior notation: S is the spot price of an asset, F is the futures price for maturity T, and R is the annually compounded risk-free interest rate for maturity T.[15] Suppose that P is the present value of the futures price when discounted from time T to time zero at the risk-free rate:

$$P = \frac{F}{(1 + R)^T}$$

By investing P at the risk-free rate, the trader can be sure of obtaining F at time T. The trader can therefore synthetically create a long position in the asset at time T by:

- Investing P at the risk-free rate, and
- Entering into a long futures contract to buy the asset for F at time T.

Let S_T be the spot price at time T. The cash flows from the trader's strategy are

Time 0: $-P$, and

Time T: $+S_T$.

Let E denote the expected value operator. Thus, the expected cash flow at time T is $E(S_T)$. Suppose the expected return from the trades we have just considered is X (with annual compounding). It follows that:

$$E(S_T) = P(1 + X)^T$$

Substituting P gives

$$E(S_T) = F\frac{(1 + X)^T}{(1 + R)^T}$$

[15] As indicated earlier, the relevant risk-free interest rate is the opportunity cost of funds to the trader. This could be an interbank borrowing/lending rate.

This shows that the relationship between $E(S_T)$ and F depends on X and R. If $X > R$, then $E(S_T) > F$. If $X < R$, however, then $E(S_T) < F$.

The variable X is the return from the synthetically produced position. The systematic risk of this investment depends on the correlation between the underlying asset return and the market return. If this correlation is positive, the systematic risk of the investment will be positive (which implies $X > R$ and therefore $E(S_T) > F$). If the correlation is negative, the systematic risk of the investment is negative (which implies $X < R$ and thus $E(S_T) < F$). When the correlation is zero, the synthetic trade has no systematic risk, and the futures price should be equal to the expected future spot price.

These results apply to the FX forwards and futures considered in Chapter 9, the financial futures considered in Chapter 10, and the commodity futures considered in this chapter. They are summarized in Table 11.2.

An example of an asset with positive systematic risk is the S&P 500 itself. We therefore expect this asset to fall into the middle category in Table 11.2 with $F < E(S_T)$. From Equation 10.3, we know the forward/future price of a stock index is given by:

$$F = S\left(\frac{1 + R}{1 + Q}\right)^T$$

where Q is the dividend yield on the index.

Suppose that the index equals the forward price at time T. If dividends are reinvested in the index, the position in the index will grow at rate Q, and value of the investment at time T is

$$F = S\left(\frac{1 + R}{1 + Q}\right)^T (1 + Q)^T = S(1 + R)^T$$

This shows that the return realized by the investor is the risk-free rate (R). However, because the index (by definition) is positively correlated to itself, the investor should expect to earn more than R. This means that the expected value of the index at time T should be greater than F (which is consistent with Table 11.2).

Many commodities have positive systematic risk because they tend to cost more when the economy is doing well. An exception may be gold, which is often thought to have negative systematic risk. When the economy is doing poorly, investors increase their holding of gold, and its price increases. When the economy starts to recover, gold is exchanged for stocks, and its price declines. Assuming this is true, theory suggests that the futures price of gold overstates the expected future spot price:

$$F > E(S_T)$$

For those assets with little or no systematic risk, CAPM argues that the futures price should be equal to the expected future spot price. But has this been confirmed empirically? The

Table 11.2 Relationship Between Futures Price and Expected Future Spot Price Suggested by CAPM

Underlying asset return and the market return are uncorrelated.	Synthetic trade has no systematic risk.	$X = R$	Futures price equals expected future spot price.
Underlying asset return is positively correlated with the market return.	Synthetic trade has positive systematic risk.	$X > R$	Futures price is less than expected future spot price.
Underlying asset return is negatively correlated with the market return.	Synthetic trade has negative systematic risk.	$X < R$	Futures price is greater than expected future spot price.

answer: The futures price probably provides an unbiased forecast of the spot price (i.e., the spot price is just as likely to be above the futures price as below it), but the forecast is not very accurate.

Normal Backwardation and Contango

When the futures price is below the expected future spot price, the situation is known as *normal backwardation*. When the futures price is above the expected future spot price, the situation is known as *contango*. However, sometimes the term *normal backwardation* is used to refer to the situation where the futures price is below the *current* spot price. Contango is then used to refer to the situation where the futures price is above the *current* spot price.

SUMMARY

Commodities differ from financial assets in several ways. For example, they are subject to transportation costs and storage costs. Commodities are also usually held in inventory so that they can be used in some way.

In order to understand a commodity's price, it is therefore necessary to understand how the commodity is used as well as how it is influenced by factors such as the seasonality of supply and demand.

There are three main categories of commodities: agricultural, metals, and energy. These commodities all come with futures contracts that are commonly traded on exchanges.

Unlike financial assets, most commodities are not held for purely investment purposes. (Exceptions include some metals such as gold and silver.) Thus, a no-arbitrage argument can only give an upper bound for a commodity futures price. The convenience yield of a commodity measures the benefits of owning the physical commodity over the holding of a corresponding futures contract. These benefits include the ability to keep a production process running and to obtain supplies of the asset on short notice.

The cost of carry for a commodity is the cost of storage plus the cost of financing the asset. For a commodity held for consumption, the relationship between the futures price and the spot price depends on the convenience yield as well as the cost of carry.

The theoretical relationship between the futures price and the expected future spot price depends on whether the return on the underlying asset is positively or negatively correlated with the return on the stock market. A positive correlation implies the futures price is less than the expected future spot price. A negative correlation implies the futures price is greater than the expected future spot price. When the correlation is zero, the futures price is equal to the expected future spot price.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

11.1 What differences between commodities and financial assets are important for explaining why their futures prices are determined differently?

11.2 What are the three main energy products on which futures contracts trade?

11.3 Explain how CDD and HDD are defined for weather futures.

11.4 Define the lease rate of a commodity, such as gold.

11.5 What is meant by the convenience yield of a commodity?

11.6 What is meant by the cost of carry of a commodity?

11.7 Explain why some equations show the relationship between the forward price F and the spot price S for a non-dividend-paying stock to be $F = S(1 + R)^T$, while others show it to be $F = Se^{RT}$.

11.8 Suppose that speculators tend to take short futures positions on an asset, while and hedgers take long futures position. What would Keynes argue about the ability of futures prices to predict expected future spot prices?

11.9 How can a long position in an asset at a future time be created synthetically?

11.10 Under what circumstances does CAPM suggest that the futures price equals the expected future spot price?

Practice Questions

11.11 What is meant by mean reversion?

11.12 Explain the main differences between agricultural and metal commodities.

11.13 A gold producer enters into a forward contract with an investment bank to sell 10,000 ounces of gold for USD 1,200 per ounce. Explain how the bank would hedge its risk without using futures contracts.

11.14 If the current spot price of gold is USD 1,250, the one-year futures price of gold is USD 1,300, and the risk-free rate is 5% per annum (annually compounded), what is an estimate of the implied gold lease rate?

11.15 Why is it not possible to calculate the futures price for a commodity such as copper in the same way as a futures price for a financial asset?

11.16 "The convenience yield is the lease rate for an asset if it were possible to borrow it." Discuss this statement.

11.17 "If news reaches the market that the price of oil will almost certainly increase by 20% in three months, the price will increase immediately. If news reaches the market that the price will almost certainly decrease by 20% in the next three months, the price does not necessarily decrease immediately." Discuss this statement.

11.18 Suppose that the storage costs for crude oil are USD 2 per barrel per year payable at the beginning of the year. The current spot price is USD 75, and the two-year futures price is USD 72. The risk-free interest rate is 8% per annum (compounded annually). Estimate the convenience yield of crude oil per year.

11.19 If the futures price equals the future spot price for a financial asset, what is the return of that asset?

11.20 Suppose that F_1 and F_2 are the futures prices on the same commodity with maturities t_1 and t_2 with $t_2 > t_1$. Storage costs are negligible. The risk-free rate is R for all maturities. Use an arbitrage argument to show that:

$$F_2 \leq F_1(1 + R)^{t_2 - t_1}$$

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

ANSWERS

Short Concept Questions

11.1 Commodities have storage costs and transportation costs. They are in most cases not held for investment purposes. Market participants who own commodities want to use them in some way.

11.2 Oil, natural gas, and electricity

11.3 The CDD, cooling degree for a day, is defined as $\max(A - 65, 0)$ where A is the average of the highest and lowest temperature in degrees Fahrenheit during the day. HDD, heating degree for a day, is defined as $\max(65 - A, 0)$.

11.4 The lease rate is the rate at which the commodity can be borrowed from its owner.

11.5 The convenience yield of an asset is the benefit of owning the product so that it is readily available for use in a production process.

11.6 The cost of carry is the cost of owning the commodity. It is sum of the cost of financing and the cost of storing it.

11.7 The first equation assumes that R is the annually compounded risk-free interest rate. The second equation assumes that R is the continuously compounded risk-free interest rate. (Continuous compounding will be discussed in Chapter 16.)

11.8 Keynes would argue that the futures price overstates the expected future spot price because speculators are taking risks and require an expected positive profit to compensate for such risk. Hedgers are reducing risk and may be satisfied with a negative expected return.

11.9 An investor can borrow the present value of the futures price from the bank and take a long position in the corresponding futures contract on one unit of the asset.

11.10 The future price equals the expected future spot price when the return from the underlying asset is uncorrelated with the market.

Solved Problems

11.11 Mean reversion is the tendency for a market variable to be pulled back to some central value over time.

11.12 Agricultural commodities have higher storage costs. Prices of agricultural products tend to be seasonal and are influenced by the weather. Metal price may depend on the success of exploration to find new mines. Also, exchange rates, extraction methods, actions by foreign governments, and cartels can all be possible driving factors of a metal price.

11.13 The bank can borrow gold from a central bank (paying the lease rate) and sell it in the market (investing the proceeds of the sale). When the forward contract matures, the bank buys the gold from the gold producer and returns it the central bank.

11.14 Let ℓ be the implied gold lease rate (annually compounded). Then:

$$1{,}300 = 1{,}250 \times \left(\frac{1.05}{1 + \ell}\right)$$

so that

$$\ell = \frac{1{,}250}{1{,}300} \times 1.05 - 1 = 0.0096$$

that is, the implied lease rate is 0.96%.

11.15 Copper is not an investment asset. Individuals and companies that own copper plan to use it in some way. That is, they need to have copper in inventory and are not prepared to sell copper and buy it back in the futures market when futures prices are relatively low compared to the spot prices.

11.16 For an asset such as copper, one can argue that the lease rate that would be charged compensates the lender for benefits of having it on hand, that is, it compensates the lender for the convenience yield.

11.17 If oil prices are expected to increase sharply, speculators will buy oil and store it. This means that the price will rise until the financing cost plus the storage cost equals the expected profit. If oil prices are expected to fall sharply, there is no similar trade involving spot oil because those who hold oil in inventory need it to feed their refineries.

11.18 The present value of the storage costs per barrel over two years is

$$2 + \frac{2}{1.08} = 3.85$$

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

The convenience yield Y is given by solving:

$$72 = (75 + 3.85) \times \left(\frac{1.08}{1 + Y}\right)^2$$

The solution is

$$Y = \left(\frac{78.85}{72}\right)^{1/2} \times 1.08 - 1 = 0.130$$

that is, the convenience yield is about 13% per year.

11.19 The return of the asset is the risk-free interest rate. This is evident from the equations for the futures price. For example, when there is no income generated by the underlying asset, the futures price is the spot price compounding forward at the risk-free rate. When there is income generated by the asset during the tenor of the futures, the futures price is adjusted for the income and then compounded forward at the risk-free rate.

11.20 A trader can enter into a long futures contract with maturity t_1 and a short futures contract with maturity t_2. At time t_1 F_1 is borrowed and the asset is bought for F_1. The loan is repaid at time t_2 and the asset is sold for F_2. The cash flows are

Time t_1: $-F_1 + F_1 = 0$, and

Time t_2: $F_2 - F_1(1 + R)^{t_2 - t_1}$.

This simple strategy is certain to lead to a profit at time t_2 if:

$$F_2 > F_1(1 + R)^{t_2 - t_1}$$

Thus, the prices will adjust such that:

$$F_2 \leq F_1(1 + R)^{t_2 - t_1}$$

12 Options Markets

Learning Objectives

After completing this reading, you should be able to:

- Describe the various types and uses of options, define moneyness.
- Explain the payoff function and calculate the profit and loss from an options position.
- Explain how dividends and stock splits can impact the terms of a stock option.
- Describe the application of commissions, margin requirements, and exercise procedures to exchange-traded options, and explain the trading characteristics of these options.
- Define and describe warrants, convertible bonds, and employee stock options.

BOX 12.1 A HISTORY OF OPTIONS

Options have been traded for thousands of years. Greek philosopher Thales of Miletus, who lived over 2,500 years ago, may have been one of the first option traders. According to legend, he predicted that a coming olive harvest would be larger than normal. Acting on his prediction, he paid olive press owners a fee for the right to use the presses at harvest time. His forecast for the harvest proved correct and he was able to sell his call options on the presses for a sizable profit.

The seventeenth century Dutch tulip bubble also involved option trading. The so-called "tulip mania" led to exorbitantly high prices for tulip bulbs. As a result, tulip wholesalers would buy call options on tulips to protect against unexpected price increases. Meanwhile, tulip growers would buy put options to protect themselves in case the price decreased.

In 1637, the bubble burst and tulip prices collapsed. Many put option sellers subsequently defaulted, giving options a bad name throughout Europe for some time. And while an options market developed in England during the late seventeenth century, options were later banned in that country for over 100 years.

In the nineteenth century, American financier Russell Sage sold put and call options for stocks listed on New York Stock Exchange. He is believed to be the first person who understood the relationship between the prices of European put options and call options. (This is called the put-call parity and it will be explained in the next chapter.)

The popularity of options gradually increased during the late nineteenth and early twentieth centuries. During this time, options were traded over-the-counter by brokers who negotiated prices and matched buyers and sellers. The Put and Call Brokers and Dealers Association was formed to help with this matching process.

The Chicago Board of Trade, which was concerned about a decline of trading interest in commodity futures, decided to create an options exchange in 1968. Five years later, the Chicago Board Options Exchange (CBOE) was formed. Options (initially only calls) were standardized for trading and (more importantly) the Options Clearing Corporation (OCC) was established. The OCC required option sellers to post collateral (in the form of margin) as way to protect buyers from default risk.

That same year, the famous Black-Scholes-Merton options valuation model was published. The model, which is used to value options under various volatility assumptions, proved to be a useful tool for traders. By 1974, 20,000 call options contracts were traded daily on the CBOE. Three years later, the exchange began trading put options.[1]

This chapter discusses how options markets are organized. Later chapters in this book cover the properties of options and options trading strategies.

There are important differences between options and forwards/futures. A forward or futures contract obligates a trader to buy or sell the underlying asset at a certain price and does not require an upfront payment.[2] By contrast, an option involves paying a certain amount (called the premium) to obtain the right to buy (or sell) an asset at a certain price in the future. However, this right does not have to be exercised.

In the U.S., tens of millions of options are traded daily on exchanges such as the CBOE, NASDAQ, the New York Stock Exchange, and the International Securities Exchange. Exchanges trading options outside of the U.S. include the Eurex, the National Stock Exchange of India, and BM&FBOVESPA. Options are also traded globally in the over-the-counter market.

12.1 CALLS AND PUTS

A European call (or put) option gives the buyer the right to buy (or sell) an asset at a certain price on a specific date. An American call (or put) option gives the buyer the right to buy (or sell) an asset at a certain price at any time before or on the specified date.[3] The date specified in the option is known as the *expiration date* (also called the *maturity date*). Most (though not all) exchange-traded options are American. By contrast, many of the options traded in the over-the-counter market are European.

The price at which an asset can be bought or sold using an option is referred to as the *strike price* (also called the *exercise price*). In the early days of options trading, the strike price was typically set as the asset's spot price at the time the option was sold. Today, it is possible to buy or sell options with a range of different strike prices.

American options are more difficult to analyze than European options because they can be exercised at any time before

[1] See www.cboe.com for a history of the exchange.

[2] However, it is necessary for traders to post margin in the case of futures trading as explained in earlier chapters. Forward contracts may also require collateral to be posted.

[3] Note that the terms European and American have nothing to do with the location of the option traders or the exchange. Amusingly, a Bermudan option is an option that is partly American. It can be exercised on pre-specified dates during its life, but not at any time.

maturity. While European options can be valued using the Black-Scholes-Merton model (which will be discussed in *Valuation and Risk Models*), American options can only be valued by using numerical procedures such as binomial trees (also discussed in *Valuation and Risk Models*). As discussed in the next chapter, however, one can argue that in certain situations an American option should never be exercised early and therefore must have the same price as its European counterpart.

Moneyness

Options can be *in-the-money*, *at-the-money*, or *out-of-the-money*. This is referred to as their *moneyness*. Under the simplest definition of these terms, it is imagined what would happen if the option could be exercised today. (This is done even if the option cannot be exercised today.) An option that would give a positive payoff if exercised today is referred to as *in-the-money*. If it would give a negative payoff, the option is referred to as *out-of-the money*. An option that is *at-the-money* would give a payoff of zero.

A call (put) option is in-the-money when the asset price is greater (less) than the strike price. This is because the option (if exercised today) would allow an asset to be purchased (sold) for less (more) than its current price. An investor who exercises an in-the-money call option can sell the asset immediately for a positive payoff equal to the excess of the current asset price over the strike price. An investor exercising an in-the-money put option could buy the asset and sell it for the strike price to make an immediate profit equal to the excess of the strike price over the current asset price.

Likewise, a call (put) option is out-of-the-money when the asset price is less (greater) than the strike price. In these situations, immediate exercise (if allowed) would not be profitable because the payoff would be negative. An option is referred to as at-the-money when the strike price equals the asset price.[4]

Profits from Call Options

As an example, suppose a trader buys an out-of-the-money European call option with a strike price of USD 60 for an asset currently worth USD 55. The option premium is USD 4 and the expiration date is six months from the day of transaction.

Let S be the asset price on the expiration date. If S is less than USD 60, then the option will not be exercised. If S is greater than USD 60, the call option will be exercised (i.e., the trader

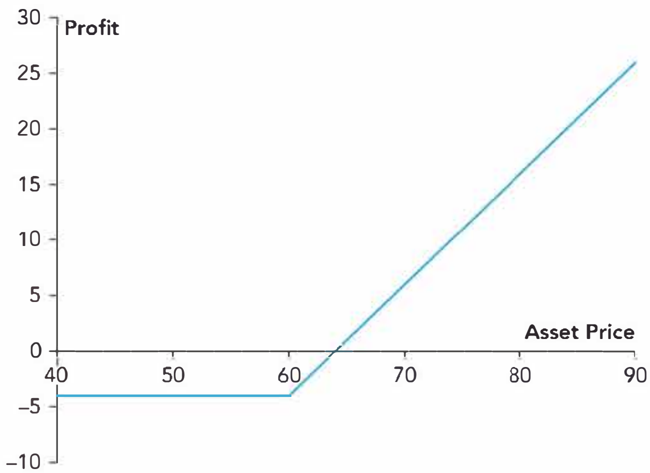


Figure 12.1 Profit as a function of asset price on the expiration date for the buyer of a European call option is shown. The cost of the call is USD 4, and the strike price is USD 60.

will buy the asset for USD 60 and immediately sell it for more than USD 60). When considering the price paid to buy the option, the trader's total profit (if the option is exercised) will be USD $S - 60 - 4$.[5] If the option is not exercised, the trader's loss will be USD 4.

The profit to the trader is shown in Figure 12.1. Note that sometimes the trader can exercise the option and still suffer a net loss. This happens when S is between USD 60 and USD 64. Would it be better for the trader not to exercise the option when the asset price is in this range? The answer is no. The trader will always lose USD 4 if the option is not exercised. Meanwhile, the loss from exercising when S is greater than USD 60 is always less than USD 4 (e.g., when $S =$ USD 61, the loss is USD 3).

Now consider the situation of the call option's seller. The option seller will make a profit of USD 4 (the option premium) if S is less than 60 because the option will not be exercised at that price. If S is between USD 60 and USD 64, the option will be exercised and the seller's profit will be less than USD 4. If S is above USD 64, the seller will suffer a loss. Figure 12.2 shows the seller's profit as a function of S. It should be no surprise that Figure 12.2 is the mirror image of Figure 12.1. Options (like other derivatives) are *zero-sum games* in the sense that one side's gain always equals the other side's loss.

[4] Other definitions of moneyness are sometimes used by traders as well. (See J. Hull, *Options, Futures and Other Derivatives*, 10th edition, section 20.4.)

[5] Note that we simplify matters by ignoring the impact of discounting. Because the cost of the option is paid at time zero and the payoff is six months later, the present value of the profit is PV($S - 60$) − 4 where PV denotes present value.

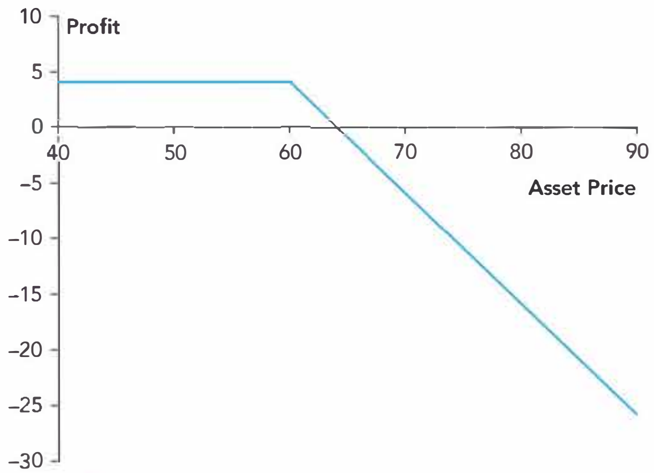


Figure 12.2 Profit as a function of asset price on the expiration date for the seller of a European call option is shown. The cost of the call is USD 4, and the strike price is USD 60.

Profits from Put Options

Now suppose a trader buys an out-of-the-money European put option with a strike price of USD 60 for an asset currently worth USD 62. The option premium is USD 5, and the expiration date is three months from the date of transaction.

If the asset price on the expiration date is less than USD 60, the put option will be exercised (i.e., the trader will buy the asset at price S when the option expires and immediately sell it at USD 60). Once again, accounting for the price paid to buy the option, the trader's profit is then USD $60 - 5 - S$.[6] If S is more than USD 60, the option will not be exercised and the trader's loss will be USD 5.

The profit to the trader is shown in Figure 12.3. As in the case of call options, sometimes the trader exercises the option and still suffers a loss. This happens when S is between USD 55 and USD 60. Exercising the option within this price range will not produce a profit, but it will reduce the trader's loss.

As in the case of call options, the profit of the corresponding put option seller is the mirror image of the profit of the option buyer. If S is greater than USD 60, then the option will not be exercised and the seller will gain the USD 5 paid by the buyer (the option premium). If S is between USD 55 and USD 60, the trader will make a profit of less than USD 5. If S is below USD 55, the trader suffers a loss. Figure 12.4 shows the trader's profit as a function of the asset price.

[6] As in the case of call options, the impact of discounting is ignored. In fact, the present value of the profit is $PV(60 - S) - 5$ where PV denotes present value.

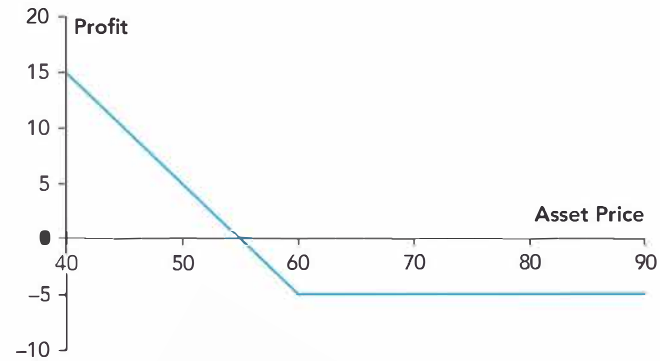


Figure 12.3 Profit as a function of asset price on the expiration date for the buyer of a European put option is shown. The cost of the put is USD 5, and the strike price is USD 60.

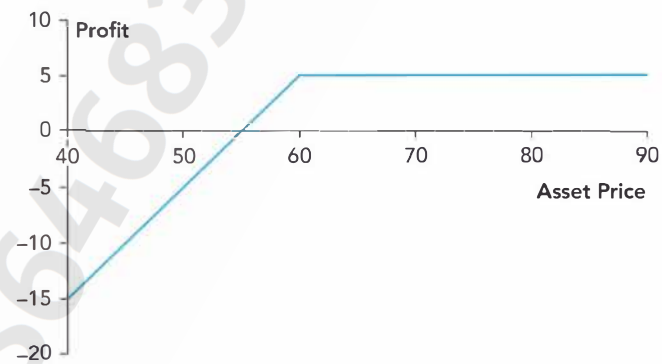


Figure 12.4 Profit as a function of asset price on the expiration date for the seller of a European put option is shown. The cost of the put is USD 5, and the strike price is USD 60.

Payoffs

Instead of characterizing a European option by its profit, we can plot the payoff as a function of the asset price on the expiration date. These plots show the value of the option at maturity and do not account for how much was paid to purchase the option. As illustrated in Figures 12.1 through 12.4, there are four possible option positions:

1. Long call (Figure 12.1),
2. Short call (Figure 12.2),
3. Long put (Figure 12.3), and
4. Short put (Figure 12.4).

Selling (or shorting) an option is also referred to as *writing the option*. Figure 12.5 shows the payoff from each of the option positions previously described when the strike price is 60.

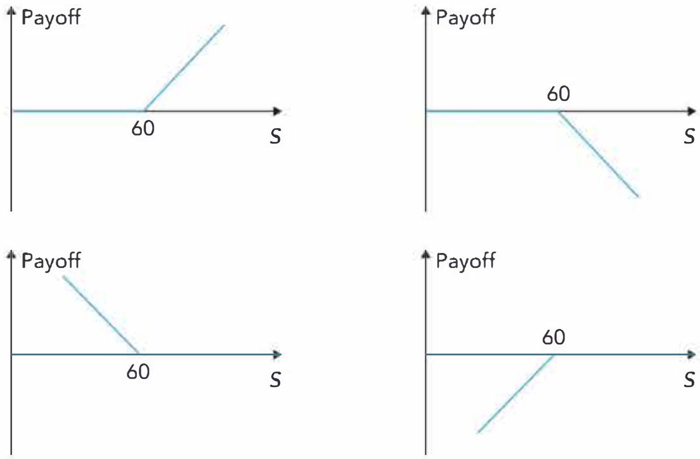


Figure 12.5 **Payoffs from option positions when strike price is 60 are shown. The asset price on the expiration date is S.**

Denoting the asset price on the expiration date by S and the strike price by K, the payoffs from the option positions are as follows:

Long Call: $\max(S - K, 0)$,

Short Call: $-\max(S - K, 0) = \min(K - S, 0)$,

Long Put: $\max(K - S, 0)$, and

Short Put: $-\max(K - S, 0) = \min(S - K, 0)$.

While the payoff from a European option is related to the asset price at expiration, what is called the *intrinsic value* is based on the current asset price (which we denote by S_0). Intrinsic value measures the value the option would have if it could only be exercised immediately. The intrinsic value of a call is $\max(S_0 - K, 0)$ and the intrinsic value of a put is $\max(K - S_0, 0)$, where S_0 is the current asset price.

12.2 EXCHANGE-TRADED OPTIONS ON STOCKS

This section describes how stock options are traded on the CBOE. This is the largest options exchange in the world and trades over a billion contracts each year. The options on individual stocks are American-style (i.e., they can be exercised at any time before their maturities). When a trader with a long position decides to exercise, the Options Clearing Corporation uses a random procedure to choose a trader with a short position to be exercised against. (This trader is referred to as being *assigned*.) Options that are in-the-money at maturity are usually exercised automatically. A single option contract on the CBOE is the right to buy or sell 100 shares.

Maturity

Stock options on the CBOE are assigned one of the following maturity cycles:

- January, April, July, and October (referred to as the January cycle);
- February, May, August, and November (referred to as the February cycle); and
- March, June, September, and December (referred to as the March cycle).

Prior to the third Friday of the current month, options trade with maturities in the current month, the following month, and the next two months in the cycle. Following the third Friday of the current month, options trade with maturities in the next month, the month after that, and the next two months in the expiration cycle.

For example, IBM is on a January cycle. This means that at the beginning of April, options trade with maturities in April, May, July, and October. After the third Friday in April, they trade with expirations in May, June, July, and October. The precise maturity date is the third Friday of the month (e.g., April 16, 2021, for an April 2021 option). Trading takes place every business day (8:30 p.m. to 3:00 p.m. Chicago time) until the maturity date.

The CBOE also offers short-term options called *Weeklys* and long-term options known as *LEAPS* (long-term equity anticipation securities). Weeklys mature on Fridays, other than the third Friday of the month, and are typically offered on several such Fridays. LEAPS on individual stocks mature on the third Friday of January and provide maturities of up to three years.

Strike Prices

The CBOE sets option strike prices as multiples of USD 2.50 when the current price of the underlying asset is between USD 5 and USD 25. When the current price is between USD 25 and USD 200, strike prices are usually multiples of USD 5. When the current price is greater than USD 200, they are usually multiples of USD 10. Strike prices are also adjusted for dividends and splits (as will be described in the next section).

Initially, the three strike prices closest to the current price are typically listed. As the stock price moves, more strike prices are introduced. For example, suppose the stock price is USD 19 when trading for a new options contract begins. Initially, strike prices of USD 17.50, USD 20.00, and USD 22.50 would be traded. If the stock price moved below USD 17.50, options with a strike price of USD 15 would start trading. If the price moved above USD 22.50, options with a strike price of USD 25 would start trading.

These rules can result in many different options contacts being available (e.g., if there are five strike prices for each of ten maturity dates, there could be 50 tradeable calls and 50 tradeable puts). All options of the same type (i.e., calls or puts) are referred to as a *class*. All options of a particular class with a specific maturity date and strike price are referred to as an option series. For example, call options on IBM maturing in April 2021 form an option series.

Dividends and Stock Splits

Cash dividends usually do not affect the terms of a stock option.[7] However, exceptions are sometimes made when the cash dividend is unusually large. Specifically, if the cash dividend is more than 10% of the stock price, the Options Clearing Corporation forms a committee to determine whether adjustments should be made.

In contrast, stock splits do lead to strike price adjustments. For example, if a company announces a 5-to-1 stock split (i.e., each share is replaced by five new shares), the strike price will be reduced to one fifth of the original price and the number of options owned by a trader is multiplied by five.

Similarly, stock dividends also lead to strike price adjustments. Note that a 10% stock dividend means shareholders receive one new share for each ten shares owned (this is identical to an 11-to-10 stock split). Using the stock split rule, the strike price will be reduced to ten-elevenths of its original level and the number of options owned by a trader is multiplied by 11/10. All these adjustments are designed to keep the positions of buyers and sellers unchanged by incidents of stock splits or stock dividends.

Index Options

In addition to options on individual stocks, the CBOE also offers options on equity indices. These indices include the S&P 500 Index, the S&P 100 Index, the Dow Jones Industrial Average, the Russell 1000, the FTSE 100, and the FTSE China 50. Many of these options are European-style rather than American-style. Final settlement is in cash and calculated with reference to the value of the index at a pre-specified time. Index Weeklys (maturing on Wednesdays and Fridays) and LEAPS (maturing in December and June) are also traded.

[7] In the early days when options were traded over-the-counter, the strike price was reduced by the dividend per share when a cash dividend was declared.

ETP Options

There are also CBOE options on many exchange-traded products (ETPs) such as exchange-traded funds (ETFs). An ETF can be designed to track an equity index, bond index, a commodity, or a currency.

ETP options are like options on individual stocks in that they are American-style and involve physical settlement (i.e., settlement by delivery of the underlying asset).

Non-Standard Products

To compete with the over-the-counter options market, the CBOE also offers some non-standard options. For example, *FLEX options* are options with non-standard terms that sometimes include

- Non-Standard strike prices,
- Non-Standard maturity dates, and
- Variations in style (American or European).

Other non-standard options on the CBOE include Asian and cliquet options on indices. Asian options provide a payoff based on the average price of the underlying asset during the life of the option. Cliquet options provide a payoff equal to the sum of the monthly capped returns provided by the asset (if this sum is positive). Further discussion of non-standard options can be found in Chapter 15.

Some non-standard contracts introduced by the CBOE have not been particularly successful. For example, credit event binary options (CEBOs) and deep-out-of-the-money (DOOM) put options were attempts to compete with the credit derivatives market. However, both options have been discontinued.[8]

12.3 TRADING

In recent years, options exchanges have replaced the open-outcry trading system with electronic matching platforms. However, exchanges still rely on market makers to provide liquidity to the options market. These market makers will quote option bid and ask prices when they are asked to do so, while exchanges set upper limits on the size of the bid-ask spread (i.e., the difference between the bid and ask quotes).

To take an option position, a trader customarily places an order through a broker, who charges a commission. The types of orders that can be placed are like those for futures described in Chapter 7

[8] CEBOs give positive payoff when the underlying company defaults before the option matures and DOOM options have strike prices that are extremely low compared with the initial price.

(e.g., market order, limit order, stop order, and so on). Commissions vary from broker to broker and can be as low as USD 5 per trade.

Like futures, exchange-traded options can be closed out by taking an offsetting position. For example, a trader owning a put with a certain strike price and maturity can exit from the position by selling a put with the same strike price and maturity. As with the futures market, the open interest in the options market measures the number of outstanding contracts.

The CBOE imposes position limits and exercise limits on traded options. These are designed to prevent the market from being unduly influenced by one investor (or a group of investors acting in concert). A position limit is defined as the number of contracts an investor can hold on the same side of the market (long calls and short puts are considered to be on one side of the market, while short calls and long puts are considered to be on the other side of the market). The exercise limit is the maximum number of contracts that can be exercised in five business days. It is usually the same as the position limit.

12.4 MARGIN REQUIREMENTS

Margin requirements have already been discussed in Chapter 5. Recall that, in the U.S., a stock can be purchased by borrowing up to 50% of its price. Such a transaction is known as *buying on margin*. Options with maturities less than nine months cannot be bought on margin, and the full price must be paid upfront. Options with maturities greater than nine months can be bought on margin, but no more than 25% of the purchase price can be borrowed.

If a trader pays cash for an option, there would be no margin requirements because the trader has no future liabilities. However, the seller of an option does have potential future liabilities. As explained in Chapter 5, the CBOE margin requirement for a short call position is 100% of the sale's proceeds plus the greater of the following two values:

1. 20% of the share price less the amount the option is out of the money, and
2. 10% of the share price.

The CBOE margin requirement for a short put is 100% of the sale's proceeds plus the greater of the following two calculations:

1. 20% of the share price less the amount the option is out of the money, and
2. 10% of the strike price.

For options on indices, the 20% in this calculation is replaced by 15%. This recognizes that the volatility of a stock index is usually less than that of an individual stock.

This calculation is repeated each day, with the option's current market price replacing the proceeds of the sale in determining the new margin requirements. If the margin requirement has increased, funds must be added to the trader's margin account. If the margin requirement has decreased, funds can be withdrawn.

The CBOE Margin Manual has special rules governing traders with portfolios consisting of long and short option positions (which are perhaps combined with positions in the underlying asset). For example, no margin is required on a covered call position (i.e., a written call plus a long position in the asset underlying the call). This is because there is no reason to suppose that the trader will default; as long as the position is maintained, the asset is available to be delivered in the event the option is exercised.

As with futures, options margins are handled by the members of the Options Clearing Corporation (OCC). All option trades must be cleared through an OCC member. Non-member brokers must arrange to clear their clients' trades with a member. Thus, all brokers maintain margin accounts with OCC members, while all end users maintain margin accounts with their respective brokers. The margin requirements for these accounts must be at least as great as those specified by the OCC for its members.

12.5 OVER-THE-COUNTER MARKET

Like the exchange-traded options market, the over-the-counter (OTC) options market is substantial. While options on individual stocks trade primarily on exchanges, options on foreign currencies, interest rates, and many other financial variables are traded actively in the OTC market.

The main advantage of the OTC market is that financial institutions can tailor options to meet the specific needs of their clients. Option characteristics in the OTC market (e.g., strike prices, maturity dates, and the times when options can be exercised) can differ from those available on the exchanges. The size of the typical options transaction in the OTC market is large and the options often last longer than those traded on exchanges. OTC options can also be exotic (i.e., have non-standard structures). Exotic options will be discussed in Chapter 15.

12.6 WARRANTS AND CONVERTIBLES

Warrants are options issued by a corporation. They are usually call options on the corporation's own stock, but they can also be options to buy or sell another asset (e.g., gold). Once issued, they are often traded on an exchange.

To exercise a warrant, the holder needs to contact the issuer. When a call warrant on the issuing company's stock is exercised, the company issues more of its stock. Once that happens, the warrant holders can then buy the stock at the strike price.

Warrants can be used by firms to make debt issuances more attractive to investors. For example, suppose a company's stock price is currently USD 40 and the firm is planning a debt issue. It might choose to add two warrants to each USD 1,000 bond where each warrant gives the holder the right to buy one share at USD 45 on the expiration date. The bondholders would then have a stake in the fortunes of the company beyond the desire to see it avoid a default.

A convertible bond (also referred to as a convertible) is similar to a warrant. Specifically, a convertible is a bond that can be converted into equity using a pre-determined exchange ratio. For example, suppose a company's current share price is USD 40. It might choose to issue ten-year bonds, each with a USD 1,000 par value, that can be converted into 20 shares at any time after four years. When an investor chooses to convert, the company simply issues more of its stock to be exchanged for the bonds. Like warrants, convertible bonds are often traded on exchanges.

12.7 EMPLOYEE STOCK OPTIONS

Employee stock options are call options granted by a company to its employees. They differ from exchange-traded options in several ways. For example:

- There is usually a vesting period during which options cannot be exercised. This period often lasts as long as four years.
- Employees may forfeit their options if they leave their jobs (voluntarily or involuntarily) during the vesting period.
- When employees leave after the vesting period, they usually forfeit options that are out-of-the-money. They may also have to exercise their in-the-money options immediately.
- Employees are not permitted to sell their stock options to a third party.

Because the employee options cannot be sold, they are usually exercised earlier than their exchange-traded counterparts. This point is discussed further in the next chapter.

It is debatable whether employee stock options align the interests of executives with those of shareholders. Note that if the stock price does well, both shareholders and executives will gain. If the stock price declines, however, the shareholders typically suffer more than executives. Thus, granting executives shares (rather than stock options) can create a better alignment of interests.

Employee stock options are commonly used by start-up companies that cannot afford to pay a competitive salary.

BOX 12.2 EMPLOYEE STOCK OPTION ACCOUNTING AND BACKDATING SCANDALS

It used to be the case that employee stock options had no effect on a firm's profits if they had a strike price set to current stock price. This made them an extremely attractive form of compensation to senior management. However, accounting rules have changed and employee stock options must now be valued (and expensed) at the time of issue. As a result, they are not as widely used as they once were.[9]

Suppose that a firm's stock price has increased from USD 42 to USD 50 over the past week. It was once tempting for company executives to backdate an issue of at-the-money employee stock options by three weeks and set the strike price to USD 42. This would increase the value of the stock options for the executives while at the same time reducing its cost to the company (because at-the-money options had no effect on the firm's profit). This practice was (and still is) illegal, but academic studies have produced evidence showing that it was once widespread.[10] As a result, the U.S. Securities and Exchange Commission now requires option grants be reported within two business days.

SUMMARY

A call (or put) option gives the holder the right to buy (or sell) an asset for a certain price by a specific date. There are four possible positions: long call, short call, long put, and short put.

Exchanges such as the Chicago Board Options Exchange (CBOE) have rules for determining option strike prices and maturities. On the CBOE, each stock option contract gives the holder the right to buy or sell 100 shares. These options are American-style and the expiration date is the third Friday of the expiration month. The CBOE sets strike prices with intervals that are multiples of USD 2.50, USD 5, or USD 10. In addition to stock options, the CBOE also trades options on market indices, exchange-traded funds, and other products.

[9] Many companies now award shares of their stock instead of options to employees.

[10] See D. Yermack, "Good timing: CEO stock option awards and company news announcements," *Journal of Finance*, 52 (1987): 449–476; E. Lie, "On the timing of CEO stock option awards," *Management Science*, 51, 5 (May 2005): 802–812; R. Heron and E. Lie, "Does backdating explain the stock option pattern around executive stock option grants," *Journal of Financial Economics*, 83, 2 (February 2007): 271–295.

Option terms are not adjusted for cash dividends, but they are adjusted for stock dividends and stock splits in a way that keeps the positions of both the option buyer and seller unchanged.

Market makers are often used to improve the liquidity of the options market. Option sellers (i.e., writers) must post margin to ensure that they will not renege on their obligations. The Options Clearing Corporation operates in a similar manner to a futures clearinghouse in that it requires margin from its members selling options. These members also require margin from brokers who are not members, and these brokers will in turn require margin from their clients.

Warrants are options issued by companies, usually on their own stock. Convertible bonds are bonds that give their holders the right to convert the bond into a certain amount of equity at pre-determined times. Employee stock options are options issued by a company on its own stock and are reserved for its own employees. These securities all have the property that, when the holder exercises his or her right to obtain shares in the company, the company issues more stock.

The following questions are intended to help candidates understand the material. They are __not__ actual FRM exam questions.

QUESTIONS

Short Concept Questions

12.1 What is the strike price of an option?

12.2 What is the difference between an American and a European option?

12.3 What is the difference between buying a call and selling a put?

12.4 What is the difference between the profit and payoff from an option?

12.5 What is the intrinsic value of an option?

12.6 What is a LEAP?

12.7 What is (a) a position limit and (b) an exercise limit in the options market? Why are they used?

12.8 Does an option writer have to keep the position until the end of the option's life? Explain your answer.

12.9 Why is there no margin requirement for a covered call?

12.10 What is the difference between a warrant and an option traded by the CBOE?

Practice Questions

12.11 Explain the meaning of the term *at-the-money*.

12.12 Suppose a European call option to buy a share for USD 50 in nine months costs USD 6. Under what circumstances will the holder of the option make a profit? Under what circumstances will the option be exercised?

12.13 If the option in Question 12.12 is a put, rather than a call, how does your answer change?

12.14 A stock price is USD 41 when options for a certain maturity are introduced. What strike prices would initially trade?

12.15 An option is on the March cycle. What options (in addition to Weeklys and LEAPs) trade on August 1?

12.16 An option has a strike price of USD 50. The company declares a 20% stock dividend. If the OCC were to adjust the strike price, what would be the new strike price?

12.17 An option has a strike price of USD 50. The company declares a 3-for-1 stock split. What effect does this have on the strike price?

12.18 An option has a strike price of USD 50. A cash dividend of USD 0.50 is announced with an ex-dividend date before the end of the life of the option. What effect does this have on the strike price?

12.19 Why do you think employee stock options cannot be traded?

12.20 How would you design an employee stock option that pays off only if the company outperforms others in the same industry?

The following questions are intended to help candidates understand the material. They are _not_ actual FRM exam questions.

ANSWERS

Short Concept Questions

12.1 The strike price of an option is the price at which the holder of a call (or put) can buy (or sell) the underlying asset at a future time.

12.2 A European option can be exercised only on its maturity date. An American option can be exercised any time up to the maturity date.

12.3 Buying a call gives a trader the right to buy an asset at a certain price in the future. Selling a put gives the other side the right to sell the asset at a certain price at a certain time in the future. Buying a call gives a payoff of $\max(S - K, 0)$ where S is the stock price when the option is exercised and K is the strike price. Selling a put gives a payoff of $\min(S - K, 0)$. Margin has to be posted when a put is sold but not when a call is bought.

12.4 A profit calculation considers the initial option price. A payoff calculation does not involve the transacted option price.

12.5 The intrinsic value of a call option is $\max(S_0 - K, 0)$ and the intrinsic value of a put option is $\max(K - S_0, 0)$ where S_0 is the current stock price and K is the strike price.

12.6 LEAPS (Long-Term Equity Anticipation Securities) are long term options (with maturities up to three years) traded by the CBOE.

12.7 The position limit is a limit on the size of the position that an investor (or group of investors acting together) can have in option contracts. For this purpose, the positions in long calls and short puts are summed and the positions in long puts and short calls are summed. An exercise limit is a limit on the number of options that can be exercised within a five-day period. The purpose of the limits is to prevent an investor (or group of investors) from unduly influencing the market.

12.8 The writer of an option contract can buy the same option contract to exit from the position.

12.9 When a trader owns a stock and has written a call option on the stock, the stock is available to be delivered whenever the call option is exercised. Thus, the credit risk is covered and no margin requirement is needed.

12.10 In a warrant, a fixed number of options are issued by a company and the company guarantees that it will satisfy investors when they exercise the options. In an exchange-traded option, the number of options expands as there are more trades and contracts as traders close out their positions. The Options Clearing Corporation administers margins and ensures that investors are able to exercise the options they own.

Solved Problems

12.11 An at-the-money option is usually defined as an option where the asset price equals the strike price.

12.12 The option will be exercised if the share price on the expiration date is greater than USD 50. The holder of the option will make a profit if the share price at expiry is greater than USD 56 and present value discounting is ignored.

12.13 The option will be exercised if the share price at expiry is less than USD 50. The holder of the option will make a profit if the share price on the expiry date is less than USD 44.

12.14 Strike prices of USD 35, USD 40, and USD 45 would normally be introduced.

12.15 Options trade with maturities in August, September, and December of the current year and March of the following year. Weeklys and LEAPS may also trade. The maturity is the third Friday of the month.

12.16 The strike price becomes USD 50 $\times$ (5/6) = USD 41.67.

12.17 The strike price becomes USD 50 $\times$ (1/3) = USD 16.67.

12.18 There is normally no effect on the strike price when a cash dividend is declared. An exception may be made when the dividend is large—more than 10% of the stock price.

12.19 Employee stock options are intended to provide an incentive for the employee to work hard and increase value for shareholders. If the employees were allowed to sell their options without constraints, this incentive would no longer exist.

12.20 One possibility is to make the strike price proportional to an index of the prices of the stocks of other companies in the same industry as the company. The company would then have to outperform its competitors for the options to move in-the-money.

13

Properties of Options

Learning Objectives

After completing this reading, you should be able to:

- Identify the six factors that affect an option's price.
- Identify and compute upper and lower bounds for option prices on non-dividend and dividend paying stocks.
- Explain put-call parity and apply it to the valuation of European and American stock options, with dividends and without dividends, and express it in terms of forward prices.
- Explain and assess potential rationales for using the early exercise features of American call and put options.

The price of stock option depends on:[1]

- The price of the underlying stock,
- The strike price,
- The risk-free rate,
- The volatility of the stock price,
- The time to maturity, and
- The dividends to be paid during the life of the option.

Chapters 14 and 15 of *Valuation and Risk Models* will discuss how these variables are combined to value European and American options under the Black-Scholes Merton assumptions. This chapter produces upper and lower bounds for options prices that do not depend on the valuation model used. It also considers factors that might lead a trader to exercise an American option before maturity.

This chapter highlights an important result known as put-call parity, which is the relationship between the price of a European call option and the price of a European put option with the same strike price and time to maturity.

For the most part, this chapter assumes that the underlying asset is a stock. However, Section 13.4 illustrates how results can be extended to options on other assets by using forward prices.

13.1 CALL OPTIONS

This section considers the properties of call options and discusses the circumstances under which American call options should be exercised before maturity. Upper and lower bounds for both American and European call option prices are also derived.

American versus European Options: No Dividends

Many exchanged-traded options (including options on individual stocks) are American. An important question is whether an American option should ever be exercised before its maturity date. To investigate this, first consider call options on stocks that pay no dividends.

Consider an American call option where the strike price is USD 30 and the stock price is USD 50. Suppose further that the underlying asset pays no dividends and there are two months until maturity.

In this case, it is tempting for an investor to exercise the call option and sell the stock. The investor could buy the stock for USD 30, sell it for USD 50, and thus obtain an immediate profit of USD 20. Because each options contract is for the right to buy 100 shares (as mentioned in the previous chapter), the potential gain is USD 2,000 per contract.

Despite this profit, however, the option should *not* be exercised before maturity if interest rates are positive. To understand why this is so, note that the option owner is in one of two situations.

Situation 1: The investor wants to have a position in the stock.

Situation 2: The investor does not want to have a position in the stock.

In Situation 1, the option holder would not sell the stock after exercising the option. He or she is better off holding the option until expiry. To understand this, consider two outcomes for Situation 1.

Outcome 1: The stock price is greater than USD 30 on the expiry date.

Outcome 2: The stock price is less than USD 30 on the expiry date.

Under Outcome 1, it is suboptimal to exercise the option early because waiting until the expiration of the option would allow the investor to earn additional interest by investing the strike price at the risk-free rate for two months.

Now consider Outcome 2. By exercising early, the investor incurs a loss of USD 30 − S (where S is the stock price at expiry). If the investor waits until maturity, the option is not exercised and thus the investor does not incur this loss.

The key point here is that the optionality (i.e., the choice provided by the option) gives the option holder insurance against the value of the stock falling below USD 30. As soon as the option is exercised, this optionality is lost.

Now suppose that the investor is in Situation 2 and does not want a position in the stock. The investor should then sell the option instead of exercising it. Exercising the option immediately would yield a profit equal to intrinsic value of the option (which in this case is USD 20). However, selling the option yields a profit of the intrinsic value plus what is called the *time value*. The time value is the value of the optionality mentioned above.

For a more formal proof that the option should not be exercised until maturity, consider two portfolios that could be held prior to the maturity of the option.

Portfolio A: A call option plus a cash amount equal to the present value of the strike price. (The present value is calculated by discounting the strike price from the maturity of the option to today at the risk-free rate.)

Portfolio B: The stock.

In Portfolio A, the cash can be invested so that it becomes the strike price K by the maturity of the option. If the stock price at maturity

[1] The price of the option does not depend on the expected return from the stock. Chapter 15 in *Valuation and Risk Models* discusses the reasons for this.

is greater than the strike price, the option in Portfolio A can be exercised so that the portfolio becomes worth the stock price. If the stock price is less than the strike price at option maturity, the option is not exercised and Portfolio A is worth the strike price.

It follows that Portfolio A is worth at least as much as (and in some circumstances more than) the stock price at the maturity of the option. On the other hand, Portfolio B is always worth the stock price at maturity of the option.

This means that at maturity of the option:

$$\text{Value of Portfolio A} \geq \text{Value of Portfolio B}$$

This must also be true today, as otherwise there would be an arbitrage opportunity. (If Portfolio B were worth more than Portfolio A today, for example, an arbitrageur could buy Portfolio A and short Portfolio B to create a position that would never lead to a loss and would sometimes lead to a profit.)

The value of Portfolio A today is

$$\text{Call Price} + PV(K)$$

where PV denotes present value, with discounting being done from the option's maturity to today at the risk-free rate. Meanwhile, the value of Portfolio B today is the stock price S. It follows that:

$$\text{Call Price} + PV(K) \geq S$$

or

$$\text{Call Price} \geq S - PV(K) \qquad (13.1)$$

Because an option's price cannot be negative, this result can be extended to:

$$\text{Call Price} \geq \max(S - PV(K), 0)$$

Assuming interest rates are positive:

$$K > PV(K)$$

Equation (13.1) thus implies

$$\text{Call Price} > S - K$$

If the option were exercised today, the call price would equal $S - K$. It follows that the call option should never be exercised early.

An upper bound to the call price is the stock price. This is because a call option can never be worth more than the stock price.

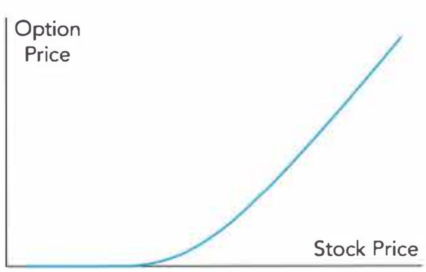


Figure 13.1 Variation of call option price with stock price.

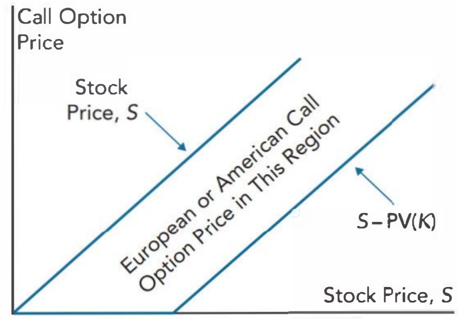


Figure 13.2 Bounds for European or American call option on non-dividend paying stock.

The general shape of a call option price as a function of maturity is shown in Figure 13.1 and the bounds we have just derived are illustrated in Figure 13.2.

Employee Stock Options

Recall that Chapter 12 introduced employee stock options. Employee stock options on non-dividend-paying stocks are frequently exercised before maturity. Is this suboptimal behavior on the part of employees? The answer is: not necessarily.

Suppose an employee has stock options that are vested and in-the-money. If the employee is in Situation 2 mentioned previously (i.e., the employee does not want to keep a position in the stock), the argument given earlier shows that selling the option would be more profitable than exercising the option and selling the stock. However, employee stock options cannot be traded and thus must be exercised for employees to gain a profit. This explains why employee stock options tend to be exercised earlier than similar exchange-traded stock options.

Impact of Dividends

Now consider exchange-traded options when underlying assets pay dividends before the option matures. Because dividends reduce the price of a stock, it can be optimal to exercise an American call option immediately before an ex-dividend date.

Whether this should be done depends on the size of the dividend and the extent to which the option is in-the-money. It can be shown that it is never optimal to exercise on an ex-dividend date, regardless of how high the stock price is, if the dividend is less than:[2]

$$K - K^*$$

where K^* is defined as the present value of K, with discounting being done at the risk-free rate from the next ex-dividend date

[2] Strictly speaking, it is the amount the stock price will decline on the ex-dividend date as a result of the dividend (see Footnote 3).

(or option maturity if there are no further ex-dividend dates) to the current ex-dividend date.

If the dividend is greater than $K - K^*$, exercise of the option is optimal for a sufficiently high stock price.

As an example, suppose a call option has a strike price of USD 40 and the current stock price is USD 45. Suppose further there is one week to maturity. A USD 2 dividend has been declared and the ex-dividend date is about to be reached.[3] An option holder may wish to exercise the option early because the dividend is (most likely) greater than the small difference between K and K^* over one week.[4]

To gain a general understanding of this analysis, suppose an option is so deep-in-the-money that it is almost certain to be eventually exercised. Exercising the option immediately before an ex-dividend date would be the optimal decision if the reduction in the stock price arising from the dividend is greater than the gain from delaying the payment of the exercise price.

Lower Bound When There Are Dividends

For a stock paying no dividends, Equation (13.1) applies

$$\text{Call Price} \geq S - \text{PV}(K)$$

This is true for an American and European call options because American call options on non-dividend-paying stocks should never be exercised early. A lower bound for the price of a call option on a non-dividend paying stock is therefore the stock price minus the present value of the strike price.

In the case where there are dividends, the lower bound for a European call option is

$$\text{European Call Price} \geq S - \text{PV}(K) - \text{PV(Divs)} \quad (13.2)$$

where PV(Divs) is the present value of the dividends where discounting is done from the ex-dividend dates to today at the risk-free rate.

To see how this works, the definitions of Portfolio A and Portfolio B are changed to:

Portfolio A: A European call option plus an amount of cash equal to PV(K) + PV(Divs), and

Portfolio B: The stock.

[3] The ex-dividend date is defined so that those who own the stock before the ex-dividend date receive the dividend. Those who become owners of the stock on or after the ex-dividend date do not receive the dividend.

[4] It might be expected that the stock will decline by USD 2 on the ex-dividend date. In practice it may decline by less than this because of tax effects: capital gains/losses and dividend income are taxed differently in some jurisdictions.

In Portfolio A, PV(K) becomes K at option maturity. If the stock price is greater than the strike price, K is exchanged for the stock. At option maturity, Portfolio A is therefore worth:

$$\max(S_T, K) + \text{FV(Divs)}$$

where S_T is the stock price at option maturity and FV(Divs) is the value to which the dividends grow until the option matures (assuming they are invested at the risk-free rate).

On the other hand, Portfolio B is worth:

$$S_T + \text{FV(Divs)}$$

It follows that:

$$\text{Portfolio A} \geq \text{Portfolio B}$$

at option maturity. For no arbitrage, it also be true today so that:

$$\text{European Call Option Price} + \text{PV}(K) + \text{PV(Divs)} \geq S$$

This shows Equation (13.2) provides a lower bound for the price of a European call option. Because an option price can never be negative, we can extend Equation (13.2) to:

$$\text{European Call Price} \geq \max(S - \text{PV}(K) - \text{PV(Divs)}, 0)$$

As an example, consider a one-year European call option where the current stock price is USD 64 and the strike price is USD 60. A dividend of USD 1 is expected in three months, six months, and nine months, and the risk-free rate is 4% per annum (with annual compounding) for all maturities. The present value of the strike price is USD 57.69 (= 60/1.04). The present value of the dividends is

$$\frac{1}{1.04^{0.25}} + \frac{1}{1.04^{0.5}} + \frac{1}{1.04^{0.75}} = 2.94$$

The call option price must therefore be at least:

$$64 - 57.69 - 2.94 = 3.37$$

When there are dividends, the lower bound in Equation (13.2) applies only to European options (i.e., USD 3.37 in the example above is not the best lower bound for the corresponding American call option). In fact, the American option could be exercised immediately and must therefore be worth at least its intrinsic value of USD 4 (= 64 − 60).

13.2 PUT OPTIONS

Now consider put options. Note that the early-exercise properties of put options differ from those of call options. For example, American put options are sometimes exercised before maturity even when there are no dividends.

American versus European Options: No Dividends

Consider a two-month put option on a non-dividend-paying stock when the current stock price is USD 1 and the strike price is USD 20. The option gives the holder the right to sell the underlying asset for USD 20. Should the option be exercised? The answer is clear: The put option should almost certainly be exercised.[5]

Now suppose an investor owns both the stock and the put option. Assuming interest rates are positive, it is better to receive the USD 20 strike price now than in two months. Notice the difference between the arguments for call options and those for put options: call option holders pay the strike price (and can benefit from holding the option and paying later) whereas put option owners receive the strike price (and can benefit from exercising the option and investing the proceeds at the risk-free rate).

What does an investor give up by exercising early? There is a small chance the stock price in our example will rise above USD 20 in two months. By exercising early, an investor gives up the opportunity to benefit from that.

The decision to exercise an American put option is therefore a trade-off between:

- Receiving the strike price early so it can be invested to earn interest, and
- Benefiting from the optionality in circumstances where the stock price moves above the strike price.

Consider the following two portfolios:

Portfolio C: A European put option plus one share, and

Portfolio D: Amount of cash equal to PV(K),

where PV(K) is defined (as before) as the present value of the strike price discounted from option maturity to today. If the strike price is greater than the stock price at the option's maturity, the option will be exercised and the share will be exchanged for the strike price so that Portfolio C is worth the strike price K. If the strike price is less than the stock price, the option is not exercised and Portfolio C is worth the stock price. Portfolio C is therefore worth:

$$\max(S_T, K)$$

at option maturity, where S_T is the stock price at maturity (as previously defined).

[5] As we will explain, the early exercise decision depends on the chance of the stock recovering and being above USD 20 in two months. The chance of this occurring is likely to be very small in this example so that exercising is almost certainly optimal.

Meanwhile, Portfolio D is worth K at option maturity. It follows that Portfolio C is worth at least as much as Portfolio D at maturity of the option. For there to be no arbitrage, the same must be true today. Hence:

$$\text{European Put Price} + S \geq \text{PV}(K)$$

so that

$$\text{European Put Price} \geq \text{PV}(K) - S \tag{13.3}$$

A European put price cannot be negative and thus (13.3) becomes

$$\text{European Put Price} \geq \max(\text{PV}(K) - S, 0)$$

Because an American put can be exercised at any time, a better lower bound we can determine is

$$\text{American Put Price} \geq \max(K - S, 0)$$

The upper bound for the price of a European put on a non-dividend-paying stock is PV(K) because we know it cannot be worth more than K at maturity. The upper bound for an American-style put option is K.

The general shape of a put option price as a function of the stock price is shown in Figure 13.3. The bounds for European and American put options on non-dividend paying stocks are illustrated in Figures 13.4 and 13.5.

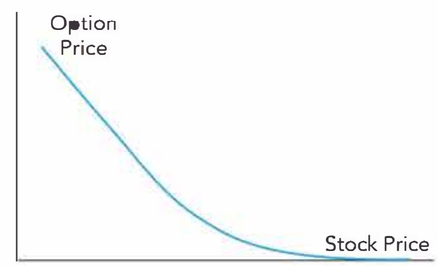


Figure 13.3 General shape of put option price as a function of stock price.

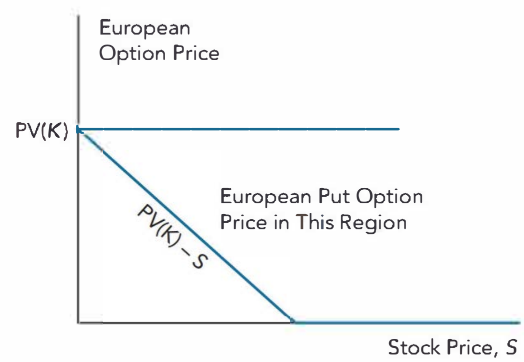


Figure 13.4 Bounds for price of a European put option on non-dividend paying stock.

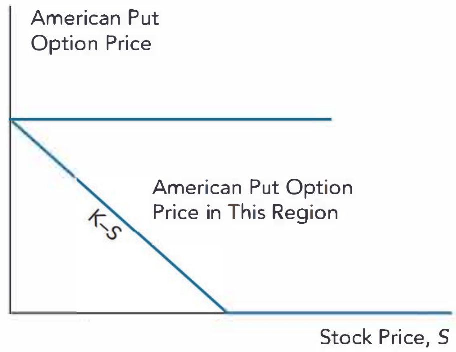


Figure 13.5 **Bounds for American put option on a non-dividend-paying stock.**

The Impact of Dividends

To incorporate dividends into the analysis of put options, Portfolios C and D are redefined as:

Portfolio C: A European put option plus one share, and

Portfolio D: An amount of cash equal to PV(K) + PV(Divs),

where PV(Divs) is (again) the dividends' present value. Portfolio C becomes

$$\max(S_T, K) + \text{FV(Divs)}$$

at maturity, where FV(Divs) is the value to which the dividends grow at option maturity. In Portfolio D, PV(K) becomes K at maturity and PV(Divs) becomes FV(Divs). Portfolio D therefore becomes

$$K + \text{FV(Divs)}$$

Because Portfolio C is worth at least as much as Portfolio D at maturity of the option, and there is no arbitrage, such inequality must also hold for their present values. Hence:

$$\text{European Put Price} + S \geq \text{PV}(K) + \text{PV(Divs)}$$

Equation (13.3) is modified to:

$$\text{European Put Price} \geq \max(\text{PV}(K) + \text{PV(Divs)} - S, 0)$$

Whereas dividends make it more likely a call option will be exercised early, they make it less likely that a put option will be exercised early. If an investor knows there will be future dividends, he or she is likely to be less inclined to exercise because that would mean selling the stock and giving up those dividends.

Options far from maturity are also less likely to be exercised early because there will be more time for the stock to move above the strike price.

In general, exercising becomes less attractive to the holder of a put option when:

- Stock price increases,
- Interest rate decreases,
- Time to maturity increases, and
- Dividends expected during the life of the option increase.

13.3 PUT–CALL PARITY

Put-call parity describes the relationship between the price of a European call option and that of a European put option with the same strike price and time to maturity. Consider the following two portfolios from previous sections:

Portfolio A: European call option plus an amount of cash equal to PV(K) + PV(Divs), and

Portfolio C: European put option plus one share.

Assume the call and put options have the same strike price and time to maturity.

As discussed in the previous sections, PV(K) in Portfolio A becomes K at the option maturity, while PV(Divs) becomes FV(Divs). If the call option is exercised (because the stock price is greater than the strike price), K will be exchanged for a share and Portfolio A will become one share plus FV(Divs). If the call option is not exercised (because the stock price is less than the strike price), Portfolio A will simply be K plus FV(Divs). Combining these two observations gives the value of Portfolio A at maturity:

$$\max(S_T, K) + \text{FV(Divs)}$$

Next consider Portfolio C. The dividends on the stock grow to FV(Divs) by investing at the risk-free interest rate. If the stock price at option maturity is greater than the strike price, the put option will not be exercised and the value of Portfolio C will become S_T + FV(Divs). If the stock price at option maturity is less than the strike price, then the put option will be exercised and the value of Portfolio C will become K + FV(Divs). Putting these results together gives the value of Portfolio C:

$$\max(S_T, K) + \text{FV(Divs)}$$

Portfolios A and C are therefore worth the same at option maturity. For there to be no arbitrage, they must be worth the same today and hence:

$$\text{European Call Price} + \text{PV}(K) + \text{PV(Divs)} = \text{European Put Price} + S \quad (13.4)$$

The above argument is summarized in Table 13.1.

Table 13.1 Values of Portfolio A and Portfolio C at Option Maturity Leading to Equation (13.4) are Shown. S_T is the Stock Price at Option Maturity

		$S_T > K$	$S_T < K$
Portfolio A	Call Option	$S_T - K$	0
	PV(K)	K	K
	PV(Divs)	FV(Divs)	FV(Divs)
	Total	S_T + **FV(Divs)**	K + **FV(Divs)**
Portfolio C	Put Option	0	$K - S_T$
	Share of Stock	S_T + FV(Divs)	S_T + FV(Divs)
	Total	S_T + **FV(Divs)**	K + **FV(Divs)**

To illustrate the existence of arbitrage opportunities when put–call parity does not hold, consider the following two examples.

Example 1:

Suppose the current stock price is USD 38. The price of a European call option with strike price USD 40 that will mature in six months is USD 5. The price of a European put with the same strike price and time to maturity is USD 4, while the six-month risk-free rate is 4% per year (annually compounded). Assume the stock on which the options are written provides no dividends.

Then the current value of Portfolio A is therefore 5 + PV(40). This can be written as:

$$5 + \frac{40}{1.04^{0.5}} = 44.22$$

Meanwhile, the value of Portfolio C is USD 42 (= 38 + 4). This means that Portfolio C is worth less than Portfolio A, and it implies that the put price is too low relative to the call price. An arbitrageur should

- Short the call,
- Buy the put, and
- Buy the stock.

The cost of this position (in USD) is

$$-5 + 4 + 38 = 37$$

If the stock price is greater than USD 40 at option maturity, the call option will be exercised against the arbitrageur and the put option will be worthless. Then, the arbitrageur will deliver the stock and will receive USD 40. The value of the portfolio is then exactly USD 40 at maturity. If the stock price is less than USD 40 at maturity, the put option will be exercised and the call option will be worthless. The arbitrageur sells the stock for USD 40, and again the value of the portfolio is exactly USD 40 at maturity.

Because the risk-free rate is 4% per year, the arbitrageur can simply borrow USD 37 at 4% per annum to set the position. As a result, he or she will have more than enough funds at maturity to repay the loan. The gain to the arbitrageur at maturity is therefore:

$$40 - 37 \times 1.04^{0.5} = 2.27$$

Example 2:

Suppose the current stock price is USD 38. The price of a European call option, with a strike price of USD 40 and time to maturity of six months is USD 5. The price of a European put option with the same strike price and time to maturity is USD 9. The risk-free rate is 4% per year (annually compounded) for all maturities and a dividend of USD 1 is expected in three months.

The present value (USD) of the dividends is

$$\frac{1}{1.04^{0.25}} = 0.99$$

The value of Portfolio A is

$$5 + \frac{40}{1.04^{0.5}} + 0.99 = 45.21$$

The value of Portfolio C is USD 47 (= 38 + 9). This means that Portfolio C is worth more than Portfolio A and implies that the put price is too high relative to the call price. An arbitrageur should

- Buy the call,
- Short the put, and
- Short the stock.

This creates a cash inflow today equal to:

$$-5 + 9 + 38 = 42$$

This is invested at the risk-free rate to grow to:

$$42 \times 1.04^{0.5} = 42.83$$

at option maturity.

If the stock price is greater than USD 40 at maturity, the call option will be exercised and the put option will be worthless. The arbitrageur will then buy the stock for USD 40 and will use it to close out the short position in the stock. The cost of doing this will be USD 40. If the stock price is less than USD 40 at maturity, the put option will be exercised against the arbitrageur and the call option will be worthless. The arbitrageur will again buy the stock for USD 40 and close out the short stock position.

Thus, the cash inflow from setting up the portfolio grows to USD 42.83 at option maturity, and the cost of the position at option maturity is always USD 40. Therefore, there is always a gain of:

$$42.83 - 40.00 = 2.83$$

Note that put-call parity applies only to European options. There is no exact relationship between the price of an American call option and that of the corresponding American put option.

13.4 USE OF FORWARD PRICES

The results we have produced can be generalized to other underlying assets besides stocks. The simplest way of doing this is to work with the corresponding forward contracts and avoid consideration of the income on the asset. We redefine Portfolios A and C as:

Portfolio A: A European call option plus cash equal to PV(K), and

Portfolio C: A European put plus a forward contract to buy asset for F at option maturity plus cash equal to PV(F),

where F is the forward price for a contract maturing at the same time as the options and PV(F) is the present value of F when discounted from the options' maturity at the risk-free rate. The call and put options have the same strike price K and the same time to maturity.

Assume PV(K) and PV(F) are invested at the risk-free rate so that they become K and F (respectively) at option maturity. The value of Portfolio A at option maturity is the European call option plus an amount of cash equal to K. If the asset price is greater than K at maturity, then the option will be exercised and Portfolio A will be worth the asset price. Otherwise, Portfolio A will be worth K.

At the same time, Portfolio C becomes the European put option plus the asset (purchased for F under the terms of the forward contract). If the asset price is less than K at maturity, the option will then be exercised, and the portfolio will be worth K. Otherwise, Portfolio C will be worth the asset price.

Because both portfolios will be worth the greater of the strike price and the asset price at maturity (and if there are no arbitrage opportunities), both portfolios must be worth the same today. The current value of Portfolio A is

$$\text{Call Price} + \text{PV}(K)$$

Because it costs nothing to enter into a forward contract, the current value of Portfolio C is

$$\text{Put Price} + \text{PV}(F)$$

The put-call parity relationship is therefore:

$$\text{European Call Price} + \text{PV}(K) = \text{European Put Price} + \text{PV}(F) \quad (13.5)$$

The arguments that lead to this equation are summarized in Table 13.2.

Table 13.2 **Values of Portfolio A and Portfolio C at Option Maturity Leading to Equation (13.5) are Shown. S_T is the Asset Price at Maturity**

		$S_T > K$	$S_T < K$
Portfolio A	Call Option	$S_T - K$	0
	PV(K)	K	K
	Total	S_T	K
Portfolio C	Put Option	0	$K - S_T$
	Forward Contract	$S_T - F$	$S_T - F$
	PV(F)	F	F
	Total	S_T	K

Because the put price cannot be negative, a lower bound for a European call price can be deduced from Equation (13.5) as:

$$\text{European Call Price} \geq PV(F) - PV(K)$$

which is an extension of the result in Equation (13.1).

Similarly, because the call price cannot be negative, the lower bound of the European put price is

$$\text{European Put Price} \geq PV(K) - PV(F)$$

which is an extension of the result in Equation (13.3).

These results can be used in conjunction with the formulas for forward prices in Chapters 9 and 10. Analysts often interpolate between forward prices observed in the market to obtain forward prices for the maturity of an option under consideration.

SUMMARY

This chapter derives several important results on option prices without using any assumptions regarding the dynamics of the underlying asset. In particular, an American-style call option on a stock paying no dividends should never be exercised before maturity. When there are dividends, it might be optimal to exercise a call option on a stock immediately before an ex-dividend date, but it is never optimal to exercise at other times.

In contrast, it can be optimal to exercise a put option before maturity regardless of whether the underlying asset pays a dividend.

Arguments can be used to derive the lower and upper bounds of option prices. If option prices in the market are less than the lower bound or greater than the upper bound, there will be straightforward arbitrage opportunities.

An important result presented in this chapter is put-call parity, which is the relationship between the price of a European call option and that of a European put option sharing the same strike price and time to maturity.

The results in this chapter can be extended to assets other than stocks. Analysts often work with forward prices because this avoids the need to make estimates about future dividends or other income on an asset.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

13.1 Give two reasons why it is not optimal to exercise an American call option on a non-dividend-paying stock before maturity. One reason should involve the time value of money. The other should involve the loss of optionality.

13.2 Why is an employee stock option on a non-dividend-paying stock sometimes exercised before maturity?

13.3 Give an example to show that a European call option's price can decrease as its maturity is increased.

13.4 What is the lower bound for the price of a European call option on a non-dividend-paying stock?

13.5 "The early exercise of an American put option on a non-dividend-paying stock is a trade-off between the time value of money and the desire to retain the optionality." Explain this statement.

13.6 Give an example to show that the price of a European put option can decrease as its maturity increases.

13.7 Does it become more likely that a put option will be exercised before maturity when (a) the interest rate increases and (b) the stock price increases?

13.8 What is the lower bound for the price of a European put option on a non-dividend-paying stock?

13.9 What is the put-call parity formula in terms of forward prices?

13.10 What is the advantage of expressing put-call parity in terms of the forward prices?

Practice Questions

13.11 Is an in-the-money American call option on a non-dividend-paying stock worth more than its intrinsic value when interest rates are positive? Explain.

13.12 A four-month European call option on a stock is currently selling for USD 2.50. The current stock price is USD 54, and the strike price is USD 50. A dividend of USD 1.50 is expected in one month. The risk-free interest rate is 3% per annum (annually compounded) for all maturities. What opportunities are there for an arbitrageur?

13.13 A seven-month call option pays dividends of USD 0.5 in three months and six months. The strike price is USD 40. Assume a constant risk-free rate of 8% per annum (annually compounded) for all maturities. Is it ever optimal to exercise the option before maturity? Explain.

13.14 What will be the lower bound for the price of a three-month European put option on a non-dividend-paying stock if the current stock price is USD 22, the strike price is USD 25, and the risk-free rate is 6% per year (annually compounded)?

13.15 As explained in Chapter 12, historically calls were traded on exchanges before puts. Explain how you can synthetically create a European put from a European call. Assume no dividends are paid and no arbitrage opportunities exist.

13.16 The current price of a non-dividend-paying stock is USD 29 and the price of a four-month call option on the stock with a strike price of USD 30 is USD 2. The risk-free rate is 4% per annum (annually compounded). What is the price of a four-month put option on the stock with a strike price of USD 30? Assume no arbitrage opportunities exist.

13.17 A European call and European put option on a stock both cost USD 5 with a common strike price USD 30 and a common time to maturity of one year. The current stock price is USD 30. What arbitrage opportunities does this create? Assume no dividend is paid and the interest rate is positive.

13.18 Use the results in Chapter 9 to determine put-call parity for a currency options on the GBP/USD exchange rate. Express your answer in terms of the USD risk-free rate, R_{USD}, the GBP risk-free rate, R_{GBP}, and the time to maturity, T.

13.19 Use the results in Chapter 10 to determine put-call parity for an index option. Express your answer in terms of the risk-free rate, R, the dividend yield, Q, and the time to maturity, T.

13.20 Under what circumstances does a European call on an asset equal the price of a European put on the asset when both have the same strike price and time to maturity? Express your answer in terms of forward prices.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

13.1 Delaying paying the strike price allows interest to be earned on the strike price for a longer time period. The call option also provides insurance against the event that the stock price becomes less than the strike price at maturity. Once the option has been exercised, this optionality will be lost.

13.2 An employee stock option cannot be traded. Thus, if the employee does not want a position in the company's stock, the only alternative will be exercising the option and selling the stock.

13.3 If a big dividend is due to be paid, the holder of the option would like to exercise immediately before the ex-dividend date. A European option that matures just before the ex-dividend date is then likely to be worth more than one that matures after the ex-dividend date.

13.4 The lower bound is $\max(S - PV(K), 0)$ where S is the stock price and K is the strike price.

13.5 By exercising before maturity, the holder of a put option receives the strike price earlier. However, delaying the exercise of the option will benefit the holder of the stock if the stock price rises above the strike price later. Once the option is exercised, however, the benefit of this optionality is lost.

13.6 An example would be a deep-in-the-money European put option that is (almost) certain to be exercised. The shorter the maturity of the option, the earlier the strike price is received, and the greater the value of the option.

13.7 (a) When interest rates increase, the put option will be more likely to be exercised before maturity because the value of investing the profit at the risk-free rate after selling the stock at the strike price is increased. (b) When the stock price increases, the put option will be less likely to be exercised because the payoff from doing so will be smaller.

13.8 The lower bound is $\max(PV(K) - S, 0)$ where S is the stock price and K is the strike price.

13.9 Call price $+$ PV(K) $=$ Put price $+$ PV(F) where K is the common strike price and F is the forward price for a contract that matures at the same time as the options.

13.10 It is not necessary to predict income on the asset as the forward prices reflect all future income before maturity.

Solved Problems

13.11 Yes. An in-the-money American call on a non-dividend paying stock is worth at least $S - PV(K)$, which is greater than $S - K$. It is never worth exercising the call early and so it must be worth more than its intrinsic value.

13.12 The lower bound for the option price is

$$S - PV(K) - PV(Divs) = 54 - \frac{50}{1.03^{1/3}} - \frac{1.5}{1.03^{1/12}} = 2.99$$

The option is selling for less than its lower bound. An arbitrageur can buy the option and short the stock for an initial cash inflow of USD 51.50. The arbitrageur has to pay dividends of USD 1.50 after one month. If the option is exercised, the cost of closing out the short position will be USD 50. If it is not exercised, the cost of closing out the short position will be less than USD 50. The worst-case scenario for the arbitrageur is therefore:

Today:	+51.50,
One month:	−1.50, and
Four months:	−50.00.

When the discount rate is zero, the sum of these cash flows will have zero present value. Any positive discount rate gives a positive sum of present values.

13.13 It is only optimal to exercise immediately before a dividend payment. Immediately before the three-month payment, the option holder should wait, because there are three months until the next dividend payment and $K - K^*$ is greater than the dividend payment:

$$K - K^* = 40 - \frac{40}{1.08^{0.25}} = 0.76 > 0.5$$

Exercise can be optimal immediately before the six-month dividend payment because there is only one month to maturity and $K - K^*$ is less than the dividend payment:

$$K - K^* = 40 - \frac{40}{1.08^{1/12}} = 0.26 < 0.5$$

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

13.14 The lower bound (USD) is

$$\frac{25}{1.06^{0.25}} - 22 = 2.64$$

13.15 From put-call parity:

$$\text{Call} + \text{PV}(K) = \text{Put} + S$$

so that:

$$\text{Put} = \text{Call} + \text{PV}(K) - S$$

The put can therefore be created by buying the call, shorting the stock, and investing PV(K) so that it grows to K at maturity.

13.16 By put-call parity:

$$\text{Put} = \text{Call} + \text{PV}(K) - S,$$

the put price (USD) is thus given by:

$$2 + \frac{30}{1.04^{1/3}} - 29 = 2.61$$

13.17 From put-call parity, the excess of the call price over the put price is S − PV(K). In this case $S = K = 30$ and so S − PV(K) is positive. The call should be worth more than the put, but they are both worth the same. An arbitrageur should buy the call, sell the put, and short the stock.

13.18 From Equation (9.1):

$$F = S\frac{(1 + R_{USD})^T}{(1 + R_{GBP})^T}$$

Substituting this into Equation (13.5) and noting that:

$$PV(K) = \frac{K}{(1 + R_{USD})^T}$$

$$PV(F) = S\frac{(1 + R_{USD})^T}{(1 + R_{GBP})^T}\frac{1}{(1 + R_{USD})^T} = \frac{S}{(1 + R_{GBP})^T}$$

$$\text{European Call Price} + \frac{K}{(1 + R_{USD})^T}$$

$$= \text{European Put Price} + \frac{S}{(1 + R_{GBP})^T}$$

13.19 From Equation (10.3):

$$F = S\left(\frac{1 + R}{1 + Q}\right)^T$$

Substituting this into Equation (13.5) and noting that:

$$PV(K) = \frac{K}{(1 + R)^T}$$

$$PV(F) = S\frac{(1 + R)^T}{(1 + Q)^T}\frac{1}{(1 + R)^T} = \frac{S}{(1 + Q)^T}$$

$$\text{European Call Price} + \frac{K}{(1 + R)^T}$$

$$= \text{European Put Price} + \frac{S}{(1 + Q)^T}$$

13.20 The put-call parity formula is

$$\text{European Call Price} + \text{PV}(K)$$
$$= \text{European Put Price} + \text{PV}(F)$$

PV(F) = PV(K) when $F = K$. The equation therefore shows that a European call has the same price as a European put when $F = K$, that is, when the strike price equals the forward price for a forward contract maturing at the same time as the option.

Trading Strategies 14

Learning Objectives

After completing this reading, you should be able to:

- Explain the motivation to initiate a covered call or a protective put strategy.
- Describe principal protected notes (PPNs) and explain necessary conditions to create them.
- Describe the use and calculate the payoffs of various spread strategies.
- Describe the use and explain the payoff functions of combination strategies.

Options can be arranged to form a wide spectrum of payoff patterns. In theory, an investor with access to European options with all strike prices for a given maturity could achieve any continuous payoff function of the underlying asset price at expiry.

The option trading strategies considered in this chapter can be divided into four groups:

1. Strategies involving an option and the underlying asset,
2. Strategies involving two or more call options,
3. Strategies involving two or more put options, and
4. Strategies involving both call and put options.

Strategies involving only call options or only put options are termed *spreads*. Those involving both call and put options are termed *combinations*.

For ease of exposition, it is assumed that (except where otherwise stated) the underlying asset provides no income and all options are European.

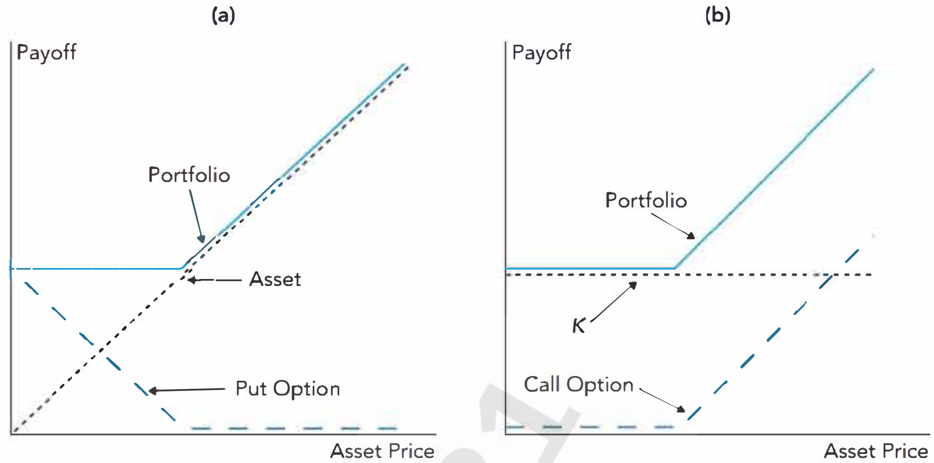


Figure 14.1 An Illustration of Equation (14.1) (a) shows the final value of a protective put (i.e., put plus the asset) and (b) shows the final value of call plus PV(K).

14.1 STRATEGIES INVOLVING A SINGLE OPTION

The put-call parity result describes the relationship between the price of a European put option and that of a European call option with the same strike price and time to maturity. As derived in Chapter 13, the put-call parity result (assuming that the underlying asset provides no income) is

$$p + S = c + PV(K) \tag{14.1}$$

where p is the price of a European put option with strike price K and c is the price of a European call option with strike price K (with both options having the same time to maturity). S is the current asset price, and PV denotes present value from option maturity to today at the risk-free rate. This result can also be written in the following ways:

$$S - c = PV(K) - p \tag{14.2}$$

$$-p - S = -c - PV(K) \tag{14.3}$$

$$c - S = p - PV(K) \tag{14.4}$$

Equations (14.1) to (14.4) capture the various ways in which an option position can be combined with a position in the asset. Consider the following specific points.

- Equation (14.1) shows that a put option combined with the asset gives a position equivalent to a call option combined with an amount of cash equal to PV(K). Figure 14.1 illustrates this by showing the values of the two portfolios at option maturity. Buying a put option when the underlying asset is held is referred to as a *protective put* strategy. A trader who sets up a protective put is usually bullish about the asset price (hence the position in the asset), but also wishes to limit losses in the event of an unexpected price decrease (hence the put option).
- Equation (14.2) shows that an asset combined with a short call is equivalent to a short put option combined with an amount of cash equal to PV(K). Figure 14.2 illustrates this by showing the values of the two portfolios at option maturity. An asset plus a short call is known as a *covered call*. The call is usually out-of-the-money, allowing the asset owner to obtain a cash inflow equal to the call option premium in exchange for giving up some of the potential upside from an increase in the asset price.
- Equation (14.3) is the reverse of a protected put. It shows that a short put combined with a short position in the asset is equivalent to a short call combined with a liability of PV(K).
- Equation (14.4) is the reverse of a covered call. It shows that a long call combined with a short position in the asset is equivalent to a long put combined with a liability of PV(K).

Principal Protected Notes

A principal protected note (PPN) is a security created from a single option such that the investor benefits from any gain in the value of a specified portfolio without the risk of losses.

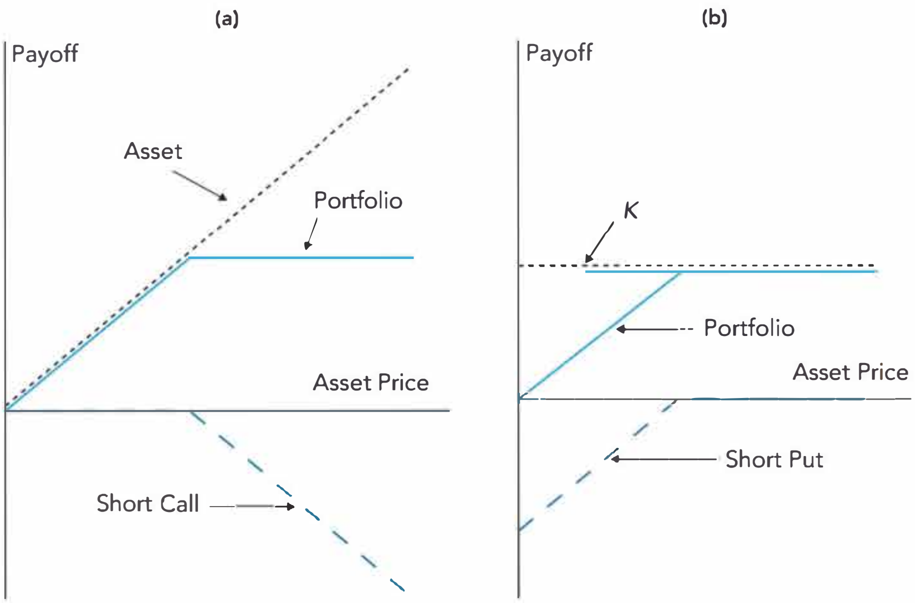


Figure 14.2 **An Illustration of Equation (14.2) (a) shows the final value of a covered call (i.e., asset plus short call) and (b) shows the final value of a short put plus PV(*K*).**

To show how a principal protected note is created, suppose the three-year interest rate is 7% (annually compounded). This means that the present value of USD 10,000 is

$$\frac{\text{USD } 10{,}000}{1.07^3} = \text{USD } 8{,}162.98$$

Suppose that Portfolio A consists of:

- A three-year zero-coupon bond that will pay USD 10,000 in three years; and
- A three-year call option on Portfolio B, which is currently worth USD 1,832.02 with a strike price of USD 10,000.

The holder of Portfolio A will to benefit from any increase in the value of Portfolio B without incurring any loss of principal if its value declines. This feature is likely to be attractive to risk-averse investors.

Note that the first item in Portfolio A costs USD 8,162.98. If the option can be purchased for less than USD 1,837.02 (= USD 10,000 − USD 8,162.98), then the portfolio will cost less than USD 10,000 and can be profitably offered to investors.

PPNs are possible because of the following.

- The investor is giving up three years of interest on the USD 10,000 investment.
- The investor does not receive any income that the owners of Portfolio B would receive during the next three years.

Full participation PPNs (i.e., the owner receives 100% of the upside) are only possible for portfolios that provide an income.

To see this, note that when there is no income Equation (14.1) gives

$$p + S = c + \text{PV}(K)$$

Because the call is currently at the money (so that $K = S$),

$$c = p + S - \text{PV}(S)$$

This implies

$$c > S - \text{PV}(S)$$

This shows that the call always costs more than the funds available to pay for it. (In our example, the call will always cost more than USD 1,837.02). However, Portfolio B's income reduces the value of the call. If the income is sufficiently high, a PPN can be created.

Returning to our example, suppose that Portfolio B provides a yield of 2%. An extension of the Black-Scholes Merton model (which will be covered in *Valuation and Risk Models*) shows that a PPN can be created if its volatility is less than 18%. (This is because the at-the-money call option can then be created for less than USD 1,837.02.)

14.2 SPREAD TRADING STRATEGIES

Now consider trading strategies involving positions in two or more call (or put) options.

Bull Spread

A bull spread (as its name implies) is a position appropriate for an investor expecting an increase in the price of an asset. To create the spread, the trader buys a European call option with strike price K_1 and sells a European call option with strike price K_2. In this situation, $K_2 > K_1$ and both options have the same time to maturity. The payoff from a bull spread is calculated in Table 14.1 and illustrated in Figure 14.3.

Table 14.1 **A Payoff from a Bull Spread is Shown. S_T is the Asset Price at Maturity**

Range for S_T	Payoff from Long Call with Price K_1	Payoff from Short Call with Strike Price K_2	Total Payoff
$S_T \leq K_1$	0	0	0
$K_1 < S_T \leq K_2$	$S_T - K_1$	0	$S_T - K_1$
$S_T > K_2$	$S_T - K_1$	$-(S_T - K_2)$	$K_2 - K_1$

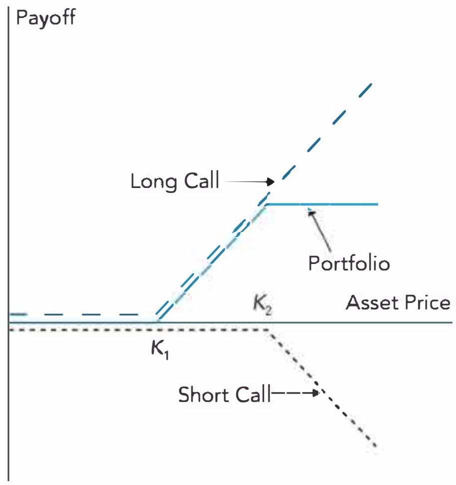


Figure 14.3 **Payoff from a bull spread created from long and short call options.**

The advantage of a bull spread (compared with simply buying a call option that has a strike price K_1) is that it is less expensive to set up. The trader chooses to give up gains from the asset price rising above K_2. In return, there is a cost savings equal to the price of an option with strike price K_2.

The choice of strike prices affects the potential returns to the bull spread creator. If both options are out-of-the-money, the bull spread costs very little to set up. There is also a small probability that both options will become in-the-money and thus generate a high return. If both options are initially in-the-money, the bull spread is more expensive to set up and there is a high probability that the strategy gives a positive return, but the return is quite modest.

Put options can also be used to set up a bull spread. Suppose c_1 and c_2 are the prices of call options with strike prices K_1 and K_2 (respectively) and that p_1 and p_2 are the prices of the corresponding put options (also respectively). By put-call parity:

$$c_1 + PV(K_1) = p_1 + S$$

$$c_2 + PV(K_2) = p_2 + S \tag{14.5}$$

Thus:

$$c_1 - c_2 = p_1 - p_2 + PV(K_2) - PV(K_1)$$

Note that the left-hand side of this equation is the cost of setting up the bull spread using calls. The right-hand side shows the cost is the same if the call options are replaced by put options providing an amount of cash equal to the present value of $K_2 - K_1$. As shown in Table 14.2, the payoff when using puts and the cash is the same as in Table 14.1 (where calls are used).

Bear Spread

A bear spread is a position where the trader buys a European put option with strike price K_2 and sells a European put option with strike price K_1. In this situation, $K_2 > K_1$ and both options have the same time to maturity. The payoff from a bear spread is calculated in Table 14.3 and is illustrated in Figure 14.4.

Like a bull spread, a bear spread has a small probability of attaining a large return if both options begin out-of-the-money and a high probability of attaining a modest return if both options begin in-the-money.

Table 14.2 **A Payoff from a Bull Spread Created from Puts Rather than Calls Is Shown. S_T Is the Asset Price at Maturity**

Range for S_T	Payoff from Long Put with Strike Price K_1	Payoff from Short Put with Strike Price K_2	Cash Amount	Total Payoff
$S_T \le K_1$	$K_1 - S_T$	$-(K_2 - S_T)$	$K_2 - K_1$	0
$K_1 < S_T \le K_2$	0	$-(K_2 - S_T)$	$K_2 - K_1$	$S_T - K_1$
$S_T > K_2$	0	0	$K_2 - K_1$	$K_2 - K_1$

Table 14.3 **A Payoff from a Bear Spread Is Shown. S_T Is the Asset Price at Maturity**

Range for S_T	Payoff from Long Put with Strike Price K_2	Payoff from Short Put with Strike Price K_1	Total Payoff
$S_T \le K_1$	$K_2 - S_T$	$-(K_1 - S_T)$	$K_2 - K_1$
$K_1 < S_T \le K_2$	$K_2 - S_T$	0	$K_2 - S_T$
$S_T > K_2$	0	0	0

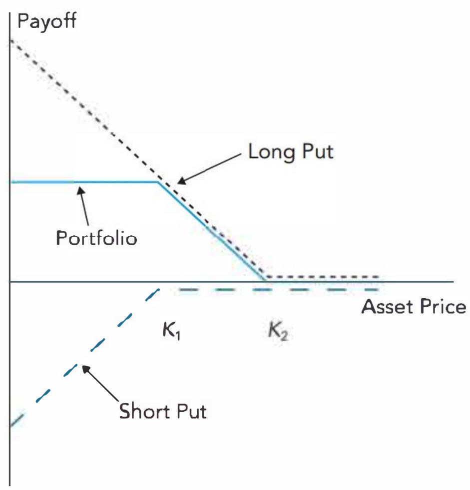


Figure 14.4 A payoff from bear spread created from long and short put options is shown.

Call options can also be used to create a bear spread. From Equation (14.5):

$$p_2 - p_1 = c_2 - c_1 + PV(K_2) - PV(K_1)$$

The left-hand side of this equation is the cost of setting up the bear spread by using put options. The right-hand side shows the cost is the same if the put options are replaced by call options along with an amount of cash equal to the present value of $K_2 - K_1$. As shown in Table 14.4, the payoff is the same as in Figure 14.4.

Note that while a bull spread (regardless of whether calls or puts are used) involves buying at the low strike price and selling at the high strike price, a bear spread always involves buying at the high strike price and selling at the low strike price.

Box Spread

A box spread is a portfolio created from a bull spread (using call options) and a bear spread (using put options). The strike prices and times to maturity used for the bull spread are the same as those used for the bear spread. The box spread aggregates the payoffs from the portfolios in Figures 14.3 and 14.4 (the solid lines) and produces a payoff that is always $K_2 - K_1$. This is demonstrated in Table 14.5.

The cost of setting up a box spread should be $PV(K_2 - K_1)$. If it costs less than this, an arbitrageur can profit by buying the box spread and earning more than the risk-free rate. If it costs more than this, an arbitrageur can short the box spread and achieve a borrowing rate that is less than the risk-free rate.

Note that only European options can be used to construct a box spread with a known payoff at a future time. Because American options can be exercised early, the final payoff generated when American options are used can differ from $K_2 - K_1$.

Butterfly Spread

A butterfly spread involves positions in three options. Like the other spreads discussed in this chapter, it can be created from either call or put options. If calls are used, a trader will take the following positions:

- One long European call with strike price K_1,
- One long European call with strike price K_3 where $K_3 > K_1$, and
- Two short European calls with strike price $K_2 = (K_1 + K_3)/2$.

Table 14.4 A Payoff from a Bear Spread Created from Calls Rather than Puts Is Shown. S_T Is the Asset Price at Maturity

Range for S_T	Payoff from Long Call with Strike Price K_2	Payoff from Short Call with Strike Price K_1	Cash Amount	Total Payoff
$S_T \leq K_1$	0	0	$K_2 - K_1$	$K_2 - K_1$
$K_1 < S_T \leq K_2$	0	$-(S_T - K_1)$	$K_2 - K_1$	$K_2 - S_T$
$S_T > K_2$	$S_T - K_2$	$-(S_T - K_1)$	$K_2 - K_1$	0

Table 14.5 Payoff from a Box Spread

Asset Price Range	Payoff from Bull Spread (Table 14.1)	Payoff from Bear Spread (Table 14.3)	Total Payoff
$S_T \leq K_1$	0	$K_2 - K_1$	$K_2 - K_1$
$K_1 < S_T < K_2$	$S_T - K_1$	$K_2 - S_T$	$K_2 - K_1$
$S_T \geq K_2$	$K_2 - K_1$	0	$K_2 - K_1$

Table 14.6 **Calculation of the Payoff from a Butterfly Spread Created from Calls Is Shown. S_T Is the Asset Price at Maturity**

Range for S_T	Payoff from Long Call with Strike Price K_1	Payoff from Two Short Calls with Strike Price K_2	Payoff from Long Call with Strike Price K_3	Total Payoff[1]
$S_T \le K_1$	0	0	0	0
$K_1 < S_T \le K_2$	$S_T - K_1$	0	0	$S_T - K_1$
$K_2 < S_T \le K_3$	$S_T - K_1$	$-2(S_T - K_2)$	0	$2K_2 - K_1 - S_T = K_3 - S_T$
$S_T > K_3$	$S_T - K_1$	$-(S_T - K_2)$	$S_T - K_3$	$2K_2 - K_1 - K_3 = 0$

All options have the same time to maturity. The payoff from a butterfly spread as a function of the asset price is shown in Table 14.6 and Figure 14.5.

Because a butterfly spread always provides a payoff that is positive or zero, it must be the case that the cost of setting up a butterfly spread is positive. This means that:

$$c_1 + c_3 > 2c_2$$

Such a condition can be proved to be true under the Black-Scholes Merton model (which will be covered in *Valuation and Risk Models*). Under any model, the price of a call option is always a convex function of the strike price (as illustrated in Figure 14.6).

Usually, K_2 is close to the current stock price. As indicated in Figure 14.5, a butterfly spread is an appropriate trading strategy when an investor anticipates a small movement in the asset price so that it is close to K_2 at maturity. A big move in either direction will lead to a payoff of zero. If a trader anticipates a big movement in the asset price and is uncertain whether the price will go up or down, he or she can short a butterfly spread. This will lead to a small profit if the trader is right and a small loss if the trader is wrong.

To see the how a butterfly spread can be created using put options, apply put-call parity three times:

$$c_1 + PV(K_1) = p_1 + S$$

$$c_2 + PV(K_2) = p_2 + S$$

$$c_3 + PV(K_3) = p_3 + S$$

Adding the first of these equations to the third and then subtracting twice the second equation gives

$$c_1 + c_3 - 2c_2 + PV(K_1 + K_3 - 2K_2) = p_1 + p_3 - 2p_2$$

Because $K_2 = 0.5(K_1 + K_3)$, the above equation reduces to:

$$c_1 + c_3 - 2c_2 = p_1 + p_3 - 2p_2$$

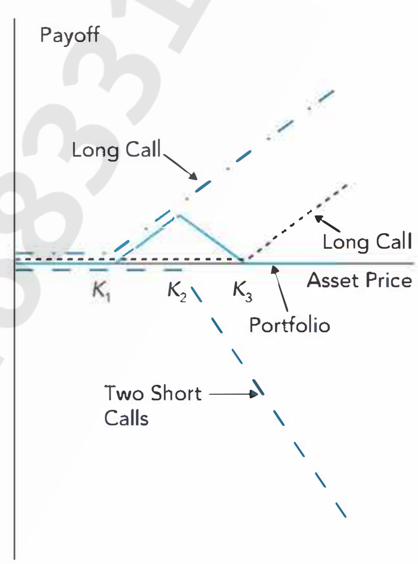


Figure 14.5 **Payoff from a butterfly spread created from call options.**

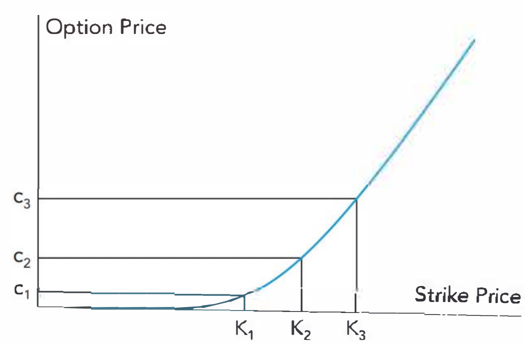


Figure 14.6 **Call option price as a function of the strike price.**

This shows that the cost of setting up the position should be the same whether the trader uses put options or call options. Furthermore, the payoffs (as shown in Table 14.7) are the same regardless of whether the trader uses puts or calls.

[1] This uses the assumption that $K_2 = 0.5(K_1 + K_3)$.

Table 14.7 Butterfly Spread Created from Put Options

Range for S_T	Payoff from Long Put with Strike Price K_1	Payoff from Two Short Puts with Strike Price K_2	Payoff from Long Put with Strike Price K_3	Total Payoff[2]
$S_T \leq K_1$	$K_1 - S_T$	$-2(K_2 - S_T)$	$K_3 - S_T$	$2K_2 - K_1 - K_3 = 0$
$K_1 < S_T \leq K_2$	0	$-2(K_2 - S_T)$	$K_3 - S_T$	$S_T + K_3 - 2K_2 = S_T - K_1$
$K_2 < S_T \leq K_3$	0	0	$K_3 - S_T$	$K_3 - S_T$
$S_T > K_3$	0	0	0	0

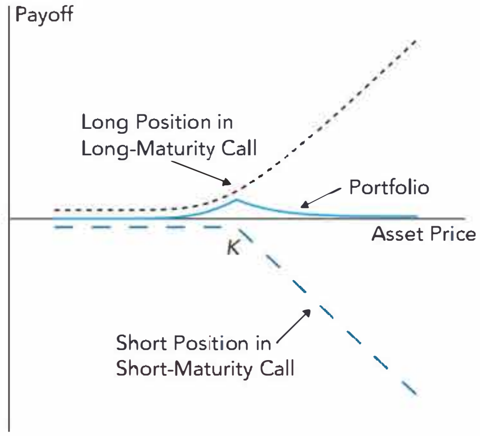


Figure 14.7 Payoff from a calendar spread as seen at the time the short-maturity option reaches maturity.

Calendar Spread

All positions considered thus far have involved options maturing at the same time. To create a calendar spread, a trader buys a call option maturing at time T^* and sells a call maturing at time T. With this position, $T^* > T$ and the two calls have the same strike price of K.

To examine the payoff from a calendar spread, consider the situation at the time when the short-maturity option matures (i.e., at time T). At this time, the option sold has just matured, while the option bought will mature at $T^* - T$. As mentioned earlier, the latter is a convex function of the asset price. The payoff is shown in Figure 14.7.

The profit from a calendar spread (as a function of the asset price) is like that of a butterfly spread. There will be a small gain if the asset price at time T_1 is near the strike price K. Otherwise, there will be a loss approximately equal to the cost of setting up the spread.

Let c and c* be the prices of call options with strike price K and maturities T and T^* (respectively). Meanwhile, let p and p^* be the prices of put options with strike price K and maturities T and T^* (also respectively). By put-call parity:

$$c + PV(K) = p + S$$

$$c^* + PV^*(K) = p^* + S$$

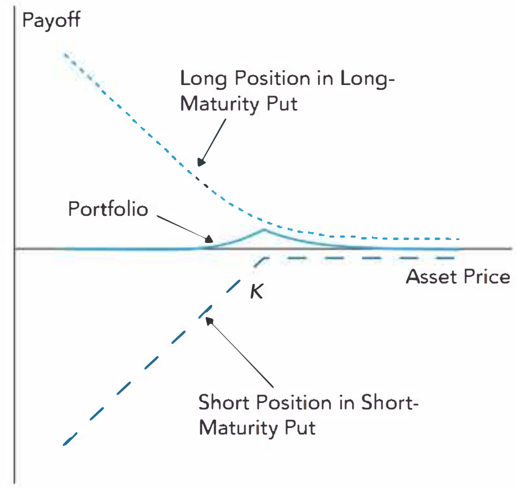


Figure 14.8 Payoff from a calendar spread created from put options as seen at the maturity of short-maturity option.

where PV denotes the present value when discounting from time T to today and PV* denotes present value when discounted from time T^* to today. It follows that:

$$c^* - c = p^* - p + PV(K) - PV(K^*)$$

Usually $PV(K) - PV^*(K)$ is small so that the cost of a calendar spread created from put options is close to that of a spread created from call options. Figure 14.8 shows the payoff pattern of a calendar spread when it is created from puts.

14.3 COMBINATIONS

This section considers positions created using both call and put options.

Straddle

A straddle is a position created from a long call and a long put with the same strike price and time to maturity. The strike price is usually close to the current asset price.

[2] Like Table 14.6, this uses $K_2 = 0.5(K_1 + K_3)$

Table 14.8 Payoff from a Straddle

Range for S_T	Payoff from Long Call with Strike Price K	Payoff from Long Put with Strike Price K	Total Payoff
$S_T \leq K$	0	$K - S_T$	$K - S_T$
$S_T > K$	$S_T - K$	0	$S_T - K$

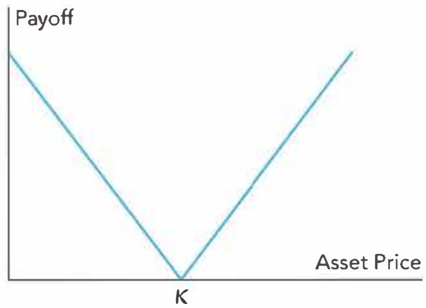


Figure 14.9 Payoff from straddle.

A straddle trader believes there will be a big move in the asset price but is unsure whether the move will be up or down.

As indicated in Table 14.8 and Figure 14.9, a straddle provides a V-shaped payoff.

A straddle can be fairly expensive to set up. This means that the asset price movement has to be quite large for a profit to be realized.

Consider a three-month straddle where the asset price and the strike price (for both the call and the put) is USD 50. If the volatility of the asset price is 20% per annum and the risk-free rate is 2% per annum, the Black-Scholes Merton model (to be covered in *Valuation and Risk Models*) shows that the prices of a three-month European call option and a European put option are USD 2.12 and USD 1.87 (respectively). This means that the asset price has to move by USD 3.99 (= 2.12 + 1.87) for there to be a profit. If the asset price is between USD 46 and USD 54, the straddle will incur a loss.

Strangle

The cost of a straddle can be reduced by making the strike price of the call greater than the strike price of the put. The position is then called a strangle. The payoff function is shown Table 14.9 and Figure 14.10.

Table 14.9 Payoff from a Strangle

Range for S_T	Payoff from Long Call with Strike Price K_2	Payoff from Long Put with Strike Price K_1	Total Payoff
$S_T \leq K_1$	0	$K_1 - S_T$	$K_1 - S_T$
$K_1 < S_T \leq K_2$	0	0	0
$S_T > K_2$	$S_T - K_2$	0	$S_T - K_2$

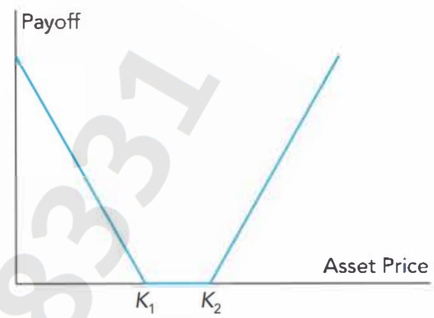


Figure 14.10 Payoff from a strangle.

Continuing with our earlier example, suppose the lower strike price (K_1) used for the put option is USD 45 and the upper strike price (K_2) used for the call option is USD 55. The Black-Scholes Merton model gives prices for the call and put options as USD 0.52 and USD 0.32 (respectively). The cost of setting up the strangle is therefore around USD 0.84, which is less than a quarter the cost of the straddle where both strike prices are USD 50. However, the asset price has to move even further for the position to be profitable. To make a profit, the asset price has to be greater than USD 55.84 or less than USD 44.16 on the expiry date.

To summarize, both a straddle and a strangle are positions designed to provide profits when there is a large move in the asset price. The larger the move, the greater the payoff. A strangle costs less than a straddle, but the asset price has to move further for it to be profitable.

14.4 MANUFACTURING PAYOFFS

If options with all strike prices could be traded, a trader could (in theory) create any payoff that is a continuous function of the asset price at a future time. To see this, note that the payoff from a butterfly spread is a *spike payoff* (see Figure 14.11). The spike can be made arbitrarily small by choosing strike prices that are close together. By constructing a portfolio from a large

Figure 14.11 **Spike payoff from a butterfly spread.**

number of small spikes, any payoff function can be replicated to arbitrary accuracy. Because a butterfly spread can be constructed from either calls or puts, any continuous payoff function can be created by using either type of option.

SUMMARY

This chapter describes some of the ways in which options can be used to produce various payoffs. Spreads are created from two or more call options as well as two or more put options. Combinations are created by using both calls and puts.

Many other strategies are used by option traders in addition to those discussed in this chapter. The following are examples.

- A *diagonal spread* is created from a long call (or put) and a short call (or put) where both the strike prices and the times to maturity are different. It can be considered as a cross between a bull/bear spread and a calendar spread.
- A *strip* is like a straddle except that two puts are purchased for every call. It is appropriate when a trader anticipates a big move in the asset price and a downward movement is considered more likely than an upward movement.
- A *strap* is like a straddle except that two calls are purchased for every put. It is appropriate when a trader anticipates a big move in the asset price and an upward movement is considered more likely than a downward movement.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

14.1 How is a covered call created? How can an equivalent position be created using a put?

14.2 How is a protective put created? Why would a trader create such a position?

14.3 How is a bull spread created from call options?

14.4 Explain two ways a bear spread can be created.

14.5 When should a trader consider creating a butterfly spread?

14.6 What options positions are necessary to build a short position in a box spread?

14.7 What is the difference between a straddle and a strangle?

14.8 A range forward contract involves buying a call and selling a put where the strike price of the call is greater than that of the put and the maturities of the two are the same. Sketch the payoff from the contract.

14.9 A trader feels that there will be a big jump in a stock price but is uncertain of the direction. Identify three trading strategies that the investor can implement to reflect his or her views.

14.10 A trader creates three six-month butterfly spreads. The first has strike prices of USD 30, USD 31, and USD 32. The second has strike prices of USD 31, USD 32, and USD 33. The third has strike prices of USD 32, USD 33, and USD 34. What is the resulting payoff pattern?

Practice Questions

14.11 "The higher interest rates are, the easier it is for a bank to create a principal protected note." Explain this statement.

14.12 Explain why two of the options in a box spread create a long forward position and another two create a short forward position.

14.13 Three put options have the same expiration date and their strike prices are USD 45, USD 50, and USD 55. The market prices of the options are USD 2, USD 4, and USD 7 (respectively). Explain how a butterfly spread can be created from these options. Provide a table showing the profit as a function of the asset price at maturity.

14.14 A trader creates a bear spread using put options with strike prices of USD 25 and USD 30 and the same time to maturity. The options cost USD 2 and USD 4.50 (respectively). Under what circumstances will the trade be profitable?

14.15 Suppose that put options on an asset with strike prices USD 30 and USD 35 and the same time to maturity cost USD 4 and USD 7 (respectively). How can these options be used to create (a) a bear spread and (b) a bull spread?

14.16 Describe the payoff from a long position in a call with strike price K_1 and a long position in a put with strike price K_2 where $K_2 > K_1$.

14.17 A long strangle is combined with a short straddle. The strike price in the straddle is halfway between the strike prices in the strangle. Describe the payoff from this position.

14.18 A call option with a strike price of USD 40 costs USD 2 and a put option with a strike price of USD 30 costs USD 3. Both have the same time to maturity. Explain how a strangle can be created using these options, and construct a table showing the profit as a function of the asset price at option maturity.

14.19 A call and a put with a common strike price of USD 70 cost USD 7 and USD 5 (respectively). How is a straddle created from these options? State the range of asset prices leading to some profit on the straddle.

14.20 Company A has just announced a takeover offer for Company B. The outcome of the takeover attempt is uncertain. If successful, a big increase in Company B's stock price can be anticipated. If unsuccessful, a big decrease can be anticipated. An investor buys a straddle on Company B's stock where the strike price is close to the current market price. Is this a good trade? Discuss.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

ANSWERS

Short Concept Questions

14.1 A covered call is created from a long position in an asset and a short position in a call. Equation (14.2) shows that it is equivalent to a short put position together with a cash position that equals the present value of the strike price.

14.2 A protective put is created from a long position in both an asset and a put option on the asset. The motivation is setting up protection against downward movements of the asset price while anticipating profit generated by upward asset price movement.

14.3 A bull spread is created by buying a call with a certain strike price and selling another call with a higher strike price. Both calls have the same time to maturity.

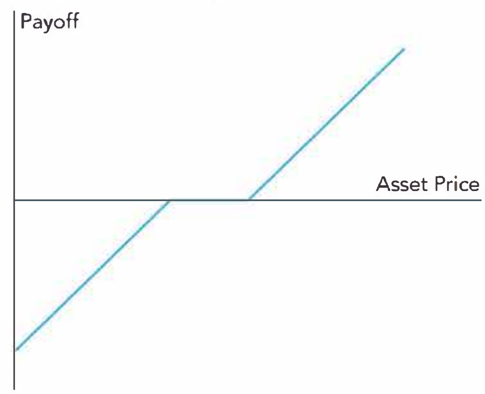
14.4 A bear spread can be created by buying a put with a certain strike price, K_2, and selling a put with a lower strike price, K_1. Both puts have the same time to maturity. It can also be created by buying a call with strike price K_2, selling a call with strike price K_1, and adding an amount of cash equal to the present value of the difference between the strike prices.

14.5 A butterfly spread is appropriate when the price of the asset at a certain future time is expected to be close to a particular level. The spread is created from options on three equally spaced strike prices where the middle strike price equals this expected level.

14.6 Suppose $K_2 < K_1$. A long box spread is created by buying a call with strike price K_1, selling a call with strike price K_2, buying a put with strike price K_2, and selling a put with strike price K_1. A short box spread is therefore created by selling a call with strike price K_1, buying a call with strike price K_2, selling a put with strike price K_2, and buying a put with strike price K_1. All options have the same time to maturity.

14.7 A straddle is constructed from a call and a put with the same strike price and time to maturity. For a strangle, the strike price of the put is lower than the strike price of the call, but the times to maturity are still the same.

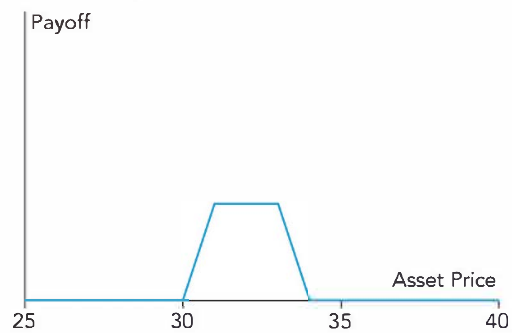
14.8 The payoff of a range forward contract is of the form:



Note that when the strike prices are identical, a range forward contract will be equivalent to a regular forward contract.

14.9 The trader can try a short butterfly spread, a short calendar spread, a strangle, or a straddle.

14.10 The payoff pattern is



Solved Problems

14.11 In a principal protected note, the price of a call option has to be greater than the difference between S and PV(S). As interest rates increase, the difference becomes greater. Thus, it is easier to create a principal protected note.

14.12 A box spread is created by buying a call with strike price K_1, selling a call with strike price K_2, buying a put with strike price K_2, and selling a put with strike price K_1. All have the same maturity and $K_1 < K_2$. Buying a call with strike price K_1 and selling a put with strike price K_1 is equivalent to a long forward contract with delivery price K_1. Selling a call with strike price K_2 and buying a put with strike price K_2 is equivalent to a short forward contract with delivery price K_2.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

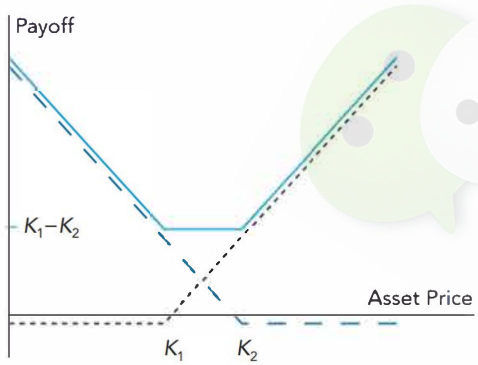
14.13 A butterfly spread is created by buying put options with strike prices of USD 45 and USD 55 and selling two put options with a strike price of USD 50. The cost is USD 1 (= 2 + 7 − (2 × 4)). A table showing the profit is as follows:

Asset Price Range	Payoff: Long 45 Put	Payoff: Two Short 50 Puts	Payoff: Long 55 Put	Profit Including the Cost of Setting Up
$S_T < 45$	$45 - S_T$	$-2(50 - S_T)$	$55 - S_T$	−1
$45 \le S_T < 50$	0	$-2(50 - S_T)$	$55 - S_T$	$S_T - 46$
$50 \le S_T < 55$	0	0	$55 - S_T$	$54 - S_T$
$S_T \ge 55$	0	0	0	−1

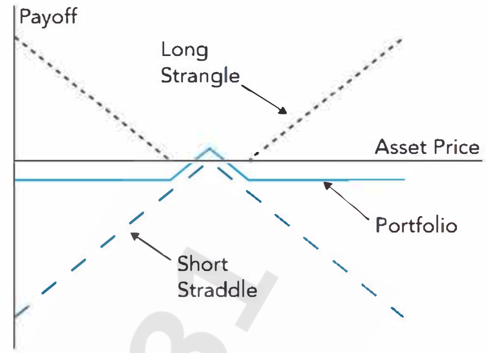
14.14 It costs USD 2.50 to set up the bear spread. If the asset price is between USD 27.50 and USD 30, the payoff from the spread will be less than USD 2.50. If it is less than USD 27.50, the payoff will be between USD 2.50 and USD 5. Thus, the trade will be profitable if the price of the asset at maturity is less than USD 27.50.

14.15 (a) A bear spread is created by buying the put option with a strike price of USD 35 and selling the put option with a strike price of USD 30. (b) A bull spread is created by buying the put option with a strike price of USD 30, selling the put option with a strike price of USD 35, and adding the present value of USD 5 to the portfolio.

14.16 This has the same payoff as a strangle except that there is an extra amount of cash equal to $K_1 - K_2$. It is more expensive to set up than a regular strangle where the put has a lower strike price than the call.



14.17 This is a butterfly spread minus some cash:



14.18 A strangle costs USD 5 and provides a profit as follows:

Asset Price Range	Long Call with Strike Price of 40	Long Put with Strike Price of 30	Profit (Including Initial Cost)
$S_T < 30$	0	$30 - S_T$	$25 - S_T$
$30 \le S_T < 40$	0	0	−5
$S_T \ge 40$	$S_T - 40$	0	$S_T - 45$

14.19 A straddle is created by buying the call and buying the put. The cost is USD 12. It leads to a profit if the asset price at maturity is above USD 82 or below USD 58.

14.20 The outcome of the takeover is uncertain, and we can expect a big move in the price of Company B's stock by the time the result of the takeover attempt is known. A straddle would therefore appear to be a good trade. However, because knowledge of the takeover is in the public domain, the prices of calls and puts will be higher than usual to reflect the fact that a big move in the stock price is likely. This emphasizes the point that having the same view as the rest of the market is not usually sufficient to make money trading options. It is necessary to take a view that is different from the market consensus (and be right!).

15

Exotic Options

Learning Objectives

After completing this reading, you should be able to:

- Define and contrast exotic derivatives and plain vanilla derivatives.
- Describe some of the reasons that drive the development of exotic derivative products.
- Explain how any derivative can be converted into a zero-cost product.
- Describe how standard American options can be transformed into nonstandard American options.
- Identify and describe the characteristics and payoff structures of the following exotic options: gap, forward start, compound, chooser, barrier, binary, lookback, Asian, exchange, and basket options.
- Describe and contrast volatility swaps and variance swaps.
- Explain the basic premise of static option replication and how it can be applied to hedging exotic options.

Standard European and American options (e.g., exchanged-traded options) are termed *plain vanilla options*. Options with non-standard properties are termed *exotic options* (or simply *exotics*). Exotic options are designed by derivatives dealers to meet the specific needs of their clients and are usually traded in the over-the-counter markets. Exotic options can be very profitable for derivatives dealers because they have relatively large bid-offer spreads.

Exotic options arise for several reasons. In some situations, they can provide more efficient hedging than plain vanilla options. Exotic options may also best reflect a firm's view on factors such as interest rates, exchange rates, and commodity prices. Occasionally, exotic options are used for tax or regulatory purposes.

15.1 EXOTICS INVOLVING A SINGLE ASSET

In this section, we describe some non-standard options that provide payoffs based on the price of a single asset.

Packages

A package is a portfolio consisting of plain vanilla options on an asset. We presented several packages in Chapter 14: bull spreads, bear spreads, butterfly spreads, calendar spreads, straddles, and strangles. Packages are sometimes regarded as exotic options because they are positions built to reflect a specific market view and risk tolerance.

For example, an investor who takes a long position in a butterfly spread is acting on his or her belief that the future asset price will be near the middle strike price. At the same time, the investor is building this position without taking on a great deal of risk. In contrast, an investor with a similar view of the market who chooses to sell a straddle or a strangle is taking on much more risk.

Zero-Cost Products

Any derivative product can be converted into a zero-cost product by arranging for it to be paid for in arrears.

Consider a derivative that matures at time T and has a premium equal to f. Rather than requiring that the premium be paid upfront (which is normally the case), the derivative can be structured so that the buyer instead pays $f(1 + R)^T$ at maturity (where R is the interest rate for maturity T).[1]

[1] The seller of the option is taking some credit risk and (assuming no collateral is posted) the interest rate R should be the interest rate that would be paid on a zero-coupon bond with maturity T by the option buyer.

For example, consider a European call option. When converted into a zero-cost product, it has a payoff:

$$\max(S_T - K - A, -A)$$

where $A = c(1 + R)^T$ and c is the normal option premium.

This is essentially is a forward contract where the holder agrees to buy the option payoff at maturity for A. One extension of this idea is to make the zero-cost product a futures contract (rather than a forward contract) on the option payoff. This is known as a *futures-style option*.

A futures-style call is marked to market in the same way as a regular futures contract with the final settlement of $\max(S_T - K, 0)$, where S_T is the final price of the underlying asset and K is the strike price. Futures-style options are offered by exchanges such as the CME Group and Eurex.

Futures-style options can be contrasted with the more usual *equity-style options*, which involve the buyer paying an upfront premium.

Non-Standard American Options

Exchange-traded American options can be exercised at any time at the pre-determined fixed strike price. Variations on this standard product are traded in the over-the-counter markets. The following are examples.[2]

- The exercise may be restricted to certain dates. Such an option is known as a *Bermudan* option. Interest rate options are sometimes Bermudan (e.g., an American-style bond option might only be exercisable on interest payment dates).
- There may be an initial *lock-out* period during which time the option cannot be exercised. As we saw in an earlier chapter, employee stock options usually have a lock-out period. Once the lock-out period is over, employee stock options are referred to as *vested*.
- The strike price may change during the life of the option. For example, when a corporate bond has a call option feature allowing the issuer to retire the bond early (i.e., buy it back from the holder), the strike price at which the issuer can exercise the option tends to decline as time passes.

Valuation and Risk Models explains how binomial trees can be used to value American options. This methodology can be adapted to accommodate all the non-standard features just mentioned.[3]

[2] Warrants (which were introduced in Chapter 12) sometimes have all the features mentioned.

[3] As we will see in *Valuation and Risk Models*, in the standard binomial tree set up there is a test for whether early exercise is optimal at each node of the tree. To value a non-standard American option, we test for the optimality of early exercise only at nodes where exercise of the option is allowed using the applicable strike price.

Forward Start Options

A forward start option is an option that will begin at a future time. It is usually stated that the option will be at-the-money at the time it starts. Employee stock options can be forward start options if an employer promises that they will be granted on future dates.

Gap Options

A gap option is a European call or put option where the price triggering a payoff is different from the price used in calculating the payoff.

Suppose the trigger price is K_2 and the price used in calculating the payoff is K_1. This means

- The payoff from a call option is $S_T - K_1$ if $S_T \geq K_2$.
- The payoff from a put option is $K_1 - S_T$ if $S_T \leq K_2$.

The payoffs from gap options are shown in Table 15.1 as well as in Figures 15.1 and 15.2. Note that the payoff from a gap option can be negative. For example, if $K_1 = 15$ and $K_2 = 10$, the payoff from a gap call option is negative when $10 \leq S_T < 15$. Similarly, if $K_1 = 15$ and $K_2 = 20$, the payoff from a gap put option is negative when $15 < S_T \leq 20$. In some circumstances, the probability of a negative payoff is sufficiently high such that the cost of a gap option is negative.

Table 15.1 Payoffs from Gap Call and Gap Put Options

Range for S_T	Payoff from Gap Call Option	Payoff from Gap Put Option
$S_T \leq K_2$	0	$K_1 - S_T$
$S_T > K_2$	$S_T - K_1$	0

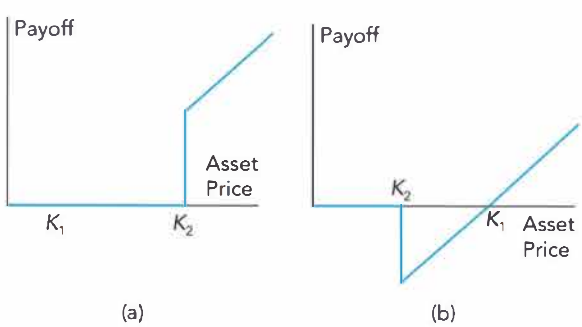


Figure 15.1 Payoff from a gap call option when (a) $K_2 > K_1$ and (b) $K_1 > K_2$.

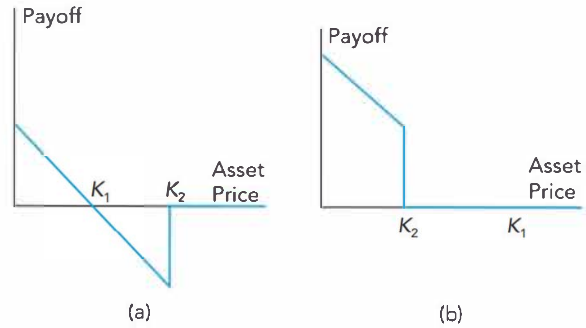


Figure 15.2 Payoff from a gap put option when (a) $K_2 > K_1$ and (b) $K_1 > K_2$.

Gap options can be used to describe some insurance contracts. For example, suppose an asset has been insured for USD 50,000. However, there is a deductible of USD 500 (i.e., the insured person must bear the first USD 500 of any loss in value). Suppose further that the insurance company incurs an administrative cost of USD 1,000 in assessing a claim. Without considering the administrative cost, the insurance company has in effect sold a put option on the value of the asset with a strike price of USD 49,500. When administrative costs are considered, the cost to the insurance company is

$$\begin{array}{ll} 1{,}000 + 49{,}500 - S_T & \text{when } S_T \leq 49{,}500, \text{ and} \\ 0 & \text{when } S_T > 49{,}500, \end{array}$$

where S_T is the value of the assets. The insurance company's position is equivalent to selling a gap put option where (using the previous notation) $K_2 = 49{,}500$ and $K_1 = 50{,}500$.

Cliquet Options

A cliquet option is a series of forward start options with certain rules for determining the strike prices. For example, a cliquet option might consist of five call options: a one-year option, a one-year option starting in one year, a one-year option starting in two years, a one-year option starting in three years, and a one-year option starting in four years. This is therefore a portfolio consisting of a regular one-year option plus four forward start options. A simple rule for the strike prices could be that each option is initially at-the-money.

Chooser Options

With a chooser option, the holder has a period of time (after purchasing the option) where he or she can choose whether it is a put option or a call option. For example, the holder of two-year European option might be allowed to choose whether it is a call option or a put option at the end of the first year.

One feature of chooser options is that they can be viewed as packages of call options and put options with different strike prices and times to maturity. To see this, suppose the time when the choice is made is T_1. Assume the value of the chooser option at this time is

$$\max(c, p)$$

where c is the price of a call option and p is the price of a put option. This can be rewritten as:

$$c + \max(0, p - c)$$

If both options are European options maturing at time T_2 (where $T_2 > T_1$) with strike price K, we know from put-call parity that if there is no income on the asset:

$$p - c = \mathrm{PV}(K) - S_1$$

where PV denotes present value from T_2 to T_1. This means that the value of the chooser option is

$$c + \max(0, PV(K) - S_1)$$

where S_1 is the asset price at time T_1. This shows that the option is a package consisting of:

- A call option with strike price K maturing at time T_2, and
- A put option with strike price PV(K) maturing at time T_1.

Binary Options

There are four types of binary options.

1. *Cash-or-nothing call:* This pays a fixed amount if the asset price is above the strike price at maturity and zero otherwise.
2. *Cash-or-nothing put:* This pays a fixed amount if the asset price is below the strike price at maturity and zero otherwise.
3. *Asset-or-nothing call:* This pays an amount equal to the asset price if it is above the strike price at maturity and zero otherwise.
4. *Asset-or-nothing put:* This pays an amount equal to the asset price if it is below the strike price at maturity and zero otherwise.

The payoffs are illustrated in Figures 15.3 and 15.4.

Traditional European options can be thought of as combinations of binary options.

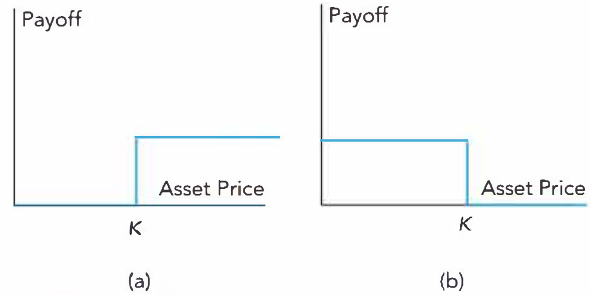


Figure 15.3 The payoff from (a) a cash-or-nothing call and (b) a cash-or-nothing put is shown. *K* is the strike price.

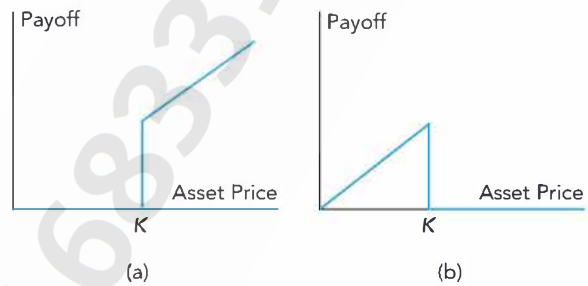


Figure 15.4 The payoff from (a) asset or nothing call and (b) an asset or nothing put is shown. *K* is the strike price.

A long position in a European call option is a combination of:

- A long position in an asset-or-nothing call, and
- A short position in a cash-or-nothing call with a payoff equal to the strike price.

Similarly, a long position in a European put option is

- A short position in an asset-or-nothing put, and
- A long position in a cash-or-nothing put with a payoff equal to the strike price.

Cash-or-nothing options are sometimes referred to as digital options.

Binary options, like gap options, have discontinuous payoffs. For example, suppose a cash-or-nothing call pays off USD 100,000 if the price of a stock at a certain time is above USD 29. If the price is USD 28.9, there is zero payoff; if it is USD 29.1, the payoff is USD 100,000. As a result, traders may have an incentive to engage in illegal price manipulation (especially if the underlying asset is thinly traded).

To see how this would work, suppose that there is only a little time until maturity and the price is just below USD 29. In this case, a trader with a long position in the option might enter buy orders to move the price above USD 29. Meanwhile, a trader that has sold the option might enter sell orders to try to keep the price below USD 29. Note that there is much greater incentive to manipulate the asset price compared to when the payoff is a continuous function of the underlying asset (as with a plain vanilla call or put).

Cash-or-nothing options are notorious for their association with fraudulent activities and are banned in several countries. This type of fraud can take several forms, but it often features firms that:[4]

1. Do not allow customers to withdraw funds from their accounts,
2. Make unauthorized use of personal data (e.g., credit cards, passports, and drivers licenses), and
3. Overstate the average expected return and then structure the product so that the expected return is negative.

Asian Options

Asian options provide a payoff dependent on an arithmetic average of the underlying asset price during the life of the option. The average price is usually calculated using periodic observations (e.g., at the end of each day).

Suppose K is the strike price, S_T is the final asset price, and S_{ave} is the average asset price. There are four types of Asian options.

1. *Average price calls:* These provide a payoff at maturity equal to $\max(S_{ave} - K, 0)$.
2. *Average price puts:* These provide a payoff at maturity equal to $\max(K - S_{ave}, 0)$.
3. *Average strike calls:* These provide a payoff at maturity equal to $\max(S_T - S_{ave}, 0)$.
4. *Average strike puts:* These provide a payoff at maturity equal to $\max(S_{ave} - S_T, 0)$.

An Asian option is less expensive than a regular option with the same strike price and can be more appropriate for hedging. For example, consider a company that repatriates its earnings in a foreign currency every week. Because its yearly profits depend on the average exchange rate during the year, an Asian put option on the value of the foreign currency can ensure that the average exchange rate will not be worse than a particular exchange rate (i.e., the option's strike price). This strategy is much less expensive than entering into 52 plain vanilla options (i.e., one option for every week of the year).

Lookback Options

The payoff from a lookback option depends on the maximum or minimum asset price reached during the life of the option. There are four types of lookback options.

1. A *floating lookback call* gives a payoff equal to the amount by which the final asset price exceeds the minimum asset price.
2. A *floating lookback put* gives a payoff equal to the amount by which the maximum asset price exceeds the final asset price.
3. A *fixed lookback call* gives a payoff equal to $\max(S_{max} - K, 0)$, where S_{max} is the maximum asset price in a given time frame and K is the strike price.
4. A *fixed lookback put* gives a payoff equal to $\max(K - S_{min}, 0)$, where S_{min} is the minimum asset price in a given time frame and K is the strike price.

Lookback options are more expensive than regular options. To illustrate this, consider options on a non-dividend paying stock when the stock price is USD 30, the strike price is USD 30, the risk-free rate is 4%, the volatility is 20%, and the time to maturity is one year.

Using the Black-Scholes Merton assumptions (as discussed in Chapter 15 of *Valuation and Risk Models*), the value of a European call is USD 2.98 and that of a European put is USD 1.80. With the same assumptions, a one-year fixed lookback call and put with the same strike price are worth USD 5.61 and USD 3.85 (respectively). A one-year floating lookback call and put are worth USD 5.03 and USD 4.44 (also respectively).

The value of a lookback option depends on how often the price of the underlying asset is observed. (These observations are used to calculate the maximum or minimum price.) Specifically, a lookback option increases in value as the observation frequency is increased.[5]

Lookback options can make an investor appear prescient. For example, a floating lookback call provides a payoff that makes it look as though an asset purchased at the end of a period was purchased at the minimum price during the period. Meanwhile, a floating lookback put provides a payoff that makes it look as though an asset sold at the end of a period was sold at the maximum price during the period. A fixed lookback call or put is akin to an American option where an investor can use hindsight to choose the best possible exercise date.

[4] Securities and Exchange Commission. (n.d.). Binary options. Retrieved from https://www.investor.gov/additional-resources/general-resources/glossary/binary-options

[5] The results just given assume that the asset price is observed continuously.

Barrier Options

Barrier options have payoffs that depend on whether the asset price reaches a particular barrier. There are four types of barrier options.

1. *Down-and-out:* This is a European (call or put) option that ceases to exist if the asset price moves down from its initial level to the barrier level during the life of the option.
2. *Down-and-in:* This is a European (call or put) option that comes into existence if the asset price moves down from its initial level to the barrier level during the life of the option.
3. *Up-and-out:* This is a European (call or put) option that ceases to exist if the asset price moves up from its initial level to the barrier level during the life of the option.
4. *Up-and-in:* This is a European (call or put) option that comes into existence if the asset price moves up from its initial level to the barrier level during the life of the option.

Options that cease to exist when a barrier is reached are sometimes referred to as knock-out options, whereas options that come into existence when a barrier is reached are referred to as knock-in options.

Barrier options can be attractive to market participants because they are less expensive than regular options. For example, suppose an asset price is currently USD 30 and a trader wants a one-year call option with strike price USD 32. Furthermore, suppose that the trader considers it unlikely that the price will fall to USD 27 during the year. In this case. he or she could be tempted to buy a down-and-out call with a strike price of USD 32 and a barrier at USD 27. This call will provide the same payoff as a (more expensive) regular call with a strike price of USD 32 as long as the stock price does not fall below USD 27.[6]

Barrier options have some interesting properties. As with binary options and gap options, their payoff is discontinuous. Consider an up-and-out call with a strike price of USD 50 and a barrier at USD 60. If the stock price stays below the barrier and reaches USD 59.90 at maturity, the payoff is USD 9.90. If the stock price is 0.1 higher, the payoff is zero.

Normally, option prices increase as volatility increases. However, this is not necessarily the case with barrier options. For example, when a knock-out option is close to the barrier, an increase in the volatility may lower the price because it increases the probability that the barrier will be hit.

[6] Using Black and Scholes' assumptions, if volatility is 20% and the risk-free rate is 2%, with no dividends, the value of the regular option is USD 1.82. When there is a down-and-out barrier at USD 27 the price reduces to USD 1.56.

Parisian options are similar to standard barrier options, but they have one important difference: The asset price must remain above or below the barrier for a specified number of days before the option is knocked in or out. Suppose the number of days specified is ten. In one type of Parisian option, the ten days must be consecutive. In another type, they can be any ten days during the life of the option.

Compound Options

A compound option is an option on another option. Thus, there are two strike prices and two maturity dates.

Suppose the maturity dates are T_1 and T_2 (with $T_2 > T_1$) and the strike prices corresponding to those maturity dates are K_1 and K_2 (respectively). Four types of compound options are as follows.

1. *Call option on call option:* The holder has the right to pay K_1 at time T_1 in order to obtain a long position in a call option on an asset. This call option allows the asset to be bought for K_2 at time T_2.
2. *Put option on call option:* The holder has the right to receive K_1 at time T_1 and obtain a short position in a call option on an asset. This call option allows the asset to be bought for K_2 at time T_2.
3. *Call option on put option:* The holder has the right to pay K_1 at time T_1 in order to obtain a long position in a put option on an asset. This put option allows the asset to be sold for K_2 at time T_2.
4. *Put option on put option:* The holder has the right to receive K_1 at time T_1 and obtain a short position in a put option on an asset. This put option allows the asset to be sold for K_2 at time T_2.

Compound options are attractive to traders who want more leverage than that provided by plain vanilla options. For example, suppose that the current price of a non-dividend paying stock is USD 50, its volatility is 20% per year, and the risk-free rate is 2% per year. The Black-Scholes Merton assumptions (covered in *Valuation and Risk Models*) show that a one-year call option on the stock with a strike price of USD 55 should cost USD 2.47. A compound option could be structured as follows.

1. At the six-month point, the investor has the right to pay USD 3 for a call option.
2. The call option, if purchased, gives the investor the right to buy the stock for USD 55 at the one-year point.

The Black-Scholes Merton assumptions give the value of this option as USD 0.99.

Under what circumstances will the call option be purchased at the six-month point in this example? Assuming no changes in the volatility or the risk-free rate, the Black-Scholes Merton model shows that the value of the call option is greater than USD 3 if the stock price at the six-month point is greater than USD 54.30. The call option will therefore be purchased if the stock price is greater than USD 54.30.

Compared to regular options prices, the prices of compound options are more sensitive to changes in volatility. For example, if the volatility increases from 20% per year to 30% per year, the Black-Scholes Merton assumptions show that the price of the compound option in the previous example increases by 160% (from USD 0.99 to USD 2.57). Meanwhile, using the same assumptions, the price of the regular option increases by only 79% (from USD 2.47 to USD 4.43).

15.2 EXOTICS INVOLVING MULTIPLE ASSETS

We now move on to consider options with payoffs that depend on the prices of two or more assets.

Asset-Exchange Options

In an asset-exchange option, the holder has the right to exchange one asset for another. Asset-exchange options can arise in several ways. From the perspective of a U.S. investor, an option to exchange X euros for Y Australian dollars is an option to exchange one asset for another. An offer by Company X to acquire Company Y through the exchange of a certain number of its own shares for shares of Company Y is another type of asset-exchange option.

Asset-exchange options are closely related to options where a trader will receive the more valuable of two assets (i.e., Asset A and Asset B). This is because the trader's position in the latter option can be regarded as a position in Asset A combined with an option to exchange Asset A for Asset B. Similarly, it can be regarded as a position in Asset B combined with an option to exchange Asset B for asset A.

Basket Options

A basket option is an option on a portfolio of assets. The portfolio can contain assets such as stocks, stock indices, and currencies. Basket options can be appropriate hedging instruments for firms seeking to reduce costs by hedging their aggregate exposure to several assets with a single trade.

Basket options are dependent on the correlation between the returns from the assets in the basket. For example, a call option on the future value of a portfolio consisting of ten assets with highly correlated returns is more expensive than a similar call option where the asset returns are uncorrelated (due to portfolio diversification effects).

15.3 EXOTICS DEPENDENT ON VOLATILITY

Some exotic options are dependent on the volatility of an asset price, rather than the asset price itself. An asset's volatility per day is generally measured as the standard deviation of its daily returns. The annualized volatility is usually calculated by multiplying the daily volatility by the square root of 252.[7]

Volatility Swap

A volatility swap is appropriate for a trader who wants to take a position dependent only on volatility. While plain vanilla options provide an exposure to volatility, they also depend on the price of the underlying asset. A potential advantage of volatility swaps is that their payoffs depend solely on realized volatility.

A volatility swap is a forward contract on the realized volatility of an asset during a certain period. A trader agrees to exchange a pre-specified volatility for the realized volatility at the end of the period, with both volatilities being multiplied by a certain amount of principal.

Suppose that L_{vol} is the principal, σ_K is the pre-specified yearly volatility, and σ is the realized yearly volatility (calculated as described above). A swap where the trader *pays fixed* provides a payoff at the end of the period equal to:

$$L_{vol}(\sigma - \sigma_K)$$

Similarly, a volatility swap where the trader *receives fixed* provides a payoff at the end of the period equal to:

$$L_{vol}(\sigma_K - \sigma)$$

Suppose that the principal (L_{vol}) for a pay-fixed volatility swap on the S&P 500 during the next three months is USD 1 million and σ_K is 15% per year. Suppose further that the realized volatility of the S&P 500 during the three-month period is 1% per day and thus the yearly volatility is 15.87% $(= 1\% \times \sqrt{252} = 15.87\%)$. The payoff from the volatility swap would then be USD 8,700 $(= (0.1587 - 0.15) \times \text{USD } 1{,}000{,}000)$. The corresponding payoff from a receive-fixed volatility swap is USD $-8{,}700$.

[7] 252 is an estimate of the number of trading days in a year. It is assumed (as will be discussed in *Valuation and Risk Models*) that the volatility of an asset price during a period is dependent on the number of trading days during the period, rather than the number of calendar days during the period.

Variance Swap

The payoff from a variance swap is calculated analogously to the payoff from a volatility swap. As a reminder, the variance rate for an asset is the square of its volatility. Thus, the payoff for a pay-fixed variance swap is

$$L_{var}(\sigma^2 - V_K)$$

where L_{var} is the principal for the variance swap and V_K is the specified fixed variance. The payoff for a receive-fixed variance swap is similarly:

$$L_{var}(V_K - \sigma^2)$$

Sometimes L_{vol} and σ_K are specified in a variance swap with the understanding[8] that $V_K = \sigma_K^2$ and $L_{var} = \frac{L_{vol}}{(2\sigma_K)}$.

15.4 HEDGING EXOTICS

Compared to regular options, some exotics are easier to hedge, while some are more difficult. For example, an average price option is easier to hedge than a regular option because as its maturity approaches, more of the asset prices used to calculate S_{ave} have been observed and thus the payoff becomes progressively more certain. The value of the option therefore has very little sensitivity to asset price movements occurring late in its life.

By contrast, barrier options can be quite difficult to hedge because when the asset price is close to the barrier, uncertainty about the payoff *increases*. Traders have developed a hedging procedure known as *static options replication* for dealing with such exotic options.[9] The key principle underlying the procedure is that two portfolios that are worth the same on some boundary, defined as a function of asset price and time, must also be worth the same at all interior points (i.e., points that could be reached before the boundary is reached). Therefore, if we can find a portfolio of plain vanilla options worth approximately the same as an exotic option on a boundary, the exotic option can be hedged by shorting the portfolio.

To illustrate the nature of static options replication, consider a one-year up-and-out call option. The strike price is USD 30, the barrier is USD 50, and the current stock price is USD 25. The natural boundary to use is the one shown in Figure 15.5.

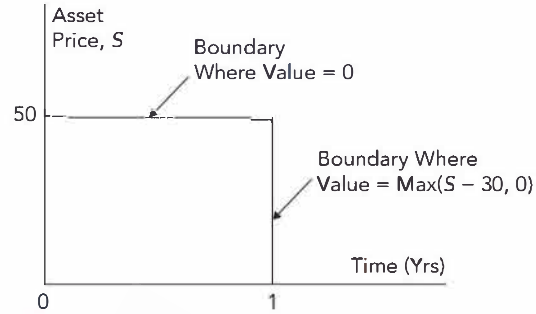


Figure 15.5 Possible boundary for an up-and-out call option when initial asset price is 30 and barrier is 50.

The up-and-out option is worth zero if the horizontal boundary is reached and max(S − 30, 0) if the vertical boundary is reached. The task in static options replication is to choose a portfolio of plain vanilla options that match the payoff as closely as possible on the boundary.

The natural first option to use is a one-year European call option with a strike price of 30. This matches the payoff on the vertical part of the barrier. To match payoffs on the horizontal part of the barrier, choose N equally spaced points on the horizontal barrier and an additional N European options with maturities of less than one year[10] and value each option.[11] Then, solve a set of simultaneous equations to ensure that the portfolio formed from the all the options has a value of zero at each of the N points on the horizontal boundary.

To express this algebraically, suppose c_j is the value of a one-year European call option with a strike price of USD 30 at the jth point on the horizontal boundary ($1 \le j \le N$) and that f_{ij} is the value of the ith additional option at the jth point on the horizontal boundary ($1 \le i,\ j \le N$). Define a_i as the size of the position in the ith additional option. The a_i are determined by solving the N simultaneous equations:

$$c_j + \sum_{i=1}^{N} a_i f_{ij} = 0 \qquad (1 \le j \le N)$$

As N is increased, the horizontal boundary is matched more precisely and the hedge created by shorting the portfolio improves.

A static options replication hedge can be left unchanged until the boundary is reached. The trader must then unwind the hedge portfolio and create a new hedge.

[8] From calculus, the change in σ^2 is approximately equal to 2σ times the change in σ. This explains why the value of L_{var} that corresponds to L_{vol} is approximately given by $L_{var} = L_{vol}/(2\sigma_K)$.

[9] See E. Derman, D. Ergener, and I. Kani, "Static Options Replication," *Journal of Derivatives*, 2, 4 (Summer 1995): 78–95.

[10] The options must have a maturity of less than one year so that they do not affect the value of the portfolio on the vertical part of the boundary. We could choose call options with strike price 50 and maturities 1/(N + 1), 2/(N + 1), . . . , N/(N + 1) years.

[11] If the option expires before a time corresponding to a barrier point, the value of the option at the barrier point is zero.

SUMMARY

Derivatives dealers have been very creative in designing exotic options. And while this chapter has covered some of the most common exotic options, it has not mentioned other derivatives with intriguing names such as shout options, Madonna options, Himalaya options, and pyramid options.

Some options (such as Asian options and basket options) can be efficient hedging instruments. However, it is not entirely clear why some of the other options we have mentioned in this chapter are used by market participants. Some traders with a particular view of the market may find a certain type of exotic option attractive. For example, if a short-term spike in an asset price is expected for some reason, a long position in a lookback put position could be considered. If a trader wants to speculate on an increase in a stock price with a highly levered transaction, a compound option might be a reasonable choice. If a trader wants to reduce the cost of a regular call option and considers it unlikely that the asset price will fall to a certain level, a down-and-out call option could be appropriate.

The adage "you get what you pay for" applies to options. However, exotic options generally are less competitively priced than regular options, so in their case it might be more accurate to say, "you get less than you pay for." Traders should bear this in mind and avoid being overly optimistic in assessing potential returns from exotic options.

In addition to trading exotics (whose future payoff depends on the price of an asset), investors can trade instruments that depend on an asset's future volatility. A volatility swap provides a payoff proportional to the difference between a realized volatility and a pre-specified volatility. A variance swap provides a payoff proportional to the difference between a realized variance rate and a pre-specified variance rate.

Some exotic options are easier to hedge than plain vanilla options, while the opposite is true for other exotics. A procedure known as static options replication can be used as an alternative to the Greek letter hedging procedures discussed in *Valuation and Risk Models*. This involves shorting a portfolio of regular options designed to replicate the value of an exotic option on a boundary.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

15.1 What is a Bermudan option?

15.2 Explain how a futures-style option works.

15.3 What is a gap option?

15.4 Explain how (a) a chooser option and (b) a forward-start option work.

15.5 List four types of compound options.

15.6 How is a floating lookback put defined? Does it become more or less valuable as the number of asset price observations (used to calculate the payoff) increases?

15.7 List four types of binary options.

15.8 Does a basket option become more or less valuable as the correlation between the asset returns in the basket increases?

15.9 How is a pay-fixed volatility swap defined?

15.10 Why are average price Asian options easier to hedge than plain vanilla options?

Practice Questions

15.11 What is equivalent to a portfolio consisting of an up-and-in put option and an up-and-out put option where the barrier, strike price, and time to maturity are the same for the two options?

15.12 What is the put-call parity relation between the following compound options: call on a call and a put on a call?

15.13 Does a down-and-out put option become more or less valuable as we increase the frequency with which we observe the asset prices in determining whether the barrier has been hit?

15.14 What is the payoff from a portfolio consisting of a long position in both a floating lookback call and a floating lookback put?

15.15 Derive expressions for the payoffs from a:

- **a.** Long position in an average price call and short position in an average price put,
- **b.** A long position in an average strike call and short position in an average strike put, and
- **c.** A long position in a plain vanilla European call and short position in a plain vanilla European put.
- **d.** All options have the same strike price and time to maturity. Use the results to derive a relationship between the prices of the six options you have considered.

15.16 Explain why a gap call option is a regular call option plus a binary option when $K_2 > K_1$, using the same notation as in the chapter

15.17 Provide an alternative decomposition of the chooser option to that given in the chapter so that it is a call maturing at time T_1 plus a put maturing at time T_2.

15.18 Is a one-year at-the-money call option on a basket of three stocks more or less valuable than a portfolio consisting of three one-year at-the-money options, one on each stock?

15.19 What view of the market would a trader have if he or she chose a receive-fixed variance swap?

15.20 What is the key theoretical result underlying static options replication?

ANSWERS

Short Concept Questions

15.1 A Bermudan option is an option that can be exercised only at certain pre-specified discrete times.

15.2 A futures-style option is a futures contract on the option's payoff. The final settlement of the futures contract equals this payoff.

15.3 A gap option is a European call or put option where the trigger determining whether there will be a payoff is different from the strike price used to calculate the payoff.

15.4 A chooser option is an option with a certain strike price and time to maturity where at some stage during the option's life the holder will choose whether it is a call or a put. A forward start option is an option that will come into existence at some future time.

15.5 The four types of compound options are call on a call, call on a put, put on a call, and put on a put.

15.6 A floating lookback put is an option that pays off the amount by which the maximum asset price during the option's life exceeds the final asset price. It becomes more valuable as the asset price is observed more frequently in the calculation of the payoff.

15.7 The four types of binary options are cash-or-nothing call, cash-or-nothing put, asset-or-nothing call, and asset-or-nothing put.

15.8 As the correlation between the returns increases, a basket option becomes more valuable because large movements in the value of the portfolio become more likely.

15.9 A pay-fixed volatility swap provides a payoff equal to $L_{vol}(\sigma - \sigma_K)$ where σ is the realized volatility per annum during a certain period, σ_K is a pre-specified volatility, and L_{vol} is a pre-specified principal amount.

15.10 An average price Asian option depends on the average asset price during the life of the option. As the end of the option's life approaches, the average becomes increasingly more certain so that there is very little risk to be hedged. The same is not true of a plain vanilla option.

Solved Problems

15.11 The portfolio is equivalent to a plain vanilla put option. If the barrier is hit, the up-and-in option provides the plain vanilla option payoff and the up-and-out option provides no payoff. If the barrier is not hit, the up-and-out option provides the plain vanilla option payoff and the up-and-in option provides no payoff.

15.12 We use the notation in the chapter and assume that the K_1, K_2, T_1, and T_2 are the same for both the call on the call and the put on the call. A long position in the call on a call plus the present value of K_1 (with discounting from T_1 to today) equals a long position in a put on the call plus the value of a European call maturing at time T_2. If c_1 is the value at time T_1 of a European call maturing at time T_2, the result can be seen from:

$$\max(c_1 - K_1, 0) + K_1 = \max(K_1 - c_1, 0) + c_1 = \max(c_1, K_1)$$

15.13 A down-and-out put becomes less valuable because the chance of the barrier being hit, so that the option is knocked out, is greater.

15.14 A floating lookback call provides a payoff of $S_T - S_{min}$ and a floating lookback put provides a payoff of $S_{max} - S_T$. The payoff from the portfolio is therefore the excess of the maximum asset price over the minimum asset price $(S_T - S_{min} + S_{max} - S_T = S_{max} - S_{min})$.

15.15 (a) A long average price call gives a payoff of $\max(S_{ave} - K, 0)$. A short average price put gives a payoff of $-\max(K - S_{ave}, 0)$. The payoff in (a) is therefore always $S_{ave} - K$ whether $S_{ave} > K$ or $S_{ave} \leq K$. Similarly, the payoff in (b) is always $S_T - S_{ave}$ and the payoff in (c) is always $S_T - K$. From this, it follows that:

$$(c_1 - p_1) + (c_2 - p_2) = (c - p)$$

where c_1 and p_1 are the prices of the average price call and put, c_2 and p_2 are the prices of the average strike call and put, and c and p are the prices of the plain vanilla call and put.

15.16 Consider a gap option where the trigger price is K_2 and the strike price for determining payoffs is K_1. The gap call option is a plain vanilla option with strike price K_2 plus a cash-or-nothing binary option that pays off $K_2 - K_1$ if the asset price is above K_2.

15.17 With the notation in the text we can write

$$\begin{aligned} max(c,\ p) &= p + max(c - p, 0) \\ &= p + max(S - PV(K), 0) \end{aligned}$$

This shows that the chooser option is a portfolio consisting of:

- A put option maturing at time T_2, and
- A call option with strike price PV(K) maturing at time T_1.

15.18 It is less valuable because the stocks are less than perfectly correlated.

15.19 Because the payoff is $L_{var}(V_K - \sigma^2)$ the trader would be taking the view that the realized variance rate (σ^2) will be less than the pre-specified variance rate (V_K) in the future period that is considered. The variance rate is the square of the volatility. The trader is therefore expecting a low volatility.

15.20 If two portfolios have the same value on some boundary in {S, t} space they must have the same value at interior points (i.e., all points that could be reached prior to the boundary being reached).

16

Properties of Interest Rates

Learning Objectives

After completing this reading, you should be able to:

- Describe Treasury rates, Libor, Secured Overnight Financing Rate (SOFR), and repo rates, and explain what is meant by the "risk-free" rate.
- Calculate the value of an investment using different compounding frequencies.
- Convert interest rates based on different compounding frequencies.
- Calculate the theoretical price of a bond using spot rates.
- Calculate the Macaulay duration, modified duration, and dollar duration of a bond.
- Evaluate the limitations of duration and explain how convexity addresses some of them.
- Calculate the change in a bond's price given its duration, its convexity, and a change in interest rates.
- Derive forward interest rates from a set of spot rates.
- Derive the value of the cash flows from a forward rate agreement (FRA).
- Calculate zero-coupon rates using the bootstrap method.
- Compare and contrast the major theories of the term structure of interest rates.

Interest rates are central to finance. This chapter discusses different types of interest rates and how they are measured, analyzed, and used.

The *interest rate term structure* describes how interest rates vary depending on their maturity. In an upward-sloping term structure, long-term interest rates are higher than short-term interest rates. In a downward-sloping term structure, the opposite is true.

Chapters 17 and 18 will discuss products that rely on interest rates. Chapter 19 will cover interest rate futures and day count conventions. Chapter 20 will explain interest rate swap markets. Some of the topics introduced in this chapter will be covered further in Chapters 9 to 13 of *Valuation and Risk Models*.

16.1 TYPES OF INTEREST RATES

An interest rate is the return earned by a lender when advancing funds to a borrower. A major determinant of an interest rate is credit risk. This is the risk that the borrower will default and not repay the lender in full. As the credit risk increases, the interest rate required by the lender from the borrower also increases. Liquidity (i.e., the ease with which an interest-bearing instrument can be sold from one investor to another at a competitive price) is also a factor in determining interest rates.

Interest rates, and the spreads between different interest rates, are often expressed in basis points. One basis point is 0.01%, so that an interest rate of 2% is 200 basis points.

Government Borrowing Rates

One important interest rate is that paid by a government on its borrowings in its own currency. In the United States, this is referred to as the *Treasury rate*. Note that it is considered highly unlikely that the government of a developed country will default on debt issued in its own currency. This is because a government can always increase the supply of the currency to meet its obligations.[1]

As a result, the borrowings of the governments of developed countries are considered to be risk-free. Thus, the rates of interest on these borrowings are generally below those on other borrowings in the same currency.

Overnight Interbank Borrowing

In many countries, banks are required to keep cash (known as reserves) with the central bank. This reserve requirement depends on outstanding liabilities. At the end of each day, some banks have excess reserves while others have a shortage. This leads to a highly liquid market where overnight interbank borrowing and lending takes place.

In the United States, the interest rate on this overnight lending is referred to as the *federal funds rate*. The weighted average of the rates in these transactions is termed the *effective federal funds rate*. This rate is monitored by the Federal Reserve, which periodically intervenes with its own trades to increase or decrease the rate. Other countries have similar procedures to those in the United States. For example, the sterling overnight index average (SONIA) is the average overnight rate in the United Kingdom; ESTER is the euro overnight rate[2]; in Japan, TONAR is the Tokyo overnight average; and so on.

Repo Rates

In a repo agreement, securities are sold by Party A to Party B for a certain price with the intention of being repurchased at a higher price by Party A from Party B at a later time.

Suppose that the initial price paid by B to A for the securities is X, and the price at which they will be repurchased is $X + e$. The effect of the repo is that Party B has lent X to Party A for an amount of interest equal to e.

If Party A fails to repurchase the securities as agreed, Party B simply keeps the securities. This means that Party B takes very little risk, as long as (a) the value of the securities equals (or is very close to) X and (b) this value is fairly stable. Note that Party B is in a better position than it would be if the securities had been merely pledged as collateral. This is because the repo leads to Party B actually owning the securities, which avoids the need for it to initiate legal action to gain possession of the securities. Overnight repos (lasting one day) are the most common, but longer-term repos exist as well. An index of overnight repo rates in the United States is the Secured Overnight Financing Rate (SOFR). SARON is an index of overnight repo rates in the Swiss market.

---

[1] It is not unknown for a government of a developing country to default on debt issued in its own currency, as is explained in Chapter 5 of Valuation and Risk Models. The countries in the European Union that use the euro are in the situation where they do not have control over the money supply of the currency they use and may therefore default.

[2] ESTER replaces the previous euro overnight benchmark, EONIA.

Reference Rates

Many financial contracts involve interest rates. Sometimes rates are determined in advance. For example, a company might agree to borrow at a fixed rate of 4% for five years. However, it is often the case that the rate in a financial contract will vary as conditions in the market change. The financial contract then specifies a reference rate and states how the applicable rate will be determined by the reference rate. In the past, Libor has been a very popular reference rate. For example, prior to the discontinuation of Libor, a borrowing rate might be expressed as three-month Libor plus 80 basis points. This would mean that the rate applicable to future three-month periods would be the Libor rate observed in the market at the beginning of the three-month period plus 0.8%.

As we explain in what follows, Libor has now been phased out as a reference rate and new reference rates have been developed to replace it.

Libor

LIBOR is an acronym for London Interbank Offered Rate. Between 2014 and 2021, Libor interest rates were compiled each day by the Intercontinental Exchange (ICE). It asked 16 global banks to provide interest rate estimates applicable to unsecured borrowing from another bank just prior to 11 a.m. (U.K. time).[3] Estimates were made for five different currencies (USD, GBP, EUR, CHF, and JPY) and seven different borrowing periods (overnight, one week, one month, two months, three months, six months, and 12 months). The highest and lowest four quotes for each estimate were discarded, and the remainder were averaged to determine Libor fixings for the day. The banks submitting estimates typically had an AA credit rating.[4] A Libor rate was therefore usually assumed to be an estimate of the unsecured borrowing rate for an AA-rated bank in a certain currency for a certain period.

Libor rates have historically served as reference rates for hundreds of trillions of dollars of transactions throughout the world. As already indicated, the interest rate in a five-year financial contract might be expressed as three-month Libor plus 80 basis points. The value of three-month Libor would then be noted at the beginning of each three-month period during the five years, and the applicable rates would be based on that value of Libor plus 0.8%. For example, if the three-month Libor rate at the beginning of a three-month period was 3%, a rate of 3.8% would be used for the three-month period.

Problems with the use of Libor as a benchmark began to emerge in 2011. Banks were unable to base their Libor quotes on actual transactions because there were not enough of these. As a result, their quotes had to be based on what they thought the borrowing rate would be if they did borrow a particular currency for a particular period of time from another bank.

In 2012, an international investigation revealed that a number of large banks had manipulated their Libor estimates and that there had been some collusion between banks. There are two reasons why it was tempting for banks to manipulate their Libor estimates. One is to make contracts that were dependent on the day's Libor fixing more profitable. For example, if a bank had contracts where it was due to receive (pay) six-month Libor plus a spread over the next six months with Libor being based on a particular day's Libor fixings, it would have an incentive to provide a high (low) quote on that day. Another reason is that it might want to signal to the market that it is creditworthy and submit a low quote. The practice of removing the highest and lowest quotes and then averaging the rest prevented manipulation to some extent. However, it appears that this did not completely solve the problem because there was some collusion between banks.

As the fraudulent behavior came to light, regulators throughout the world levied billions of dollars in fines on banks. Many lawsuits were filed. Traders and others involved in providing fraudulent quotes lost their jobs, and some served jail time.

Regulators decided that Libor should be discontinued and replaced by rates based on actual transactions as follows:

- For the US dollar (USD): SOFR
- For the British pound (GBP): SONIA
- For the Swiss franc (CHF): SARON
- For the euro (EUR): ESTER
- For the Japanese yen (JPY): TONAR

These rates were all mentioned earlier. SOFR and SARON are calculated from overnight repo rates and are therefore secured lending rates. SONIA, ESTER, and TONAR are calculated from unsecured overnight borrowing rate between banks.

[3] Prior to 2014, Libor rates were compiled in a similar way by the British Bankers' Association.

[4] Credit ratings are discussed further in Chapter 4 of Valuation and Risk Models.

All the rates are subject to virtually no risk and are referred to as risk-free rates (RFRs).

At this stage, we explain one important difference between Libor and the overnight (RFR) reference rates that will replace it. Libor rates are known at the beginning of the period they apply to. However, the rate implied by overnight rates for a period such as one month or three months is known only at the end of the period. Libor is a forward-looking rate. The rates derived from overnight rates are backward-looking rates.

The rates implied by overnight rates for a period such as three months can be calculated by observing the overnight rates each business day during the period and compounding them. Suppose that there are n business days during the period, the overnight rate on the ith of these business days is r_i, and the rate applies to d_i calendar days. The rate for the period can be calculated at the end of the period as:[5]

$$R = [(1 + r_1\hat{d}_1)(1 + r_2\hat{d}_2)\ldots..(1 + r_n\hat{d}_n) - 1] \times \frac{360}{D}$$

where $\hat{d}_i = d_i/360$ and D is the total number of days in the period so that

$$D = \sum_i d_i$$

For most i, $d_i = 1$, but weekends and holidays lead to the overnight rate being applied to more than one day. For example, on a Friday d_i will normally be equal to 3.

Credit Sensitive Rates

Libor was a credit-sensitive reference rate. The spread between three-month USD Libor and a three-month rate based on overnight rates was normally about ten basis points, but it can be much higher in stressed market conditions. For example, it spiked to an all-time high of 364 basis points (3.64%) in October 2008 during the financial crisis. The new (RFR) reference rates are not credit sensitive and, as already mentioned, are backward looking. This has led the market to look for a forward-looking credit sensitive reference rate to supplement SOFR. Several have been suggested. One is the Bloomberg Short-Term Bank Yield Index (BSBY). This is a proprietary index calculated daily and published at 8:00 a.m. (ET). It seeks to measure the average yields at which large global banks access unsecured wholesale funding by observing actual transactions. It uses a curve fitting procedure to provide overnight, one-month, three-month, six-month, and 12-month maturities. Another proposed credit sensitive index is AMERIBOR, which was created by the American Financial Exchange. This reflects the actual borrowing costs of thousands of community and regional banks throughout the United States. How widely alternative reference rates such as BSBY and AMERIBOR will be used remains to be seen.

[5] We cover the compounding of interest rates later in this chapter. The equation given is not the only one that can be used. One alternative is to assume a gap between a daily interest rate being observed and the day to which it is applied.

Derivatives

The development of an active derivatives market dependent on overnight rates is important. It enables a complete term structure of interest, based on overnight (RFRs), to be determined. This in turn can be used to determine forward-looking term rates based on overnight RFRs.

We will discuss interest rate derivatives in Chapters 19 and 20. The most successful exchange-traded interest rate derivative used to be the Eurodollar futures contract. This provided a payoff based on three-month USD Libor fixings. The CME group now provides a similar contract based on SOFR. As we will see in Chapter 19, this is proving to be popular.

The most common over-the-counter interest rate derivative has historically been an interest rate swap where Libor is exchanged for a fixed rate. We discuss the contracts that have replaced it in Chapter 20.

16.2 COMPOUNDING FREQUENCY

In order to fully understand a quoted interest rate, it is necessary to know the compounding frequency with which it is measured. The compounding frequency defines the extent to which interest is earned on interest. Up to now, this book has usually assumed annual compounding (i.e., compounding that occurs once per year).

For example, suppose USD 100 is invested at 5% for five years and that the interest rate is expressed with annual compounding. This means that this USD 100 grows to:

$$100 \times 1.05 = 105$$

at the end of one year. Funds are then reinvested so that it becomes

$$105 \times 1.05 = 110.25$$

at the end of two years. Further reinvestment means that it becomes

$$110.25 \times 1.05 = 115.76$$

at the end of three years. At the end of year five, the investor will receive

$$100 \times 1.05^5 = 127.63$$

Now suppose that the 5% interest rate is expressed with semi-annual compounding. This means that half of 5% (i.e., 2.5%) is earned every six months with the funds being reinvested. By the end of six months, USD 100 grows to:

$$100 \times 1.025 = 102.5$$

This is then reinvested for a further six months to become

$$102.5 \times 1.025 = 105.12$$

This is then reinvested for a further six months to become

$$105.12 \times 1.025 = 107.75$$

and so on.

At the end of five years, the investment becomes

$$100 \times 1.025^{10} = 128.01$$

From an investor's perspective, it is better that funds be invested at 5% with semi-annual compounding than at 5% with annual compounding (because it is better to receive USD 128.01 than to receive USD 127.63). The opposite is true from a borrower's perspective (because it is better to pay USD 127.63 than to pay USD 128.01).

Quarterly compounding means that reinvestment is assumed to take place every three months. If the interest rate was 5% with quarterly compounding, 1.25% (= 0.05/4) would be earned each quarter and USD 100 would grow to:

$$100\left(1 + \frac{0.05}{4}\right)^{20} = 128.20$$

at the end of five years. Monthly, weekly, and even daily compounding frequencies are defined similarly. Table 16.1 shows how the compounding frequency affects the value to which USD 100 grows to at the end of five years when it is invested at 5% per annum.

The general formula for the future value of an amount A when it is invested at rate R for T years when R is expressed with a compounding frequency of m times per year is

$$\text{Future Value} = A\left(1 + \frac{R}{m}\right)^{mT} \qquad (16.1)$$

Table 16.1 Value to Which USD 100 Grows at 5% as the Compounding Frequency is Varied

Compounding Frequency	Value of USD 100 at the End of Five Years (USD)
Annual	127.63
Semi-annual	128.01
Quarterly	128.20
Monthly	128.34
Weekly	128.39
Daily	128.40

Just as pounds and grams are alternative units of measurement for weight, annual compounding and semi-annual compounding can be considered as alternative units of measurement for interest rates.

Suppose R_1 is an interest rate when compounding is m_1 times per year and R_2 is the equivalent rate when compounding is m_2 times per year. From Equation (16.1), the future values are the same if:

$$A\left(1 + \frac{R_1}{m_1}\right)^{m_1 T} = A\left(1 + \frac{R_2}{m_2}\right)^{m_2 T}$$

This means that:

$$\left(1 + \frac{R_1}{m_1}\right)^{m_1} = \left(1 + \frac{R_2}{m_2}\right)^{m_2}$$

This can be rewritten as:

$$R_2 = m_2\left[\left(1 + \frac{R_1}{m_1}\right)^{\frac{m_1}{m_2}} - 1\right] \qquad (16.2)$$

For example, suppose an interest rate is 8% per year with annual compounding and we want to know what it would be if it had been expressed with quarterly compounding. In this case, $R_1 = 0.08$, $m_1 = 1$, and $m_2 = 4$. The rate R_2 (i.e., with quarterly compounding) is

$$4\left[\left(1 + \frac{0.08}{1}\right)^{1/4} - 1\right] = 0.0777$$

or 7.77% per year.

As a second example, suppose a rate is 10% per year with monthly compounding and we want to know how it would be expressed with semi-annual compounding. In this case,

$R_1 = 0.1$, $m_1 = 12$, and $m_2 = 2$. The rate R_2 (i.e., with semi-annual compounding) is

$$2\left[\left(1 + \frac{0.10}{12}\right)^{12/2} - 1\right] = 0.1021$$

or 10.21%.

Usual Conventions

The usual convention is that interest rates are expressed so that the compounding frequency corresponds to the frequency with which payments are made. Instruments that last one year or less (from date of issue) are termed money market instruments; their principal and interest are both paid at maturity. An interest rate on a three-month money market instrument is therefore typically quoted with quarterly compounding, the interest rate on a one-year money market instrument is typically quoted with annual compounding, and so on.[6]

This convention can be misleading. If the one-year money market interest rate is 10% and the one-month money market interest rate is 9.8%, it may appear as though the one-year rate is greater than the one-month rate. However, the 9.8% is measured with monthly compounding, while the 10% is measured with annual compounding. Equation (16.1) shows 9.8% with monthly compounding becomes 10.25% when measured with annual compounding. This means that the one-month rate is actually *higher* than the one-year rate.

In the U.S. and some other countries, bonds which last longer than one year (from date of issue) normally pay interest every six months. The yield provided by a bond (which is defined later in this chapter) is therefore normally expressed with semi-annual compounding.

Sometimes the compounding frequency does not reflect the frequency of payments. In Canada, for example, mortgage interest rates are expressed with semi-annual compounding even though payments are made every month (or every two weeks). In the U.S., this mortgage interest rate would be expressed with monthly compounding; in the United Kingdom, it would be expressed with annual compounding. These varied ways of expressing interest rates are regulatory requirements designed to allow borrowers to compare interest rates within a country with as little confusion as possible.

[6] There are other conventions associated with the way interest rates are quoted on money market instruments. These will be discussed in Chapter 19.

16.3 CONTINUOUS COMPOUNDING

Table 16.1 shows the impact of daily compounding. We can imagine increasing the compounding frequency further so that we compound every hour, every minute, or every second. In the limit, we obtain *continuous compounding*. With continuous compounding, Equation (16.1) becomes

$$\text{Future Value} = Ae^{RT} \tag{16.3}$$

where e is the mathematical constant is approximately equal to 2.71828. The future value of USD 100 at the end of five years when the interest rate is 5% with continuous compounding is (in USD):

$$100e^{0.05\times 5} = 128.40$$

From Table 16.1, we see that this is (to two decimal places) the same as the future value with daily compounding.

Continuous compounding is used in the valuation of options and other, more complex derivatives. It can also be convenient to work with continuously compounded interest rates and yields when considering futures prices (see Section 11.5 of this book). The formulas involving interest rates expressed with continuous compounding are actually simpler than those involving periodic compounding (e.g., compare Equation (16.3) with (16.1)).

Suppose R_m is an interest rate expressed with compounding m times per year and R_c is the same rate expressed with continuous compounding. From Equations (16.1) and (16.3), we must have

$$A\left(1 + \frac{R_m}{m}\right)^{mT} = Ae^{R_cT}$$

so that

$$\left(1 + \frac{R_m}{m}\right)^{m} = e^{R_c}$$

and

$$R_c = m\ln\left(1 + \frac{R_m}{m}\right) \tag{16.4}$$

$$R_m = m\left(e^{\frac{R_c}{m}} - 1\right) \tag{16.5}$$

The continually compounded interest rate R_c that corresponds to a semi-annually compounded rate of 6% per year is therefore:

$$2 \times \ln\left(1 + \frac{0.06}{2}\right) = 0.0591$$

or 5.91%. The quarterly compounded rate corresponding to a continuously compounded 9% per year is

$$4 \times (e^{0.09/4} - 1) = 0.0910$$

or 9.10% per year.

16.4 ZERO RATES

A zero-coupon interest rate (also called a zero rate or spot rate) for maturity T is the applicable interest rate when an investor receives the total return (interest and principal) at the end of T years.

As mentioned previously, instruments with maturities of one year or less usually provide all principal and interest at the ends of their lives. They therefore provide zero rates directly. Most instruments that (when issued) last longer than one year provide regular interest payments known as coupons.[7] Zero rates must then be determined using a procedure we will describe later in this chapter.

16.5 DISCOUNTING

In Section 16.2 and 16.3, we saw how the future value of an amount A invested today can be calculated. As an alternative, we can ask how much a certain sum of money received in the future is worth today. This is known as discounting (or determining the present value). The present value of an amount A received at time T when the interest rate R is expressed with a compounding frequency of m times per year is

$$\text{Present Value} = A\left(1 + \frac{R}{m}\right)^{-mT} \qquad (16.6)$$

For example, USD 500 received in three years when the interest rate is 4% (compounded semi-annually) has a present value of:

$$500\left(1 + \frac{0.04}{2}\right)^{-2\times 3} = 443.99$$

The *discount factor* is the amount by which the cash flow is multiplied to get the present value. In this case it is

$$\left(1 + \frac{0.04}{2}\right)^{-2\times 3} = 0.8880$$

When the interest rate R is compounded continuously we obtain

$$\text{Present Value} = Ae^{-RT}$$

[7] An exception is provided by STRIPS, which are explained in Chapter 9 of *Valuation and Risk Models*.

16.6 BOND VALUATION

The valuation of a bond involves identifying its cash flows and discounting them at the interest rates corresponding to their maturities. As an example, suppose that zero-coupon interest rates with semi-annual compounding are those shown in Table 16.2. (How they can be calculated will be discussed later in this chapter.)

Consider a three-year bond with a principal of USD 100. The principal is also known as the *par value* (also known as the *face value*) and it is the amount that must be repaid at maturity. In this example, the maturity is three years.

Suppose further that the bond provides semi-annual coupons at the rate of 6% per year every six months. This means that the bond provides interest of USD 3 every six months. These cash flows are as shown in Table 16.3. To value the bond, we discount the cash flows at the interest rates corresponding to the times when they are received and then sum the results.

The present value of the first cash flow (received at time 0.5 years) can be calculated using Equation (16.6) and is

$$3 \times \left(1 + \frac{0.03}{2}\right)^{-2\times 0.5} = 2.9557$$

Table 16.2 Zero-Coupon Interest Rates (Compounded Semi-Annually)

Maturity (Years)	Zero-Coupon Interest Rate (Semi-Annually Compounded)
0.5	3.0%
1.0	3.8%
1.5	4.4%
2.0	4.8%
2.5	5.1%
3.0	5.3%

Table 16.3 Cash Flows from the Bond

Time (Years)	Cash Flow (USD)
0.5	3
1.0	3
1.5	3
2.0	3
2.5	3
3.0	103

Table 16.4 Calculation of Bond Value

Time (Years)	Cash Flow (USD)	Discount Factor	Present Value (USD)
0.5	3	0.9852	2.9557
1.0	3	0.9631	2.8892
1.5	3	0.9368	2.8104
2.0	3	0.9095	2.7285
2.5	3	0.8817	2.6451
3.0	103	0.8548	88.0407
Total			102.0695

The present value of the second cash flow is

$$3 \times \left(1 + \frac{0.038}{2}\right)^{-2\times 1} = 2.8892$$

These and other results are summarized in Table 16.4. Using these present values, the value of the bond itself can be shown to be USD 102.0695. This is greater than USD 100 because the interest rate earned on the bond is 6% per year (which is higher than the rates at which the bond's cash flows are being discounted).

Bond Yield

The return earned by an investor on a bond is often described by what is termed the *bond yield*. This is the discount rate that equates the present value of all the cash flows to the market price.

Suppose that the market price of the bond we have been considering is the value calculated in Table 16.4 (i.e., USD 102.0695). The bond yield y (semi-annually compounded) is the solution to:

$$\frac{3}{1 + y/2} + \frac{3}{(1 + y/2)^2} + \frac{3}{(1 + y/2)^3} + \frac{3}{(1 + y/2)^4} + \frac{3}{(1 + y/2)^5} + \frac{103}{(1 + y/2)^6} = 102.0695$$

The solution can be found using an iterative trial-and-error procedure. (Solver in Excel is a useful resource.) The solution for our example is $y = 5.2455\%$.

We would therefore refer to the bond in Table 16.3 as providing a yield of 5.2455% (with semi-annual compounding).

Par Yield

The par yield of a bond is the coupon rate that would cause the value of the bond to equal its par value.

For example, consider the rates in Table 16.2 with a par value of USD 100. If the par rate (paid semi-annually) on a three-year bond is c, the par yield would require

$$\frac{c/2}{1 + 0.03/2} + \frac{c/2}{(1 + 0.038/2)^2} + \frac{c/2}{(1 + 0.044/2)^3} + \frac{c/2}{(1 + 0.048/2)^4} + \frac{c/2}{(1 + 0.051/2)^5} + \frac{100 + c/2}{(1 + 0.053/2)^6} = 100$$

This can be solved to obtain $c = 5.2517\%$. A bond with a coupon of 5.2517% paid semi-annually would therefore be worth par.

Defining terms:

- m: The frequency with which coupons are paid ($m = 2$ in our example),
- A: The value of an annuity that pays USD 1 on each coupon payment date, and
- d: The value of USD 1 received at bond maturity.

The value of coupons is $A(c/m)$. The value of the final principal is $100d$. If c is the par yield, we therefore require

$$\left(\frac{c}{m}\right)A + 100d = 100 \qquad (16.7)$$

This can be rewritten as

$$c = \frac{(100 - 100d)m}{A}$$

In our example, $m = 2$, $A = 5.5310$, and $d = 0.8548$, so that the par yield is

$$\frac{(100 - 100 \times 0.8548) \times 2}{5.5310} = 5.2517$$

which is the same as that given by solving the bond pricing equation.

16.7 DURATION

Yield duration[8] measures the sensitivity of a bond's price to a change in its yield. Defining terms:

- B: The price of a bond,
- D: The bond's duration,
- Δy: The change in the bond's yield, and
- ΔB: The change in the bond's price resulting from the change in the yield.

[8] We will cover duration in more detail in Chapter 12 of *Valuation and Risk Models*.

The approximate duration relationship is[9]

$$\Delta B = -DB\Delta y \tag{16.8}$$

This is equivalent to:

$$\frac{\Delta B}{B} = -D\Delta y$$

The value of the duration (D) depends (slightly) on the compounding frequency with which the yield (y) is measured. The *Macaulay duration* is the correct value of D if the yield is measured with continuous compounding. The *modified duration* is the correct duration to use when D is measured with some other compounding frequency.

We will first consider the situation where y is measured with continuous compounding (and therefore the Macaulay duration applies). The correct D to use in Equation (16.8) is then the average time the bondholder must wait before receiving the present value, where the interest rate used in the present value calculations[10] is the yield y.

For example, if a bond has a present value of USD 106 with a cash flow in one year providing a present value of USD 6 and a cash flow in two years providing a present value of USD 100, the Macaulay duration would be

$$\frac{6}{106} \times 1 + \frac{100}{106} \times 2 = 1.9434$$

Equation (16.8) therefore indicates that a 5-basis point (0.05% or 0.0005) change in the continuously compounded yield to give rise to a price change of:

$$-1.9434 \times 106 \times 0.0005 = -0.103$$

so that the price decreases from USD 106 to USD 105.897.

For another example, consider the bond in Table 16.3. The price of the bond is USD 102.0695, and the yield is 5.2455% with semi-annual compounding. With continuous compounding, the yield is obtained from Equation (16.4) is

$$2 \times \ln(1 + 0.052455/2) = 0.051779$$

or 5.1779%.

The Macaulay duration calculation is illustrated in Table 16.5. When discounted at the yield, the first cash flow has a present value of USD 2.9233. This is 2.864% of the total present value.[11] The amount of weight given to $T = 0.5$ in the calculation of Macaulay duration is therefore 2.864% (= 2.9233/102.0695). The second cash flow has a present value of USD 2.8486 and weighting given to $T = 1.0$ is therefore 2.791%. Doing this calculation for each of the cash flows, Table 16.5, shows that the total Macaulay duration is 2.79238.

Table 16.5 Calculation of Duration for the Bond in Table 16.3

Time (Years)	Cash Flow	Present Value Using Yield as Discount Rate	Weight	Time × Weight
0.5	3	2.9233	0.02864	0.01432
1.0	3	2.8486	0.02791	0.02791
1.5	3	2.7758	0.02720	0.04079
2.0	3	2.7049	0.02650	0.05300
2.5	3	2.6357	0.02582	0.06456
3.0	103	88.1811	0.86393	2.59180
Total		102.0695	1.00000	2.79238

To test the accuracy of Equation (16.8) when the yield is expressed with continuous compounding, we set $D = 2.79238$. When there is a 10-basis point increase in the continuously compounded yield (i.e., from 5.1779% to 5.2779%), the bond's new price can be calculated as:

$$3e^{-0.5\times0.052779} + 3e^{-1.0\times0.052779} + 3e^{-1.5\times0.052779} + 3e^{-2.0\times0.052779} + 3e^{-2.5\times0.052779} + 103e^{-3.0\times0.052779} = 101.7849$$

Meanwhile, Equation (16.4) predicts the bond price will be

$$102.0695 - 2.7924 \times 102.0695 \times 0.001 = 101.7845$$

Thus, the duration relationship provides a good prediction of the change in the bond price resulting from a change in yield.

Modified Duration

The calculations so far have assumed that the yield (y) is measured with continuous compounding. When rates are compounded m times per year, Equation (16.8) provides an

[9] The relationship is approximate because the bond price is a non-linear function of its yield.

[10] This explains the terminology *duration*. Roughly speaking, duration is an estimate of how long in years an investor has to wait to receive returns.

[11] Note that for the purposes of producing Table 16.5, it does not matter whether the yield is expressed with continuous compounding, semi-annually compounding, or any other compounding frequency, provided that present values are calculated in a way that reflects the yield's compounding frequency. However, the sensitivity of the bond price to the yield in Equation (16.4) does depend on how the yield is measured.

appropriate approximation if the duration (when calculated as in Table 16.5) is divided by $1 + y/m$. This is referred to as the *modified duration* and hence:

$$\text{Modified Duration} = \frac{\text{Macaulay Duration}}{1 + y/m}$$

To illustrate the use of modified duration, consider again the bond in Table 16.3. The yield (measured with semi-annual compounding) is 5.2455% and the duration is 2.79238. The modified duration is therefore:

$$\frac{2.79238}{1 + 0.052455/2} = 2.7210$$

To test the accuracy of Equation (16.8) when D is set equal to the modified duration, suppose the semi-annually compounded yield is increased by 10 basis points from 5.2455% to 5.3455%. The bond price then becomes

$$\frac{3}{1 + 0.053455/2} + \frac{3}{(1 + 0.053455/2)^2} + \frac{3}{(1 + 0.053455/2)^3} + \frac{3}{(1 + 0.053455/2)^4} + \frac{3}{(1 + 0.053455/2)^5} + \frac{103}{(1 + 0.053455/2)^6}$$

$$= 101.7923$$

Meanwhile, the duration relationship in Equation (16.8) predicts

$$102.0695 - 2.7210 \times 102.0695 \times 0.001 = 101.7918$$

Again, note that Equation (16.8) is fairly accurate.

Dollar duration is another measure that is sometimes used. This is the product of the modified duration and the price of the bond. If dollar duration is $D_\$$, the relationship in Equation (16.4) becomes

$$\Delta B = -D_\$ \Delta y$$

Whereas D provides the sensitivity of *proportional* changes in a bond's price to a change in its yield, $D_\$$ provides the sensitivity of *actual* changes in a bond's price to changes in its yield.

Limitations of Duration

Duration provides a good approximation of the effect of a small parallel shift in the interest rate term structure. When all interest rates change by a certain amount, the yield on a bond changes by almost the same amount. In the previous example, the change in the yield was 10 basis points (or 0.1%). If all interest rates increased by 10 basis points, the yield would increase by almost exactly 10 basis points and Equation (16.4) would provide a good estimate of the decrease in the bond price.

Equation 16.8 cannot be relied upon, however, if the change in the bond yield arises from a non-parallel shift in the interest rate term structure or the change being considered is large.

16.8 CONVEXITY

The impact of parallel shifts in the interest rate term structure can be measured more accurately by considering convexity in addition to duration. If C is convexity, the approximation in Equation (16.4) can be refined to:

$$\Delta B = -DB\Delta y + \frac{1}{2}CB(\Delta y)^2 \qquad (16.9)$$

This equation allows relatively large parallel shifts to be considered.

The calculation of convexity is like the calculation of duration in Table 16.5, except that in the final column we calculate the weight multiplied by the square of time (rather than just the time). Table 16.6 shows that the convexity of the bond in Table 16.3 is 8.13904.

The convexity calculated in Table 16.6 can be used in Equation (16.9) with the Macaulay duration calculated in Table 16.5 when rates are expressed with continuous

Table 16.6 **Calculation of Convexity for the Bond in Table 16.3**

Time (Years)	Square of Time	Cash Flow	PV Using Yield as Discount Rate	Weight	Square of Time × Weight
0.5	0.25	3	2.9233	0.02864	0.00716
1.0	1.00	3	2.8486	0.02791	0.02791
1.5	2.25	3	2.7758	0.02720	0.06119
2.0	4.00	3	2.7049	0.02650	0.10600
2.5	6.25	3	2.6357	0.02582	0.16139
3.0	9.00	103	88.1811	0.86393	7.77539
Total			102.0695	1.00000	8.13904

compounding. When rates are expressed with compounding m times per year, we can define

$$\text{Modified Convexity} = \frac{\text{Convexity}}{(1 + y/m)^2}$$

and use it together with modified duration in Equation (16.5).

In our example, the modified convexity is

$$\frac{8.1390}{(1 + 0.52455/2)^2} = 7.7283$$

Now consider a 2% increase the semi-annually compounded yield for the bond in the previous example. This reduces the bond price to USD 96.6951. The duration approximation in Equation (16.8) is

$$102.0695 - 2.720 \times 102.0695 \times 0.02 = 96.5149$$

The duration plus convexity approximation in Equation (16.9) is

$$102.0695 - 2.7210 \times 102.0695 \times 0.02 + \frac{1}{2} \times 7.7283 \times 102.0695 \times 0.02^2 = 96.6726$$

This is a more accurate estimate.

16.9 FORWARD RATES

Forward rates are the future interest rates implied by today's zero-coupon interest rates. Consider the rates in Table 16.2. The six-month rate is 3% and the one-year rate is 3.8% (both semi-annually compounded). The forward rate for the period between six months and one year is the rate implied by these two rates. Specifically, it is the six-month rate that, when compounded with the current six-month rate of 3%, gives a one-year rate of 3.8%.

The general procedure for calculating the forward rate between time T_1 and time T_2 (where $T_2 > T_1$) is as follows:

- Calculate the value to which one dollar grows by time T_1 ($= V_1$).
- Calculate the value to which one dollar grows by time T_2 ($= V_2$).
- Calculate what one dollar at time T_1 is equivalent to at time T_2 ($= V_2/V_1$).
- Calculate the interest rate between times T_1 and T_2 that equates one dollar at time T_1 to V_2/V_1 dollars at time T_2.

Consider again the rates in Table 16.2. When we are calculating the forward rate for the period between six months and one year:

$$V_1 = \left(1 + \frac{0.03}{2}\right) = 1.015$$

$$V_2 = \left(1 + \frac{0.038}{2}\right)^2 = 1.038361$$

so that:

$$\frac{V_2}{V_1} = \frac{1.038361}{1.015} = 1.023016$$

If the forward rate F is expressed with semi-annual compounding, we have

$$1 + \frac{F}{2} = 1.023016$$

so that $F = 0.04603$. Therefore, the forward rate is 4.603%.

When we are calculating the forward rate for the period between one year and 1.5 years using the data in Table 16.2, we have

$$V_1 = \left(1 + \frac{0.038}{2}\right)^2 = 1.038361$$

$$V_2 = \left(1 + \frac{0.044}{2}\right)^3 = 1.067463$$

so that:

$$\frac{V_2}{V_1} = \frac{1.067463}{1.038361} = 1.028027$$

Again, we assume that the forward rate F is expressed with semi-annual compounding and obtain

$$1 + \frac{F}{2} = 1.028027$$

so that $F = 0.05605$. The forward rate here is 5.605%.

Forward rates for later six-month periods can be calculated similarly. The results are shown in Table 16.7.

When interest rates are expressed with continuous compounding, the equations for determining forward rates are less

Table 16.7 Forward Rates Calculated from the Interest Rates in Table 16.2

Period	Forward Rate
0.5 year to 1.0 year	4.603%
1.0 year to 1.5 years	5.605%
1.5 years to 2.0 years	6.005%
2.0 years to 2.5 years	6.304%
2.5 years to 3.0 years	6.303%

cumbersome. If R_1 and R_2 are the zero-coupon interest rates for the period between times T_1 and T_2:

$$V_1 = e^{R_1T_1}$$
$$V_2 = e^{R_2T_2}$$

so that:

$$\frac{V_2}{V_1} = e^{R_2T_2 - R_1T_1}$$

If F is the continuously compounded forward rate:

$$e^{F(T_2 - T_1)} = \frac{V_2}{V_1} = e^{R_2T_2 - R_1T_1}$$

so that:

$$F(T_2 - T_1) = R_2T_2 - R_1T_1$$

or:

$$F = \frac{R_2T_2 - R_1T_1}{T_2 - T_1}$$

For example, suppose the three-year zero-coupon interest rate is 5% and the four-year zero-coupon interest rate is 6% (both expressed with continuous compounding). The continuously compounded forward rate for the fourth year is

$$\frac{0.06 \times 4 - 0.05 \times 3}{4 - 3} = 0.09$$

or 9%. Note that if rates were expressed with annual compounding rather than continuous compounding, $V_1 = 1.05^3$, $V_2 = 1.06^4$, and $V_2/V_1 = 1.0906$ so that the forward rate with annual compounding is 9.06%.

Forward Rate Agreement

A forward rate agreement (FRA) can be thought of as an agreement that a certain interest rate will apply to a certain principal for a certain period in the future. Formally, it is an agreement that a pre-specified fixed rate of interest will be exchanged for a floating rate, with both being applied to a certain principal for a certain period. The floating rate of interest was Libor. The replacement rates mentioned earlier are now used.

An interest rate swap, which will be discussed in Chapter 20, can be regarded as a portfolio of FRAs. Each exchange in an interest rate swap is an agreement to exchange a pre-specified fixed interest rate for a floating interest rate with both applied to a certain principal.

Suppose an FRA involves paying a fixed interest rate of 4% and receiving three-month floating rate, with both being applied to a principal of USD 1 million for the three-month period beginning in one year. If the three-month floating rate in one year proves to be 4.8%, the FRA will involve interest being paid at a rate of 4% and received at a rate of 4.8%. Assume that the rates are expressed with quarterly compounding and that day count issues are not considered. The USD fixed payment (referred to as the payment on the fixed leg) is therefore:

$$0.25 \times 0.04 \times 1{,}000{,}000 = 10{,}000$$

The USD floating payment is similarly:

$$0.25 \times 0.048 \times 1{,}000{,}000 = 12{,}000$$

Because the FRA involves a fixed rate being paid and a floating rate being received, the holder of the FRA is due:

$$\text{USD } 12{,}000 - \text{USD } 10{,}000 = \text{USD } 2{,}000$$

Interest is normally paid at the end of the period to which it applies. The USD 2,000 in this example is therefore due to be paid after 15 months rather than after 12 months. By convention, an FRA is usually settled at the beginning of the period covered by the FRA. The settlement amount is the present value of the difference in interest amounts with the discount rate being the floating rate. In our example, this means that the FRA will lead to a USD amount of:

$$\frac{2{,}000}{1 + 0.048/4} = 1{,}976.28$$

being received at the one-year point. Expressed algebraically, the payoff is

$$\frac{(R - R_K)\tau L}{1 + R\tau} \tag{16.10}$$

where R is the realized floating rate, R_K is the fixed rate, L is the principal to which the rates are applied, and τ is the length of the time period.

To the party on the other side of the transaction (who pays floating and receives fixed) the payoff is

$$\frac{(R_K - R)\tau L}{1 + R\tau}$$

In our example this payoff would (in USD) be −1,976.28.

An FRA can be valued by assuming that the forward rates will occur. At the initiation of the transaction we are considering, we can assume that the forward value of floating rate is 4% so that the transaction is worth zero to both sides. Suppose that six months later, the forward rate for the period covered by the FRA (which would then be the period between six months and nine months in the future) is 4.4%. The USD payoff from FRA at the beginning of the three-month period covered by FRA can be estimated by substituting 4.4% for R into equation (16.10):

$$\frac{(0.044 - 0.04) \times 0.25 \times 1{,}000{,}000}{1 + 0.044 \times 0.25} = 989.12$$

The value of the FRA is the present value of 989.12 with discounting at the risk-free rate. Expressed algebraically, the value of an FRA where floating is received and fixed is paid is

$$\text{PV}\left(\frac{(R_F - R_K)\tau L}{1 + R_F\tau}\right)$$

where R_F is the forward rate for the period underlying the FRA and PV denotes present value at the risk-free rate from the beginning of the period covered by the FRA to today. Similarly, the value of an FRA where floating is paid and fixed is

$$\text{PV}\left(\frac{(R_K - R_F)\tau L}{1 + R_F\tau}\right)$$

The risk-free rate when derivatives such as this are valued is normally assumed to be the rate from calculated overnight index rates and swaps based on overnight index rates.

16.10 DETERMINING ZERO RATES

Analysts are tasked with determining zero-coupon interest rates from the market prices of traded instruments. As explained previously, money market instruments lasting less than one year provide their entire return at the end of their lives. Thus, they provide information about their zero-coupon interest rates in a direct way. However, instruments lasting longer than one year usually make regular payments prior to maturity. It is then necessary to calculate the zero-coupon rates implied by these instruments.

One way of doing this is by working forward and fitting the zero-coupon rates to progressively longer maturity instruments. This is called the *bootstrap* method.

As a simple example of how this can be done, suppose the zero-coupon interest rates (semi-annually compounded) for maturities of 0.5, 1.0, and 1.5 years in Table 16.8 have already been determined. Suppose further that a two-year bond with a par value of USD 100 has a market value of USD 102.7 when it pays a (semi-annual) coupon at the rate of 4% per year.

If the two-year zero-coupon interest rate is R, the equation that equates the value of the bond to USD 102.7 is

$$\frac{2}{1 + 0.02/2} + \frac{2}{(1 + 0.023/2)^2} + \frac{2}{(1 + 0.025/2)^3} + \frac{102}{(1 + R/2)^4} = 102.7$$

This can be solved to give $R = 0.0261$ and the two-year rate is therefore 2.61% with semi-annual compounding.

Table 16.8 Zero-Coupon Interest Rates for the First 1.5 Years

Maturity (Years)	Zero-Coupon Interest Rate Semi-Annually Compounded
0.5	2.0%
1.0	2.3%
1.5	2.5%

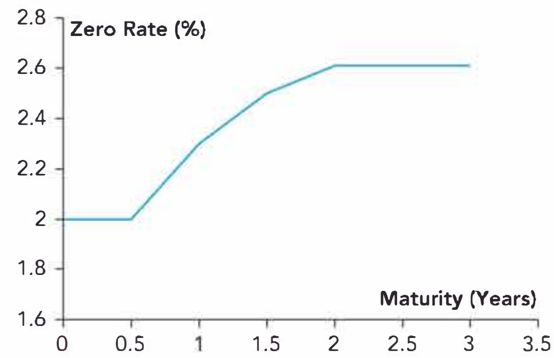


Figure 16.1 Zero curve calculated from the data in Table 16.8 and a two-year bond price.

A zero curve defines the relationship between zero-coupon interest rates and their maturities. The zero-coupon interest rates in a zero curve are usually assumed to be constant until the first maturity for which data is available. They are also usually assumed to be constant beyond the longest maturity for which data is available.

The zero curve we have calculated would therefore be as in Figure 16.1.

The calculations to produce a piecewise linear zero-rate curve are not always as straightforward as in the case above. For example, suppose the next price available was for a 2.4-year bond (instead of a two-year bond). We would set the 2.4-year zero rate equal to R. Because a coupon is paid on the bond at the 1.9-year point, the zero rate for a maturity of 1.9 years has to be determined in terms of R. By interpolating between the 1.5-year rate of 0.025 and the 2.4-year rate of R, the 1.9-year rate becomes

$$\frac{0.025 \times 0.5 + R \times 0.4}{0.9}$$

The other coupons are at times 1.4, 0.9, and 0.4 and can be determined by interpolation from the rates that have already been determined. An iterative search can then be used to determine the value of R that matches the bond price.

16.11 THEORIES OF THE TERM STRUCTURE

There are several theories about what determines the shape of the zero-coupon interest rate term structure. One is known as the *market segmentation theory*. This argues that short-, medium-, and long-maturity instruments attract different types of traders. While short-maturity rates would be determined by the activities of traders who are only interested in short-maturity investments, medium- and long-maturity instruments would only attract traders interested in those investments.

It is now generally recognized that the market segmentation theory is unrealistic. Market participants do not focus on just one segment of the interest rate term structure. For example, pension plans are primarily interested in trading long-maturity bonds, but will move to medium- or short-maturity instruments if their yields seem more attractive.

Another theory is known as the *expectations theory*. This argues that the interest rate term structure reflects where the market is expecting interest rates to be in the future. If the market expects interest rates to rise, the term structure of interest rates will be upward-sloping (with long-maturity rates being higher than short-maturity interest rates). If the market expects interest rates to decline, then the term structure will be downward-sloping (with long-maturity rates being lower than short-maturity rates). Expectations theory argues that forward rates should be equal to expected future spot rates.

In practice, interest rate term structures are upward-sloping much more often than they are downward-sloping. This calls the expectations theory into question, because we can reasonably hypothesize that the market expects interest rates to decrease as often as it expects interest rates to increase.[12] If the expectations theory were correct, we would expect to see downward-sloping term structures occurring as often as upward-sloping term structures.

A third theory, known as the *liquidity preference theory*, can explain this disparity. The theory argues that if the interest rate term structure reflects what the market expects interest rates to be in the future (i.e., the expectations theory holds), most investors will choose a short-term investment over a long-term investment. This is because of liquidity considerations. A short-maturity investment means that the funds will be available earlier to meet any needs (anticipated or unforeseen) that arise.

We can also look at the choice between long-maturity and short-maturity instruments from the perspective of the borrower. Short maturity borrowing usually has to be rolled over into new borrowing at the end of its life. This also entails liquidity risks. If (rightly or wrongly) the market's view of the financial health of the borrower declines, it may not be possible for the company to roll over short-term borrowing at a competitive interest rate.

Liquidity considerations therefore lead to lenders wanting to lend for short periods of time and borrowers wanting to borrow for long periods of time. In order to match borrowers and lenders, financial intermediaries (such as banks) must increase long-term rates relative to the market's expectations about future short-term rates. As they do this:

- Long-term borrowing becomes less attractive because short-term rates are more attractive to borrowers, and
- Short-term lending becomes less attractive because long-term rates are more attractive to lenders.

Liquidity preference theory therefore argues that the process of matching borrowers and lenders leads to long rates being higher than the market's expectations would suggest. To be precise, it argues that forward rates consistently overstate expected future spot rates.

To take a simple example, suppose that only two rates are offered in the market: a three-month rate and a five-year rate. Suppose further that both rates are 2.5% per year and that this reflects the market's expectations (so that all expected future three-month rates are 2.5% per year). If the term structure of interest rates is flat at 2.5% (consistent with expectations theory), liquidity considerations will lead lenders to choose to commit funds for only three months, while borrowers will choose the five-year maturity. This will lead to a mismatch. As financial intermediaries try to match borrowers and lenders, market forces will lead to the five-year rates being pushed above 2.5%. For example, it might be found that making the five-year rate 3.5% (while keeping the three-month rate at 2.5%) will cause some borrowers to switch from five-year borrowing to three-month borrowing and some lenders to switch in the other direction. The result is that supply and demand are matched at both maturities.

SUMMARY

In any given currency, several different interest rates are monitored. These include government borrowing rates, repo rates, and interbank borrowing rates such as Libor and overnight rates.

The way in which an interest rate is measured is defined by its compounding frequency. The difference between measuring an interest rate with annual compounding and one with monthly compounding is analogous to the difference between measuring a temperature in degrees centigrade and degrees Fahrenheit. Continuous compounding is commonly used when valuing options and similar derivatives.

[12] Short-term interest rates are mean reverting. This means that they tend to get pulled back toward a long-run average level. Over a long period of time they are below the long-run average (and therefore expected to move up) roughly half the time and above the long-run average (and therefore expected to move down) roughly half the time. For a further discussion of this see J. Hull, *Options, Futures and Other Derivatives*, 10th edition.

Zero-coupon interest rates (also referred to as zero rates or spot rates) are rates where all the return is received at the end of the instrument's life. Zero-coupon interest rates must usually be implied from the values of instruments that pay interest on a regular basis. A common approach is known as the bootstrap method; this is where an analyst works forward from short-maturity instruments to longer-maturity instruments to determine the rates that match prices.

Zero-coupon interest rates can be used to value bonds. Each payment (coupon or principal) on the bond is discounted at the zero-coupon interest rate corresponding to the time it will be received. The par yield for a bond with a certain maturity is the coupon that leads to the value of the bond equaling its par value. The yield on a bond is the discount rate that (when applied to all payments received by the investor) leads to a present value equal to the market price of the bond.

Duration is a measure of the sensitivity of a bond's price to its yield.

Forward rates are the future spot rates implied by the current interest rate term structure. Forward rate agreements (FRAs) are agreements where a certain rate observed in the future will be exchanged for a pre-determined fixed rate (with both rates being applied to the same principal).

One explanation for how interest rate term structures are determined by the market is the liquidity preference theory. This theory argues that if rates merely reflect the market's expectations about future interest rates, borrowers will borrow at a fixed rate for long periods, whereas lenders will prefer to lend for short periods. To match borrowers and lenders, it is necessary to increase forward rates relative to the expected future short rates.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

16.1 Why was Libor discontinued?

16.2 What is a repo rate?

16.3 Which is better for an investor: 6% per year compounded annually or 6% per year compounded monthly?

16.4 What is the formula for calculating the value at a future time T of an investment of amount A when the interest rate is R and the compounding frequency is m?

16.5 What is the formula for the present value of an amount A received at time T when the continuously compounded interest rate is R?

16.6 How is the par yield on a bond defined?

16.7 How is a bond yield defined?

16.8 What is the difference between modified duration and dollar duration?

16.9 What is a forward rate agreement? Explain how it is usually settled.

16.10 Explain what liquidity preference theory and expectations theory imply about forward interest rates and expected future spot rates.

Practice Questions

16.11 How is the effective federal funds rate determined?

16.12 The six-month and one-year zero rates are 3% and 4% (both compounded semi-annually) and a 1.5 year-bond paying a coupon of 4% per annum semi-annually has a yield of 5%. What is the 1.5-year zero-coupon interest rate?

16.13 A rate is 8% with semi-annual compounding. What would it be if expressed with monthly compounding?

16.14 A rate is 7% per year with continuous compounding. What is it when measured with quarterly compounding?

16.15 A two-year bond that pays coupons semi-annually at the rate of 7% per year has a market price of 103. What is the bond's yield measured with semi-annual compounding?

16.16 The six-month, 12-month, 18-month, and 24-month rates are 5%, 5.5%, 6%, and 6.5% (all measured with semi-annual compounding), respectively. What is the two-year par yield for a bond paying coupons every six months?

16.17 For the interest rates in Problem 16.16, what are the (semi-annually compounded) forward rates for six-month period beginning in 6 months, 12 months, and 18 months?

16.18 A three-year bond with a face value of USD 100 pays coupons annually at the rate of 10% per year. Its yield is 7% with annual compounding. What are (a) the Macaulay duration, (b) the convexity, (c) the modified duration, and (d) the modified convexity?

16.19 For the bond in 16.18, estimate the price change if the annually compounded yield changes from 7% to 8.5% using both the duration and duration plus convexity approximations.

16.20 In an FRA, an annualized rate of 3% will be received and a six-month floating rate will be paid on a principal of USD 5,000,000 for a six-month period starting in 18 months. If the annualized six-month forward rate in 18 months proves to be 3.5%, what is the settlement on the FRA? When is it made?

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

16.1 Libor was to a large extent based on estimates rather than actual transactions. Studies uncovered some manipulation of Libor rates.

16.2 A repo rate is a short-term secured lending rate. The borrower agrees to sell securities to the lender and buy them back for a slightly higher price a little later.

16.3 6% per year compounded monthly is better. 6% per year compounded monthly is

$$(1 + 0.06/12)^{12} - 1 = 0.0617$$

or 6.17% per year compounded annually.

16.4 The formula is

$$\text{Future Value} = A\left(1 + \frac{R}{m}\right)^{mT}$$

16.5 The formula is

$$\text{Present Value} = Ae^{-RT}$$

16.6 A bond's par yield is the coupon rate on the bond that causes the bond's price to equal its par value.

16.7 A bond's yield is the discount rate that when used for all of the bond's cash flows causes the total present value of the cash flows to equal the current market price of the bond.

16.8 Modified duration measures the relationship between proportional decreases (increases) in a bond's price and increases (decreases) in its yield. Dollar duration measures the relationship between the actual decreases (increases) in a bond's price and increases (decreases) in its yield.

16.9 A forward rate agreement (FRA) is an agreement to exchange

(a) An amount obtained when a pre-determined fixed rate of interest is applied to a certain principal for a certain future period, and

(b) An amount obtained when the actual (floating) rate observed in the market is applied to the same principal for the same period.

An FRA is normally settled at the beginning of the underlying period. The settlement amount is the difference between the two interest amounts discounted from the end of the period to the beginning of the period at the floating rate.

16.10 Expectations theory assumes that forward rates equal expected future spot rates. In liquidity preference theory, forward rates are greater than expected future spot rates.

Solved Problems

16.11 Banks borrow and lend between themselves to meet their reserve requirements with the central bank. Some of the transactions are brokered. The weighted average of the interest rate in brokered transactions is the effective federal funds rate.

16.12 The price of the 1.5-year bond with a face value of 100 is

$$\frac{2}{1 + 0.05/2} + \frac{2}{(1 + 0.05/2)^2} + \frac{102}{(1 + 0.05/2)^3} = 98.5720$$

If the 1.5-year zero rate is R, we must have

$$\frac{2}{1 + 0.03/2} + \frac{2}{(1 + 0.04/2)^2} + \frac{102}{(1 + R/2)^3} = 98.5720$$

The solution to this equation is $R = 0.05027$. The 1.5-year zero rate is therefore 5.027%.

16.13 From equation (16.2) rate with monthly compounding is

$$12\left[\left(1 + \frac{0.08}{2}\right)^{2/12} - 1\right] = 0.07870$$

or 7.870%.

16.14 From equation (16.5) the rate is

$$4(e^{0.07/4} - 1) = 0.07062$$

or 7.062% per year.

16.15 To determine the yield, y, we must solve

$$\frac{3.5}{1 + y/2} + \frac{3.5}{(1 + y/2)^2} + \frac{3.5}{(1 + y/2)^3} + \frac{103.5}{(1 + y/2)^4} = 103$$

By trial and error (or using Solver) the solution is $y = 0.053975$. The yield is 5.3975%.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

16.16 The par yield is the coupon rate c satisfying

$$\frac{c/2}{1+0.05/2}+\frac{c/2}{(1+0.055/2)^2}+\frac{c/2}{(1+0.06/2)^3}+\frac{100+c/2}{(1+0.065/2)^4}=100$$

It is 6.46%. Alternatively, we can use equation (16.7). In this case, $m = 2$, $d = 0.8799$, and $A = 3.7179$.

16.17 The forward rates are

$$2\times\left(\frac{1.0275^2}{1.025}-1\right)=0.060012$$

$$2\times\left(\frac{1.03^3}{1.0275^2}-1\right)=0.070037$$

$$2\times\left(\frac{1.0325^4}{1.03^3}-1\right)=0.080073$$

If all rates were continuously compounded, the forward rates would be 6%, 7%, and 8%. Because we are dealing with a semi-annually compounded rate, they are slightly different: 6.0012%, 7.0037%, and 8.0073%.

16.18 The calculations are in the table below:

Time (Yrs.) T	Cash Flow	PV with 7% Disc Rate	Weight W	$T \times W$	$T^2 \times W$
1	10	9.3458	0.0866	0.0866	0.0866
2	10	8.7344	0.0810	0.1619	0.3239
3	110	89.7928	0.8324	2.4972	7.4915
Total		107.8729	1.0000	2.7458	7.9021

The Macaulay's duration is 2.7458. The convexity is 7.9021, and the modified duration is

$$\frac{2.7458}{1.07}=2.5661$$

The modified convexity is

$$\frac{7.9021}{1.07^2}=6.9020$$

16.19 Using duration the price change is

$$-2.5661\times 107.8729\times 0.015=-4.1522$$

Using duration and convexity it is

$$-2.5661\times 107.8729\times 0.015+\frac{1}{2}\times 6.9020\times 107.8729\times 0.015^2=-4.0685$$

The actual bond price decline is 4.0419, showing that duration plus convexity gives a better estimate than convexity alone.

16.20 The USD settlement in 18 months is

$$\frac{(0.03-0.035)\times 0.5\times 5{,}000{,}000}{1+0.035/2}=-12{,}285.$$

It is settled in 18 months.

Corporate Bonds

17

Learning Objectives

After completing this reading, you should be able to:

- Describe features of bond trading and explain the behavior of bond yield.
- Describe a bond indenture and explain the role of the corporate trustee in a bond indenture.
- Define high-yield bonds and describe types of high-yield bond issuers and some of the payment features unique to high-yield bonds.
- Differentiate between credit default risk and credit spread risk.
- Describe event risk and explain what may cause it to manifest in corporate bonds.
- Describe different characteristics of bonds such as issuer, maturity, interest rate, and collateral.
- Describe the mechanisms by which corporate bonds can be retired before maturity.
- Define recovery rate and default rate, and differentiate between an issue default rate and a dollar default rate.
- Evaluate the expected return from a bond investment and identify the components of the bond's expected return.

A bond is a debt instrument sold by the bond issuer (the borrower) to bondholders (the lenders). The bond issuer agrees to make payments of interest and principal to bondholders. The principal of a bond (also called its face value or par value) is the amount the issuer has promised to repay at maturity.

Bonds perceived to be riskier than others available require higher interest rates to attract investors. The interest rate on a bond is termed the coupon rate. In the U.S., coupons are usually paid every six months. In some other countries, coupons are paid with other frequencies (e.g., monthly, quarterly, or annually). A U.S. bond with a principal of USD 1,000 paying a coupon of 8% would pay interest to the holder equal to 4% of USD 1,000 (i.e., USD 40) every six months. For example, the interest might be payable on February 15 and August 15 of each year during the life of the bond.

The face value of a bond in the U.S. is usually USD 1,000 and bond prices are typically quoted per USD 100 of principal. The value of the global bond market is approximately USD 100 trillion (as measured by the value of outstanding bonds), which is larger than the global equity market.

This chapter focuses on bonds issued by corporations. Bonds issued by the U.S. government are discussed in some detail in Chapters 9–13 of *Valuation and Risk Models*, while the credit risk discussion in this chapter is continued in Chapters 4 and 6 of that book.

17.1 BOND ISSUANCE

Corporate bond issuances are typically arranged by investment banks. These banks have connections with many potential bond purchasers and will work with the issuing corporation to determine the appropriate terms for the issuance.

The issuing corporation can choose between a *private placement* and a *public issue*. In a private placement, bonds are placed with a small number of large institutions (e.g., pension funds and insurance companies). Sometimes one institution will buy the entire bond issuance; on other occasions, bonds are sold to several different institutions. In either case, the bonds are not offered to the general public.

A private placement has several advantages to the issuer.

- There are fewer registration requirements. In the U.S., for example, private placement issuances do not need to register with the Securities and Exchange Commission.
- Rating agencies are not involved because they don't usually rate non-public issuances.
- The issuance cost is lower.
- The issuance can be completed quickly.
- The issuance can be relatively small.

Interest rates for private placement bonds are generally higher than those for equivalent publicly issued bonds. The issuer must therefore weigh the benefits of private placement against the payment of a higher interest rate.

In a public issue, the investment bank acts as the underwriter. This means that it buys the bonds from the corporation and then tries to sell them to investors. The investment bank's profit is earned from the difference between the price it pays to buy the bonds from the corporation and the price at which it sells the bonds to investors. Once issued, the bonds are traded and given ratings by rating agencies (as will be discussed later in this chapter).

The risks taken by the underwriter are defined in its contract with the issuer. Among these risks include the possibility that interest rates will increase, reducing the value of the bonds before the underwriter can to sell them to investors.

17.2 BOND TRADING

Bonds issued via private placements are often not traded. Instead, they are held by the original purchasers until maturity. Bonds issued in a public offering, on the other hand, are typically traded in the over-the-counter market.[1] This is in contrast to stocks, which are typically traded on exchanges.

There is a network of dealers who buy and sell bonds, either for their own portfolios or for their clients. Dealers aim to make a profit from the difference between the prices at which they buy and the prices at which they sell. Bond prices (like those of any other financial security) are determined by supply and demand. If more investors want to buy a bond than want to sell it, the price increases. If the reverse is true, the price decreases.

In Chapter 16, we explained how a bond's yield is defined. Bond yield can be thought of as the return earned on a bond over its life assuming that all interest and principal payments are paid as promised. As indicated in Figure 17.1, the yield is composed of a risk-free return (which is the return that would be earned on a similar risk-free instrument) and a credit spread (which is the extra return to compensate the investor for the possibility

[1] Convertible bonds (i.e., bonds that can be converted to equity in certain pre-defined ways) are sometimes traded on exchanges.

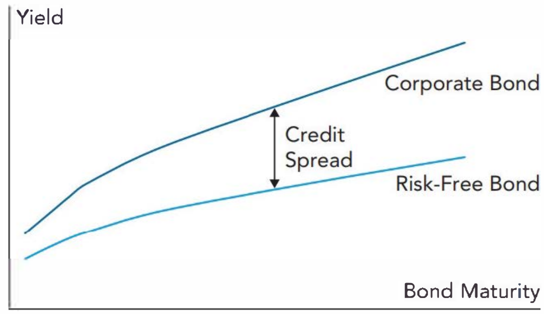


Figure 17.1 Illustration of the relationship between the yield on a corporate bond and the yield on a risk-free bond.

of a default).[2] As the maturity of the bond increases, the credit spread for a bond with a good credit rating tends to increase. This is because the chance that the issuer will experience financial difficulties in a year increases with the passage of time.

If the price of a bond increases, its yield declines (and vice versa). When it is considered likely that interest rates will increase, investors will demand higher yields. This will lead to selling pressure in the bond market, and the prices of all bonds should decline. When it is considered likely that interest rates will decline, the yields required by investors will also decline and the prices of all bonds will increase. News about the financial health of a bond issuer may cause the market to adjust the required credit spread. This can also increase or decrease the price of a bond.

Bond dealers are market makers who quote bid and ask prices for bonds on request as well as maintain inventory to facilitate trading. If risk-free interest rates and/or credit spreads increase, bond prices decrease and dealers lose money on their inventories.

Liquidity is an important issue in bond trading. It can be defined as the ability to turn an asset into cash within a reasonable time period at a reasonable price. Markets where the trading volume is high tend to be highly liquid and have low bid-ask spreads. Markets where the trading volume is low are usually less liquid and have higher bid-ask spreads.

Liquidity in the bond market varies from bond to bond. Some bonds trade only a few times a year, while others trade several times a day.[3] Part of the yield on a bond is compensation for its liquidity risk. As a bond's liquidity declines, investors require a greater yield. The Volcker rule (part of the Dodd–Frank legislation in the U.S.) now restricts the extent to which banks can trade bonds (or other securities) for their own account. Some banks in the U.S. have had concerns that the Volcker rule is reducing bond liquidity because it restricts their ability to keep bonds in their inventories for market-making purposes.[4]

The price quoted for a bond is referred to as the *clean price*. The cash price for a bond is called the *dirty price* and is the clean price plus accrued interest (i.e., the interest earned by the seller since the last coupon payment date). For corporate bonds in the U.S., this interest is calculated with a 30/360-day convention (where it is assumed that there are 30 days in a month and 360 days in a year).[5]

Suppose that the coupon rate is 8% per year, today is June 4, and the last coupon payment date was February 15. The number of days between February 15 and June 4 using the 30/360-day convention is 109 (= 15 + 30 + 30 + 30 + 4). The accrued interest on June 4 is therefore:

$$\frac{109}{360} \times 8\% = 2.422\%$$

If a dealer quotes the price of the bond as USD 96.00, the price that would be paid for a bond with principal of USD 1,000 is

$$960 + 24.22 = 984.22$$

Day count conventions can vary across countries.

[2] Debt issued by a government in its own currency is often considered to be risk-free because governments can always print more of their currency to avoid defaulting on the debt. However, developing countries do sometimes default on debt issued in their own currency. (The reasons for this are discussed in Chapter 5 of *Valuation and Risk Models*.) In practice, bonds issued by governments of the OECD member countries in their own currency are usually considered to be risk-free (see www.oecdwatch.org for information about the OECD and a list of member countries). An exception is bonds issued in euros by countries in the European Union, because these countries are not able to meet their obligations by printing euros. (The European Central Bank makes decisions on the euro money supply.)

[3] This may depend on whether the bond issuer is well known and whether the credit rating of a bond is sufficiently high to attract institutional investors. (Some institutional investors are restricted to buying *investment grade bonds*. We explain the meaning of this term later in the chapter.)

[4] This seems to be supported by research. See for example J. Bao, M. O'Hara, and A. Zhou. "The Volcker Rule and Market-Making in Times of Stress," *Journal of Financial Economics*, 2018.

[5] Day count conventions and accruals are also discussed in Chapter 19 of this book and in Chapter 9 of *Valuation and Risk Models*.

17.3 BOND INDENTURES

A bond indenture is a legal contract between a bond issuer and the bondholder(s) defining the important features of a bond issue. These features include the maturity date (i.e., when the principal must be repaid), the amount and timing of interest payments, callable and convertible features (if any), and the rights of bondholders in the event of contract violations. The bond indenture may also include several covenants. These can be categorized as:

- *Negative covenants* (also known as *restrictive covenants*), which limit the issuer's ability to engage in further debt financing, asset sales, dividend payouts, and share buybacks;
- *Positive covenants*, which require the issuer to produce regular financial statements, maintain properties, carry insurance, and use the money raised by the bond issuance in the manner stated in the offering document; and
- *Financial covenants*, which may require the issuer to maintain certain financial ratios (such as interest coverage or leverage ratios).[6]

There may also be other covenants that outline what happens if there is a change of control (referred to as CoC) and a credit rating downgrade.

Bonds issued by highly creditworthy firms generally contain few covenants, while those issued by riskier firms often have an extensive list of covenants.

Corporate Bond Trustee

In a public bond issue, it is unrealistic to expect individual bondholders to monitor the issuer and ensure it follows the bond indentures. The issuer therefore appoints (and pays for) a corporate bond trustee. This is a financial institution (usually a bank or trust company)[7] that looks after the interests of the bondholders and ensures that the issuer complies with the bond indentures.

The trustee reports periodically to the bondholders and will act on behalf of bondholders if the need arises. It may also act as paying agent and registrar (i.e., it may handle record keeping and the disbursement of interest and principal to the bondholders). Its specific duties are itemized in the bond indentures and the trustee is under no obligation to exceed those duties. For example, sometimes the indenture states that trustee can rely on the issuer and the issuer's attorneys for information on whether some covenants are being adhered to. In such cases, the trustee is not required to conduct its own investigations.

[6] Interest coverage is the ratio of earnings before interest and taxes to interest expense. Leverage ratios can be defined in many ways and reflect the percentage of a business financed with debt.

[7] In the U.S., a federal statute requires that trustees be appointed for bond issues over USD 5 million involving interstate commerce.

17.4 CREDIT RATINGS

Ratings agencies such as Moody's, S&P, and Fitch provide opinions on the creditworthiness of bond issuers. The scales used by these rating agencies are shown in Table 17.1. The best rating assigned by Moody's is Aaa; bonds with this rating are considered to have virtually no chance of default. The ratings continue downwards as follows: Aa, A, Baa, Ba, B, Caa, Ca, and C. Meanwhile, S&P and Fitch use a slightly different rating scale (e.g., AAA is used instead of Aaa, AA instead of Aa, BBB instead of Baa, and so on).

Moody's subdivides its ratings (except Aaa and Ca) to create a finer scale. For example, the Aa rating is subdivided into Aa1, Aa2, and Aa3; the A rating is divided into A1, A2, and A3; and so on. S&P and Fitch have similar subdivisions (e.g., AA rating category is divided into AA+, AA, and AA−; the A rating category is divided into A+, A, and A−; and so on).

Table 17.1 **Rating Scales for Corporate Bonds that Are Not in Default**

Moody's	S&P and Fitch
Investment Grade Bonds	
Aaa	AAA
Aa1	AA+
Aa2	AA
Aa3	AA−
A1	A+
A2	A
A3	A−
Baa1	BBB+
Baa2	BBB
Baa3	BBB−
Non-Investment (Speculative) Grade Bonds	
Ba1	BB+
Ba2	BB
Ba3	BB−
B1	B+
B2	B
B3	B−
Caa1	CCC+
Caa2	CCC
Caa3	CCC−
Ca	CC/C

Public bond issuers normally pay rating agencies to rate their bonds (although they are not required to do so). In theory, this would appear to be a conflict of interest; agencies have an incentive to produce overly positive ratings to attract and retain paying clients (bond issuers) instead of the objective ratings sought by the non-paying users (bond investors).[8] In practice, however, rating agencies know that their reputations would suffer if they allowed their relationships with bond issuers to influence their ratings; this creates an incentive for rating agencies to be as objective as possible.

High-Yield Bonds

Bonds rated above a threshold (as seen in Table 17.1) are referred to as investment grade. Bonds below this threshold are given various names: high-yield, non-investment grade, speculative grade, or simply *junk*.

There are several circumstances that can give rise to high-yield bonds.

- High-yield bonds may be sold by young and growing companies. These firms may have good prospects, but they lack the track record and strong financial statements of more established companies.
- Investment-grade bonds may become high-yield bonds as the financial situation of the issuing firm deteriorates. Such companies are sometimes referred to as *fallen angels*.
- A company with stable cash flows may increase its debt burden to benefit shareholders.[9]

High yield bonds sometimes have unusual features, such as the following examples.

- A *deferred-coupon bond* is a bond that pays no interest for a specified time period, after which time a specified coupon is paid in the usual manner.
- A *step-up bond* is a bond where the coupon increases with time.
- A *payment-in-kind* bond is a bond where the issuer has the option of providing the holder with additional bonds in lieu of interest.[10]
- An *extendable reset bond* is a bond where the coupon is reset annually (or more frequently) to maintain the price of the bond at some level (e.g., USD 101).[11]
- The issuer may have rights to call the bond from the proceeds of an equity issue.

17.5 BOND RISK

Credit ratings measure default risk. Rating agencies produce tables showing:

- The probability that a bond will default within *n* years after being given a certain rating for various values of *n, and*
- The probability that a bond will move from one rating category to another during a certain period (e.g., one year or five years).

We discuss these tables further and provide examples in Chapter 4 of *Valuation and Risk Models*.

Another risk faced by bondholders arises from changes in how the market prices credit risk (i.e., the credit spread). During normal times, the credit spread for a seven-year A-rated bond is approximately 100 basis points. This means that if the seven-year risk-free rate is 3%, the yield on a seven-year A-rated corporate bond is 4%. (See Figure 17.1). During stressed periods (i.e., when investors are particularly averse to taking risks), this credit spread could rise to 2% or even 3%.[12]

Another measure sometimes used by analysts is spread duration. This is (approximately) the percentage change in the bond price for a 100-basis point increase in the credit spread (assuming that the risk-free rate remains unchanged). For example, a spread duration of four indicates that a 100-basis point increase in the credit spread will reduce the bond price by 4%.

Event Risk

There are many events (e.g., natural disasters or the death of a CEO) that could adversely affect bonds. An important type of event risk is that of a large increase in leverage.

Firms can increase their leverage through activities such as leveraged buyouts,[13] share buybacks,[14] certain mergers and acquisitions, and other types of restructurings that benefit shareholders at the expense of bondholders.

An example of an event that hurt bondholders (and had a negative effect on the whole bond market) is the 1988 leveraged buyout of RJR Nabisco by Kohlberg, Kravis, Roberts & Co. The USD 25 billion buyout (and the resulting increase in leverage) increased the credit spread on existing bonds from 100 basis points to 350 basis points.

---

[8] One can make a similar argument about bond trustees, who are paid by the issuer but act on behalf of the bondholders.

[9] The RJR Nabisco leveraged buyout is an example of a situation where high-yield bonds were created from a buyout by a private equity firm.

[10] The issuer might do this if prohibited from paying interest by covenants in other bond issues.

[11] This structure has a flaw. If the company experiences financial difficulties, the coupon necessary to maintain a specified price may be extremely large. This will create cash flow problems for the company and the need for an even higher coupon.

[12] This phenomenon is referred to as a *flight to quality*.

[13] This is where a private group of shareholders borrows funds to buy out the current shareholders.

[14] This is where shares are repurchased from investors.

Sometimes bond indentures for lower-rated issues anticipate the possibility of increased leverage and include a *maintenance of net worth* clause. This clause requires the firm to keep its equity value above a prescribed level. If it fails to do this, it must begin to retire its debt at par until the equity moves above the prescribed level.

In some cases, the company merely has to offer to retire debt at par; bondholders would then choose to accept or decline the offer. However, the bondholders will normally accept the offer because it is unlikely that the market price of the bonds would be above par if the maintenance of net worth covenant has been breached.

Defaults

Default occurs when a bond issuer fails to make the agreed upon payments to the bondholders. Those who are owed money by the bond issuer have a claim against the issuer's assets. The issuing company may then reorganize itself (in negotiation with its creditors) or sell its assets to meet creditor (including bondholder) claims.

The bankruptcy laws in the U.S. facilitate reorganizations. For example, a Chapter 11 bankruptcy filing gives a company time to negotiate a re-organization with bondholders and other creditors. During this period, the firm's executives remain in control of the business. They are prevented from taking certain actions, however, such as selling fixed assets, arranging new loans, and stopping (or expanding) business operations. The reorganization may involve the sale of all or part of the business, a reduction in the amount owed on loans, and/or a reduction in the interest rate charged on the loans. Reorganizations can also result in debtholders becoming equity holders. Many large U.S. companies, including General Motors, Kmart, and United Airlines have made Chapter 11 filings and survived.

An important consideration when a default occurs is the ranking of claimants (i.e., which claims get satisfied first from available funds). Bondholders always rank above equity holders. The holders of some bond issues may rank ahead of others, while some bonds may rank ahead of trade creditors (e.g., suppliers of goods to the company that are owed money). This is discussed in more detail in the next section.

17.6 CLASSIFICATION OF BONDS

In this section, we consider the ways in which corporate bonds can be classified.

Issuer

Issuers can be put into five broad categories:

1. *Utilities:* Examples include electric, gas, water, and communications companies,
2. *Transportation companies:* Examples include airlines, railroads, and trucking companies,
3. *Industrials:* Examples include manufacturing, retailing, mining, and service companies,
4. *Financial institutions:* Examples include banks, insurance companies, brokerage firms, and asset management firms, and
5. *Internationals:* Examples include supranational organizations such as the European Investment Bank, foreign governments, and other non-domestic entities. The bonds that they issue are referred to as *Yankee Bonds* in the U.S.

Maturity

Corporate bonds have an original maturity of at least one year. (Instruments with an original maturity of less than one year are referred to as commercial paper.) Bonds with maturities of up to five years are usually referred to as short-term notes, those with maturities between five and 12 years are referred to as medium-term notes, and those with maturities of greater than 12 years are referred to as long-term bonds. As explained later, there are instances when all (or part) of the bond principal is repaid prior to maturity.

Interest Rate

Bonds can also be categorized by how they structure their interest rates.

Fixed-rate bonds pay the same rate of interest throughout their lives. Occasionally, the interest is payable in a foreign currency. For example, entities outside the U.S. sometimes issue bonds in U.S. dollars.

Floating-rate bonds, also known as *floating-rate notes* (FRNs) or *variable rate bonds*, are bonds where the coupon equals a floating reference rate plus a spread.

To see how a floating rate bond might work, consider a bond that pays a coupon every three months equal to a floating reference rate plus 20 basis points. The floating rate for each period would be observed and the fixed spread of 20 basis points added. Some floating bonds specify maximum or minimum (or both) levels for their coupons. As we will see in Chapter 20, a floating-rate bond can be created from a fixed-rate bond and an interest rate swap.

Zero-coupon bonds (as their name implies) pay no coupons to the holder. Instead, they sell at a discount to the principal amount.

For example, consider a five-year, zero-coupon bond that sells for USD 80. This means that an investor could pay USD 800 at the outset and get USD 1,000 in five years. The interest rate with annual compounding would be

$$\left(\frac{100}{80}\right)^{1/5} - 1 = 0.0456$$

or 4.56%.

The holder of a coupon-bearing bond in the U.S. usually has a claim on the principal in the event of a bankruptcy. In contrast, the holder of a zero-coupon bond has a claim on the original price paid plus accrued interest.

One of the attractions of zero-coupon bonds is that they can turn one form of income into another under certain tax regimes. If the USD 200 difference between the price at which the bond in the previous example is bought and the final repayment of its par value is treated as a capital gain, the investor will have essentially converted what would normally be interest income into capital gain income.[15] This is advantageous if capital gains are taxed at a lower rate.

Another advantage of zero-coupon bonds is that there is no reinvestment risk. By contrast, the coupons received from coupon-bearing bonds must be reinvested. If interest rates decline (increase), investors will be forced to reinvest at a relatively low (high) interest rate and thus a coupon bearing bond will lead to a worse (better) result than a zero-coupon bond.

Collateral

Bonds can also be classified by the collateral provided, which becomes important if a company defaults on its debt payments. As previously mentioned, a default leads to either a reorganization or an asset liquidation. In either situation, a bondholder with collateral should fare better than one without collateral. In the case of a liquidation, bonds with collateral will be paid first from the proceeds of the sale of the collateral; in the case of a reorganization, bonds with collateral will be in a stronger negotiating position than bonds without collateral.

[15] The U.S. now taxes zero-coupon bonds on imputed interest that accrues each year.

A mortgage bond provides specific assets (e.g., homes and commercial property) as collateral. In the event of a default, the bondholders have the right to sell the assets to satisfy unpaid obligations (although it is usually necessary to get permission of the courts first). The mortgage bondholder usually wants to ensure that its position will not be worsened by future bond issues. It may therefore impose conditions concerning future bond issues and the extent to which assets acquired in the future can be used as collateral for future bond issues.

Sometimes there is an *after-acquired clause* that requires property acquired after the bonds are issued to be used as collateral for the bonds. This effectively prevents the collateral from being used for other mortgage bond issues.

A collateral trust bond is a bond where shares, bonds, or other securities issued by another company are pledged as collateral. Usually, the other company is a subsidiary of the issuer. A company that has pledged shares of a subsidiary as collateral would like to be able to vote the shares when key decisions are made at shareholder meetings. Usually, the issuer has the right to vote shares in the subsidiary if there has not been a default. But if there has been a default, the corporate bond trustee votes the shares.

Note that the corporate bond trustee will act in the best interests of the bondholders (which may not always be in the best interest of the company's shareholders). Sometimes, there are provisions requiring additional collateral to be provided if the appraised value of the collateral falls below a certain level.

An equipment trust certificate (ETC) is a debt instrument used to finance the purchase of an asset. (They are commonly used to fund aircraft purchases.) The title to the property vests with the trustee, who then leases it to the borrower for an amount sufficient to provide the lenders with the return they have been promised. When the debt is fully repaid, the borrower obtains the title to the asset.

An advantage an ETC to the lender is that the asset is already owned by the trustee. Thus, legal proceedings are not necessary to take possession of the asset in the event of a default. Instead, the trustee (acting in the interest of the investors) can simply lease the asset to another company.

Debentures are unsecured bonds (i.e., bonds where no collateral has been posted by the issuer). They rank below mortgage bonds and collateral trust bonds and are likely to pay a higher interest rate. Often, a debenture bond's indenture will include provisions that limit the extent to which the issuing company can issue more debentures in the future. These provisions are needed because a new debenture issue weakens the position of existing debenture holders.

For example, debenture holders might get 30 cents on the dollar (i.e., an amount equal to 30% of their principal) in the event of a default or liquidation. If the company had been allowed to issue twice as many debentures (and if there are no other significant general creditors), the 30 cents on the dollar would become 15 cents on the dollar. Debentures sometimes include a *negative pledge clause* preventing the issuer from pledging assets as security for new bond issues if doing so weakens the debenture holder's position.

A subordinated debenture, as its name implies, ranks below other debentures and other general creditors in the event of a bankruptcy (which means that other debentures get paid first from available funds). Subordinated debentures require a higher rate of interest than unsubordinated debentures to compensate the holders for their inferior standing in the event of a default by the issuer.

Sometimes a bond issued by one company is guaranteed by another company (e.g., the company's parent). The bondholder should then receive the promised interest and principal (unless both the issuer and the guarantor default). The value of the guarantee depends on the correlation between the financial performance of the issuer and the guarantor. As the correlation increases, the guarantee becomes less valuable.

17.7 DEBT RETIREMENT

Bonds often last until the specified maturity date and are then retired using the proceeds from a new bond issuance. If interest rates decline during the life of the bond, an issuer would prefer to replace an older bond issue with a new one before maturity. This benefits the bond issuer because its interest payments will be lowered. However, the bondholder is worse off because the funds received from the early retirement must be reinvested at a lower rate.

Bond covenants can also prompt an issuer to replace an old bond issue with a new one. Changes in the nature of the issuer's business or improvements in its financial health can cause bond covenants to be unnecessarily burdensome (even if those covenants were reasonable at the time the bond was issued). Repaying bondholders early is a way to eliminate highly restrictive covenants.

In this section, we examine ways in which bonds can be retired early by the issuer.

Call Provisions

Bond indenture can sometimes allow the issuer to call the bond (i.e., buy it back from the bondholder) at a certain price at a certain time. Specifically, a schedule in the bond indenture specifies when bonds can be called and at what price. The price that the issuer must pay for the bond is known as the call price. Usually, the call price is greater than par when the bond is issued and declines towards par as the bond approaches maturity. Typically, a bond is not callable for the first few years of its life; this gives bondholders some protection against the possibility of early interest rate declines.

Some bonds are convertible into equity on pre-agreed terms. The conversion option is usually combined with a call feature. Without the call feature, it is likely to be optimal for bondholders to delay conversion for as long as possible.[16] To force conversion (so that bonds become equity and new debt can be raised), the issuer typically calls the bond as soon as (or shortly after) the price of its equity has risen to the point where it is better for a bondholder to convert the bond rather than sell it back to the issuer.

A *make-whole* call provision occurs when the call price is calculated instead of being set in advance. The call price is typically set equal to the present value of the remaining interest and principal cash flows owed to the bondholder. The discount rate is typically the risk-free rate plus a certain spread.

For example, the call price for a U.S. bond with five years remaining could be the present value of the remaining cash flows discounted at the five-year Treasury rate plus 10-basis points. Make-whole bonds have been growing in popularity in recent years. The advantage of a make-whole call provision is that the call provision has no financial cost to bondholders and therefore they do not need to demand a higher return as compensation.

A sinking fund is an arrangement where it is agreed that bonds will be retired periodically before maturity. The issuer may provide funds to the bond trustee so that the trustee can retire the bonds (usually at par value). Alternatively, the issuer can buy bonds in the open market and present them to the trustee. The latter is likely to be attractive to the issuer when bonds are selling below par. Accelerated sinking fund provisions allow the issuer to retire more bonds than the amount specified in the sinking fund.

Certain assets pledged as collateral (such as plant and equipment) can depreciate in value. If the principal amount of a bond issue remains the same, it will become less well covered by the collateral as time passes. One advantage to sinking fund provisions is that they can be made so that the amount borrowed declines in lockstep with the declining value of the collateral.

As an alternative to retiring bonds, sinking fund requirements can sometimes be satisfied by adding to the property held as collateral. Normally, the property added is greater than the collateral required so the security backing the debt improves. Another

[16] If the bondholder delays conversion, he or she keeps protection against the equity price declining. As soon as the bond is converted, this protection is lost.

approach is the use of *maintenance and replacement funds*, which requires the issuer to maintain the value of the collateral with property additions. If property additions are not made, cash can be used to retire debt. Whereas satisfying sinking fund requirements by adding property will usually *increase* the collateral supporting the corresponding debt, maintenance and replacement funds are usually designed to *maintain* the value of the collateral.

There are times that a company wants to sell assets that have been pledged as collateral. The bond indenture will normally allow a company to do this as long as the proceeds from the sale are used to retire the bonds. Selling property can therefore be a way to retire debt early.

A final way in which bonds can be redeemed early is by the company making a tender offer to bondholders. A tender offer is simply an offer to purchase the bonds. The offer can be at a fixed price or it can be calculated as the present value of future cash flows. As in the case of a make-whole call price, the discount rate is the risk-free rate plus a pre-specified spread.

17.8 DEFAULT RATE AND RECOVERY RATE

Two important statistics published by rating agencies are the default rate and the recovery rate. The default rate for a year can be measured in two ways.

1. *Issuer default rate*: This is the number of bonds that have defaulted in a given year divided by the number of issues outstanding.
2. *Dollar default rate*: This is the total par value of bonds that have defaulted in a given year divided by the total par value of all bonds outstanding.

The issuer default rate does not consider the size of the issues that defaulted, whereas the dollar default rate does. For example, suppose that there are 100 bonds with a total par value of USD 1 billion and that two bonds, each with a par value of USD 50 million, default. The issuer default rate is 2% and the dollar default rate is 10%.

When a bond defaults, the bondholder typically does not lose everything and some recovery is usually made. Because it is difficult to track the value of what is eventually received by claimants in the event of a default, the recovery rate is calculated as the value of the bond a few days after default as a percentage of its par value.

For example, if a bond sells for USD 40 per USD 100 of face value immediately after a default, the recovery rate would be calculated as 40%. The *loss given default* is one minus the recovery rate (60% in this example).

Credit risk and recovery rates will be discussed in more detail in *Valuation and Risk Models*. At this stage, note some research findings:[17]

- The average recovery rate is 38%.
- The distribution of recovery rates is bimodal.
- There is no relationship between the average recovery rate and the default rate in a given year.
- There is a negative relationship between recovery rates and default rates.
- Recovery rates are lower in an economic downturn or in a distressed industry.
- Tangible asset-intensive industries have higher recovery rates.

The bimodal distribution of recovery rates arises from the different seniority of the bonds. Consider the situation where there are bond issues X and Y, with X being senior to Y. Suppose the holders of X are owed USD 10 million while the holders of Y are owed USD 20 million. If the issuer defaults and the assets are sold for USD 8 million, the X bonds will receive 80 cents per dollar owed, while the Y bonds will receive 0 cents per dollar owed.

The *expected loss rate* on a bond in a given year can be defined as:

$$\text{Probability of Default} \times (1 - \text{Expected Recovery Rate})$$

For example, if the probability of default in a year is 0.5% and the recovery rate is 40%, the expected loss rate is 0.3%.

17.9 EXPECTED RETURN FROM BOND INVESTMENTS

The expected return from a bond is

$$\text{Risk-Free Rate} + \text{Credit Spread} - \text{Expected Loss Rate}$$

It might be thought that a bond with a credit spread of 100 basis points (i.e., 1%) would have an expected loss rate of 1% (and therefore the expected return on the bond is the risk-free rate). This is not usually the case, however, and the expected loss rate is lower than the credit spread.

Table 17.2 presents some results from Hull (2018)[18] showing that the excess of the spread over the loss rate tends to increase as the credit quality of the issuer decreases. (An exception is that Ba issuers tend to have an excess greater than that of B issuers.)

[17] See D. B. Madan, G. S. Bakshi, and F. X. Zhang, "Understanding the Role of Recovery in Default Risk Models: Empirical Comparisons and Implied Recovery Rates," ssrn id:285940.

[18] See J. Hull, *Risk Management and Financial Institutions*, 5th edition, Wiley, 2018.

Table 17.2 Loss Rates and Credit Spreads by Rating Category for Bonds with a Maturity of about Seven Years

Rating	Spread Over Treasuries (%)	Loss Rate (%)	Excess of Spread Over Loss Rate (%)
Aaa	0.78	0.02	0.76
Aa	0.86	0.05	0.81
A	1.11	0.12	0.99
Baa	1.69	0.25	1.44
Ba	3.22	1.30	2.92
B	5.23	3.40	1.83
Caa	11.46	7.50	3.96

It can be argued that while the Treasury rate is used as a risk-free benchmark by bond traders, it is not the right reference point for corporate bond yields. Rather, research suggests that a higher risk-free benchmark (e.g., the inter-bank borrowing rate) should be used instead. However, changing the risk-free benchmark does not change the nature of the results: bond traders still have an expected return greater than the risk-free rate and the credit spread they earn is always greater than their expected loss from defaults.

The poor liquidity of corporate bonds is one explanation for the results shown in Table 17.2. However, while bond traders do require some compensation for the illiquidity of bonds, research shows that the compensation provided by the market for bond illiquidity does not account for the results in Table 17.2.[19]

The main explanation for these results is that bonds do not default independently of one another. When the economy is doing well, the default rate on bonds tends to be low. However, the default rate increases when there is a recession. This means that bonds have systematic risk (i.e., risk related to the overall performance of the market) that cannot be diversified away. Thus, bond traders require compensation for taking systematic risk.

It is also worth noting that the non-systematic (i.e., idiosyncratic) risk of bonds can (in theory) be diversified away. But in practice, a very large portfolio is necessary to eliminate non-systematic risk because the return from a bond includes only a small probability of a default.[20] Bond traders may therefore require compensation for both non-systematic and systematic risks.

SUMMARY

Public bond issues are usually underwritten by investment banks. The bonds are then traded in the over-the-counter market. The yield on a bond is comprised of the risk-free return and a credit spread.

Bond indentures describe the promises of the issuer and the rights of the bondholder. A corporate bond trustee looks after the interests of bondholders by ensuring that the issuer follows the terms of the indenture.

Credit rating agencies are paid by the issuer to provide an opinion on the credit quality of the bonds it issues. Investment grade bonds have a credit rating of BBB (Baa) or above. High-yield bonds have a lower credit rating. Bondholders take several risks:

1. The risk that interest rates in the market will increase,
2. The risk that there will be a general increase in credit spreads,
3. The risk that the issuer's credit quality will decline, and
4. The risk that there will be decisions taken that improve the positions of equity holders at the expense of bondholders.

Bonds can be classified by the type of issuer, the bond maturity, the nature of the interest rate paid, and the collateral provided. Sometimes the issuer can take advantage of call provisions to retire a bond early. Other factors that can lead to early retirement of some or all of a bond issue are sinking funds, maintenance and replacement funds, the sale of assets, and tender offers by the issuer.

Two measures of default rates include the issuer default rate and the dollar default rate. Rating agencies monitor recovery rates in addition to default rates. The expected loss rate is the default rate multiplied by one minus the expected recovery rate. The expected loss rate per year is less than the credit spread. This means that the expected return of an investor on a portfolio of bonds is greater than the risk-free rate.

[19] For example, J. Dick-Nielsen, P. Feldhütter, and D. Lando, "Corporate Bond Liquidity Before and After the Onset of the Subprime Crisis," *Journal of Financial Economics*, 103, 3 (2012): 471–492, uses several different liquidity measures and a large database of bond trades. It shows that the liquidity component of credit spreads is relatively small.

[20] Whereas 30 well-chosen stocks are sufficient to eliminate all non-systematic risk in equity markets, thousands of bonds are necessary to do the same thing in bond markets.

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

QUESTIONS

Short Concept Questions

17.1 What are the advantages to a bond issuer of a Chapter 11 bankruptcy filing compared with a filing that involves the immediate liquidation of assets?

17.2 Give two other words used to describe the principal of a bond.

17.3 What happens to a bond's yield when its price increases?

17.4 What is the relationship between a bond's clean price and its cash price?

17.5 What is the role of the corporate bond trustee?

17.6 What is the difference between default risk, spread risk, and event risk?

17.7 What is a debenture?

17.8 Under what circumstances is an issuer likely to exercise its right to call a bond?

17.9 What are the two different ways of measuring default rates?

17.10 Why do bonds have systematic risk?

Practice Questions

17.11 Which of the following is likely to have a lower coupon: (a) private placement or (b) a public bond issue? Why do issuers not always choose the lower coupon alternative?

17.12 Why do banks argue that the Volcker rule will reduce bond liquidity?

17.13 A corporate bond in the U.S. pays a coupon of 10% per year. Payment dates are January 15 and July 15. What is the accrued interest on a USD 1,000 bond on the start of trading on April 1 in a (a) leap year and (b) a year that is not a leap year?

17.14 Explain three types of covenants in bond indentures.

17.15 How is an investment-grade bond defined? What are four names given to bonds that are not investment grade?

17.16 A seven-year zero-coupon bond sells for USD 70. What is the bond's yield?

17.17 Explain what is meant by a collateral trust bond and an equipment trust certificate.

17.18 Explain how the depreciation of property that has been pledged as security is handled in bond indentures.

17.19 How is the recovery rate on a bond usually calculated?

17.20 A bond has a credit spread over the risk-free rate of 1.5%, a default probability per year of 0.6%, and an expected recovery rate of 30%. What is the expected return in excess of the risk-free rate?

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

17.1 A Chapter 11 bankruptcy filing is a filing where a company has a period of time to negotiate a reorganization with its creditors.

17.2 Face value and par value.

17.3 The yield decreases. (Bondholders have to pay more to receive the same cash flows.)

17.4 The cash price is the clean price plus accrued interest.

17.5 The corporate bond trustee looks after the interests of bondholders.

17.6 Default risk is the risk that the bondholder will not receive the promised payments. Spread risk is the risk that the credit spread required by the market for the default risk will increase. Event risk is the risk that actions are taken that increase leverage and worsen the position of bondholders.

17.7 A debenture is a bond with no collateral.

17.8 The issuer will exercise its rights to call a bond if interest rates have declined or if it wants to eliminate restrictive covenants in the bond's indenture.

17.9 The issuer default rate is the percentage of issuers defaulting. The dollar default rate is the dollar value of bonds that defaulted as a percentage of the dollar value of outstanding bonds.

17.10 The default rate on bonds is higher when the economy is performing poorly and the stock market is producing poor returns.

Solved Problems

17.11 Public issues usually have a lower coupon but higher issuance costs. Issuers do not always choose the lower coupon because of the higher issuance costs.

17.12 The Volcker rule restricts a bank's ability to trade for its own account. This makes it more difficult for a bank to keep an inventory of bonds and act as a market maker.

17.13 Whether it is a leap year or not makes no difference. The day count convention for bonds is 30/360 so that it is always assumed that there are 30 days in a month. In this case, there are 75 days in the accrual period, the interest is USD 100 per year and the accrued interest (USD) is

$$\frac{75}{360} \times 100 = 20.83$$

17.14 A negative or restrictive covenant limits the issuer's ability to do such things as raise more debt, sell assets, pay dividends, and buy back shares. A positive covenant requires the issuer to do certain things (e.g., produce financial statements, maintain properties, and carry insurance). A financial covenant requires that certain financial ratios be maintained.

17.15 An investment grade bond is a bond with a rating from Moody's of Baa3 or above or a rating of BBB− or above from S&P or Fitch. The four names given to bonds that do not meet the investment grade criterion are high-yield bonds, non-investment grade bonds, speculative bonds, and junk bonds.

17.16 The yield (with annual compounding) is

$$\left(\frac{100}{70}\right)^{1/7} - 1 = 0.0523$$

or 5.23%.

17.17 A collateral trust bond is a bond backed by securities that are owned by the bond issuer (e.g., shares of a subsidiary). An equipment trust certificate is a leasing arrangement for a large asset such as an airplane. The trustee owns the asset and leases it to the borrower for a period of time for an amount sufficient to provide the lenders with their required return. At the end of the lease period, the borrower obtains the title to the asset.

17.18 It can be handled by arranging for bonds to be retired as the value of the collateral declines. It can also be handled by requiring more assets to be pledged as collateral as the value of the original collateral declines. Sinking funds and maintenance and replacement funds are used to ensure that bondholders are not adversely affected by the depreciation of collateral.

17.19 It is usually calculated as the value of a bond a few days after a default as a percentage of its face value.

17.20 The expected loss rate is $0.6\% \times (1 - 0.3) = 0.42\%$. The expected return in excess of the risk-free rate is therefore $1.5\% - 0.42\%$ or 1.08%.

18

Mortgages and Mortgage-Backed Securities

Learning Objectives

After completing this reading, you should be able to:

- Describe the various types of residential mortgage products.
- Calculate a fixed-rate mortgage payment and its principal and interest components.
- Summarize the securitization process of mortgage-backed securities (MBS), particularly the formation of mortgage pools, including specific pools and to-be-announced (TBAs).
- Calculate the weighted average coupon, weighted average maturity, single monthly mortality rate (SMM), and conditional prepayment rate (CPR) for a mortgage pool.
- Describe the process of trading pass-through agency MBS.
- Explain the mechanics of different types of agency MBS products, including collateralized mortgage obligations (CMOs), interest-only securities (IOs), and principal-only securities (POs).
- Describe a dollar roll transaction and how to value a dollar roll.
- Describe the mortgage prepayment option and factors that affect it; explain prepayment modeling and its four components: refinancing, turnover, defaults, and curtailments.
- Describe the steps in valuing an MBS using Monte Carlo simulation.
- Define Option Adjusted Spread (OAS) and explain its uses and challenges.

This chapter covers the properties of mortgages and mortgage-backed securities. Mortgages are used to finance residential and commercial property. Mortgage-backed securities (MBSs) are investments created from the cash flows provided by portfolios of mortgages. This discussion will focus on the U.S. mortgage market because it is a large market that is very important to fixed-income investors.

Residential mortgages in the U.S. typically last 15 or 30 years. Their interest rates can be fixed or variable. Variable-rate mortgages are termed adjustable-rate mortgages (ARMs). In an ARM, the interest rate is typically fixed for several years and is then tied to an interest rate index. Among the most common indices are the one-year Treasury rate (which is referred to as the constant maturity Treasury rate) and a cost-of-funds index (which is the average interest expense incurred by financial institutions in a region). ARMs are less risky than fixed-rate mortgages for lenders and riskier than fixed-rate mortgages for borrowers. For this reason, ARMs typically have lower initial interest rates than comparable fixed-rate mortgages.

Most of this chapter is concerned with fixed-rate mortgages. An interesting aspect of fixed-rate mortgages in the U.S. is that borrowers have an American-style option to pay off their outstanding mortgage balances. This is referred to as the borrower's *prepayment option*. Sometimes a mortgage is prepaid because the mortgaged property is being sold. Often, it will be prepaid because interest rates have declined and thus the property can be refinanced at a lower interest rate. There is usually no penalty for exercising the prepayment option, and it can be quite valuable to the borrower. On the other hand, the prepayment option can be quite costly to those who invest in MBSs. This is because the prepayment amounts must be reinvested even though interest rates tend to be low when prepayments occur. Analysts must take the prepayment option into consideration when valuing mortgage portfolios.

This chapter discusses how mortgage payments can be calculated and then explains how pools of mortgages are formed to create tradeable investment vehicles. It also outlines how prepayments can be modeled to calculate the value of an investment in a mortgage pool.

18.1 CALCULATING MONTHLY PAYMENTS

As mentioned in earlier chapters, the compounding frequency with which a mortgage interest rate is expressed does not always match the frequency of payments.

To calculate the monthly interest payments on a fixed-rate mortgage, it is necessary to first convert the quoted rate to a rate with monthly compounding. This can be done using Equation (16.2). For example, if a mortgage rate in Canada is quoted as 4% with semi-annual compounding, the rate with monthly compounding is

$$12 \times \left[\left(1 + \frac{0.04}{2}\right)^{2/12} - 1\right] = 0.03967$$

or 3.967%.

In the U.S., mortgage rates are quoted with monthly compounding and thus such a calculation is not necessary.

Consider a 30-year U.S. mortgage where the fixed rate is 6% with monthly compounding. The monthly payments (constant each month for $30 \times 12 = 360$ months) must totally amortize the mortgage. This means that the monthly payments must provide the lender with a return of principal as well as interest at 0.5% (= 6%/12) per month on the outstanding principal.

If the USD payment made at the end of each month is X, it must be the case that the present value of the summed payments equals the amount borrowed. If the amount borrowed is USD 250,000, we must have

$$\sum_{i=1}^{360} \frac{X}{(1 + 0.06/12)^i} = 250{,}000$$

Using the formula for the sum of a geometric series, this becomes[1]

$$\frac{\dfrac{X}{1 + 0.005}\left[1 - \dfrac{1}{1.005^{360}}\right]}{1 - \dfrac{1}{1 + 0.005}} = 250{,}000$$

or

$$\frac{X}{0.005}\left[1 - \frac{1}{1.005^{360}}\right] = 250{,}000$$

This equation can be solved to give $X = 1{,}498.88$. Payments of USD 1,498.88 per month therefore fully amortize (i.e., repay) borrowings of USD 250,000 over 30 years.

In general, the relationship between the amount borrowed A, the interest rate R (compounded monthly), and the monthly payment X is

$$\frac{X}{R/12}\left[1 - \frac{1}{(1 + R/12)^{12T}}\right] = A \qquad (18.1)$$

where T years is the life of the mortgage.

[1] The sum of a geometric series: $a + ab + ab^2 + \cdots ab^{n-1}$ is

$$\frac{a(1 - b^n)}{1 - b}$$

Table 18.1 An Amortization Table for a USD 250,000 30-Year Mortgage when the Interest Rate Is 6% with Monthly Compounding Is Shown. The Monthly Payment Is USD 1,498.88

Month	End-of-Month Interest Payment	End-of-Month Principal payment	Balance at End-of-Month
0			250,000.00
1	1,250.00	248.88	249,751.12
2	1,248.76	250.12	249,501.00
3	1,247.51	251.37	249,249.63
4	1,246.25	252.63	248,997.00
... ...	... ...	... ...	... ...
... ...	... ...	... ...	... ...
356	36.92	1,461.96	5,921.30
357	29.61	1,469.27	4,452.03
358	22.26	1,476.62	2,975.42
359	14.88	1,484.00	1,491.42
360	7.46	1,491.42	0.00

Amortization Tables

An amortization table shows the monthly principal and interest payments on a mortgage (assuming no prepayment of principal). At the beginning of the mortgage, most of the monthly payment is interest. Toward the end of the mortgage, most of the monthly payment is principal. In the USD 250,000 mortgage in the previous example, the interest rate is 6% per year (i.e., 0.5% per month) and the monthly payment is USD 1,498.88. In the first month the interest on the mortgage is

$$0.005 \times \text{USD } 250{,}000 = \text{USD } 1{,}250.00$$

which leaves

$$\text{USD } 1{,}498.88 - \text{USD } 1{,}250.00 = \text{USD } 248.88$$

as repayment of principal. At the beginning of the second month, the principal is therefore:

$$\text{USD } 250{,}000 - \text{USD } 248.88 = \text{USD } 249{,}751.12$$

The interest during the second month is

$$0.005 \times \text{USD } 249{,}751 = \text{USD } 1{,}248.76$$

The repayment of principal is therefore:

$$\text{USD } 1{,}498.88 - \text{USD } 1{,}248.76 = \text{USD } 250.12$$

An extract from the full amortization table is shown in Table 18.1.[2]

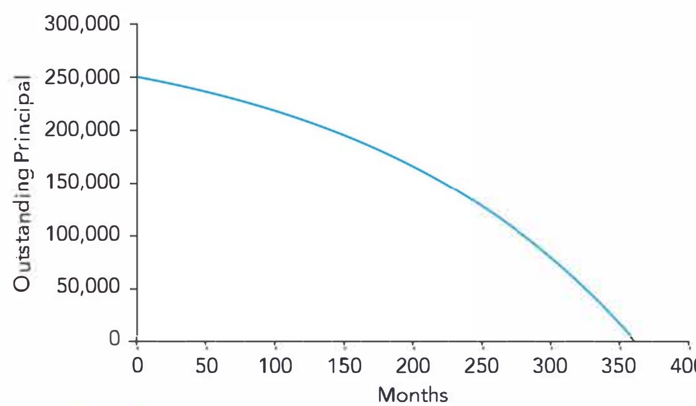


Figure 18.1 Decline of outstanding principal with time in a 30-year mortgage.

Figure 18.1 shows the way in which the principal declines with maturity in Table 18.1.

As an alternative to using an amortization table, analysts can calculate the outstanding principal by discounting the remaining cash flows using Equation (18.1). For example, when there are ten years (i.e., 120 months) remaining in the life of the mortgage, the outstanding USD principal (assuming there have been no prepayments) is[3]

$$\frac{1{,}498.88}{0.005}\left[1 - \frac{1}{1.005^{120}}\right] = 135{,}008.97$$

[2] Note that the numbers in Table 18.2 are based on the exact solution to Equation (18.1), that is, a mortgage payment, *X*, of USD 1,498.876313.

[3] This calculation uses the exact solution to Equation (18.1), that is, a monthly mortgage payment of USD 1,498.876313.

It should be noted that the calculations in this section are not for the valuation of mortgages. Instead, they are focused on the way the principal is amortized over the life of the mortgage when a rate is quoted. The value of a mortgage's cash flows to an investor at a specific time depends on the interest rate term structure at that time along with the expected costs from the prepayment option being exercised. This will be discussed later.

18.2 MORTGAGE POOLS

Mortgage portfolios (also known as *mortgage pools*) can be created for investment purposes. The mortgages in a pool are usually similar in terms of loan type, interest rate, and origination date. The *weighted-average coupon* (WAC) is the weighted-average interest rate on the mortgages in the pool, with the weight assigned to each mortgage being proportional to its outstanding principal.

To take a simple example, suppose a pool consists only of a USD 200,000 mortgage with an interest rate (referred to as a coupon) of 4% and a USD 400,000 mortgage with an interest rate of 5%. It would have a WAC of:

$$\text{WAC} = \frac{200{,}000}{600{,}000} \times 4\% + \frac{400{,}000}{600{,}000} \times 5\% = 4.667\%$$

The *weighted-average maturity* (WAM) is similarly calculated as the weighted-average of the number of months to maturity, with the weight assigned to each mortgage being proportional to its outstanding principal.

Suppose that the two mortgages we have just considered have maturities of 340 and 280 months (respectively). A pool consisting of the two mortgages would have a WAM (in months) of:

$$\text{WAM} = \frac{200{,}000}{600{,}000} \times 340 + \frac{400{,}000}{600{,}000} \times 280 = 300$$

In general:

$$\text{WAC} = \sum_{i=1}^{n} w_i c_i$$

$$\text{WAM} = \sum_{i=1}^{n} w_i L_i$$

where n is the number of mortgages in the pool, c_i is the coupon of the ith mortgage, L_i is the remaining life of the ith mortgage, and:

$$w_i = \frac{P_i}{\sum_{i=1}^{n} P_i} \tag{18.2}$$

where P_i is the remaining principal of the i^{th} mortgage.

In addition to WAC and WAM, mortgage pool statistics of interest are as follows.

- *The* average *loan balance* is the total current outstanding principal of the mortgages in the pool divided by the number of mortgages.
- *The pool's factor* is the total current outstanding pool principal as a percentage of the original pool principal. A pool's factor declines with time because part of each monthly payment by the borrower is a repayment of principal. It also declines as a result of prepayments.
- *The weighted-average FICO score* (FICO is short for Fair Isaac Corporation) is a FICO score that measures the creditworthiness of the borrower. It can range from 300 to 850, with a score above 650 being considered acceptable by many lenders. The weights are given by Equation (18.2), but with P_i usually set equal to the original principal of the ith mortgage.
- *The weighted-average loan-to-value ratio (LTV)* for a mortgage is the principal amount of the loan divided by the assessed value of the mortgaged property at the time of the loan. The weights are given by Equation (18.2), but with P_i usually set equal to the original principal of the ith mortgage.
- *The geographical distribution of the loans* describes the location of the properties being financed by the mortgages.

Prepayments are monitored in mortgage pools. In any given month, some mortgages may totally prepay and some may curtail (i.e., partially prepay). The *single monthly mortality rate* (SMM) is the percentage of the outstanding principal that was prepaid during a given month. This does not include scheduled payments of principal, such as those shown in Table 18.1. The *constant prepayment rate* (CPR), also known as the *conditional prepayment rate*, is the annualized SMM. The CPR is calculated as:

$$\text{CPR} = 1 - (1 - \text{SMM})^{12}$$

After one month, the proportion of the original principal remaining (excluding the impact of scheduled repayments of principal) is $1 - \text{SMM}$.

After two months, it is

$$(1 - \text{SMM})^2$$

After 12 months, it is

$$(1 - \text{SMM})^{12}$$

The amount of principal lost because of prepayments during the 12 months (if an observed SMM continues for 12 months) is therefore:

$$1 - (1 - \text{SMM})^{12}$$

18.3 AGENCY MORTGAGE-BACKED SECURITIES (MBSS)

In the U.S., there are three agencies that buy mortgages from banks and create mortgage pools. The agencies are

1. Government National Mortgage Association, referred to as Ginnie Mae (GNMA);
2. Federal National Mortgage Association, referred to as Fannie Mae (FNMA); and
3. Federal Home Loan Mortgage Corporation, referred to as Freddie Mac (FHLMC).

GNMA is a government agency, whereas FNMA and FHLMC are private companies known as *government-sponsored enterprises*. FNMA and FHLMC loans are not explicitly guaranteed by the U.S. government, but most market participants consider that there is an implicit guarantee.[4]

The three agencies enable banks to make long-term loans to home buyers without keeping the loans on their books. The funds banks receive from selling their mortgages can then be used to make new loans to home buyers. From the U.S. government's perspective, the agencies perform a useful function because they ensure that banks always have funds available for new home buyers to obtain mortgages.

The mortgages bought by these agencies must satisfy certain criteria in areas such as size and credit quality. Once the mortgages are purchased, cash flows from the mortgage pools are used to create mortgage-backed securities that are then sold to investors.

The three agencies guarantee their mortgages for a fee so that pool investors have protection against mortgage defaults. However, there is no protection for prepayment risk. If a mortgage is prepaid when interest rates are low (as is often the case), the investor must reinvest the funds and earn a lower than expected rate of interest.

The securities created by the agencies are known as agency mortgage-backed securities (agency MBSs). They can be contrasted with non-agency MBSs, which are issued by private corporations (typically financial institutions) and are not protected against defaults by government-sponsored institutions.

The simplest agency MBSs are *pass-throughs*, with all investors in a pool receiving the same return. Specifically, investors get their share of the cash flows from the mortgages in the pool minus the agency's fees for guaranteeing and servicing the mortgages. The pools are structured so that they offer returns in 50-basis point increments (e.g., 3%, 3.5%, and 4%). Although the returns are referred to as *coupons*, they are different from the coupons provided by Treasury bonds and corporate bonds. MBS coupons are made at the end of each month (rather than semi-annually) and the payments are a blend of interest and principal on the underlying mortgages.[5]

Trading of Pass-Throughs

Pass-throughs are characterized by their issuers, their coupons, and their maturities. For example, one pass-through security could be an investment in a GNMA 30-year 4% pool. The "30-year" descriptor refers to the original lives of the mortgages in the pool (rather than their current lives).

Pass-throughs are different from other risk-free investments in that they have prepayment risk. Prepayment behavior depends upon several factors, including interest rates and mortgage balances. (This will be discussed later.)

Pass-through agency securities trade as *specified pools* and *to-be-announced* (TBAs). In the specified pools market, buyers and sellers agree to trade a certain amount of a specified pool at a specified price. In the TBA market, buyers and sellers agree on:

- Issuer (e.g., FNMA);
- Maturity (original maturity) of the mortgages (e.g., 30 years);
- Coupon (e.g., 4.5%);
- Price per USD 100 of par value (e.g., USD 104.50);
- Par value (e.g., USD 100 million); and
- Settlement month (e.g., August).

The TBA market is a forward market and attracts more trading than the specified pools market.

For example, consider a TBA market trade where the seller must deliver mortgages from a FMNA 30-year 4.5% pool with a par value of USD 100 million for USD 104.50 million (plus accrued interest from the beginning of the month) in August. The particular mortgage pool that will be delivered is not specified. Instead, the seller has what is termed a *cheapest-to-deliver option*. This means the seller can choose to deliver those securities in August from any FNMA 30-year pool where the coupon is 4.5%. The settlement dates within the settlement month for TBAs are provided by the Securities Industry and Financial Markets Association (SIFMA).[6]

[4] After the 2007–2008 financial crisis, these entities experienced financial difficulties and there was an injection of capital from the U.S. Treasury.

[5] The way in which interest and principal are blended for an individual mortgage is illustrated in Table 18.1.

[6] See https://www.sifma.org/resources/general/mbs-notification-and-settlement-dates/ The settlement date is usually the twelfth or thirteenth of the month.

Two business days before the settlement day, the seller has to announce from which pool (or pools) the MBSs will be delivered.[7] The seller then receives the agreed price plus accrued interest from the beginning of the month. (Accrued interest is calculated assuming 30 days in a month.)

The remaining maturity of the mortgages in the delivered pool is generally less than 30 years and there are rules that define the range of maturities that are acceptable. For example, the mortgages in a delivered 30-year pool must have remaining maturities between 15 and 30 years.

Dollar Roll

A trade known as a *dollar roll* involves selling a TBA for one settlement month and buying a similar TBA for the following settlement month. For example, a trader could sell a USD 100 million 30-year FNMA pool with a 4.5% coupon for August settlement and buy a USD 100 million 30-year FNMA pool with a 4.5% coupon for September settlement.

Chapter 16 covered repo transactions (i.e., where one party sells securities to another party and agrees to buy them back at a future time for a slightly higher price). A dollar roll is similar (in some respects) to a repo. But there are two important differences.

1. The securities purchased in the second month may not be the same as the securities provided in first month. The party on the other side of the transaction can sell back the same securities, but it may also deliver securities with worse pre-payment properties.
2. No interest is added to the price at which the securities are repurchased. The dollar roll transaction involves the initiating party losing one month of interest payments from a pool with the specified coupon, while the party on the other side gains one month of interest.

Defining terms:

A: The price at which the pool is sold during the first month (including accrued interest),

B: The price at which the pool is purchased during the second month (including accrued interest),

C: The interest earned on the proceeds of the sale for one month, and

D: The coupon and the principal repayment that would have been received on the pool sold during the first month.

[7] Typically, the securities delivered can be from at most three pools that have the specified coupon, maturity, and issuer.

What is termed the *value of the roll* is thus calculated as:

$$A - B + C - D$$

For example, suppose that a USD 1 million par value of a 4.5% pool is sold for USD 102.50 in March and repurchased for USD 102.00 in April. We suppose that the payment date is the twelfth of the month for both months. This means that the accrued interest is USD 1,500 (= (12/30) × (0.045/12) × 1,000,000) in total for both transactions. It follows that A = USD 1,026,500 and B = USD 1,021,500.

Now assume that the proceeds of the sale in the first month can be invested at 0.1% for the month so that C = USD 1,026.5. In calculating D, we assume that if the pool had been not been sold, interest and principal payments on the pool during the month of the roll would have amounted to 0.45% of the par value. This means that D = USD 4,500.0.

In this case, the value of the roll (USD) is

$$1{,}026{,}500 - 1{,}021{,}500 + 1{,}026.5 - 4{,}500 = 1{,}526.50$$

Other Agency Products

The agency securities mentioned thus far have been pass-throughs. Another type of product is called a *collateralized mortgage obligation* (CMO). In a CMO, classes of securities that bear different amounts of prepayment risk are created. These classes are referred to as *tranches.*

As a simple example, suppose that there are Tranches A, B, and C with the following properties:

- Tranche A investors finance 30% of the MBS principal,
- Tranche B investors finance 50% of the MBS principal, and
- Tranche C investors finance the remaining 20% of the MBS principal.

Cash flows to the tranches are defined so that each tranche gets interest on its outstanding principal. However, there are special rules for principal payments (both scheduled and prepayments). Payments are first channelled to Tranche A. When the principal of Tranche A has been fully repaid, principal payments are then channelled to Tranche B. When the principal of Tranche B has been fully repaid, all remaining principal payments are channelled to Tranche C.

In this example, most of the prepayment risk is borne by Tranche A and very little is borne by Tranche C. However, the distribution of prepayment risk can be adjusted by changing the percentage of the pool financed by the different tranches.

Two other agency securities are *interest-only* securities (IOs) and *principal-only* securities (POs). These are also called stripped MBSs. As their names imply, all the interest payments from a mortgage pool go to the IOs, while all the principal payments go to the POs. Both IOs and POs are risky instruments.

As prepayments increase, a PO becomes more valuable because cash flows are received earlier than expected. By contrast, IOs become less valuable because fewer interest payments are made overall. As prepayments decrease, the reverse happens.

Non-Agency MBSs

Agency MBSs should be distinguished from MBSs that are not issued by GNMA, FNMA, or FHLMC. In a typical non-agency securitization, a mortgage portfolio is sold by a bank to a special purpose vehicle (SPV), which in turn passes the cash flows to the various securities it creates. In this case, there is usually no guarantee protecting investors against defaults. Indeed, default risk (rather than prepayment risk) is the major risk being taken by investors. Typically, the SPV creates several tranches that are subject to different amounts of default risk and promised returns.[8] This securitization model played a key role in the 2007–2008 crisis.[9]

18.4 MODELING PREPAYMENT BEHAVIOR

Calculating the cost of the prepayment option to investors is more complicated than valuing an interest rate option. This is because prepayment behavior depends on more than just interest rates. There are generally four reasons for prepayments:

1. Refinancing,
2. Turnover,
3. Defaults, and
4. Curtailments.

Refinancing

Refinancing arises when a borrower prepays a mortgage in order to refinance the underlying property. The most likely reason for this is a decline in interest rates. By refinancing in such an environment, the borrower can reduce his or her monthly payments. However, there can be other reasons. For example, the borrower's credit rating may have improved so that he or she is able to obtain a lower rate even when interest rates have not changed. Another reason for refinancing can be that the value of the property has increased so that a higher loan can be negotiated. (This is referred to as *cash-out refinancing*.)

The extent to which refinancing is likely to occur is measured by the *incentive function*. A simple incentive function I for a pool could be

$$I = \text{WAC} - R \tag{18.3}$$

where WAC is the weighted-average coupon and R is the current mortgage rate available to borrowers. The incentive function is then the amount by which refinancing allows borrowers to reduce their interest rates.

A slightly more elaborate incentive function is

$$I = (\text{WAC} - R) \times \text{ALS} \times A - K$$

where ALS is the average loan size, K is the estimated cost of refinancing a loan, and A is an annuity factor giving the present value of one dollar of payments per year for a period equal to the weighted-average maturity. This incentive function reflects the amount by which refinancing allows borrowers to reduce the total present value of their remaining payments. It reflects the empirical evidence that the prepayment rate increases as the average loan size increases.

For a given incentive function I, the annualized prepayment rate is sometimes modeled as:

$$\frac{1}{a + be^{-cI}}$$

where a, b, and c are parameters estimated from empirical data. It is illustrated in Figure 18.2 using the incentive function in Equation (18.2) with interest rates measured in basis points and assuming $a = 4$, $b = 0.02$, and $c = 25$. Note that this is an S-shaped function.

The shape of the CPR in Figure 18.2 is broadly consistent with prepayment experience. It shows that refinancing creates very little prepayment when mortgage rates have increased (so that WAC $- R$ is negative). As mortgage rates decrease, the prepayment rate increases quite fast and then levels off.

The parameters a, b, and c may depend on the economic environment. For example, we mentioned earlier that increases in housing prices make prepayments more likely. Another phenomenon that may affect the parameters is termed *burnout*. When a mortgage pool has been in existence for some time and interest rates have declined, the mortgage holders most likely to refinance (e.g., those who are financially sophisticated, have a good credit rating, or high loan balances)

[8] See Figure 1.2 of Chapter 1 for a simple example.

[9] Banks relaxed their lending standards and the tranches created provided lower returns than expected because the default rates on the underlying mortgages were much higher than expected. It was also the case that in the run-up period prior to the crisis, Fannie Mae and Freddie Mac guaranteed riskier mortgages than previously and they experienced severe financial difficulties as a result of the high default rates during the crisis.

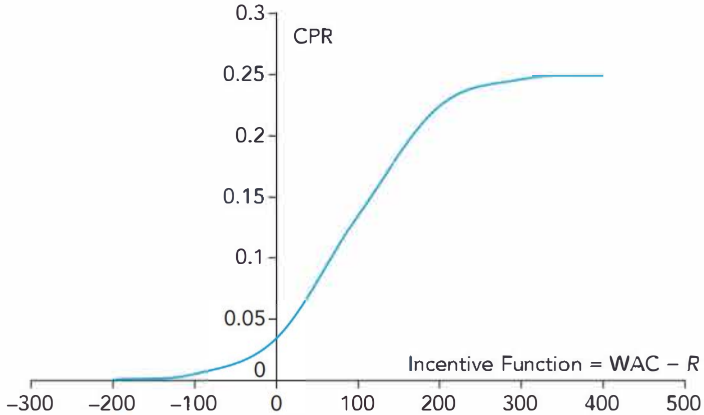


Figure 18.2 **Relationship between annualized conditional prepayment rate, CPR, from refinancing and the incentive function $I = \text{WAC} - R$.**

will tend to have already done so. The remaining mortgage holders are (on average) less likely to refinance. The burnout phenomenon shows that a prepayment function can be path dependent; prepayments on a mortgage pool that is a few years old can depend on where interest rates were in the past as well as where they are today.

Turnover

Turnover prepayments arise when a borrower sells the house. Turnover is higher in the summer months than the winter months. It is also lower early in the life of a mortgage because homeowners usually do not relocate immediately after taking out a mortgage.

In the prepayment model developed by the Public Securities Association (PSA) for analyzing American MBSs, increasing prepayment rates are assumed for the first 30 months after mortgage origination and constant prepayment rates are assumed thereafter. In the standard model (called the 100% PSA), the annualized prepayment rate is 0.2% in month one. It then increases by 0.2% each month until reaching 6% in month 30. After month 30, it remains constant at 6%. In the 150% PSA model, these numbers are 50% higher so that the annualized prepayment rate is 0.3% in the first month and rises to 9% in month 30.

The turnover rate is liable to depend on the geographical location of the properties and on the average age of the mortgage holders. It does not depend on interest rates, except to the extent that a homeowner may be less inclined to move if he or she is paying below-market rates on his or her mortgage. (A reluctance to move when the mortgage rate is low is referred to as the *lock-in* effect.)

Defaults

When a mortgage holder defaults and the mortgage is part of an agency pool, the agency pays the outstanding balance on the mortgage. This is treated as a prepayment and therefore defaults are relevant to the calculation of prepayments despite agency guarantees. Defaults added considerably to the prepayment experience of mortgage pools during the 2007–2008 crisis. Models use average FICO scores, LTVs, and the history of housing price movements to predict the default component of prepayments.

Curtailments

Curtailments are partial prepayments. These tend to occur when loans are relatively old and balances are relatively low. Prepayments from curtailments can rise as high as 5% when the loans in a mortgage pool have only one or two years to maturity.

18.5 VALUATION OF AN MBS POOL

The first step in valuing an MBS pool is to develop a prepayment model. The previous section indicates (in a general way) what the different components of such a model might be. Two variables that future prepayments may depend on are as follows.

1. Level of interest rates: The burnout phenomenon shows that the complete pattern of interest rates since mortgage origination is relevant to determining prepayments in a given month.
2. Housing prices: The history of housing prices since the mortgage origination may also be relevant. Sharp increases in housing prices may lead mortgage holders to refinance (i.e., use cash-out refinancing). Sharp decreases in housing prices may lead to defaults. Both are liable to increase prepayments.

It is necessary to develop models describing the uncertain future behavior of these variables in order to calculate future prepayment behavior.[10]

In addition, there are several relevant parameters describing the mortgage pool. The following are examples.

- The prepayment rate tends to increase as the average loan size increases.
- The geographical distribution of the loans may affect the model used to project housing prices and expected turnover.

[10] The models should be in a risk-neutral world. This is explained in Chapter 16 of *Valuation and Risk Models*.

Table 18.2 Interest and Principal Cash Flows When There Are No Prepayments

Month	End-of-Month Interest Payment	End-of-Month Principal Payment	Balance at End-of-Month
0			100,000
1	333.33	406.35	99,593.65
2	331.98	407.71	99,185.94
3	330.62	409.07	98,776.87
4	329.26	410.43	98,366.44
... ...	... ...	... ...	... ...
... ...	... ...	... ...	... ...
177	9.78	729.91	2,204.35
178	7.35	732.34	1,472.01
179	4.91	734.78	737.23
180	2.46	737.23	0.0

- Average FICO scores and average LTVs affect default predictions.
- Average loan age affects curtailment estimates.

Once a prepayment model has been specified, a technique known as Monte Carlo simulation is used to value an MBS. This involves the following steps.

1. Randomly sample from probability distributions to determine a hypothetical month-by-month path for risk-free interest rates and housing prices. The interest rate path will be accompanied by estimates of the spread of the mortgage rate over the risk-free rate. The housing price path may depend on interest rates and will reflect the geographical distribution of the mortgage holders.
2. For each month, determine prepayment rates using the specified prepayment model, the path for interest rates, and housing prices up to that month, as well as relevant parameters describing the mortgage pool.
3. Use the prepayment rates to calculate month-by-month cash flows from the MBS.
4. Starting at the end of the life of the MBS, discount cash flows month-by-month back to today. The discount rate for a month is the risk-free interest rate sampled for that month.
5. Repeat steps 1 to 4 many times.
6. Calculate the value of the MBS pool as the average of the calculated present values.

The advantage of Monte Carlo simulation is that it can consider what is referred to as path dependence. This means that a given month's prepayments can depend on the history of interest rates and the history of housing prices, as well as on their current levels. Other analytical tools, such as the use of trees, cannot easily accommodate this feature.[11]

As a simple example to show the effect of prepayments on the valuation of a mortgage pool, suppose that a pool contains new 15-year mortgages with payments being made at the end of each month. Suppose further that the coupon (compounded monthly) is 4%. The mortgage payment for an MBS with USD 100,000 of the pool's principal is from Equation (18.1):

$$\frac{\text{USD } 100{,}000 \times 0.04/12}{\left[1 - \dfrac{1}{(1 + 0.04/12)^{180}}\right]} = \text{USD } 739.69$$

We suppose that the risk-free discount rate is always 0.5% less than the mortgage rate for all maturities. If interest rates do not change and there are no prepayments, we can value the MBS by discounting USD 739.69 per month for 180 months at 3.5% per year (or 3.5%/12 = 0.2917% per month). This gives USD 103,470. The interest cash flows and principal cash flows are shown in Table 18.2. The present value of the interest cash flows with a 3.5% per year discount rate is USD 27,759, while the present value of the principal cash flows is USD 75,711.

[11] See Chapter 14 of *Valuation and Risk Models* for the use of trees to value options.

Now suppose that there is uncertainty about the level of interest rates over the next five years and that this uncertainty will be resolved soon. We suppose the following.

- There is a 50% chance that the mortgage rate will rise to 6% and there will be no prepayments.
- There is a 50% chance that mortgage rates will drop to 2%. Assume that the prepayment at the end of a given month will be 2% of the end-of-year mortgage balance after the scheduled repayment has been considered.

Assume that there are no further changes to the interest rate environment once this change has taken place and that the change takes place almost immediately after the pool of new mortgages is created.

If the mortgage interest rate is 6%, the cash flows will be as shown in Table 18.2 except that the discount rate will be 5.5% per year (i.e., 5.5%/12 = 0.4583% per month). The present value of interest cash flows will be USD 25,259 and the present value of the principal cash flows will be USD 65,269. The combined present value will therefore be USD 90,528.

If there were no prepayments when the interest rate dropped to 2%, the total present value of the MBS cash flows (at 1.5% discount rate) would be USD 119,162 and the value of the MBS would be USD 104,845 (= 0.5 × (90,528 + 119,162)).

This is greater than the USD 103,470 that was calculated assuming no change in interest rates because of the non-linear relationship between the present value of cash flows and the discount rate. Moving from a world where interest rates are certain to one where they are uncertain but have the same expected value increases the expected present value of cash flows.

Table 18.3 shows the impact of the 2% (per month) prepayment rate. At the end of the first month, the scheduled principal payment is USD 406.35 and the principal remaining after scheduled payments is USD 99,593.65 (as in Table 18.2). There is then an additional principal repayment equal to 2% of USD 99,593.65 (i.e., USD 1,991.87) so that the total principal repayment is USD 2,398.23 (= USD 1,991.87 + USD 406.35) and the remaining principal is USD 97,601.77 (= USD 100,000 − USD 2,398.23). Interest in the second month is USD 325.34 (= (0.04/12) × USD 97,601.77). The mortgage payment in the second month is 2% lower than in the first month (i.e., USD 724.89) because 2% of the mortgage principal has been lost from the pool. Because the scheduled principal payment is 2% lower than in Table 18.2 (i.e., USD 399.55), the total principal at the end of month two (after scheduled repayments) is USD 97,202.22 (= USD 97,601.77 − USD 399.55). An extra 2% of this principal is then repaid, bringing the principal at the end of month two down to USD 95,258.17. Calculations for the remaining months continue in a similar manner.

The present value of the interest payments in Table 18.3 discounted at 1.5% per year is USD 12,346, while that of the principal payments is USD 93,089 (for a total of USD 105,434). The results we have produced are summarized in Table 18.4. As indicated in the final row, the value of the MBS is

$$0.5 \times (90{,}528 + 107{,}716) = \text{USD } 99{,}122$$

Table 18.3 The Cash Flows in the 2% Mortgage Interest Rate Environments Are Shown. Prepayments Are 2% per Month

Month	End-of-Month Scheduled Payment	End-of-Month Interest Payment	End-of-Month Scheduled Principal Payment	Additional End-of-Month Principal Payment	Remaining Principal at Month End
0					100,000.00
1	739.69	333.33	406.35	1,991.87	97,601.77
2	724.89	325.34	399.55	1,944.04	95,258.17
3	710.40	317.53	392.87	1,897.31	92,968.00
4	696.19	309.89	386.30	1,851.63	90,730.07
......	......	......		......	......
......	......	......		......	......
177	21.13	0.28	20.85	1.26	61.70
178	20.70	0.21	20.50	0.82	40.38
179	20.29	0.13	20.16	0.40	19.82
180	19.88	0.07	19.82	0.00	0.00

Table 18.4 Summary of Results (USD)

	Present Value of Interest Payments	Present Value of Principal Payments	Total Present Value
No Interest Rate Change	27,759	75,711	103,470
Mortgage Rate Increase to 6%	25,259	65,269	90,528
Mortgage Rate Decrease to 2%	12,326	95,370	107,716
Average of 6% and 2% Mortgage Environments	18,802	80,320	99,122

Interest rate uncertainty, together with the possible prepayments, reduce the price from USD 103.470 to USD 99.122 (per USD 100 of principal).

Table 18.4 also illustrates the general phenomenon that interest rate uncertainty (and the resulting prepayment uncertainty) reduces the value of IOs while increasing the value of POs. Note that the price of an IO constructed from the pool decreases from USD 27.759 to USD 18.802, whereas that of a PO constructed from the pool increases from USD 75.711 to USD 80.320.

18.6 OPTION ADJUSTED SPREAD

The option-adjusted spread (OAS) is the excess of the expected return provided by a fixed-income instrument over the risk-free return adjusted to account for embedded options. The return on an MBS is adjusted for prepayment options as follows

OAS = Expected MBS Return – Return on Treasury Instruments

We have outlined a procedure involving Monte Carlo simulation for determining the value of an MBS. This procedure can be adjusted to determine the OAS provided by the MBS. The procedure is as follows.

1. Make an initial estimate of the OAS.
2. Carry out a Monte Carlo simulation as described in the previous section but using discount rates equal to the Treasury rate plus the current estimate of the OAS.[12]
3. Compare the price obtained with the market price.
4. If the market price is higher than the simulated price, the OAS estimate is reduced. If the market price is lower than the simulated price, the OAS estimate is increased.
5. Continue changing the OAS estimate until the simulated price equals the market price.

[12] When the MBS is valued, an analyst might use a different proxy for the risk-free rate other than the Treasury rate. The OAS, however, is calculated relative to the Treasury rate.

The method of successive bisection can be used to create a workable algorithm. First, initial high and low OAS estimates are produced. This can be done by reducing OAS until the simulated price is higher than the market price and then increasing OAS until the simulated price is lower than the market price. The average of the high and low prices is then used in step four of the simulation. If this proves to be too high, it becomes the new high OAS. If it proves to be too low, it becomes the new low OAS. The procedure is then repeated. Note that the ranges between the high and low OAS are halved on each iteration.[13]

Assume that the pool in the previous example sells for USD 98.00. The OAS is therefore the spread that must be added to the discount rates to give a present value of USD 98,000 per USD 100,000 of par value. This turns out to be 24.67 basis points. If we discount at 5.7467% and 1.7467% (instead of 5.5% and 1.5%) in these two scenarios (respectively), the value of the pool changes from USD 99.12 to USD 98.00.

OAS can be useful in determining the relative valuation of different MBS pools. For example, an MBS with an OAS of 80 basis points (i.e., with a return of 0.8% over the Treasury rate) should be a better buy than one that provides an OAS of 40 basis points (i.e., with a return of 40 basis points over the Treasury rate). Of course, the OAS calculated depends on the extent to which the underlying model correctly accounts for prepayments. If the model is incorrect or has not been calibrated properly, the results cannot be relied upon. When an analyst finds a high-OAS pool, he or she should look for institutional or technical reasons why that pool might trade differently from the rest of the market. It may be that the analyst can find an assumption in the model being used that leads to the high OAS. He or she should then critically examine the validity of that assumption.

The dependence of the prepayment model on interest rates is critical. If an analyst feels that the model correctly describes prepayments, it should be possible to hedge the interest rate exposure (at least approximately) to lock in an expected profit.

[13] The sets of random number samples used to determine paths should be the same on each iteration.

The natural instruments to hedge interest rate exposure are Treasuries. However, mortgage rates are not perfectly correlated with Treasury rates and so even if an analyst has a perfect prepayment model that depends only on interest rates, there will always be some residual interest rate risk when Treasury instruments are used as hedges.

SUMMARY

The mortgage market in the U.S. presents unique challenges for analysts. The first point to note is that a mortgage is different from a bond in that the principal is not paid at the end. Instead, the payments are made by the mortgage holder monthly and are a blend of interest and principal so that by the end of the life of the mortgage there is no principal outstanding.

Mortgages in the U.S. usually last 15 or 30 years. The holder has the option to prepay the mortgage at any time (usually without penalty). This creates uncertainty about when the lender will receive cash flows. In practice, analysts develop models for estimating the rate at which prepayments will occur on in mortgage pools. Two main reasons why homeowners prepay are as follows.

1. Interest rates have declined so that the house can be refinanced at a lower cost.
2. The house has been sold.

Three agencies create mortgage pools and provide investors with protection against defaults (but not against prepayments): Government National Mortgage Association (Ginnie Mae, GNMA), Federal National Mortgage Association, (Fannie Mae, FNMA), and Federal Home Loan Mortgage Corporation (Freddie Mac, FHLMC).

There is an active forward market in mortgage pools created by these agencies, known as the TBA market. A feature of this market is that the precise pool that will be delivered is not specified at the time the forward contract is agreed. Parameters describing the pool, such as the maturity of the mortgages and coupon rate, are specified and the seller chooses which pool with the specified parameters will be delivered. In practice, the seller is likely to use a prepayment model to determine the least valuable pool that can be delivered.

A popular trade is the dollar roll, where a TBA for one month is sold and a similar TBA for the next month is purchased. This can lead to funds being borrowed at an attractive interest rate. However, the pool received during the second month may be different from (and less highly valued than) the pool sold during the first month.

The following questions are intended to help candidates understand the material. They are _not_ actual FRM exam questions.

QUESTIONS

Short Concept Questions

18.1 What is the difference between the pattern of cash flows provided by a fixed-rate mortgage and the pattern of cash flows provided by a U.S. Treasury bond? Assume no prepayments and no defaults on the mortgage.

18.2 How is weighted-average maturity for a pool defined? Provide a formula.

18.3 What is a pool's factor?

18.4 Which of FNMA, GNMA, and FHLMC are private companies?

18.5 What is a pass-through

18.6 Explain how a dollar roll works and why it is different from a repo as a way of using securities that are owned to borrow money.

18.7 List four reasons for prepayments.

18.8 Explain what is meant by burnout.

18.9 Why are mortgage defaults relevant to the prepayment rate calculated for an agency MBS?

18.10 What is a curtailment?

Practice Questions

18.11 What is the monthly payment on a 15-year mortgage where the rate is 5% with monthly compounding and the amount borrowed is USD 100,000?

18.12 How much is outstanding on the mortgage in Question 18.11 after seven years? Assume there is no prepayment.

18.13 A pool consists of three mortgages:

1. USD 100,000 mortgage with 330 months to maturity and a 4% rate
2. USD 200,000 mortgage with 310 months to maturity and a 5% rate
3. USD 300,000 mortgage with 290 months to maturity and a 6% rate

Calculate the WAC and WAM.

18.14 What is the range of FICO scores? What is mentioned in the text as a typical minimum acceptable FICO score?

18.15 Explain how CPR is defined.

18.16 Explain how a TBA works.

18.17 What tends to happen to the values of IOs and POs as prepayments increase?

18.18 How is a prepayment function used in the valuation of an MBS?

18.19 How is OAS defined for an MBS?

18.20 In a dollar roll a USD 10,000,000 30-year 6% pool is sold for 103.25 and bought for 102.80. The settlement is on the fourteenth of the month for both months. The proceeds of the sale can be invested at 4% per year (compounded monthly), and interest and principal payments on the pool during the month of the roll would amount to 0.4% of the par value. What is the value of the roll?

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

ANSWERS

Short Concept Questions

18.1 A fixed-rate mortgage provides monthly cash flows that are constant for the whole life of the mortgage. A bond provides semi-annual coupons and a final payment of principal.

18.2 The weighted-average maturity for the mortgages in a pool is

$$\sum_{i=1}^{n} w_i L_i$$

where n is the number of mortgages, L_i is the remaining life of the ith mortgage, and w_i is the proportion of the total principal provided by the ith mortgage.

18.3 A pool's factor is the total current outstanding principal divided by the total original principal for the mortgages in the pool.

18.4 FNMA and FHLMC are private companies.

18.5 A pass-through is an MBS where the cash flows from a pool of mortgages are passed to investors after subtracting the costs of guaranteeing and servicing the mortgages. All investors in the pool receive the same return.

18.6 A dollar roll involves raising funds by selling a pool in one month and buying a similar pool back the following month. Properties such as coupon, principal, and maturity are specified but the party with a short position chooses the actual pool that will be delivered. This means that the pool that is received during the second month may be different from the one sold during the first month. This is one difference between a dollar roll and a repo. The other difference is that interest does not have to be added to the price paid in the second month because the trader who sells and then buys loses one month of MBS income, while the trader on the other side gains one month of MBS income.

18.7 Refinancing, turnover, defaults, and curtailments

18.8 Burnout occurs when interest rates decline over a period of time. Those mortgage holders who are most likely to prepay (e.g. because their mortgage balances are high, they have good credit ratings, or are financially sophisticated) do so during the first part of the period and leave the pool. The prepayment rate for the mortgage holders who are left is then lower.

18.9 When a mortgage holder defaults, the guarantor pays the amount owing and the mortgage is treated as a prepayment.

18.10 A curtailment is a partial prepayment.

Solved Problems

18.11 The monthly payment, X, satisfies

$$\frac{X}{0.05/12}\left[1 - \frac{1}{(1 + 0.05/12)^{12\times 15}}\right] = 100{,}000$$

Solving this gives X equal to USD 790.79.

18.12 The remaining principal (USD) is

$$\frac{790.79}{0.05/12}\left[1 - \frac{1}{(1 + 0.05/12)^{12\times 8}}\right] = 62{,}464$$

18.13 WAC is

$$\frac{100}{600} \times 4\% + \frac{200}{600} \times 5\% + \frac{300}{600} \times 6\% = 5.33\%$$

WAL (months) is

$$\frac{100}{600} \times 330 + \frac{200}{600} 310 \times \frac{300}{600} \times 290 = 303.33$$

18.14 FICO scores range from 300 to 850. 650 is considered the minimum acceptable score by many mortgage lenders.

18.15 CPR is an annualized prepayment rate. If SMM is the prepayment for a month, the CPR is defined as:

$$\text{CPR} = 1 - (1 - \text{SMM})^{12}$$

18.16 A TBA is a forward contract to deliver an MBS pool with particular characteristics in a future month. The following are specified: the issuer, the original maturity of the mortgages, the coupon, the price per USD 100 of par value, the par value, and the settlement month. The delivery day during the settlement month is specified by SIFMA and is usually close to the twelfth day. Two days before the delivery day the seller chooses which pool (or pools) will be delivered.

The following questions are intended to help candidates understand the material. They are _not_ actual FRM exam questions.

18.17 As prepayments increase, a PO tends to increase in value, whereas an IO tends to decrease in value.

18.18 A Monte Carlo simulation is used. Paths are sampled for interest rates and other relevant variables, such as house prices. Prepayments are calculated for each month and the cash flows from the MBS are calculated. These are discounted at the interest rates along the path to obtain a present value. This procedure is repeated many times and the value of the MBS is estimated as the average of the present values.

18.19 The OAS is the spread between the discount rate and the applicable Treasury rate that leads to the value of an MBS equaling its market price.

18.20 USD selling price is 10,325,000 + 10,000,000 × (14/30) × (0.06/12) = 10,348,333

USD buying price is 10,303,000 + 10,000,000 × (14/30) × (0.06/12) = 10,303,333

Investment of proceeds yields in USD 10,348,333 × 0.04/12 = 34,494

Payments on the pool would be USD 40,000.

The value of the roll (USD) is 10,348,333 − 10,303,333 + 34,571 − 40,000 = 39,494

19

Interest Rate Futures

Learning Objectives

After completing this reading, you should be able to:

- Identify the most commonly used day count conventions, describe the markets that each one is typically used in, and apply each to an interest calculation.
- Calculate the conversion of a discount rate to a price for a U.S. Treasury bill.
- Differentiate between the clean and dirty price for a U.S. Treasury bond; calculate the accrued interest and dirty price on a U.S. Treasury bond.
- Explain and calculate a U.S. Treasury bond futures contract conversion factor.
- Calculate the cost of delivering a bond into a Treasury bond futures contract.
- Describe the impact of the level and shape of the yield curve on the cheapest-to-deliver Treasury bond decision.
- Calculate the theoretical futures price for a Treasury bond futures contract.
- Calculate the duration-based hedge ratio and create a duration-based hedging strategy using interest rate futures.
- Explain the limitations of using a duration-based hedging strategy.

Earlier chapters have examined futures on assets such as commodities, stock indices, and foreign currencies. This chapter considers interest rate futures. These instruments are actively traded and have several features that deserve special attention.

This chapter will focus on the Eurodollar, SOFR, and Treasury bond/note futures contracts traded on the Chicago Mercantile Exchange (CME). Eurodollar futures contracts, which were based on US Libor, have traditionally been used to hedge an exposure to (or speculate on) short-term U.S. interest rates. As discussed in Chapter 16, Libor has been phased out in the US and replaced by the Secured Overnight Funding Rate (SOFR). It is therefore expected that SOFR futures will replace Eurodollar futures as the most popular contract involving short-term rates. Meanwhile, contracts on Treasury bonds and notes can be used to hedge exposures to longer-term U.S. interest rates. Note that similar contracts to those discussed in this chapter involving non-U.S. interest rates are traded on exchanges in many other countries.

19.1 DAY COUNT CONVENTIONS

Many bonds pay interest semi-annually in arrears. This means that the first coupon payment is after six months and the last coupon payment is at the end of the bond's life.

For example, suppose a bond pays coupons on March 15 and September 15 at the rate of 10% per year. If a bond is purchased on July 1, it is important to know how much of the next coupon (to be received in September 15) has already been earned. This is referred to as *accrued interest,* and it factors into the price paid for the bond.

There are several ways to calculate accrued interest. These methods, called day count conventions, depend on the underlying interest rate product.

Bonds

Two popular day count conventions for bonds are

1. Actual/actual, and
2. 30/360.

Note that both methods have the form X/Y. The first part of the day count (X) describes how the number of days during a period of interest is determined. The second part of the day count (Y) describes how the number of days during some reference period is determined.

The natural reference period for a bond is the time between coupon payment dates. In the example given above, the reference period is from March 15 to September 15. The interest earned during this period on a bond with a face value of USD 100 is USD 5 (i.e., half of the annual coupon rate of 10% applied to USD 100). What we want to know is how much interest accrues in the period between the last coupon date (March 15) and the date when the bond is sold (July 1).

Consider first the actual/actual convention, which is used for Treasury instruments in the U.S. Under this convention, the number of days in the period of interest is the actual number of calendar days, while the number of days in the reference period is the actual number of days between coupon payments. The number of days between March 15 and September 15 is 184 (16, 30, 31, 30, 31, 31, and 15 in March, April, May, June, July, August, and September, respectively). The number of days between March 15 and July 1 is 108 (16, 30, 31, 30, and 1 in March, April, May, June, and July, respectively). The accrued interest from March 15 to July 1 is therefore:

$$5 \times \frac{108}{184} = 2.9348$$

Now consider the 30/360 convention, which (as mentioned in Chapter 17) is used for corporate bonds in the U.S.[1] This method assumes 30 days per month for the period of interest and 360 days in a year. This would mean that there are 180 days between March 15 and September 15 (15, 30, 30, 30, 30, 30, and 15 in March, April, May, June, July, August, and September, respectively) and 106 days between March 15 and July 1 (15, 30, 30, 30, and 1 in March, April, May, June, and July, respectively). The accrued interest between March 15 and July 1 is therefore calculated as:

$$5 \times \frac{106}{180} = 2.9444$$

The 30/360 convention has some interesting features. For example, the number of days calculated between February 28 and March 1 is three, even though (in a non-leap year) only one calendar day has passed. This means that a bond can accrue three days of interest in just one calendar day. By the same token, a bond using the 30/360 convention would earn no interest on the thirty first of January, March, May, July, August, October, and December. (Note that the 30/360 convention was developed many years ago to simplify the arithmetic associated with day counts.)

The 30/360 convention also applies to mortgage payments. This was implicitly assumed when calculating mortgage payments in Chapter 18. Recall that the accrued monthly interest (and the interest paid by the mortgage holder) for the examples in Chapter 18 did not depend on the number of calendar days in a month.

[1] For corporate bonds denominated in euros and British pounds the day count is usually actual/actual rather than 30/360.

Money Market Instruments

In the case of money market instruments, interest is usually expressed per year. Thus, the "Y" in X/Y describes the number of days in a year while the "X" describes how the number of days in the period over which accrued interest is calculated. For Treasury bills in the U.S., the day count convention is actual/360. This means that the quoted interest rate applies to 360 days, but the days in any holding period considered is the actual number of days.

For example, suppose the interest rate is 6% per year on an actual/360 basis. This means that the interest earned in a full year of 365 days is

$$\frac{365}{360} \times 6\% = 6.0833\%$$

Similarly, the interest earned in 90 days is 1.5%.

Day count conventions vary widely from country to country. For example, money market instruments in Australia, Canada, and New Zealand are quoted on an actual/365 basis, rather than actual/360 as in the U.S.

19.2 PRICE QUOTES FOR TREASURY BONDS AND BILLS

There are several conventions for quoting the prices of bonds and money market instruments. This section explains the conventions used in the U.S. for Treasury bonds and Treasury bills. There is additional discussion of U.S. Treasury instruments in Chapters 9 to 13 of *Valuation and Risk Models*.

Bonds

The bonds issued by the U.S. government with original maturities of ten years or less are referred to as *Treasury notes*, whereas bonds with longer maturities are referred to as *Treasury bonds*[2]. However, Treasury notes and bonds are collectively referred to as bonds.

U.S. Treasury bond prices are quoted per USD 100 of principal in dollars and 32nds of a dollar. Thus, a price quoted as 105-07 would be interpreted as USD $105\frac{7}{32}$ and would mean that the cost of a bond with a face value of USD 100,000 is USD 105,218.75.[3]

The quoted price is referred to as the *clean price*. The cash price paid by the purchaser (and received by the seller) is referred to as the *dirty price*. The dirty price is calculated by adding accrued interest to the clean price:

$$\text{Dirty Price} = \text{Clean Price} + \text{Accrued Interest Since the Last Coupon Date}$$

For example, suppose a trade is settled on May 8, 2021, for a bond paying semi-annual coupons at the rate of 8% per year.[4] Suppose further that coupons are paid on April 1 and October 1.

As mentioned in the previous section, the actual/actual day count convention applies to U.S. Treasury bonds. There are 183 days between the two coupon dates (April 1, 2021, and October 1, 2021) and 37 days between the last coupon payment and the settlement date (April 1, 2021, and May 8, 2021). The payments on each coupon date (per USD 100 of principal) are USD 4 and so the accrued interest is

$$4 \times \frac{37}{183} = 0.8087$$

If the quoted price (clean price) is 105-08 (i.e., USD 105.25), the dirty price is

$$105.25 + 0.8087 = 106.0587$$

The price paid for a bond with a face value of USD 100,000 would therefore be USD 106,058.70.

Treasury Bills

Recall that interest rates on U.S. Treasury bills are calculated using an actual/360 convention. However, the interest rate is also expressed on a *discount basis*. This means that it is expressed as a percentage of the final value received by the investor (rather than the initial amount). The final value received by the holder of a Treasury bill is the face value of the bill.

For example, consider a 120-day Treasury bill with a face value of USD 100,000 that is quoted as bid 4.08, ask 4.05. The bid price indicates that interest will be paid at a rate of 4.08% per

[2] Recall from Chapter 17 that the convention is a little different for corporate bonds. Corporate bonds with a maturity of less than five years are referred to as short-term notes; those with maturities of between five and 12 years are referred to as medium-term notes; those with maturities of greater than 12 years are referred to as long-term bonds.

[3] Some very active issues are quoted to the nearest 1/64. If a plus sign follows a quoted price it means that 1/64 must be added. Thus "$105\frac{7}{32}+$" should be interpreted as $105\frac{15}{64}$.

[4] The settlement date is two days after the trade date, and it is the accrued interest to the settlement date rather than the trade date that is used.

360 days on USD 100,000. Because there are 120 days to maturity, the USD interest earned is

$$\frac{120}{360} \times 0.0408 \times 100{,}000 = 1{,}360$$

The USD bid price of the Treasury bill is the face value minus the interest:

$$100{,}000 - 1{,}360 = 98{,}640$$

The Treasury bid price of 4.08 therefore means that the Treasury bill could be sold for USD 98,640. Similarly, the ask quote of 4.05 means that the Treasury bill could be bought for USD 98,650. If Q is the quote and C is the corresponding cash price per USD 100 of face value:

$$Q = \frac{360}{n}(100 - C)$$

or

$$C = 100 - \frac{n}{360}Q$$

In this example, we can use this formula to turn the bid quote into a dollar price by setting $n = 120$ and $Q = 4.08$ to get

$$C = 100 - \frac{120}{360} \times 4.08 = 98.64$$

Note that we have the seemingly strange situation where the bid quote is higher than the ask quote. This is because the bid quote is an interest rate and not a price. When interest rates decrease, prices increase (and vice versa).

19.3 TREASURY BOND FUTURES

Table 19.1 lists some of the Treasury contracts that trade on the CME. Delivery can take place at any time during the delivery month. This party with the short position chooses which bond will be delivered and when (during the delivery month) the delivery will occur.

The price received for a delivered bond is determined (in part) by what is called the *conversion factor*. Roughly speaking, the conversion factor is hypothetical clean price for a bond with a face value of one dollar given that all interest rates are 6% with semi-annual compounding. Thus, the price received for a bond is

$$Sf + A$$

where S is the most recent settlement price in the futures contract, f is the bond's conversion factor (as explained in more detail below), and A is the accrued interest calculated on an actual/actual basis (as described earlier in this chapter).

Table 19.1 Selection of the Treasury Bond and Treasury Note Futures that Trade on the Chicago Mercantile Exchange

Futures Contract	Deliverable
Two-Year Treasury Note	Any Treasury note with an original maturity not greater than five-years and three-months. The remaining maturity must not be less than one-year and nine-months on the first day of the delivery month and not greater than two-years on the last day of the delivery month.
Five-Year Treasury Note	Any Treasury note with an original maturity not greater than five years and three-months and a current maturity of not less than four-years and two-months on the first day of the delivery month
Ten-Year Treasury Note	Any Treasury note with a remaining maturity between 6.5 years and ten-years
Treasury Bond	Any bond with a remaining maturity between 15 and 25 years
Ultra Treasury Bond	Any Treasury bond with a remaining maturity greater than 25 years

As an example, suppose the most recent settlement price for a Treasury bond futures contract is USD 105.50, the conversion factor is 1.1542, and the accrued interest on the bond at the time of delivery is USD 1.12 per USD 100 of face value. The cash (USD) received by the party with a short position (who is delivering the bond) is

$$105.50 \times 1.1542 + 1.12 = \text{USD } 122.888$$

For nearly all Treasury bonds/notes, one futures contract is for the delivery of bonds with a face value of USD 100,000.[5] Therefore, the party with a short position in this example delivers bonds with a face value of USD 100,000 and receives USD 122,888 in return.

To calculate the conversion factor for bonds that are deliverable in the last three contracts in Table 19.1:

1. Calculate the time to maturity from the first day of the delivery month to the maturity of the bond,
2. Round the time to maturity calculated in Step 1 down to the nearest three months, then
3. Calculate the clean price per one dollar of face value for a bond with the maturity calculated in Step 2 and a yield of 6% per annum (with semi-annual compounding).

[5] In the case of the two-year note futures, one contract is for the delivery of USD 200,000 face value.

For example, consider a September 2021 Treasury bond futures contract on a 5% coupon bond maturing on May 15, 2040. The time to maturity on the first day of the delivery month (September 1, 2021) is 18 years (i.e., September 1, 2021, to September 1, 2039) and 8.5 months (i.e., September 1, 2039, to May 15, 2040). This is rounded down to 18 years and six months. The price of a bond with a maturity of 18 years and six months[6] with an annual yield of 6%, and a face value of USD 100 is

$$\sum_{i=1}^{37} \frac{2.5}{1.03^i} + \frac{100}{1.03^{37}} = 88.92$$

This amount represents both the clean and the dirty price (because there is no accrued interest), and the conversion factor is therefore 0.8892.

Now consider the contract for December 2021. A 5% coupon bond maturing on May 15, 2040, is also deliverable in this contract. In this case, the time to maturity (measured from the first day of the delivery month) is 18 years and 5.5 months. This is rounded down to 18 years and three months. The dirty price of the bond just *before* the next coupon is received is

$$\sum_{i=0}^{36} \frac{2.5}{1.03^i} + \frac{100}{1.03^{36}} = 91.5839$$

when the yield is 6%. While this bond has 18 years to maturity, the dirty price when the bond is 18 years and three months from maturity can be calculated by discounting this price for a further three months. The discount factor is $1/\sqrt{1.03}$ or 0.98533 so that the dirty price is USD 90.24 (= 91.5839 × 0.98533). After subtracting the accrued interest of USD 1.25, the clean price of the bond is USD 88.99. The conversion factor is therefore 0.8899.[7]

The conversion factors for bonds delivered in the two- and five-year note futures contracts are calculated similarly except that the time to maturity is rounded down in Step 2 to the nearest month.[8]

Quotes

Bond futures prices are quoted similarly to bond prices. However, the shorter-maturity contracts are quoted with greater precision. The two-year contract is quoted to the nearest 1/256, the five-year contract is quoted to the nearest 1/128, and the ten-year contract is quoted to the nearest 1/64.

Table 19.2 March 2020 Treasury Bond Futures Quoted Prices on February 26, 2020

Contract	Quoted Price
Two-Year Treasury Note	108–125
Five-Year Treasury Note	121–007
Ten-Year Treasury Note	132–305
Treasury Bond	167–12
Ultra Treasury Bond	200–26

Source: www.cmegroup.com

Specifically, a 5 in the third decimal place indicates 0.5/32; a 7 in the third decimal place indicates 0.75/32, a 2 in the third decimal place indicates 0.25/32; a 1 in the third decimal place indicates 0.125/32; a 3 in the third decimal place indicates 0.375/32; a 6 in the third decimal place indicates 0.625/32; an 8 in the third decimal place indicates 0.875/32.

Table 19.2 shows quoted prices for some March 2020 bond futures contracts in Table 19.1 on February 26, 2020. The two-year quoted price (108–125) should be interpreted as:

$$108\frac{12.5}{32} = 108.390625$$

Meanwhile, the five-year futures price (121-007) should be interpreted as:

$$121\frac{0.75}{32} = 121.023438$$

Futures prices increase with the maturity of the contract in Table 19.2. This can be understood by noting that the conversion factor system means that the underlying asset is a bond with a coupon of 6% per year (because that bond has a conversion factor of 1.0000). However, interest rates for all maturities were much less than 6% on February 26, 2020. When this is the case, the value of a bond paying a coupon of 6% per year increases with maturity (because the higher 6% coupon is received for a longer period).

Cheapest-to-Deliver Bond Option

The cost of delivering a bond or note is

$$\text{Market Price} - \text{Price Received}$$

Not surprisingly, the party with a short position will choose the bond for which this cost is least. This bond is referred to as the *cheapest-to-deliver bond*. As mentioned, the price received for a bond is

[6] This is assumed to be immediately after a coupon payment so that there is no accrued interest.

[7] Both conversion factors we have calculated correspond exactly with those published by the CME Group.

[8] Suppose for example that, when the time to maturity is rounded down to the nearest month, there are three years and eight months to maturity. We calculate the value of a 6% coupon bond with three years and six months to maturity just before the payment of the coupon. We then divide by $1/1.03^{2/6}$ to add an extra two months to the life. Four months of accrued interest are then subtracted to move from the dirty price to the clean price.

$$Sf + A$$

where S is the most recent settlement price in the futures contract, f is the conversion factor, and A is the accrued interest. The cash market price of the bond is

$$Q + A$$

where Q is the quoted price. The cheapest-to-deliver bond is therefore the one for which:

$$Q - Sf$$

is the least.

Suppose three bonds that can be delivered are those shown in Table 19.3 and that the most recent futures settlement price is USD 115.75. As indicated by the calculations in the table, the first bond is the cheapest-to-deliver.

Several factors can determine which bond is the cheapest to deliver. For example, if bond yields are greater than 6%, low-coupon long-maturity bonds will tend to be the cheapest to deliver. If yields are less than 6%, however, high-coupon short-maturity bonds will tend to be the cheapest to deliver. An upward-sloping yield curve tends to favor long-maturity bonds, whereas a downward-sloping yield curve tends to favor short-maturity bonds.

The party with the short position chooses when during the delivery month to deliver. This gives rise to an option known as the *wild card play*. Note that the settlement price is the price at which the futures contract trades at 2 p.m. Chicago time. However, a notice of intention to deliver can be issued later in the day. The party with the short position can therefore attempt to wait for a day when bond prices decline after 2 p.m. and therefore reduce the delivery cost. For example, the bond can be bought at the 3:30 p.m. price and sold at the 2 p.m. futures price.

The wild card play and the ability to use the cheapest-to-deliver bond make the futures contract more attractive to the party with the short position (because they provide that party with a high degree of choice). Additionally, the party with the short position can choose to deliver on any day during the delivery month. Together, these features tend to reduce the futures price; as the contract becomes more attractive to the party with the short position, that party becomes more prepared to accept a lower price.

Calculating the Futures Price

The fact that the party with the short position has so many delivery choices makes it difficult to determine futures prices (such as those in Table 19.1). However, assuming that both the cheapest-to-deliver bond and the bond delivery time are known, the Treasury bond futures contract is simply a futures contract on an asset providing known income. As explained in Chapter 10, the futures price is

$$(S - I)(1 + R)^T$$

where I is the present value of the coupons that will be received during the life of the futures contract, S is the current price of the bond, T is the life of the futures contract, and R is the (annually compounded) risk-free rate for maturity T. As explained in Chapter 16, if R is expressed with continuous compounding, the futures price becomes

$$(S - I)e^{RT} \qquad (19.1)$$

The current price (S) used these equations must be the dirty cash price of the bond. (Thus, the calculated futures price is also a dirty price). The steps in calculating the futures price are therefore as follows.

1. Calculate the current dirty price of the delivered bond from the quoted clean price by adding accrued interest.
2. Calculate the present value of coupons that will be received between the current time and the time when the bond will be delivered.
3. Obtain a dirty futures price by subtracting the amount in Step 2 from the amount in Step 1 and compounding forward at the risk-free rate to the time when the bond will be delivered.
4. Convert the bond's dirty futures price to a clean futures price by subtracting the accrued interest at the time of the delivery.
5. The clean futures price for the delivered bond equals the quoted futures price multiplied by the conversion factor. The quoted futures price is therefore calculated as the value in Step 4 divided by the bond's conversion factor.

Table 19.3 Calculation of Cost of Delivery for Three Bonds When the Futures Price Is 115.75

Bond	Quoted (Clean) Bond Price (in USD)	Conversion Factor	Cost of Delivery (in USD)
1	98.75	0.8384	$98.75 - 115.75 \times 0.8384 = 1.7052$
2	117.25	0.9874	$117.25 - 115.75 \times 0.9874 = 2.9585$
3	145.75	1.2325	$145.75 - 115.75 \times 1.2325 = 3.0881$

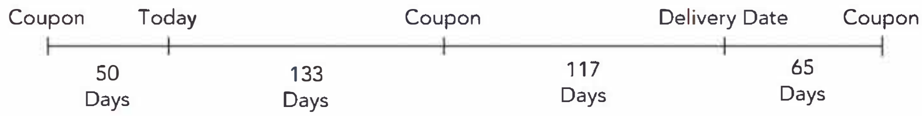


Figure 19.1 **Timeline for example.**

Suppose that it is known that a bond will be delivered in 250 days under the terms of a futures contract. The timing of coupon payments is shown in Figure 19.1. The last coupon on the bond was paid 50 days ago, and the next coupon will be paid in 133 days. The coupon after the next one will be paid in 315 days (i.e., 65 days after delivery).

Assume that the risk-free interest rate for all maturities is 5% with continuous compounding and that the delivered bond pays a coupon of 7% semi-annually with a current quoted (clean) price of USD 108.00 and a conversion factor of 1.0400. The cash (dirty) price is calculated by noting that the next coupon will be USD 3.5 (per USD 100 face value) and that that 50 of the 183 days between coupon payments have passed. It is therefore:

$$108.00 + \frac{50}{183} \times 3.5 = 108.9563$$

This is the variable S in Equation (19.1). The present value of the coupon that will be received after 133 days (133/365 = 0.3644 years) is

$$3.5e^{-0.05 \times 0.3644} = 3.4368$$

This is the variable I in Equation (19.1).

The time to delivery for the futures contract is 250 days (i.e., 250/365 = 0.6849 years). From Equation (19.1), the cash (dirty) futures price is

$$(108.9563 - 3.4368)e^{0.05 \times 0.6849} = 109.1957$$

To obtain the clean futures price, we must subtract the accrued interest. On the delivery date, 117 of the 182 days between payments will have passed and so the clean futures price is

$$109.1957 - \frac{117}{182} \times 3.5 = 106.9457$$

Dividing by the conversion factor, we obtain an estimated price of the futures contract:

$$\frac{106.9457}{1.0400} = 102.83$$

19.4 SOFR FUTURES

The three-month Eurodollar futures contract (often just referred to as the Eurodollar futures contract) has historically been a very popular contract traded by the CME Group. The settlement of the contract is in cash and based on the three-month USD Libor fixing two days before the third Wednesday of the contract month. The contract trades for maturity dates up to ten years into the future.

SOFR Futures Details

As discussed in Chapter 16, Libor has in the past been used for trillions of dollars of transactions throughout the world. LIBOR has now been discontinued and has been replaced by SOFR in the United States. Eurodollar futures, which traded very actively have been replaced by SOFR futures. As explained in Chapter 16, a difference between Libor and SOFR is that Libor is known at the beginning of the period to which it relates while SOFR is not known until the end of the period.

Table 19.4 shows settlement prices and contracts traded for the three-month SOFR futures on April 28, 2023. Note that the March 2023 contract still traded on April 28 because not all the overnight rates that determine the settlement in June 2023 had been observed. Contracts are for periods up to ten years in the future. The table indicates that contracts that last up to five years trade actively. SOFR futures are proving to be as popular as Eurodollar futures used to be.

SOFR futures can be used to hedge exposures to SOFR interest rates. Suppose that at the end of April 2021 an investor has agreed to pay the three-month SOFR rate plus 200 basis points on $100 million of borrowing for three months beginning in December 2021. From Table 20.2, the December 2021 quote is 99.94. This corresponds to a SOFR interest rate of 6 basis points or 0.06%. If the investor thinks that rate may increase and wants to lock in a borrowing rate of 2.06%, the investor can short 100 December contracts. If rates increase, futures contracts will decline and the gain on the futures contracts will offset the extra interest being paid.

For example, suppose that the final settlement of the December 2021 contract (in March 2022) is 99.20. This corresponds to a SOFR rate for the three months of 0.8%. The rate paid on the loan would be 2.8% so that the total interest would be 0.25 × $100,000,000 × 0.028 or $700,000. The futures contract has declined by 74 basis points. There is a gain of $25 per basis point per contract. The total gain is therefore:

$$100 \times \$25 \times 74 = \$185{,}000$$

When this is taken into account, the amount of interest paid is reduced to $515,000, which corresponds to a rate of 2.06% (0.25 × 0.0206 × $100,000,000 = $515,000).

Table 19.4 Volume of trading and settlement prices of SOFR three-month futures contract on April 28, 2023. The gain/loss on one contract is $25 per basis point move

Contract Month	Final Settlement price	No.,of Contracts Traded
March 2023	95.0725	111,247
June 2023	94.8750	288,402
September 2023	95.0950	210,797
December 2023	95.4650	200,941
March 2024	95.9550	141,565
June 2024	96.4000	140,272
September 2024	96.7200	90,176
December 2024	96.9000	98,460
March 2025	96.9950	75,031
June 2025	97.0350	59,855
September 2025	97.0500	42,203
December 2025	97.0550	56,563
March 2026	97.0350	33,614
June 2026	97.0150	22,182
September 2026	96.9800	18,417
December 2026	96.9500	19,024
March 2027	96.9150	17,067
June 2027	96.8850	15,537
September 2027	96.8400	17,133
December 2027	96.8000	21,053
March 2028	96.7650	17,053
June 2028	96.7300	270
September 2028	96.7000	156
December 2028	96.6650	181
March 2029	96.6300	162
June 2029	96.5900	386
September 2029	96.5550	297
December 2029	96.5300	238
March 2030	96.5100	185
June 2030	96.4850	105
September 2030	96.4500	2
December 2030	96.5650	0
March 2031	96.5600	0
June 2031	96.5450	0
September 2031	96.5250	0
December 2031	96.5150	0
March 2032	96.5050	0
June 2032	96.4950	0
September 2032	96.4700	0
December 2032	96.4600	0
March 2033	96.4500	0

Calculating Zero Curves

The futures rate calculated from a SOFR futures contract is not the same as the corresponding forward rate because the futures contact is subject to daily settlement.[9] Adjustments to convert a futures rate to the corresponding forward rate have been proposed and are referred to as convexity adjustments.[10]

As explained in Section 16.9, the relationship between forward rates and spot rates when all interest rates are expressed with continuous compounding is

$$F = \frac{R_2T_2 - R_1T_1}{T_2 - T_1}$$

where F is the forward rate for the period between T_1 and T_2, R_1 is the zero-coupon interest rate for a maturity of T_1, and R_2 is the zero-coupon interest rate for maturity T_2.

This formula can be rearranged to give

$$R_2 = \frac{F(T_2 - T_1) + R_1T_1}{T_2} \tag{19.2}$$

This means that if we know the zero-coupon interest rate for a maturity of T_1 and the forward rate between T_1 and T_2, we can deduce the zero-coupon interest rate for a maturity of T_2. For example, if the two-year zero rate is 4% and the forward rate for

[9] Daily settlement means that the hedge just discussed is not perfect.

[10] See, for example, M.P.A Henrard, Overnight Futures: Convexity Adjustment (February 2018) ssrn 3134346.

the period between two and 2.5 years is 4.4%, we can deduce that the 2.5-year zero rate is

$$\frac{0.044 \times 0.5 + 0.04 \times 2}{2.5} = 0.0408$$

or 4.08%.

Equation (19.2) has historically been used to help provide a complete term structure of interest rates for Libor. In the future it can do the same for SOFR.

19.5 DURATION-BASED HEDGING

It will be recalled that we introduced the duration measure in Chapter 16. If D is the modified duration of a bond portfolio worth B and Δy is a small parallel shift in the yield curve, the change in the value of the portfolio is approximately:

$$-BD\Delta y$$

Suppose that a futures contract is used to hedge an exposure when the maturity of the instrument to be hedged does not match the maturity of the instrument underlying the futures contract. If it is assumed that only small parallel shifts in the yield curve will occur, duration can be used to calculate an appropriate hedge ratio. Define

E_F: increase in value of one futures contract for a one basis point downward parallel shift in the zero curve.

E_V: increase in value of a trader's position for a one basis point downward parallel shift in the zero curve.

The number of long contracts that should be traded (with a negative number indicating a short position) for a hedge should be

$$-\frac{E_V}{E_F} \tag{19.3}$$

The gain (loss) on the futures position will then be offset by the loss (gain) on the trader's portfolio for a small parallel shift in the term structure of interest rates.

Suppose that a trader has a USD 2 million position in a nine-month money market instrument. The duration of the instrument is 0.75 because it is a zero-coupon instrument. From the duration relationship, the trader's gain (in USD) from a one-basis point downward parallel shift in the zero curve is approximately:

$$E_V = 2{,}000{,}000 \times 0.75 \times 0.0001 = 150$$

The three-month SOFR futures contract is designed so that a one basis downward parallel shift in the yield curve gives rise to a gain of USD 25. This means that $E_F = 25$. Equation (19.3) gives the number of contracts required as:

$$-\frac{150}{25}$$

A short position in six contracts will therefore provide an approximate hedge.[11]

As a second example, suppose that a fund manager has USD 5 million position in bonds with a duration of 12. The cheapest-to-deliver bond in a Treasury bond futures contract is estimated to have a duration of 16 at maturity of a bond futures contract. Also assume that the current futures price is 103.25. The duration relationship estimates the impact of a one basis point downward parallel shift in the yield curve on the fund manager's position as:

$$E_V = 5{,}000{,}000 \times 12 \times 0.0001 = 6{,}000$$

One futures contract involves the delivery of bonds with a face value of USD 100,000. The value of one futures contract is therefore USD 103,250. Assume that we know which bond will be delivered and that it will have a duration of D at maturity of the futures. When there is a parallel downward shift of Δy in the yield curve, the futures price can be expected to change by approximately:[12]

$$FD\Delta y$$

The impact of a one basis point downward parallel shift in the zero curve on the futures price in our example is therefore:

$$E_F = 103{,}250 \times 16 \times 0.0001 = 165.2$$

Equation (19.3) gives the number of contracts required as:

$$-\frac{6{,}000}{165.2} = -36.3$$

Rounding to the nearest whole number, 36 contracts should be shorted.

[11] There are a number of approximations. 0.75 is not a modified duration. It will therefore be evident from the discussion in Chapter 16 that the E_V calculation is theoretically correct only when we are considering continuously compounded rates. We have not considered the impact of day count conventions. However, all these points are likely to have a relatively small effect on the performance of the hedge.

[12] Because there will be a proportional increase $D\Delta y$ in the bond that will be delivered, we can expect the same proportional increase in the futures price.

It should be emphasized that the duration-based hedges we have considered in this section provide protection only against small parallel shifts in the term structure of interest rates. They also assume that the interest rate to which the hedger is exposed is perfectly correlated to the interest rate underlying the futures contract and (in the case of hedging using bond futures) that the bond that will be cheapest-to-deliver is known with certainty.

SUMMARY

The SOFR futures contract provides payoffs dependent on movements in short-term interest rates, while Treasury note and Treasury bond futures contracts provide payoffs dependent on movements in longer-term rates. Final settlement in the SOFR futures contract is in cash. By contrast, Treasury note and bond futures are settled by delivering a bond or a note. The party with the short position can choose which bond/note to deliver and decide when the delivery will take place. A conversion factor determines the price received by the party with the short position from the party with a long position for the delivered instrument.

When futures contracts are used for hedging, the duration of the portfolio being hedged may not be the same as the duration of the assets underlying the futures contract that is used for hedging. The number of futures contracts used for hedging can be adjusted by multiplying by the ratio of the duration of the portfolio being hedged to the duration of the assets underlying the futures contract.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

QUESTIONS

Short Concept Questions

19.1 What is the day convention for (a) Treasury bonds, (b) corporate bonds and, (c) Treasury bills in the United States?

19.2 What is the difference between the clean price and the dirty price of a bond?

19.3 How is a Treasury note defined in the United States?

19.4 In an upward sloping yield curve environment where interest rates are above 6% what types of bonds tend to be cheapest to deliver?

19.5 How should a quote of 103–132 for a Treasury bond futures be interpreted?

19.6 How much does someone with a short contract gain or lose when the three-month SOFR futures price changes from 94.555 to 94.715?

19.7 Why does the September SOFR futures contract for a certain year still trade in October of that year?

19.8 Why were Eurodollar futures discontinued?

19.9 What is SOFR. How is the three-month futures contract on SOFR settled?

19.10 What protection is obtained when a Treasury bond futures contract is used to hedge a bond portfolio using duration analysis? What assumptions are necessary?

Practice Questions

19.11 A Treasury bond pays coupons at the rate of 7% per year on June 1 and December 1. What is the accrued interest between June 1 and July 31 per USD 100 of face value?

19.12 What is the answer to Question 19.11 if the bond is a corporate bond rather than a Treasury bond?

19.13 What is the cash price of the Treasury bill that lasts for 200 days and has a quoted price of 4.12?

19.14 A bond that can be delivered in the December 2018 ten-year Treasury note futures contract is a bond with maturity on April 15, 2026, that pays a coupon of 4% per annum. Calculate the conversion factor for the bond.

19.15 A futures price is 115.00. Three bonds that can be delivered have quoted prices of 98–125, 103–127, and 120–230. The conversion factors of the bonds are 0.8456, 0.8844, and 1.0267, respectively. What is the cost of delivering each bond? Which is the cheapest to deliver?

19.16 Assume that the bond that will be cheapest to deliver in a Treasury bond futures contract pays semiannual coupons at the rate of 10% per annum on May 1 and November 1 and will be delivered on September 1. The bond's quoted price on August 1 is 130.00 and its conversion factor is 1.2341. Estimate the futures price on August 1 assuming that all interest rates are 4% (continuously compounded).

19.17 The SOFR futures quote for June of a certain year is 98.36. How should it be interpreted? What does it indicate about interest rates?

19.18 If in Question 19.17 the continuously compounded three-year zero rate is 4.12%, what is the continuously compounded 3.25 year rate?

19.19 Approximately how many three-month SOFR futures contracts are necessary to hedge the six-month interest that will be paid on a USD 20 million bond? Assume that the six-month period starts at the maturity of the futures contract that will be used.

19.20 It is March 10, 2019. The cheapest-to-deliver bond in the December 2019 Treasury bond futures contract is expected to be a bond with a duration of 12.5 at that time. The current futures price is 104–127. A bond portfolio has a duration of 15 and is worth USD 30 million. How many bond futures contracts are necessary to hedge the risk in the portfolio?

The following questions are intended to help candidates understand the material. They are not actual FRM exam questions.

ANSWERS

Short Concept Questions

19.1 (a) actual/actual, (b) 30/360, (c) actual/360.

19.2 The clean price is the quoted price. The dirty price is the cash price that is paid. The dirty price equals the clean price plus accrued interest.

19.3 A Treasury note is a bond with an original maturity less than or equal to ten years.

19.4 Low coupon long maturity bonds will tend to be cheapest to deliver.

19.5 $103\frac{13.25}{32} = 103.414063$

19.6 The change is an increase of 16 basis points ($94.715 - 94.555 = 0.16$). A party with a short contract loses $16 \times$ USD 25 = USD 400.

19.7 The September SOFR contract settlement depends on overnight interest rates during a three-month period beginning in September. The settlement interest rate is therefore uncertain in October.

19.8 Eurodollar futures provided a payoff dependent on Libor and Libor was discontinued.

19.9 SOFR is the overnight repo rate. The three-month futures is settled by calculating the rate that would have been earned over the previous three months by rolling an investment forward day-by-day at the SOFR rate.

19.10 The hedge protects against small parallel shifts in the zero curve. The following assumptions must be made: The cheapest-to-deliver bond is known and movements in the rates to which the portfolio is exposed are very similar to movements in the corresponding Treasury rates.

Solved Problems

19.11 The number of days between June 1 and December 1 is 183 (29, 31, 31, 30, 31, 30, and 1 in June, July, August, September, October, November, and December, respectively). The number of days between June 1 and July 31 is 60 (29 and 31 in June and July, respectively). The accrued interest is therefore:

$$3.5 \times \frac{60}{183} = 1.1475$$

19.12 Using a 30/360 day count, the number of days between June 1 and December 1 is 180 (29, 30, 30, 30, 30, 30, and 1 in June, July, August, September, October, November and December, respectively). The number of days between June 1 and July 31 is 59 (29 and 30 in June and July, respectively). The accrued interest is therefore:

$$3.5 \times \frac{59}{180} = 1.1472$$

19.13 The cash price per USD 100 of face value is

$$100 - \frac{200}{360} \times 4.12 = 97.7111$$

19.14 The bond's time to maturity on the first day of the delivery months is seven years (December 2018 to December 2025) and 4.5 months (January 2026 to mid-April 2026). This is rounded to seven years and three months. The dirty price of a seven-year and three-month bond immediately before the coupon payable in three months is

$$\sum_{i=0}^{14} \frac{2}{1.03^i} + \frac{100}{1.03^{14}} = 90.7039$$

when the yield is 6%. The dirty price of the bond three months earlier is

$$\frac{90.7039}{\sqrt{1.03}} = 89.3732$$

Subtracting the accrued interest of 1, we get a clean price of 88.3732 and the conversion factor is 0.8837.

19.15 The prices of the three bonds are

$$98\frac{12.5}{32} = 98.3906, \quad 103\frac{12.75}{32} = 103.3984,$$

and $120\frac{23}{32} = 120.7187$

The costs of delivering the bonds are

$$98.3906 - 0.8456 \times 115 = 1.1466$$
$$103.3984 - 0.8844 \times 115 = 1.6924$$
$$120.7187 - 1.0267 \times 115 = 2.6482$$

The first bond is the cheapest to deliver.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

19.16 There are 92 days between May 1 and August 1 (30, 30, 31, and 1 in May, June, July and August, respectively) and 184 days between May 1 and November 1 (30, 30, 31, 31, 30, 31, and 1 in May, June, July, August, September, October and November, respectively). The dirty price of the bond is therefore:

$$130 + 5 \times \frac{92}{184} = 132.5$$

No coupons will be paid in the 31-day period between August 1 and September 1. The time to delivery is 31/365 = 0.0849 years. The dirty futures price is therefore:

$$132.5\, e^{0.0849 \times 0.04} = 132.9509$$

The accrued interest on September 1 is 5 × 123/184 = 3.3424. The clean futures price is therefore:

$$132.9509 - 3.3424 = 129.6085$$

Dividing by the conversion factor we obtain the estimated futures price as:

$$\frac{129.6085}{1.2341} = 105.0227$$

19.17 The SOFR quote is indicative of a three-month futures rate of 100 – 98.36 or 1.64%. The rate considered is calculated by compounding overnight rates for three months beginning June.

19.18 Using equation (19.17) we get

$$\frac{0.04255 \times 0.25 + 0.0412 \times 3}{3.25} = 0.0413$$

or 4.13%.

19.19 The change in the value of the instrument for a one basis point parallel shift in the interest rate is

USD 20,000,000 × 0.5 × 0.0001 = USD 1,000

This is 40 times USD 25. It follows that 40 contracts should be shorted.

19.20 The futures price of the bond is

$$104\frac{12.75}{32} = 104.3984$$

and the value of one futures contract is USD 104,398.4. The number of contracts necessary to hedge the portfolio is

$$\frac{30{,}000{,}000 \times 15}{104{,}398.4 \times 12.5} = 344.8$$

or 345 when rounded to the nearest whole number.

20

Swaps

Learning Objectives

After completing this reading, you should be able to:

- Explain the mechanics of a plain vanilla interest rate swap and compute its cash flows.
- Explain how a plain vanilla interest rate swap can be used to transform an asset or a liability and calculate the resulting cash flows.
- Explain the role of financial intermediaries in the swaps market.
- Describe the role of the confirmation in a swap transaction.
- Describe the comparative advantage argument for the existence of interest rate swaps and evaluate some of the criticisms of this argument.
- Explain how the discount rates in a plain vanilla interest rate swap are computed.
- Calculate the value of a plain vanilla interest rate swap based on two simultaneous bond positions.
- Calculate the value of a plain vanilla interest rate swap from a sequence of FRAs.
- Explain how a currency swap can be used to transform an asset or a liability and calculate the resulting cash flows.
- Calculate the value of a currency swap based on two simultaneous bond positions.
- Calculate the value of a currency swap based on a sequence of forward exchange rates.
- Identify and describe other types of swaps, including commodity, volatility, credit default, and exotic swaps.
- Describe the credit risk exposure in a swap position.

Swaps are over-the-counter (OTC) derivatives contracts where the parties agree to exchange certain cash flows in the future. The exchanges of cash flows generally depend on the values that variables such as interest rates, exchange rates, equity prices, and commodity prices will have in the future.[1] As a result, there is some uncertainty associated with the exchanges.

A forward contract is a swap where there will be a cash flow exchange on just one future date. Swaps often feature exchanges on many future dates. And while the value of a swap is normally zero (or very close to zero) when it is initiated, the value of each exchange made in the swap is typically not zero. At the time the swap is initiated, it is usually the case that some exchanges have positive values while others have negative values.

In this chapter, we examine interest rate and currency swaps in some detail and then briefly introduce some of the many other types of swaps that are traded. As statistics produced by the Bank for International Settlements show, interest rate swaps are the most popular OTC derivatives contracts.[2]

20.1 MECHANICS OF INTEREST RATE SWAPS

An interest rate swap involves a fixed rate of interest being exchanged for a floating rate such as SOFR or SONIA. Party A might agree to pay Party B a fixed rate of interest of 2% (quarterly compounded) on USD 100 million for three years while in return Party B agrees to pay interest at three-month SOFR on the same principal over the same period. In this example interest would be exchanged every three months. The principal of USD 100 million is referred to as the notional principal because it is never actually exchanged. It is merely used to calculate the interest payments that are exchanged.

When the reference rate is calculated from realized overnight rates the swap is referred to as an overnight indexed swap (OIS). OISs that last for a year or less typically lead to a single exchange at the end of the life of the contract. For example, in a six-month OIS, the rate implied for the six months by the overnight rates would be calculated (see Section 16.1) and this would lead to an exchange at the end of the contract.

[1] The variables are often referred to as "risk factors" by market participants.

[2] See www.bis.org.

Overnight Indexed Swaps

As Libor is being phased out, swaps based on overnight rates are becoming more popular. These are referred to as overnight indexed swaps (OISs). A key difference between Libor swaps and OISs concerns when the rate paid is known. The Libor rate for a period is known at the beginning of the period. The OIS rate for the period is not known until the end of the period. In both cases, the rate determines an exchange of payments at the end of the period.

OISs that last for a year or less typically lead to a single exchange at the end of the life of the contract. Consider for example a six-month OIS. At the end of the six months, the rate implied by overnight rates would be calculated. This would usually be done, as described in Section 16.1, using daily compounding.

OISs that last longer than a year typically involve quarterly payments. Consider a hypothetical two-year OIS initiated on March 8, 2022, between Apple and Citigroup. We suppose Apple agrees to pay to Citigroup interest at the rate of 3% per annum every three months on a notional principal of $100 million, and in return, Citigroup agrees to pay Apple the three-month SOFR floating reference rate on the same notional principal. (Apple is referred to as the fixed-rate payer, while Citigroup is referred to as the floating-rate payer.) For ease of exposition, we assume that rates are quoted with quarterly compounding and we ignore the impact of day count conventions and holiday conventions (which are explained later in this chapter).

In total there are eight exchanges on the swap. One possible outcome is shown in Table 20.1. The first exchange of payments would take place on June 8, 2022. Apple would pay Citigroup $750,000 (one quarter of 3% applied to $100 million). Citigroup would pay Apple the three-month rate calculated from daily SOFR rates over the previous three months. The three-month SOFR rate calculated on June 8, 2022, (from overnight rates between March 8, 2022 and June 8, 2022) is assumed to be 2.2% per annum or 0.55% per three months. The floating payment from Citigroup to Apple is therefore $550,000. The fixed and floating payments are netted with the result that Apple pays Citigroup $200,000 on June 8, 2022. The second exchange of payments is on September 8, 2022. Again, the fixed payment from Apple to Citigroup is $750,000. We assume that the three-month SOFR rate on September 8, 2022, (calculated from overnight rates between June 8, 2022, and September 8, 2022) is 2.6% per annum or 0.65% per three months. The floating payment is therefore $650,000 and the net payment from Apple to Citigroup is therefore $100,000.

Table 20.1 Cash Flows to Apple for One Possible Outcome of an OIS Initiated on March 8, 2022. The Swap Lasts Two Years, and The Notional Principal is $100 Million. The SOFR Rate is Calculated by Compounding Daily Rates Over The Previous Three Months, as Described in Section 16.1

Date	SOFR Rate (%)*	Floating Cash Flow Received ($'000)	Fixed Cash Flow Paid ($'000s)	Net Cash Flow ('000s)
June 8, 2022	2.20	550	750	−200
Sept 8, 2022	2.60	650	750	−100
Dec 8, 2022	2.80	700	750	−50
Mar 8, 2023	3.10	775	750	+25
Jun 8, 2023	3.30	825	750	+75
Sept 8, 2023	3.40	850	750	+100
Dec 8, 2023	3.60	900	750	+150
Mar 8, 2024	3.80	950	750	+200

Day Counts and Holidays

The cash flows in Table 20.1 are approximations because they do not consider day counts. (See Section 19.1 for a discussion of day counts). Floating reference rates such as SOFR are quoted on an actual/360 basis. The first floating payment in Table 20.1 is based on a rate of 2.2%. Because there are 92 days between March 8, 2022, and June 8, 2022, USD floating payment would be

$$\$100{,}000{,}000 \times 0.022 \times \frac{92}{360} = 562{,}222$$

This is a little different from the approximate estimate of USD 550,000 in Table 20.1. The fixed interest rate of 3% is also expressed with a particular day-count convention. Among the possibilities here are actual/365 and 30/360. If the day count is 30/360, the fixed-rate cash flows in Table 20.1 are correct (except for the impact of holidays and weekends). If the day count is actual/365 the first fixed payment would be

$$\$100{,}000{,}000 \times 0.03 \times \frac{92}{365} = \$756{,}164$$

rather than the USD 750,000 in Table 20.1. It will be evident from this that fixed and floating rates are often not exactly comparable because one applies to 360 days, while the other applies to a whole year.

A further issue concerns weekends and holidays. A rule must be specified for determining what happens if the day specified for a payment happens to be a holiday or a weekend. One common rule is the *next business day* convention, where payment takes place on the next business day after the specified day. Examples of other conventions are the *preceding business day* convention, where payment is on the business day preceding the specified day, and the *modified following business day* convention, which becomes the preceding business day convention if the following business day is in a different month from the specified day.

Confirmations

The agreement in an OTC derivatives transaction is known as a *confirmation* or a *confirm*. The confirmation specifies the dates when payments will be exchanged, how the payments will be calculated, what day count conventions will be used, which country's holiday calendars will apply, and so on.

As discussed in earlier chapters, a standard interest rate swap between two financial institutions must be cleared by a central counterparty (CCP).[3] Each side then has an agreement with the CCP and will have to post the required initial margin and variation margin. If one of the parties to a transaction is not a CCP member, it must arrange to clear the transaction through a member.

If one party to an interest rate swap transaction is an end user, the transaction can be cleared bilaterally as described in earlier chapters. The confirmation agreement will then typically state that a master agreement already entered into between the two

[3] Material on CCPs was presented in Chapters 5 and 6.

Table 20.2 Example of Bid and Ask Quotes that Might Be Made by a Financial Institution for a Particular Type of Swap

Maturity (Years)	Bid (% Per Year)	Ask (% Per Year)	Swap Rate (% Per Year)
2	2.95	2.97	2.960
3	3.06	3.09	3.075
5	3.08	3.12	3.100
7	3.11	3.15	3.130
10	3.14	3.18	3.160
30	3.20	3.26	3.230

sides will apply to the swap. As discussed in Chapters 5 and 6, a master agreement applies to all bilaterally cleared transactions.[4]

Quotes

Large financial institutions are market makers in interest rate swaps. An example of swaps' bid and ask quotes that might be made by a financial institution, where exchanges take place every three months, is shown in Table 20.2. The *swap rate* is the average of the bid and ask quotes. The two-year quotes indicate that a financial institution is prepared to enter into a swap where it pays a fixed rate of 2.95% and receives a floating rate. It is also prepared to enter into a swap where it receives a fixed rate of 2.97% and pays a particular floating rate. (For the sake of generality, we have not specified which particular floating rate is considered by Table 20.2.) The treasurer of a non-financial corporation might accept the bid quote for maturity of five years and a notional principal of USD 100 million. This means that it will receive 3.08% and pay the floating rate for five years. If the financial institution is lucky, it will find another end user who wants to enter into the opposite (pay fixed, receive floating) transaction. If not, it could choose to hedge its risk by trading a swap where it is receiving-fixed and paying floating with another financial institution.

20.2 THE RISK-FREE RATE

The risk-free rate is an important input to valuing swaps and other derivatives. As discussed in Chapter 16, Treasury rates might seem to be good proxies for risk-free rates, but for various reasons they are too low.

OIS rates (i.e., rates calculated from OISs) are usually used to determine risk-free rates. The fixed rates in OISs with maturities up to one year provide a direct estimate of the corresponding zero-coupon risk-free rates. Swaps with longer maturities provide estimates of par yields, and the bootstrap method described in Chapter 16 can be used to estimate zero-coupon rates from these par yields.

20.3 REASONS FOR TRADING INTEREST RATE SWAPS

Interest rate swaps are popular products because they can be used to transform either an asset or a liability. A company with a floating-rate loan can use a swap to convert it to a fixed-rate liability. Similarly, a company with fixed-rate loan can use a swap to convert it to a floating-rate liability. A company that has invested in a fixed-rate bond but thinks that interest rates will increase can use a swap to exchange its fixed-rate income to floating-rate income. Similarly, a company with floating-rate income can exchange it for fixed-rate income with a swap.

For example, suppose that a company has borrowed USD 10 million for five years at a fixed interest rate of 3.5% (compounded quarterly) with interest being paid quarterly. The company can use the five-year bid quote of 3.08% in Table 20.3 to convert a fixed-rate borrowing into a floating-rate borrowing. When it enters into the swap contract (in conjunction with its borrowings), it has three sets of cash flows:

- It pays a fixed rate of 3.5% on its borrowings of USD 10 million.
- It pays the floating rate on USD 10 million under the terms of the swap.
- It receives 3.08% on USD 10 million under the terms of the swap.

These three sets of cash flows are shown in Figure 20.1 and net out to an interest payment of 3.5% plus floating minus 3.08%, or

[4] The International Swaps and Derivatives Association (ISDA) is the industry association for derivatives market participants and provides the documentation used in many master agreements.

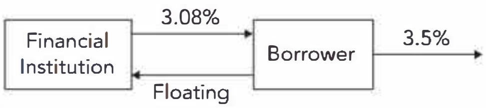


Figure 20.1 **Conversion of fixed-rate borrowing to floating-rate borrowing.**

floating plus 0.42%. The swap enables a 3.5% fixed-rate liability to be exchanged for a floating-rate liability.

Now suppose that a company has borrowed USD 10 million for two years at a three-month floating rate plus 30 basis points.[5] The company can use the ask quote of 2.97% in Table 20.3 to convert the floating-rate liability to a fixed-rate liability. When it enters the swap (in conjunction with its borrowings), it has three sets of cash flows:

- It pays floating plus 30 basis points on its borrowings of USD 10 million.
- It receives the floating rate on USD 10 million under the terms of the swap.
- It pays 2.97% on USD 10 million under the terms of the swap.

These three sets of cash flows are shown in Figure 20.2 and net out to an interest payments of floating plus 30 basis points minus floating plus 2.97%, or 3.27%. The swap thus enables the floating-rate liability to be exchanged for a 3.27% fixed-rate liability.

We now move on to consider how swaps can be used to manage assets that provide cash inflows. Suppose that a company has a ten-year USD 10,000,000 investment paying interest at 2.8%. If it accepts the ten-year ask quote in Table 20.3, it will have the following interest cash flows:

- It receives 2.8% on its investment.
- It receives the floating rate.
- It pays 3.18%.

These three sets of cash flows are shown in Figure 20.3 and net out to floating + 2.8% − 3.18%, or floating minus 0.38%. The swap enables a ten-year investment paying 2.8% to be exchanged for one paying the floating rate minus 38 basis points.

Finally, suppose that a company has a three-year USD 10,000,000 investment paying the floating rate minus 0.1%. If it accepts the three-year bid quote in Table 20.3, it will have the following interest cash flows:

[5] This means that Libor is observed every three months and interest at an annual rate that is 0.3% higher than Libor is applied to the USD 10 million borrowings and paid three months later.

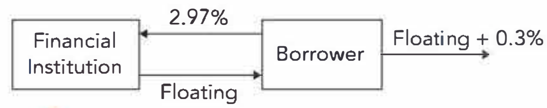


Figure 20.2 **Conversion of floating-rate borrowing to fixed-rate borrowing.**

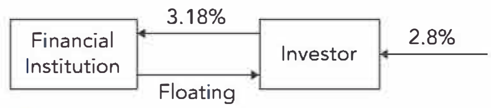


Figure 20.3 **Conversion of fixed-rate investment to floating-rate investment.**

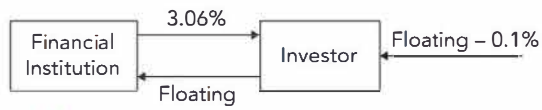


Figure 20.4 **Conversion of floating-rate investment to fixed-rate investment.**

- It receives the floating rate minus 0.1% on its investment.
- It pays the floating rate.
- It receives 3.06%

These three sets of cash flows are shown in Figure 20.4 and net out to floating minus 0.1% − floating +3.06%, or 2.96%. The swap enables a ten-year investment paying the floating rate minus ten basis points to be exchanged for one paying 2.96%.

Comparative Advantage Arguments

It is sometimes argued that highly creditworthy companies find it much easier to borrow at fixed rates than other companies. Table 20.3 shows the five-year rates that might be offered to a company with a high credit rating (Company A) and a company with a lower credit rating (Company B). We suppose that both rates involve quarterly payments. The interesting aspect of these rates is that the spread between them in the two markets are not the same. Company A pays 1.4% less than Company B in fixed-rate markets, but only 0.8% less in floating-rate markets. Thus, Company A appears to have a comparative advantage in

Table 20.3 **Five-Year Rates Offered to Companies A and B in Fixed and Floating Markets. Rates Are Paid Quarterly**

	Fixed	Floating
Company A	3.5%	Floating + 0.1%
Company B	4.9%	Floating + 0.9%

fixed-rate markets because it pays 1.4% less than Company B in this market and only 0.8% less than Company B in the floating-rate market. Similarly, it can be argued that Company B has a comparative advantage in floating-rate markets. This is because it pays more than company A in both markets but the amount more that it pays is less in floating-rate markets.

An interest rate swaps trader would argue that a company should always raise money in the market where it has a comparative advantage and then swap to the market it wants. Suppose that Company A wants to borrow at a floating rate while Company B wants to borrow at a fixed rate. Let us first imagine that Company A and Company B get in touch with each other directly to do a swap. Suppose that Company B pays a fixed rate of X% to Company A and that Company A pays the floating rate. This swap is shown in Figure 20.5.

If Company A borrows at the fixed rate of 3.5% indicated in Table 20.3, it has three sets of interest rate cash flows:

- It pays 3.5%
- It receives X%
- It pays the floating rate

These cash flows net out to floating +3.5% — X%. Company A can borrow at the floating rate + 0.1% directly. The interest rate improvement in the rate paid by Company A is therefore:

$$(F + 0.1\%) - (F + 3.5\% - X\%) = (X - 3.4)\%$$

where F is the floating rate. If Company B borrows at the floating rate indicated in Table 20.3, it has three sets of interest rate cash flows:

- It pays the floating rate+0.9%
- It pays X%
- It receives the floating rate

These cash flows net out to X+0.9%. Company B can borrow at 4.9% directly. The interest rate improvement in the rate paid by Company B is therefore:

$$4.9\% - (X + 0.9\%) = (4 - X)\%$$

The total interest rate improvement to both sides from the swap is

$$(X - 3.4)\% + (4 - X)\% = 0.6\%$$

Where does this improvement come from? It is the difference between (a) the spread between the rates offered to A and B in fixed-rate markets (=1.4%) and (b) the spread between the rates offered to them in floating-rate markets (=0.8%).

Both sides obtain the same improvement in the rate they pay if:

$$(X - 3.4)\% = (4 - X)\%$$

or $X = 3.7$%.

Figure 20.6 shows the swap when X=3.7% together with the borrowing by the two companies in the markets where they have a comparative advantage. The swap leads to Company A borrowing at the floating + 3.5% − 3.7% = floating − 0.2%. This is 0.3% less than the floating + 0.1% it would pay if it went directly to floating-rate markets. Company B borrows at 3.7% + (floating+0.9%) – floating = 4.6%. This is 0.3% less than the 4.9% rate it would pay if it went directly to fixed-rate markets. As mentioned above the total apparent gain to the two sides is the difference between (a) the spread between the fixed rates offered to the two companies (1.4% in our example) and (b) the spread between the floating rates offered to the two companies (0.8% in our example).

In practice, if Company A and B are non-financial companies, they are not likely to get together directly. In this case, a swap trader working for a financial institution could act as intermediary by standing between the companies in the way indicated in Figure 20.7. The figure shows that the swap trader would gain 0.04% and the interest rate paid by each company is improved by 0.28% instead of 0.3% (Company A pays floating minus 18 basis points, and Company B pays 4.62%). The total gain to all three parties (Company A, Company B, and the intermediary) is 0.6%, as before.

Why does this opportunity exist in financial markets? The answer is that the result of the swap arrangement is not quite as simple as we have portrayed it so far. Company B in our example is not assured of paying a fixed rate. If Company B accepts the fixed rate in Table 20.4, it is certain that it will be able to borrow at 4.9% for five years. If it accepts the floating rate, the 0.9% spread over the floating reference rate is only guaranteed for three months. If Company B's credit worthiness changes or conditions in the market change, the spread over the floating reference rate that it will pay is also likely to change. For example, there is a possibility that during the five-year period Company B's credit rating declines so that by year two it is paying floating

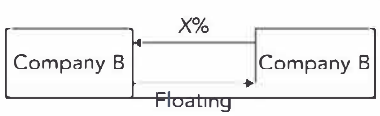


Figure 20.5 Trial swap between Company A and Company B.

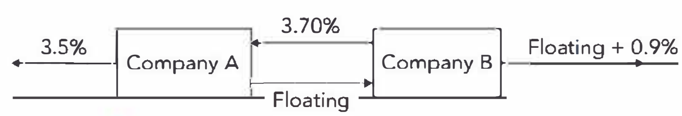


Figure 20.6 Swap between Company A and Company B Together with borrowing.

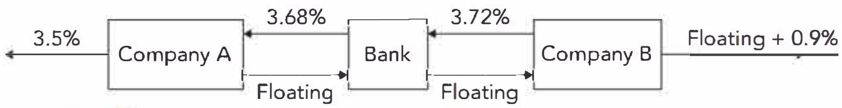


Figure 20.7 Swap in Figure 20.6 when a bank stands between the two sides.

plus 2% instead of floating plus 0.9%. The swap in Figure 20.7 will then lead to it paying 5.72% at that time (this is more than the 4.9% it would pay if it went directly to fixed-rate markets).

Note that Company A does not have the same type of uncertainty as Company B. If it arranged the swap in Figure 20.7, it would have guaranteed borrowing at floating minus 0.18% for five years.

Because of the observations we have made in connection with the rates in Table 20.4, we should modify our earlier conclusion and say that the swap in Figure 20.5 succeeds in exchanging a floating-rate liability for a fixed-rate liability if the company's creditworthiness does not change so that it can borrow at the same spread above he floating reference rate for the whole five years.

20.4 VALUATION OF INTEREST RATE SWAPS

At the time a swap is entered, its value is usually zero or close to zero.[6] At later times, it may have a positive or negative value due to changes in interest rates. A swap can be valued by comparing it to a new swap that we know would be worth zero.

Consider a swap that was entered some time ago and now has two years to maturity. We suppose that a fixed rate of 4% is received and floating is paid every three months on a principal of USD 100 million. We also suppose that the quotes in Table 20.3 are the ones made by derivatives dealers today. The swap rate (average of bid and offer quotes) for a new two-year swap is 2.96%. This is the mid-market quote. It is therefore reasonable to assume that a swap where 2.96% is paid and the floating rate is received is worth zero today. For valuation purposes, we can imagine a trader taking two positions:

- A two-year swap where a fixed rate of 4% is received and three-month floating rate is paid on a principal of USD 100 million
- A two-year swap where a fixed rate of 2.96% is paid and three-month floating is received on a principal of USD 100 million

[6] Bid-ask spreads mean that it has a slightly positive value to one side (the provider of the bid and ask quotes) and a slightly negative value to the other side (the side that accepts either a bid or an ask quote).

These two swaps net out to a position where 1.04% is received. Specifically, every three months for the next two years $0.25 \times 0.0104 \times$ USD 100,000,000 = USD 260,000 would be received. Suppose that the risk-free rate is 2.4% (compounded quarterly) for all maturities. This is equivalent to 0.6% per each three months. The (USD) value of the position is therefore:

$$\frac{260{,}000}{1.006}+\frac{260{,}000}{1.006^2}+\frac{260{,}000}{1.006^3}+\frac{260{,}000}{1.006^4}+\frac{260{,}000}{1.006^5}+\frac{260{,}000}{1.006^6}+\frac{260{,}000}{1.006^7}+\frac{260{,}000}{1.006^8}=2{,}024{,}945$$

As we have explained, the value of the second swap is zero. It follows that the value of the first swap (the one we are interested in) is USD 2,024,945. This approach can be used to value swaps involving any floating rate providing quotes are available from a dealer. The risk-free rates used for discounting can be estimated from OIS quotes as indicated in Section 20.2.

It is sometimes necessary to interpolate between dealer quotes to value a swap in the way described. For example, if a swap has a remaining life of 2.5 years and the quotes in Table 20.3 apply, one could assume that a 2.5-year swap where the fixed rate is $0.5 \times (2.96\% + 3.075\%) = 3.0175\%$ has a value of zero.

If forward rates are known, an alternative way to value a fixed-for floating interest rate swap is to regard it as a portfolio of forward rate agreements (FRAs). As explained in Chapter 16, an FRA can be value by assuming that forward rates will be realized.

20.5 CURRENCY SWAPS

We now move on to discuss currency swaps. We start by considering a fixed-for-fixed currency swap. This is a swap where a fixed interest rate in one currency is exchanged for a fixed interest rate in another currency. Two principal amounts must be specified, one for each currency.

One difference between a currency swap and an interest rate swap is that principal amounts are usually exchanged in a currency swap, whereas they are not exchanged in an interest rate swap. Specifically, principal amounts are exchanged in the opposite direction to the interest rate payments at the beginning of the life of the swap and in the same direction as the interest rate payments at the end of the life of the swap.

Consider a swap where interest at 5% in USD on a principal of USD 10 million is received and interest at 4% in euros on a principal of 8 million euros is paid. Both interest rates are annually compounded, and payments are exchanged every year for four years.

The cash flows are shown in Table 20.5. The interest cash flows in USD are 0.05 × USD 10 million or USD 0.5 million. The interest cash flows in euros are 0.04 × 8 million euros or 0.32 million euros.

A currency swap can be used to transform liabilities and assets in an analogous way to interest rates swaps. For example, the swap in Table 20.5 can be used to transform borrowings at 5% on USD 10 million to borrowings at 4% on 8 million euros. The initial exchange converts the USD principal received at the beginning to a euro principal. Subsequent exchanges provide the USD interest and principal payments that are necessary with euro payments. In each year, USD 0.5 million of interest is required on the USD loan. USD 0.5 million is received to make this payment under the terms of the swap and the euro interest rate is paid.

The swap can also be used to convert an 8 million euro investment earning 4% to a USD 10 million investment earning 5%. The euro outflow is converted to a USD outflow at the beginning, and subsequent euro inflows are converted to USD inflows.

When we discussed comparative advantage arguments for borrowers earlier, we argued that, in the case of interest rate swaps, those advantages may be largely illusory because the spreads charged over a floating reference rate are liable to change with the creditworthiness of the borrower. In the case of currency swaps, tax issues can create a (non-illusory) comparative advantage. For example, a company might have a low marginal tax rate in France and a high marginal tax rate in the United States. If it wants to borrow in euros, it could make sense for the company to borrow in USD (deducting the USD interest for tax purposes) and swap the borrowing to euro borrowings using the swap presented in Table 20.4.

Table 20.4 Cash Flows in a Fixed-for-Fixed Currency Swap Where Interest at 5% in USD on a Principal of USD 10 Million Is Exchanged for Interest at 4% on a Principal of 8 Million Euros

Time (Years)	USD Cash Flow (Millions)	Euro Cash Flow (Millions)
0	−10	+8
1	+0.5	−0.32
2	+0.5	−0.32
3	+0.5	−0.32
4	+10.5	−8.32

Valuation

The two sets of cash flows in a swap are referred to as legs. Thus, the swap in Table 20.4 has a USD leg and a euro leg. Currency swaps are usually designed so that they have a value close to zero when first negotiated. Later, the remaining cash flows can be valued by considering each leg of the swap separately.

Suppose that a fixed-for-fixed swap involves exchanging principal and interest in currency X for principal and interest in currency Y and that the valuation is required in currency X. The valuation procedure is as follows:

1. Calculate the present value of the remaining currency X cash flows in currency X.
2. Calculate the present value of the remaining currency Y cash flows in currency Y.
3. Convert the present value of the currency Y cash flows to currency X at the current exchange rate.
4. Calculate the swap value as the difference between the amount calculated in step 1 and the amount calculated in step 3.

In the case of the swap presented in Table 20.4, suppose one year after initiation (just after the first exchange of payments) that the risk-free interest rate in USD is 4.5% for all maturities, the risk-free interest rate in euros is 3.5% for all maturities, and the exchange rate (USD per euro) is 1.15. The remaining cash flows (with time being measured from the valuation date, not from the start of the swap) are shown in Table 20.5. The value in millions of USD of the USD cash flows is

$$\frac{0.5}{1.045} + \frac{0.5}{1.045^2} + \frac{10.5}{1.045^3} = 10.137$$

The value in millions of euros of the euro cash flows is

$$\frac{0.32}{1.035} + \frac{0.32}{1.035^2} + \frac{8.32}{1.035^3} = 8.112$$

The value of the swap in millions of USD is therefore:

$$10.137 - 8.112 \times 1.15 = 0.809$$

Table 20.5 Cash Flows in Swap in Table 20.4 as Seen One Year After the Swap Is Initiated

Time (Years)	USD Cash Flow (Millions)	Euro Cash Flow (Millions)
1	+0.5	−0.32
2	+0.5	−0.32
3	+10.5	−8.32

The swap can also be valued in terms of forward exchange rates. We know from earlier chapters that the forward exchange rates (USD per euro) for year one, two, and three are

$$\text{Year 1: } 1.15 \times \frac{1.045}{1.035} = 1.1611$$

$$\text{Year 2: } 1.15 \times \frac{1.045^2}{1.035^2} = 1.1723$$

$$\text{Year 3: } 1.15 \times \frac{1.045^3}{1.035^3} = 1.1837$$

The USD equivalent of the first exchange, when USD 0.5 million is received and 0.32 million is paid, is (in USD millions)

$$0.5 - 0.32 \times 1.1611 = 0.1284$$

The USD equivalent of the second exchange, when USD 0.5 million is received and 0.32 million is paid, is (in USD millions)

$$0.5 - 0.32 \times 1.1723 = 0.1249$$

The USD equivalent of the final exchange, when USD 10.5 million is received and 8.32 million is paid, is (in USD millions)

$$10.5 - 8.32 \times 1.1837 = 0.6520.$$

The value of the contract is the present value of the three exchanges. It is

$$\frac{0.1284}{1.045} + \frac{0.1249}{1.045^2} + \frac{0.6520}{1.045^3} = 0.809$$

This is the same value as the one obtained earlier. This illustrates that valuing each leg separately as bonds gives the same result as valuing each exchange separately using forward exchange rates.

Other Currency Swaps

Two other types of currency swaps are those where:

- A floating rate in one currency is exchanged for a fixed rate in another currency
- A floating rate in one currency is exchanged for a floating rate in another currency.

The first type of swap can be used to convert a floating liability (asset) in one currency to a fixed liability (asset) in another currency. The second type of swap can be used to convert a floating liability (asset) in one currency to a floating liability (asset) in another currency.

The swaps can be valued by valuing each leg in its own currency. A fixed leg can be valued in the same way as the fixed legs of a fixed-for-fixed swap. A floating leg can be valued by assuming than forward interest rates in the currency will be realized.

20.6 OTHER SWAPS

The exchanges in swaps can be defined in many ways. A variation on the interest rate swaps we have discussed is when the notional principal changes in a pre-determined way through time. This could be useful for a company embarking on a multi-year project when it knows that its borrowings will increase with time. With this kind of swap, it can transform its borrowings from fixed to floating or floating to fixed.

An *equity* swap is a swap where fixed return is exchanged for the return generated when the notional principal is invested in pre-specified equity. For example, it might be agreed that the return from USD 10 million invested in the S&P 500[7] during each six-month period for the next five years will be exchanged for USD 10 million invested at a fixed interest rate of 4%. In some cases, the equity return is swapped for a return tied to a floating interest rate reference rate instead of a fixed return.

A *commodity* swap is a swap where a certain amount of a commodity at a pre-determined fixed price is periodically exchanged for the same amount of the commodity at a floating price (i.e. the market price at the time of the exchange). This is equivalent to a portfolio of forward contracts on the price of the commodity.

In a *volatility* swap, an exchange takes place at the end of each period of the life of the swap just as it does in an interest rate swap. The historical volatility observed during the period is exchanged for a pre-determined fixed volatility. Both are multiplied by the notional principal. A *variance* swap is defined similarly: The historical variance rate (square of volatility) is exchanged for a pre-determined variance rate with both being multiplied by the notional principal.

A *credit default* swap (CDS) provides insurance against a default by a company. The buyer of this protection makes periodic fixed payments to the seller of protection for an agreed period of time. If the company that is the subject of the protection (known as the *reference entity*) defaults, there is a payment from the seller of protection to the buyer of protection. Otherwise the seller of protection does not have to make any payments. In an *index credit default* swap, a portfolio of companies is defined and in return for regular payments, the buyer of protection is compensated for defaults by any of the companies in the portfolio.

There are many other types of swaps and the exchanges in a swap can be defined in almost any way. Some can be overly

[7] The S&P index used to define the return in this type of swap would typically be the total return index, which is the index formed when dividends on the underlying stocks are reinvested in the index.

complicated. An exotic deal that was defined by Bankers Trust[8] and sold to several clients in 1993 is known as the 5/30 swap. For this type of swap, payments depended in a complex way on the 30-day commercial paper rate, a 30-year Treasury bond price, and the yield on a five-year Treasury bond. In 2007, a Spanish bank, Santander, defined swaps where the payments depended, in a complex way, on whether a three-month reference rate stayed within the 2% to 6% range.[9] In both cases, it was alleged that the risks in the swaps were not fully understood by the banks' clients.

20.7 CREDIT RISK

Swaps have the potential to give rise to credit risk. If Company X gets into financial difficulties and defaults when its outstanding swaps with Company Y have a positive value to company Y, Company Y may experience a loss.[10]

We have discussed key aspects of credit risk in earlier chapters and will summarize some of the main points here. Interest rate swaps and index credit default swaps between financial institutions must be cleared through a central counterparty (CCP). The CCP requires both initial margin and variation margin from both sides. This greatly reduces credit risks. In Chapter 6, we examined the risks associated with a possible default by the CCP.

Other swaps between financial institutions are cleared bilaterally and regulations require both initial margin and variation margin to be posted. This largely eliminates the credit risk in these transactions. The variation margin is transferred directly from one side to the other. Initial margin must be posted by each side with a third-party trustee.

There is no requirement that collateral (i.e., margin) be posted for transactions between a financial institution and non-financial companies. When no collateral is posted, or when transactions are only partly collateralized, the credit exposure must be monitored carefully. The initial pricing of transactions should consider expected credit losses by both sides.[11]

SUMMARY

In an interest rate swap, one party agrees to pay interest at a floating rate on a notional principal and receive interest at a fixed rate on the same notional principal. Interest payments are exchanged periodically, and the swap has a pre-agreed life. In a currency swap, interest payments on a pre-agreed principal in one currency are exchanged for interest payments on a pre-agreed principal in another currency. Principal amounts in the two currencies are exchanged in the opposite direction to interest payments at the beginning of the life of a currency swap and in the same direction as interest payments at the end of the life of the currency swap.

The popularity of interest rate and currency swaps arises from their ability to transform exposures connected to a company's assets or liabilities. An interest rate swap can transform a fixed-rate liability to a floating-rate liability, or vice versa. It can also transform a fixed-rate asset to a floating-rate asset, or vice versa. Currency swaps can similarly transform assets (liabilities) in one currency to assets (liabilities) in another currency.

There are several approaches to valuing swaps. We have showed how an interest rate swap can be valued by comparing it with a similar swap that is known to have a value of zero. Sometimes, a swap is valued as the difference between the value of the stream of payments that will be made by one side and the value of the stream of payments that will be made by the other side. Another approach is to value each exchange of payments and sum the results. We used both approaches when valuing currency swaps and illustrated this with an example in which they give the same result.

[8] Bankers Trust was sold for USD 10.1 billion to Deutsche Bank in 1998 after it had suffered losses from its positions in Russian bonds.

[9] See J. Hull "Risk Management and Financial Institutions", 5th edition, Chapter 27 for a description and discussion of the Santander deal, which was described in a May 2014 Risk magazine article as the worst trade of all time.

[10] As discussed in Chapter 5, credit risk is not considered on a transaction-by-transaction basis. Transactions whether cleared through a CCP or bilaterally are netted.

[11] Credit value adjustment (CVA) is a downward adjustment in the pricing of derivatives made by a bank to reflect expected losses from a default by the counterparty. Debit value adjustment (DVA) is an upward adjustment to the price by the bank reflecting the possibility of a default by the bank (and therefore credit losses) by the counterparty. The calculations of these quantities involve estimating the probability of defaults occurring during future periods and expected credit losses conditional on a default.

The following questions are intended to help candidates understand the material. They are <u>not</u> actual FRM exam questions.

QUESTIONS

Short Concept Questions

20.1 An OIS starts on June 1, 2021, and involves the three-month floating rate being paid and a fixed rate of 5% being received every six months on a principal of USD 5 million. When does the first exchange take place? How is it calculated?

20.2 Explain what a five-year swap rate is?

20.3 Why might day counts mean that floating and fixed rates in an interest rate swap are not directly comparable?

20.4 An interest rate swap is based on an overnight rate, and exchanges take place every three months for three years. How are the payments calculated?

20.5 What is the difference between the discount rate used to value swaps (a) prior to the 2007–2008 crisis and (b) after the 2007–2008 crisis?

20.6 "Borrowing at a floating rate for five years and then swapping floating-rate payments for fixed-rate payments does not guarantee a fixed rate of interest on the borrowings." Explain this statement.

20.7 Are principal payments exchanged in (a) interest rate swaps and (b) currency swaps?

20.8 Explain the two ways a fixed-for-fixed currency swap can be valued.

20.9 Explain how an equity swap works.

20.10 Explain how a volatility swap works.

Practice Questions

20.11 The quotes for a five-year interest rate swap are bid 3.20, ask 3.24. What swap would be entered into by a company that can borrow for five years at 4.2% per year but wants to borrow at a floating rate? What rate of interest does the company end up borrowing at?

20.12 Suppose that in Problem 20.11 a company can borrow at a floating rate plus 50 basis points but wants to borrow for five years at a fixed rate. What swap should the company enter into? What rate of interest does the company end up borrowing at? (Assume that the spread above floating at which the company borrows does not change.)

20.13 Company A can borrow at a fixed rate of 4.3% for five years and at floating plus 30 basis points. Company B can borrow for five years at a fixed rate of 5.9% and at a floating rate of floating plus 100 basis points. As a swaps trader, you are in touch with both companies and know that Company A wants to borrow at the floating rate and that Company B wants to borrow at a fixed rate. Both companies want to borrow the same amount of money. Design a swap where you will earn ten basis points and that will appear equally attractive to both sides.

20.14 You are required to estimate the value of an overnight indexed swap that has three years left in its life and involves paying a fixed rate of 5% at the end of each quarter and receiving the rate implied by the overnight rate when it is compounded day-by-day during the quarter. The notional principal is USD 20 million. The current quote for three-year overnight index swap is bid 3.80, ask 3.88. The risk-free rate is 3.6% for all maturities. All rates are compounded quarterly.

20.15 Suppose that a six-month floating rate is 5%, the forward floating rate for the period between 0.5 and 1 year is 5.6%, and the forward floating rate for the period between 1.0 and 1.5 years is 6.0. The two-year floating swap rate is 5.7%. All risk-free rates are 4.5%. What is the forward floating rate for the period between 1.5 and 2 years? All rates are expressed with semi-annual compounding. (Hint: Regard the swap as a portfolio of FRAs.)

20.16 In Problem 20.15, what is the 1.5-year zero rate, expressed with semiannual compounding, for the term structure corresponding to the floating rate?

20.17 Consider a currency swap where interest on British pounds at the rate of 3% is paid and interest on euros at 2% is received. The British pound principal is 1.0 million pounds, and the euro principal is 1.1 million euros. Interest is exchanged every six months, and two years are remaining. The current exchange rate is 1.15 euro/pound. The risk-free rates in pounds and euros are 2.5% and 1.5%. Value the swap by considering it as the difference between two bonds. All rates are compounded semi-annually.

20.18 Value the swap in Problem 20.17 by considering it as a portfolio of forward contracts.

20.19 A bank trades a swap where a fixed rate of 5% in currency A is paid and Floating in currency B is received. Show that the swap can be considered as a fixed-for-fixed currency swap plus an interest rate swap.

20.20 A bank is taking floating-rate deposits and using the funds to make five-year fixed-rate loans. What risks is it taking? How can they be hedged?

ANSWERS

Short Concept Questions

20.1 The first exchange will take place three months after June 1, 2021, on September 1, 2021. It will be calculated as:

$$0.5 \times (0.05 - R) \times \text{USD } 5{,}000{,}000$$

where R is the floating rate calculated on September 1 from overnight rates between June 1 and September 1.

20.2 The five-year swap rate is the average of the bid and ask quotes for five-year swaps.

20.3 Often the floating rate refers to the rate per 360 days, whereas the fixed rate is for a whole year.

20.4 The payments on a given date are the exchange of a predetermined fixed rate for the three-month rate implied by the overnight rates observed during the previous three months. As explained in Chapter 16, if there are n business days during the period, the overnight rate on the ith of these business days is r_i, and the rate applies to d_i calendar days. The rate for the period can be calculated at the end of the period as:[12]

$$R = [(1 + r_1\hat{d}_1)(1 + r_2\hat{d}_2) \ldots ..(1 + r_n\hat{d}_n) - 1] \times \frac{360}{D}$$

where $\hat{d}_i = d_i/360$ and D is the total number of days in the period so that

$$D = \sum_i d_i$$

For most i, $d_i = 1$, but weekends and holidays lead to the overnight rate being applied to more than one day. For example, on a Friday, d_i will normally be equal to 3.

20.5 Libor was used as a proxy for the risk-free rate pre-crisis. Interbank overnight rates have been used to determine discount rates since then.

20.6 The spread added to floating rate at which a company borrows is liable to change if the creditworthiness of the company or conditions in the market change. This means that the fixed rate calculated using the current spread may not be what applies for all periods.

20.7. Principal payments are not exchanged in an interest rate swap (exchanging principals would not of course make any difference when the principals used to determine interest payments on both sides of the swap are the same). Principal payments are exchanged in a currency swap (there are two different principals, one for each currency).

20.8 A fixed-for-fixed currency swap can be valued as the difference between two bonds or as a portfolio of forward contracts, one corresponding to each exchange.

20.9 A principal is specified, and the return obtained per period by investing the principal in a specified equity portfolio is swapped for the return obtained by investing the principal at a fixed rate (or possibly a floating rate).

20.10 At the end of each period, a pre-specified volatility multiplied by a pre-specified principal is exchanged for the observed historical volatility during the same period multiplied by the same principal.

Solved Problems

20.11 The company should arrange to receive fixed and pay floating to convert the fixed-rate loan to a floating-rate loan. It will accept the bid quote of 3.20. Its cash flows will be as follows:

- Receive 3.2%
- Pay 4.2%
- Pay the floating rate

These net to floating plus 1%.

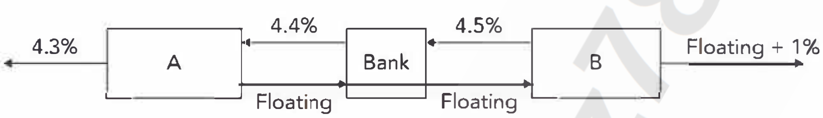
[12] We cover the compounding of interest rates later in this chapter. The equation given is not the only one that can be used. One alternative is to assume a gap between a daily interest rate being observed and the day to which it is applied.

20.12 The company should arrange to pay fixed and receive floating to convert the floating-rate loan to a fixed-rate loan. It will accept the ask quote of 3.24. Its cash flows will be as follows:

- Pay 3.24%
- Receive the floating rate
- Pay floating + 0.5%

These net to 3.74%.

20.13 The spread between the fixed rates offered to companies A and B is 5.9% — 4.3% or 1.6%. The spread between the floating rates is 70 basis points or 0.7%. The difference between these two spreads is 1.6% — 0.7% or 0.9%. It should be possible to design a swap where the parties are in aggregate 0.9% better off. The bank (intermediary) wants 0.1%. This leaves 0.4% for each side. We should therefore be able to design a swap where Company A borrows at Floating + 0.3% — 0.4% or Floating — 0.1% and Company B appears to borrow at 5.9% — 0.4% = 5.5%. If the bank pays X% to A, we require 4.3% + Floating – X% = Floating – 0.1% so that X = 4.4. Similarly, if B pays Y% to the bank, we require Y% + 1% = 5.5% so that Y = 4.5. The swap arrangement is



20.14 The swap rate is the average of 3.80 and 3.88 or 3.84%. The swap involves paying 5% when the market rate is 3.84%. The swaps value is the present value of:

$$0.25 \times (0.0384 - 0.05) \times \text{USD } 20{,}000{,}000 = -\text{USD } 58{,}000$$

on every payment date for the next three years. Because the risk-free rate is 3.6%/4 = 0.9% per quarter, the value is

$$-\sum_{i=1}^{12} \frac{58{,}000}{1.009^i} = -656{,}938$$

20.15 The swap can be valued as a portfolio of FRAs, and each FRA can be valued by assuming that forward rates are realized. A swap where 5.7% is paid and floating is received is worth zero. Per 100 of principal, first FRA is worth:

$$\frac{0.5 \times (0.05 - 0.057) \times 100}{1 + 0.045/2} = -0.342$$

The second FRA is worth:

$$\frac{0.5 \times (0.056 - 0.057) \times 100}{(1 + 0.045/2)^2} = -0.048$$

The third FRA is worth:

$$\frac{0.5 \times (0.060 - 0.057) \times 100}{(1 + 0.045/2)^3} = 0.140$$

If the required forward rate is R then:

$$\frac{0.5 \times (R - 0.057) \times 100}{(1 + 0.045/2)^4} - 0.342 - 0.048 + 0.140 = 0$$

This can be solved to give R = 0.0625. The forward rate for the period between 1.5 and 2 years is 6.25% (semi-annually compounded).

20.16 Using six-month forwards, 100 would grow to:

$$100 \times (1 + 0.05/2) \times (1 + 0.056/2) \times (1 + 0.06/2) = 108.5311$$

If R is the 1.5 year zero rate with semi-annual compounding:

$$100 \times (1 + R/2)^3 = 108.5311$$

This can be solved to give $R = 0.055329$. The zero rate is 5.5329%.

20.17 The swap involves exchanging $0.5 \times 0.03 \times 1{,}000{,}000 = 15{,}000$ pounds with $0.5 \times 0.02 \times 1{,}100{,}000 = 11{,}000$ euros with a final exchange of principal. The value of the British pound bond in British pounds is

$$\frac{15{,}000}{1 + 0.5 \times 0.025} + \frac{15{,}000}{(1 + 0.5 \times 0.025)^2} + \frac{15{,}000}{(1 + 0.5 \times 0.025)^3} + \frac{1{,}015{,}000}{(1 + 0.5 \times 0.025)^4} = 1{,}009{,}695$$

The value of the euro bond in euros is

$$\frac{11{,}000}{1 + 0.5 \times 0.015} + \frac{11{,}000}{(1 + 0.5 \times 0.015)^2} + \frac{11{,}000}{(1 + 0.5 \times 0.015)^3} + \frac{1{,}111{,}000}{(1 + 0.5 \times 0.015)^4} = 1{,}110{,}797$$

The value of the swap in British pounds is therefore $1{,}110{,}797/1.15 - 1{,}009{,}695 = -43{,}785$.

20.18 The forward rates corresponding to the exchanges at times 0.5, 1.0, 1.5, and 2 years are

$$1.15 \times \frac{1.0075}{1.0125} = 1.1443$$

$$1.15 \times \frac{1.0075^2}{1.0125^2} = 1.1387$$

$$1.15 \times \frac{1.0075^3}{1.0125^3} = 1.1330$$

$$1.15 \times \frac{1.0075^4}{1.0125^4} = 1.1275$$

The exchanges are

Time	Pounds Paid	Euros Received	Value in Pounds
0.5	15,000	11,000	11,000/1.1443 − 15,000 = −5,387
1	15,000	11,000	11,000/1.1387 − 15,000 = −5,340
1.5	15,000	11,000	11,000/1.1330 − 15,000 = −5,292
2	1,015,000	1,111,000	1,111,000/1.1275 − 1,015,00 = −29,592

The USD value of the swap is

$$-\frac{5{,}387}{1 + 0.5 \times 0.025} - \frac{5{,}340}{(1 + 0.5 \times 0.025)^2} - \frac{5{,}292}{(1 + 0.5 \times 0.025)^3} - \frac{29{,}592}{(1 + 0.5 \times 0.025)^4} = -43{,}785$$

This agrees with the answer in Problem 20.17.

20.19 The swaps are as follows:

- Fixed rate in currency A is paid and a fixed rate in currency B is received.
- The same fixed rate in currency B is paid and Floating in currency B is received.

20.20 The bank is taking the risk that interest rates will increase because, if that happens, it will have to pay a higher interest rate on the floating-rate deposits but will not receive a higher rate of interest on the fixed-rate assets. It can hedge it risk by entering into a swap where it is receiving floating and paying fixed.

INDEX

C

E

F

G

H

I

J

K

L

M

N

O

P

T

U

V

W

Y

Z