Financial Risk Manager

Handbook

▶ Learn the essentials of managing market, credit, operational, and liquidity risk

▶ Learn the essentials of investment management and hedge fund risk

▶ Learn about structured products, futures, options, and other derivative instruments

▶ Identify regulatory and legal issues

▶ Ideal for self-instruction and in-house training in financial risk management

▶ The official reference book for GARP's FRM® certification program

PHILIPPE JORION

Includes CD-ROM

GARP

GLOBAL ASSOCIATION OF RISK PROFESSIONALS

RISK MANAGEMENT LIBRARY

TRADING SOFTWARE

FOR SALE & EXCHANGE

www.trading-software-collection.com

Financial Risk Manager Handbook

Fifth Edition

PHILIPPE JORION

GARP

WILEY

John Wiley & Sons, Inc.

Contents

The *Financial Risk Manager Handbook* provides the core body of knowledge for financial risk managers. Risk management has evolved rapidly over the past decade and has become an indispensable function in many institutions.

This *Handbook* was originally written to provide support for candidates taking the FRM examination administered by GARP. As such, it reviews a wide variety of practical topics in a consistent and systematic fashion. It covers quantitative methods and capital markets, as well as market, credit, operational, and integrated risk management. It also discusses regulatory and legal issues essential to risk professionals.

This edition has been thoroughly updated to reflect recent developments in financial markets. The unprecedented losses incurred by many institutions have raised questions about risk management practices. These issues are now addressed in various parts of the book, which also include lessons from recent regulatory reports. The securitization process and structured credit products are critically examined. A new chapter on liquidity risk has been added, given the importance of this risk during the recent crisis. Finally, this *Handbook* incorporates the latest questions from the FRM examinations.

Modern risk management systems cut across the entire organization. This breadth is reflected in the subjects covered in this *Handbook*. The book was designed to be self-contained, but only for readers who already have some exposure to financial markets. To reap maximum benefit from this book, readers should have taken the equivalent of an MBA-level class on investments.

Finally, I want to acknowledge the help received in writing this *Handbook*. In particular, I thank the numerous readers who shared comments on previous editions. Any comment or suggestion for improvement will be welcome. This feedback will help us to maintain the high quality of the FRM designation.

Philippe Jorion
February 2009

Philippe Jorion is a Professor of Finance at the Paul Merage School of Business at the University of California at Irvine. He has also taught at Columbia University, Northwestern University, the University of Chicago, and the University of British Columbia. He holds an M.B.A. and a Ph.D. from the University of Chicago and a degree in engineering from the University of Brussels. He is also a managing director at Pacific Alternative Asset Management Company (PAAMCO), a global fund of hedge funds.

Dr. Jorion is the author of more than 90 publications directed to academics and practitioners on the topics of risk management and international finance. He has also written a number of books, including *Big Bets Gone Bad: Derivatives and Bankruptcy in Orange County*, the first account of the largest municipal failure in U.S. history, and *Value at Risk: The New Benchmark for Managing Financial Risk*, which is aimed at finance practitioners and has become an industry standard.

Philippe Jorion is a frequent speaker at academic and professional conferences. He is on the editorial board of a number of finance journals and was editor in chief of the *Journal of Risk*.

Founded in 1996, the Global Association of Risk Professionals (GARP) is the leading not-for-profit association for world-class financial risk certification, education, and training with close to 100,000 members representing 167 countries. With deep expertise and a strong reputation, GARP sets global standards and creates risk management programs valued worldwide. All GARP programs are developed with input from experts around the world to ensure that concepts and content reflect globally accepted practices.

GARP is dedicated to advancing the risk profession. For more information about GARP, please visit www.garp.com.

FINANCIAL RISK MANAGER (FRM®) CERTIFICATION

The benchmark FRM designation is the globally accepted risk management certification for financial risk professionals. The FRM objectively measures competency in the risk management profession based on globally accepted standards. With a compound annual growth rate of 25 percent over the past seven years, the FRM program has experienced significant growth in every financial center around the world. Now 16,000+ individuals hold the FRM designation in over 90 countries. In addition, organizations with five or more FRM registrants grew from 105 in 2003 to 424 in 2008, further demonstrating the FRM program's global acceptance.

The FRM Continuing Professional Education (CPE) program, to be offered starting in 2009 exclusively for certified FRM holders, provides the perspective and framework needed to further develop competencies in the ever-evolving field of risk management.

For more information about the FRM program, please visit www.garp.com/frmexam.

OTHER GARP CERTIFICATIONS

International Certificate in Banking Risk and Regulation (ICBRR)

The ICBRR allows individuals to expand their knowledge and understanding of the various risks, regulations, and supervisory requirements banks must face in today's economy, with emphasis on the Basel II Accord. This certificate is ideal for employees who are not professional risk managers but who have a strong need to understand risk concepts. The ICBRR program is designed for employees in

nonrisk departments such as internal audit, accounting, information technology (IT), legal, compliance, and sales, acknowledging that everyone in the organization is a risk manager!

Certificate in Energy Risk Management

The Certificate in Energy Risk Management provides individuals with a comprehensive and cross-product understanding of the physical and financial marketplaces relating to crude oil, natural gas, liquefied natural gas, and electricity/power. This program is valuable for anyone working in or servicing the energy field and requiring an understanding of the physical and financial markets, how they interrelate, and the risks involved. This program will launch in 2Q 2009.

Certificate in Risk Management for Islamic Financial Institutions

This certificate is under development by a practice oversight committee of Islamic finance experts from around the globe. The program will cover the risk management methodologies specific to Sharia'a-compliant financial products and will be the only one of its kind anywhere in the world.

GARP DIGITAL LIBRARY

As the world's largest digital library dedicated to financial risk management, the GARP Digital Library (GDL) is the hub for risk management education and research material. The library's unique iReadings™ allow users to download individual chapters of books, saving both time and money. There are over 1,000 readings available from 12 different publishers. The GDL collection offers readings to meet the needs of anyone interested in risk management.

For more information, please visit www.garpdigitallibrary.org.

GARP EVENTS AND NETWORKING

GARP hosts major conventions throughout the world, where risk professionals come together to share knowledge, network, and learn from leading experts in the field. Conventions are bookended with interactive workshops that provide practical insights and case studies presented by the industry's leading practitioners.

GARP regional chapters provide an opportunity for financial risk professionals to network and share new trends and discoveries in risk management. Each one of our 52 chapters holds several meetings each year, in some locations more often, focusing on issues of importance to the risk management community, either globally or locally.

GARP's formal mission is to be the leading professional association for financial risk managers, managed by and for its members and dedicated to the advancement of the risk profession through education, training, and the promotion of best practices globally. As a part of delivering on that mission, GARP has again teamed with Philippe Jorion to produce the fifth edition of the *Financial Risk Manager Handbook*.

The *Handbook* follows GARP's FRM Committee's published FRM Study Guide, which sets forth primary topics and subtopics covered in the FRM exam. The topics are selected by the FRM Committee as being representative of the theories and concepts utilized by risk management professionals as they address current issues.

Over the years the Study Guide has taken on an importance far exceeding its initial intent of providing guidance for FRM candidates. The Study Guide is now being used by universities, educators, and executives around the world to develop graduate-level business and finance courses, as a reference list for purchasing new readings for personal and professional libraries, as an objective outline to assess the risk management qualifications of an employee or a job applicant, and as guidance on the important trends currently affecting the financial risk management profession.

Given the expanded and dramatically growing recognition of the financial risk management profession globally, the *Handbook* has similarly assumed a natural and advanced role beyond its original purpose. It has now become the primary reference manual for risk professionals, academicians, and executives around the world. Professional risk managers must be well versed in a wide variety of risk-related concepts and theories, and must also keep themselves up-to-date with a rapidly changing marketplace. The *Handbook* is designed to allow them to do just that. It provides a financial risk management practitioner with the latest thinking and approaches to financial risk-related issues. It also provides coverage of advanced topics with questions and tutorials to enhance the reader's learning experience.

This fifth edition of the *Handbook* includes revised coverage of the primary topic areas covered by the FRM examination. Importantly, this edition also includes the latest lessons from the recent credit crisis, as well as new and more recent sample FRM questions.

The *Handbook* continues to keep pace with the dynamic financial risk profession while simultaneously offering serious risk professionals an excellent and cost-effective tool to keep abreast of the latest issues affecting the global risk management community.

Developing credibility and global acceptance for a professional certification program is a lengthy and complicated process. When GARP first administered its FRM exam in 1997, the concept of a professional risk manager and a global certification relating to that person's skill set was more theory than reality. That has now completely changed, as the number of current FRM holders exceeds 16,000.

The FRM is now the benchmark for a financial risk manager anywhere in the world. Professional risk managers having earned the FRM credential are globally recognized as having achieved a level of professional competency and a demonstrated ability to dynamically measure and manage financial risk in a real-world setting in accordance with global standards.

GARP is proud to continue to make this *Handbook* available to financial risk professionals around the world. Philippe Jorion, a preeminent risk management professional, has again compiled an exceptional reference book. Supplemented by an interactive test question CD, this *Handbook* is a requirement for any risk professional's library.

<div align="right">

Global Association of Risk Professionals
February 2009

</div>

PART
One

Quantitative Analysis

Bond Fundamentals

Risk management starts with the pricing of assets. The simplest assets to study are regular, fixed-coupon bonds. Because their cash flows are predetermined, we can translate their stream of cash flows into a present value by discounting at a fixed interest rate. Thus the valuation of bonds involves understanding compounded interest, discounting, as well as the relationship between present values and interest rates.

Risk management goes one step further than pricing, however. It examines potential changes in the price of assets as the interest rate changes. In this chapter, we assume that there is a single interest rate, or yield, that is used to price the bond. This will be our fundamental risk factor. This chapter describes the relationship between bond prices and yields and presents indispensable tools for the management of fixed-income portfolios.

This chapter starts our coverage of quantitative analysis by discussing bond fundamentals. Section 1.1 reviews the concepts of discounting, present values, and future values. Section 1.2 then plunges into the price-yield relationship. It shows how the Taylor expansion rule can be used to relate movements in bond prices to those in yields. This Taylor expansion rule, however, covers much more than bonds. It is a building block of risk measurement methods based on local valuation, as we shall see later. Section 1.3 then presents an economic interpretation of duration and convexity.

The reader should be forewarned that this chapter, like many others in this handbook, is rather compact. This chapter provides a quick review of bond fundamentals with particular attention to risk measurement applications. By the end of this chapter, however, the reader should be able to answer advanced FRM questions on bond mathematics.

1.1 DISCOUNTING, PRESENT, AND FUTURE VALUE

An investor considers a zero-coupon bond that pays $100 in 10 years. Assume that the investment is guaranteed by the U.S. government, and that there is no credit risk. So, this is a default-free bond, which is exposed to market risk only. Because the payment occurs at a future date, the current value of the investment is surely less than an up-front payment of $100.

To value the payment, we need a discounting factor. This is also the interest rate, or more simply the yield. Define C_t as the cash flow at time t and the

3

discounting factor as y. We define T as the number of periods until maturity, e.g., number of years, also known as tenor. The present value (PV) of the bond can be computed as

$$PV = \frac{C_T}{(1+y)^T} \tag{1.1}$$

For instance, a payment of $C_T = \$100$ in 10 years discounted at 6 percent is only worth $\$55.84$ now. So, all else fixed, the market value of zero-coupon bonds decreases with longer maturities. Also, keeping T fixed, the value of the bond decreases as the yield increases.

Conversely, we can compute the future value (FV) of the bond as

$$FV = PV \times (1+y)^T \tag{1.2}$$

For instance, an investment now worth $PV = \$100$ growing at 6 percent will have a future value of $FV = \$179.08$ in 10 years.

Here, the yield has a useful interpretation, which is that of an internal rate of return on the bond, or annual growth rate. It is easier to deal with rates of returns than with dollar values. Rates of return, when expressed in percentage terms and on an annual basis, are directly comparable across assets. An annualized yield is sometimes defined as the effective annual rate (EAR).

It is important to note that the interest rate should be stated along with the method used for compounding. Annual compounding is very common. Other conventions exist, however. For instance, the U.S. Treasury market uses semi-annual compounding. Define in this case y^S as the rate based on semiannual compounding. To maintain comparability, it is expressed in annualized form, i.e., after multiplication by 2. The number of periods, or semesters, is now $2T$. The formula for finding y^S is

$$PV = \frac{C_T}{(1+y^S/2)^{2T}} \tag{1.3}$$

For instance, a Treasury zero-coupon bond with a maturity of $T = 10$ years would have $2T = 20$ semiannual compounding periods. Comparing with (1.1), we see that

$$(1+y) = (1+y^S/2)^2 \tag{1.4}$$

Continuous compounding is often used when modeling derivatives. It is the limit of the case where the number of compounding periods per year increases to infinity. The continuously compounded interest rate y^C is derived from

$$PV = C_T \times e^{-y^C T} \tag{1.5}$$

where $e^{(\cdot)}$, sometimes noted as $\exp(\cdot)$, represents the exponential function.

Note that in all of these Equations (1.1), (1.3), and (1.5), the present value and future cash flows are identical. Because of different compounding periods, however, the yields will differ. Hence, the compounding period should always be stated.

Example: Using Different Discounting Methods

Consider a bond that pays \$100 in 10 years and has a present value of \$55.8395. This corresponds to an annually compounded rate of 6.00% using $PV = C_T/(1+y)^{10}$, or $(1+y) = (C_T/PV)^{1/10}$.

This rate can be transformed into a semiannual compounded rate, using $(1+y^S/2)^2 = (1+y)$, or $y^S/2 = (1+y)^{1/2} - 1$, or $y^S = ((1+0.06)^{(1/2)} - 1) \times 2 = 0.0591 = 5.91\%$. It can be also transformed into a continuously compounded rate, using $\exp(y^C) = (1+y)$, or $y^C = \ln(1+0.06) = 0.0583 = 5.83\%$.

Note that as we increase the frequency of the compounding, the resulting rate decreases. Intuitively, because our money works harder with more frequent compounding, a lower investment rate will achieve the same payoff at the end.

KEY CONCEPT

For fixed present value and cash flows, increasing the frequency of the compounding will decrease the associated yield.

EXAMPLE 1.1: FRM EXAM 2002—QUESTION 48

An investor buys a Treasury bill maturing in 1 month for \$987. On the maturity date the investor collects \$1,000. Calculate effective annual rate (EAR).

a. 17.0%

b. 15.8%

c. 13.0%

d. 11.6%

EXAMPLE 1.2: FRM EXAM 2002—QUESTION 51

Consider a savings account that pays an annual interest rate of 8%. Calculate the amount of time it would take to double your money. Round to the nearest year.

a. 7 years

b. 8 years

c. 9 years

d. 10 years

1.2 PRICE-YIELD RELATIONSHIP

1.2.1 Valuation

The fundamental discounting relationship from Equation (1.1) can be extended to any bond with a fixed cash-flow pattern. We can write the present value of a bond P as the discounted value of future cash flows:

$$P = \sum_{t=1}^{T} \frac{C_t}{(1+y)^t} \tag{1.6}$$

where: C_t = the cash flow (coupon or principal) in period t

t = the number of periods (e.g., half-years) to each payment

T = the number of periods to final maturity

y = the discounting factor per period (e.g., $y^S/2$)

A typical cash-flow pattern consists of a fixed coupon payment plus the repayment of the principal, or face value at expiration. Define c as the coupon *rate* and F as the face value. We have $C_t = cF$ prior to expiration, and at expiration, we have $C_T = cF + F$. The appendix reviews useful formulas that provide closed-form solutions for such bonds.

When the coupon rate c precisely matches the yield y, using the same compounding frequency, the present value of the bond must be equal to the face value. The bond is said to be a par bond. If the coupon is greater than the yield, the price must be greater than the face value, which means that this is a premium bond. Conversely, if the coupon is lower, or even zero for a zero-coupon bond, the price must be less than the face value, which means that this is a discount bond.

Equation (1.6) describes the relationship between the yield y and the value of the bond P, given its cash-flow characteristics. In other words, the value P can also be written as a nonlinear function of the yield y:

$$P = f(y) \tag{1.7}$$

Conversely, we can set P to the current market price of the bond, including any accrued interest. From this, we can compute the "implied" yield that will solve this equation.

Figure 1.1 describes the price-yield function for a 10-year bond with a 6% annual coupon. In risk management terms, this is also the relationship between the payoff on the asset and the risk factor. At a yield of 6%, the price is at par, $P = \$100$. Higher yields imply lower prices. This is an example of a payoff function, which links the price to the underlying risk factor.

Over a wide range of yield values, this is a highly nonlinear relationship. For instance, when the yield is zero, the value of the bond is simply the sum of cash
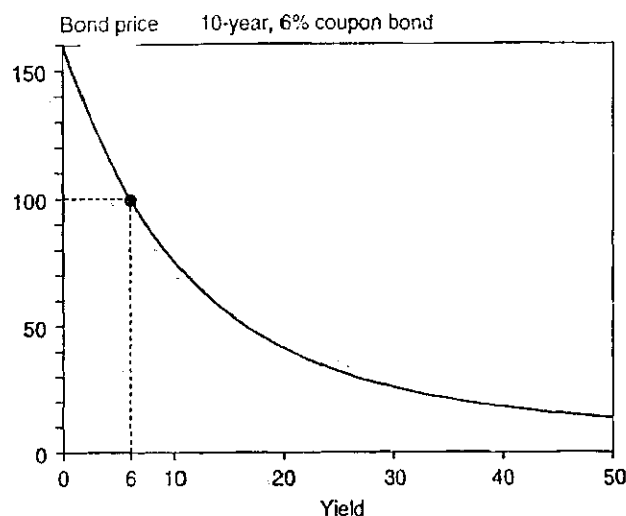
FIGURE 1.1   Price–Yield Relationship

flows, or $160 in this case. When the yield tends to very large values, the bond price tends to zero. For small movements around the initial yield of 6%, however, the relationship is quasilinear.

There is a particularly simple relationship for consols, or perpetual bonds, which are bonds making regular coupon payments but with no redemption date. For a consol, the maturity is infinite and the cash flows are all equal to a fixed percentage of the face value, $C_t = C = cF$. As a result, the price can be simplified from Equation (1.6) to

$$P = cF \left[\frac{1}{(1+y)} + \frac{1}{(1+y)^2} + \frac{1}{(1+y)^3} + \cdots \right] = \frac{c}{y} F \qquad (1.8)$$

as shown in the appendix. In this case, the price is simply proportional to the inverse of the yield. Higher yields lead to lower bond prices, and vice versa.

Example: Valuing a Bond

Consider a bond that pays $100 in 10 years and a 6% annual coupon. Assume that the next coupon payment is in exactly one year. What is the market value if the yield is 6%? If it falls to 5%?

The bond cash flows are $C_1 = \$6, C_2 = \$6, \ldots, C_{10} = \$106$. Using Equation (1.6) and discounting at 6%, this gives the present value of cash flows of $5.66, $5.34, ..., $59.19, for a total of $100.00. The bond is selling at par. This is logical because the coupon is equal to the yield, which is also annually compounded. Alternatively, discounting at 5% leads to a price of $107.72.

1.2.2 Taylor Expansion

Let us say that we want to see what happens to the price if the yield changes from its initial value, called y_0, to a new value, $y_1 = y_0 + \Delta y$. Risk management is all about assessing the effect of changes in risk factors such as yields on asset values. Are there shortcuts to help us with this?

We could recompute the new value of the bond as $P_1 = f(y_1)$. If the change is not too large, however, we can apply a very useful shortcut. The nonlinear relationship can be approximated by a Taylor expansion around its initial value[1]

$$P_1 = P_0 + f'(y_0)\Delta y + \frac{1}{2} f''(y_0)(\Delta y)^2 + \cdots \qquad (1.9)$$

where $f'(\cdot) = \frac{dP}{dy}$ is the first derivative and $f''(\cdot) = \frac{d^2 P}{dy^2}$ is the second derivative of the function $f(\cdot)$ valued at the starting point.[2] This expansion can be generalized to situations where the function depends on two or more variables. For bonds, the first derivative is related to the *duration* measure, and the second to *convexity*.

Equation (1.9) represents an infinite expansion with increasing powers of Δy. Only the first two terms (linear and quadratic) are ever used by finance practitioners. They provide a good approximation to changes in prices relative to other assumptions we have to make about pricing assets. If the increment is very small, even the quadratic term will be negligible.

Equation (1.9) is fundamental for risk management. It is used, sometimes in different guises, across a variety of financial markets. We will see later that this Taylor expansion is also used to approximate the movement in the value of a derivatives contract, such as an option on a stock. In this case, Equation (1.9) is

$$\Delta P = f'(S)\Delta S + \frac{1}{2} f''(S)(\Delta S)^2 + \cdots \qquad (1.10)$$

where S is now the price of the underlying asset, such as the stock. Here, the first derivative $f'(S)$ is called *delta*, and the second $f''(S)$, *gamma*.

The Taylor expansion allows easy aggregation across financial instruments. If we have x_i units (numbers) of bond i and a total of N different bonds in the portfolio, the portfolio derivatives are given by

$$f'(y) = \sum_{i=1}^{N} x_i f_i'(y) \qquad (1.11)$$

---

[1] This is named after the English mathematician Brook Taylor (1685–1731); who published this result in 1715. The full recognition of the importance of this result only came in 1755 when Euler applied it to differential calculus.

[2] This first assumes that the function can be written in polynomial form as $P(y + \Delta y) = a_0 + a_1 \Delta y + a_2 (\Delta y)^2 + \cdots$, with unknown coefficients a_0, a_1, a_2. To solve for the first, we set $\Delta y = 0$. This gives $a_0 = P_0$. Next, we take the derivative of both sides and set $\Delta y = 0$. This gives $a_1 = f'(y_0)$. The next step gives $2a_2 = f''(y_0)$. Here, the term "derivatives" takes the usual mathematical interpretation, and has nothing to do with *derivatives products* such as options.

1.3 BOND PRICE DERIVATIVES

For fixed-income instruments, the derivatives are so important that they have been given a special name.[3] The negative of the first derivative is the dollar duration (DD):

$$f'(y_0) = \frac{dP}{dy} = -D^* \times P_0 \tag{1.12}$$

where D^* is called the modified duration. Thus, dollar duration is

$$DD = D^* \times P_0 \tag{1.13}$$

where the price P_0 represent the *market* price, including any accrued interest. Sometimes, risk is measured as the dollar value of a basis point (DVBP),

$$DVBP = DD \times \Delta y = [D^* \times P_0] \times 0.0001 \tag{1.14}$$

with 0.0001 representing an interest rate change of one basis point (bp) or one hundredth of a percent. The DVBP, sometimes called the DV01, measures can be easily added up across the portfolio.

The second derivative is the dollar convexity (DC):

$$f''(y_0) = \frac{d^2 P}{dy^2} = C \times P_0 \tag{1.15}$$

where C is called the **convexity**.

For fixed-income instruments with known cash flow the price yield function is known, and we can compute analytical first and second derivatives. Consider, for example, our simple zero-coupon bond in Equation (1.1) where the only payment is the face value, $C_T = F$. We take the first derivative, which is

$$\frac{dP}{dy} = \frac{d}{dy}[\frac{F}{(1+y)^T}] = (-T)\frac{F}{(1+y)^{T+1}} = -\frac{T}{(1+y)}P \tag{1.16}$$

Comparing with Equation (1.12), we see that the modified duration must be given by $D^* = T/(1+y)$. The conventional measure of duration is $D = T$, which does not include division by $(1+y)$ in the denominator. This is also called Macaulay duration. Note that duration is expressed in periods, like T. With annual compounding, duration is in years. With semiannual compounding, duration is in semesters. It then has to be divided by two for conversion to years. Modified

---

[3] Note that this chapter does not present duration in the traditional textbook order. In line with the advanced focus on risk management, we first analyze the properties of duration as a sensitivity measure. This applies to any type of fixed-income instrument. Later, we will illustrate the usual definition of duration as a weighted average maturity, which applies for fixed-coupon bonds only.

duration D^* is related to Macaulay duration D

$$D^* = \frac{D}{(1 + y)}$$ (1.17)

Modified duration is the appropriate measure of interest rate exposure. The quantity $(1 + y)$ appears in the denominator because we took the derivative of the present value term with discrete compounding. If we use continuous compounding, modified duration is identical to the conventional duration measure. In practice, the difference between Macaulay and modified duration is usually small.

Let us now go back to Equation (1.16) and consider the second derivative, which is

$$\frac{d^2 P}{dy^2} = -(T + 1)(-T)\frac{F}{(1 + y)^{T+2}} = \frac{(T + 1)T}{(1 + y)^2} \times P$$ (1.18)

Comparing with Equation (1.15), we see that the convexity is $C = (T + 1)T/(1 + y)^2$. Note that its dimension is expressed in period squared. With semiannual compounding, convexity is measured in semesters squared. It then has to be divided by 4 for conversion to years squared.[4] So, convexity must be positive for bonds with fixed coupons.

Putting together all these equations, we get the Taylor expansion for the change in the price of a bond, which is

$$\Delta P = -[D^* \times P](\Delta y) + \frac{1}{2}[C \times P](\Delta y)^2 + \cdots$$ (1.19)

Therefore duration measures the first-order (linear) effect of changes in yield and convexity the second-order (quadratic) term.

Example: Computing the Price Approximation[5]

Consider a 10-year zero-coupon Treasury bond trading at a yield of 6 percent. The present value is obtained as $P = 100/(1 + 6/200)^{20} = 55.368$. As is the practice in the Treasury market, yields are semiannually compounded. Thus all computations should be carried out using semesters, after which final results can be converted into annual units.

Here, Macaulay duration is exactly 10 years, as $D = T$ for a zero coupon bond. Its modified duration is $D^* = 20/(1 + 6/200) = 19.42$ semesters, which is 9.71 years. Its convexity is $C = 21 \times 20/(1 + 6/200)^2 = 395.89$ semesters

---

[4] This is because the conversion to annual terms is obtained by multiplying the semiannual yield Δy by two. As a result, the duration term must be divided by 2 and the convexity term by 2^2, or 4, for conversion to annual units.

[5] For such examples in this handbook, please note that intermediate numbers are reported with fewer significant digits than actually used in the computations. As a result, using rounded off numbers may give results that differ slightly from the final numbers shown here.

squared, which is 98.97 in years squared. Dollar duration is $DD = D^* \times P = 9.71 \times \$55.37 = \$537.55$. The DVBP is $DVBP = DD \times 0.0001 = \$0.0538$.

We want to approximate the change in the value of the bond if the yield goes to 7%. Using Equation (1.19), we have $\Delta P = -[9.71 \times \$55.37](0.01) + 0.5[98.97 \times \$55.37](0.01)^2 = -\$5.375 + \$0.274 = -\$5.101$. Using the linear term only, the new price is $\$55.368 - \$5.375 = \$49.992$. Using the two terms in the expansion, the predicted price is slightly higher, at $\$55.368 - \$5.375 + \$0.274 = \$50.266$.

These numbers can be compared with the exact value, which is $50.257. The linear approximation has a relative pricing error of −0.53%, which is not bad. Adding a quadratic term reduces this to an error of 0.02% only, which is very small, given typical bid-ask spreads.

More generally, Figure 1.2 compares the quality of the Taylor series approximation. We consider a 10-year bond paying a 6 percent coupon semiannually. Initially, the yield is also at 6 percent and, as a result, the price of the bond is at par, at $100. The graph compares three lines representing

1. The actual, exact price $\qquad P = f(y_0 + \Delta y)$
2. The duration estimate $\qquad P = P_0 - D^* P_0 \Delta y$
3. The duration and convexity estimate $\quad P = P_0 - D^* P_0 \Delta y + (1/2) C P_0 (\Delta y)^2$

The actual price curve shows an increase in the bond price if the yield falls and, conversely, a depreciation if the yield increases. This effect is captured by the tangent to the true price curve, which represents the linear approximation based on duration. For small movements in the yield, this linear approximation provides a reasonable fit to the exact price.
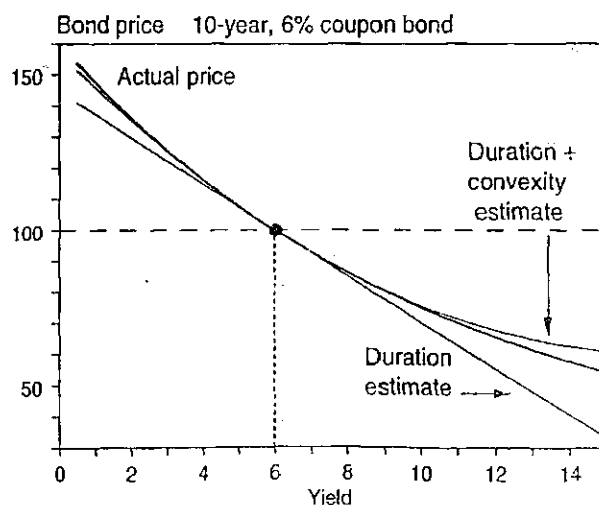

FIGURE 1.2 Price Approximation

> **KEY CONCEPT**
>
> Dollar duration measures the (negative)slope of the tangent to the price-yield curve at the starting point.

For large movements in price, however, the price-yield function becomes more curved and the linear fit deteriorates. Under these conditions, the quadratic approximation is noticeably better.

We should also note that the curvature is away from the origin, which explains the term *convexity* (as opposed to concavity). Figure 1.3 compares curves with different values for convexity. This curvature is beneficial since the second-order effect $0.5[C \times P](\Delta y)^2$ *must* be positive when convexity is positive.

As the figure shows, when the yield rises, the price drops but less than predicted by the tangent. Conversely, if the yield falls, the price increases faster than along the tangent. In other words, the quadratic term is always beneficial.

> **KEY CONCEPT**
>
> Convexity is always positive for regular coupon-paying bonds. Greater convexity is beneficial both for falling and rising yields.

The bond's modified duration and convexity can also be computed directly from numerical derivatives. Duration and convexity cannot be computed directly for some bonds, such as mortgage-backed securities, because their cash flows are uncertain. Instead, the portfolio manager has access to pricing models that can be used to reprice the securities under various yield environments.
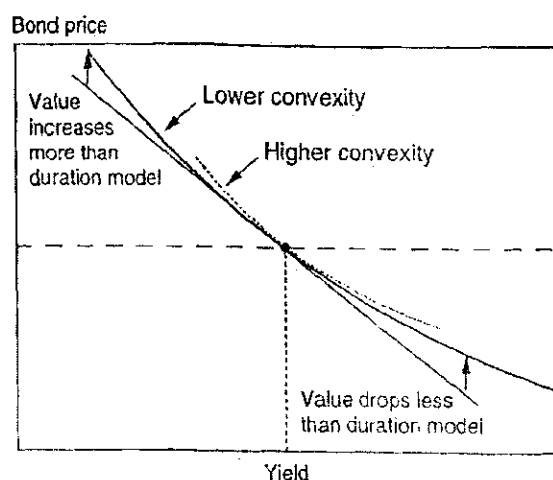

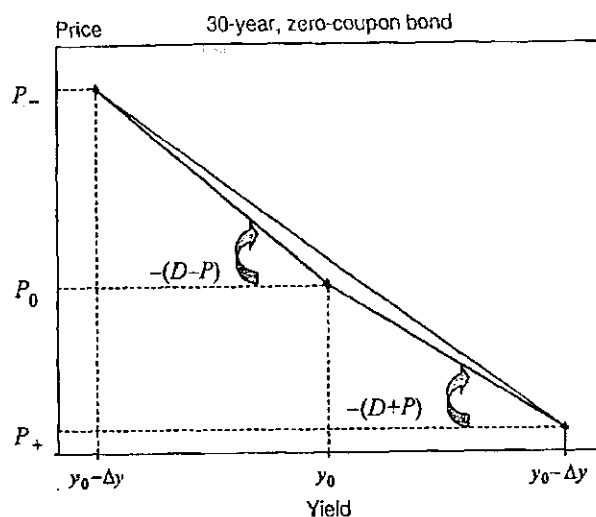
FIGURE 1.3   Effect of Convexity

FIGURE 1.4 Effective Duration and Convexity

As shown iri Figure 1.4, we choose a change in the yield, Δy, and reprice the bond under an upmove scenario, $P_+ = P(y_0 + \Delta y)$, and downmove scenario, $P_- = P(y_0 - \Delta y)$. Effective duration is measured by the numerical derivative. Using $D^* = -(1/P)dP/dy$, it is estimated as

$$D^E = \frac{[P_- - P_+]}{(2P_0\Delta y)} = \frac{P(y_0 - \Delta y) - P(y_0 + \Delta y)}{(2\Delta y)P_0} \qquad (1.20)$$

Using $C = (1/P)d^2P/dy^2$, effective convexity is estimated as

$$C^E = [D_- - D_+]/\Delta y = \left[\frac{P(y_0 - \Delta y) - P_0}{(P_0\Delta y)} - \frac{P_0 - P(y_0 + \Delta y)}{(P_0\Delta y)}\right]/\Delta y \qquad (1.21)$$

To illustrate, consider a 30-year zero-coupon bond with a yield of 6%, semi-annually compounded. The initial price is $16.9733. We revalue the bond at 5% and 7%, with prices shown in the table. The effective duration in Equation (1.20) uses the two extreme points. The effective convexity in Equation (1.21) uses the difference between the dollar durations for the upmove and downmove. Note that convexity is positive if duration increases as yields fall, or if $D_- > D_+$.

The computations are detailed in Table 1.1, which shows an effective duration of 29.56. This is very close to the true value of 29.13, and would be even closer if the step Δy was smaller. Similarly, the effective convexity is 869.11, which is close to the true value of 862.48.

Finally, this numerical approach can be applied to get an estimate of the duration of a bond by considering bonds with the same maturity but different coupons. If interest rates decrease by 1%, the market price of a 6% bond should go up to a value close to that of a 7% bond. Thus we replace a drop in yield of Δy with an increase in coupon Δc and use the effective duration method to find

TABLE 1.1 Effective Duration and Convexity

State	Yield (%)	Bond Value	Duration Computation	Convexity Computation
Initial y_0	6.00	16.9733		
Up $y_0 + \Delta y$	7.00	12.6934		Duration up: 25.22
Down $y_0 - \Delta y$	5.00	22.7284		Duration down: 33.91
Difference in values			−10.0349	8.69
Difference in yields			0.02	0.01
Effective measure			29.56	869.11
Exact measure			29.13	862.48

the coupon curve duration[6]

$$D^{CC} = \frac{[P_+ - P_-]}{(2P_0\Delta c)} = \frac{P(y_0; c + \Delta c) - P(y_0; c - \Delta c)}{(2\Delta c)P_0} \tag{1.22}$$

This approach is useful for securities which are difficult to price under various yield scenarios. It only requires the market prices of securities with different coupons.

Example: Computation of Coupon Curve Duration

Consider a 10-year bond that pays a 7% coupon semiannually. In a 7% yield environment, the bond is selling at par and has modified duration of 7.11 years. The prices of 6% and 8% coupon bonds are $92.89 and $107.11, respectively. This gives a coupon curve duration of $(107.11 - 92.89)/(0.02 \times 100) = 7.11$, which in this case is the same as modified duration.

EXAMPLE 1.3: FRM EXAM 2006—QUESTION 75

A zero-coupon bond with a maturity of 10 years has an annual effective yield of 10%. What is the closest value for its modified duration?

a. 9

b. 10

c. 99

d. 100

---

[6] For a more formal proof, we could take the pricing formula for a consol at par and compute the derivatives with respect to y and c. Apart from the sign, these derivatives are identical when $y = c$.

EXAMPLE 1.4: FRM EXAM 2007—QUESTION 115

A portfolio manager has a bond position worth USD 100 million. The position has a modified duration of eight years and a convexity of 150 years. Assume that the term structure is flat. By how much does the value of the position change if interest rates increase by 25 basis points?

a. USD −2,046,875
b. USD −2,187,500
c. USD −1,953,125
d. USD −1,906,250

EXAMPLE 1.5: FRM EXAM 2007—QUESTION 55

Consider the following three methods of estimating the profit and loss (P&L) of a bullet bond: full repricing, duration (PV01), and duration plus convexity. Rank the methods to estimate the P&L impact of a large negative yield shock from the lowest to the highest.

a. Duration, duration plus convexity, full repricing
b. Duration, full repricing, duration plus convexity
c. Duration plus convexity, duration, full repricing
d. Full repricing, duration plus convexity, duration

1.3.1 Interpreting Duration and Convexity

The preceding section has shown how to compute analytical formulas for duration and convexity in the case of a simple zero-coupon bond. We can use the same approach for coupon-paying bonds. Going back to Equation (1.6), we have

$$\frac{dP}{dy} = \sum_{t=1}^{T} \frac{-tC_t}{(1+y)^{t+1}} = -[\sum_{t=1}^{T} \frac{tC_t}{(1+y)^t}]/P \times \frac{P}{(1+y)} = -\frac{D}{(1+y)}P \qquad (1.23)$$

which defines duration as

$$D = \sum_{t=1}^{T} \frac{tC_t}{(1+y)^t}/P \qquad (1.24)$$

The economic interpretation of duration is that it represents the average time to wait for each payment, weighted by the present value of the associated cash flow. Indeed, replacing P, we can write

$$D = \sum_{t=1}^{T} t \frac{C_t/(1+y)^t}{\sum C_t/(1+y)^t} = \sum_{t=1}^{T} t \times w_t \qquad (1.25)$$

where the weights w_t represent the ratio of the present value of each cash flow C_t relative to the total, and sum to unity. This explains why the duration of a zero-coupon bond is equal to the maturity. There is only one cash flow and its weight is one.

KEY CONCEPT

(Macaulay) duration represents an average of the time to wait for all cash flows.

Figure 1.5 lays out the present value of the cash flows of a 6% coupon, 10-year bond. Given a duration of 7.80 years, this coupon-paying bond is equivalent to a zero-coupon bond maturing in exactly 7.80 years.

For bonds with fixed coupons, duration is less than maturity. For instance, Figure 1.6 shows how the duration of a 10-year bond varies with its coupon. With a zero coupon, Macaulay duration is equal to maturity. Higher coupons place more weight on prior payments and therefore reduce duration.
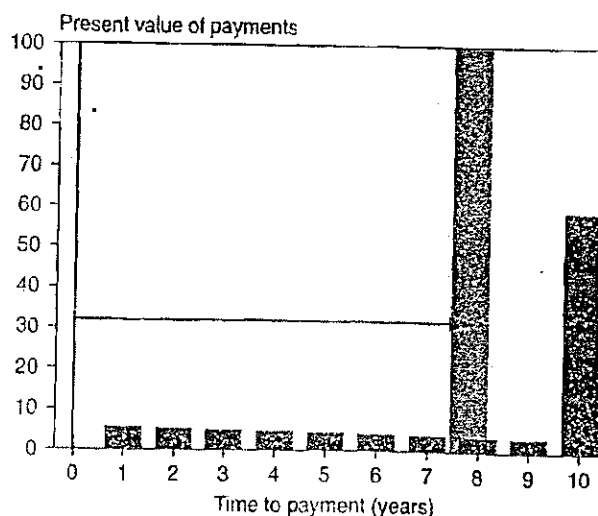

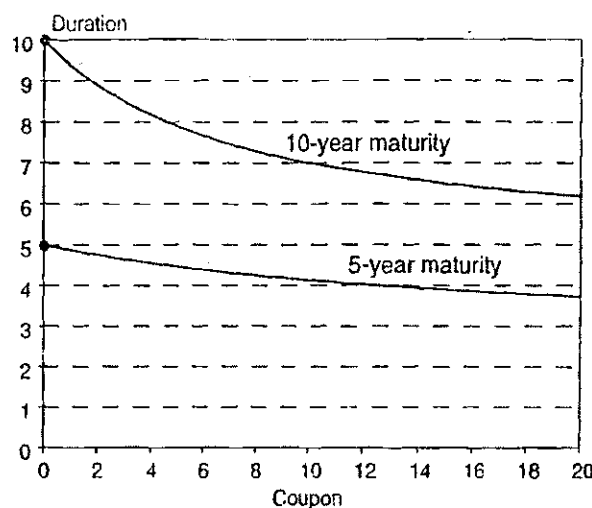**FIGURE 1.5** Duration as the Maturity of a Zero-Coupon Bond

FIGURE 1.6 Duration and Coupon

Duration can be expressed in a simple form for consols. From Equation (1.8), we have $P = (c/y)F$. Taking the derivative, we find

$$\frac{dP}{dy} = cF\frac{(-1)}{y^2} = (-1)\frac{1}{y}[\frac{c}{y}F] = (-1)\frac{1}{y}P = -\frac{D_C}{(1+y)}P \qquad (1.26)$$

Hence the Macaulay duration for the consol D_C is

$$D_C = \frac{(1+y)}{y} \qquad (1.27)$$

This shows that the duration of a consol is finite even if its maturity is infinite. Also, this duration does not depend on the coupon.

This formula provides a useful rule of thumb. For a long-term coupon-paying bond, duration should be lower than $(1+y)/y$. For instance, when $y = 6\%$, the upper limit on duration is $D_C = 1.06/0.06$, or 17.7 years. In this environment, the duration of a par 30-year bond is 14.25, which is indeed lower than 17.7 years.

KEY CONCEPT

The duration of a long-term bond can be approximated by an upper bound, which is that of a consol with the same yield, $D_C = (1+y)/y$.

Figure 1.7 describes the relationship between duration, maturity, and coupon for regular bonds in a 6% yield environment. For the zero-coupon bond, $D = T$, which is a straight line going through the origin. For the par 6% bond, duration increases monotonically with maturity until it reaches the asymptote of D_C. The
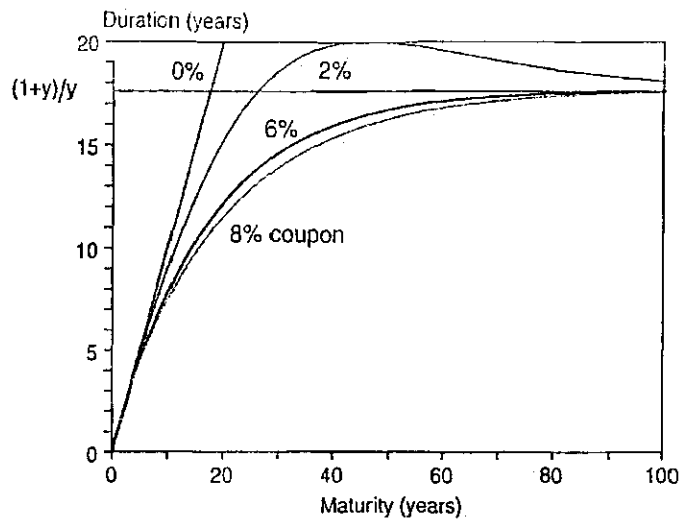
FIGURE 1.7 Duration and Maturity

8% bond has lower duration than the 6% bond for fixed T. Greater coupons, for a fixed maturity, decrease duration, as more of the payments come early.

Finally, the 2% bond displays a pattern intermediate between the zero-coupon and 6% bonds. It initially behaves like the zero, exceeding D_C initially then falling back to the asymptote, which is the same for all coupon-paying bonds.

Taking now the second derivative in Equation (1.23), we have

$$\frac{d^2 P}{dy^2} = \sum_{t=1}^{T} \frac{t(t+1)C_t}{(1+y)^{t+2}} = \left[\sum_{t=1}^{T} \frac{t(t+1)C_t}{(1+y)^{t+2}} / P \right] \times P \qquad (1.28)$$

which defines convexity as

$$C = \sum_{t=1}^{T} \frac{t(t+1)C_t}{(1+y)^{t+2}} / P \qquad (1.29)$$

Convexity can also be written as

$$C = \sum_{t=1}^{T} \frac{t(t+1)}{(1+y)^2} \times \frac{C_t/(1+y)^t}{\sum C_t/(1+y)^t} = \sum_{t=1}^{T} \frac{t(t+1)}{(1+y)^2} \times w_t \qquad (1.30)$$

Because the squared t term dominates in the fraction, this basically involves a weighted average of the square of time. Therefore, convexity is much greater for long-maturity bonds because they have payoffs associated with large values of t. The formula also shows that convexity is always positive for such bonds, implying that the curvature effect is beneficial. As we will see later, convexity can be negative for bonds that have uncertain cash flows, such as mortgage-backed securities (MBSs) or callable bonds.
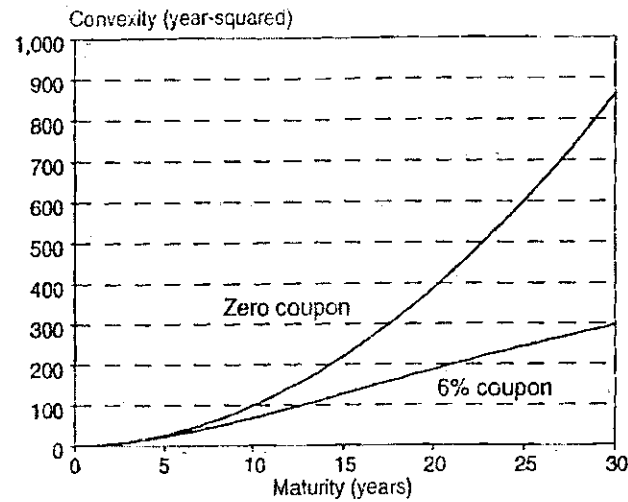
FIGURE 1.8 Convexity and Maturity

Figure 1.8 displays the behavior of convexity, comparing a zero-coupon bond with a 6% coupon bond with identical maturities. The zero-coupon bond always has greater convexity, because there is only one cash flow at maturity. Its convexity is roughly the square of maturity, for example about 900 for the 30-year zero. In contrast, the 30-year coupon bond has a convexity of about 300 only.

KEY CONCEPT

All else equal, duration and convexity both increase for longer maturities, lower coupons, and lower yields.

As an illustration, Table 1.2 details the steps of the computation of duration and convexity for a two-year, 6% semiannual coupon-paying bond. We first convert the annual coupon and yield into semiannual equivalent, $3 and 3% each. The PV column then reports the present value of each cash flow. We verify that these add up to $100, since the bond must be selling at par.

Next, the duration term column multiplies each PV term by time, or more precisely the number of half years until payment. This adds up to $382.86, which divided by the price gives $D = 3.83$. This number is measured in half years, and we need to divide by two to convert to years. Macaulay duration is 1.91 years, and modified duration $D^* = 1.91/1.03 = 1.86$ years. Note that, to be consistent, the adjustment in the denominator involves the semiannual yield of 3%.

Finally, the right-most column shows how to compute the bond's convexity. Each term involves PV_i times $t(t+1)/(1+y)^2$. These terms sum to 1,777.755, or divided by the price, 17.78. This number is expressed in units of time squared and must be divided by 4 to be converted in annual terms. We find a convexity of $C = 4.44$, in year-squared.

TABLE 1.2 Computing Duration and Convexity

Period (half-year) t	Payment C_t	Yield (%) (6 mo)	PV of Payment $C_t/(1+y)^t$	Duration Term tPV_t	Convexity Term $t(t+1)PV_t \times (1/(1+y)^2)$
1	3	3.00	2.913	2.913	5.491
2	3	3.00	2.828	5.656	15.993
3	3	3.00	2.745	8.236	31.054
4	103	3.00	91.514	366.057	1725.218
Sum:			100.00	382.861	1777.755
(half-years)				3.83	17.78
(years)				1.91	
Modified duration				1.86	
Convexity					4.44

EXAMPLE 1.6: FRM EXAM 2003—QUESTION 13

Suppose the face value of a three-year option-free bond is USD 1,000 and the annual coupon is 10%. The current yield to maturity is 5%. What is the modified duration of this bond?

a. 2.62
b. 2.85
c. 3.00
d. 2.75

EXAMPLE 1.7: FRM EXAM 2002—QUESTION 118

A Treasury bond has a coupon rate of 6% per annum (the coupons are paid semiannually) and a semiannually compounded yield of 4% per annum. The bond matures in 18 months and the next coupon will be paid 6 months from now. Which number below is closest to the bond's Macaulay duration?

a. 1.023 years
b. 1.457 years
c. 1.500 years
d. 2.915 years

EXAMPLE 1.8: DURATION AND COUPON

A and B are two perpetual bonds, that is, their maturities are infinite. A has a coupon of 4% and B has a coupon of 8%. Assuming that both are trading at the same yield, what can be said about the duration of these bonds?

a. The duration of A is greater than the duration of B.

b. The duration of A is less than the duration of B.

c. A and B both have the same duration.

d. None of the above.

EXAMPLE 1.9: FRM EXAM 2004—QUESTION 16

A manager wants to swap a bond for a bond with the same price but higher duration. Which of the following bond characteristics would be associated with a higher duration?

I. A higher coupon rate

II. More frequent coupon payments

III. A longer term to maturity

IV. A lower yield

    a. I, II, and III

    b. II, III, and IV

    c. III and IV

    d. I and II

EXAMPLE 1.10: FRM EXAM 2001—QUESTION 104

When the maturity of a plain coupon bond increases, its duration increases

a. Indefinitely and regularly

b. Up to a certain level

c. Indefinitely and progressively

d. In a way dependent on the bond being priced above or below par

---

EXAMPLE 1.11: FRM EXAM 2000—QUESTION 106

Consider the following bonds:

Bond Number	Maturity (yrs)	Coupon Rate	Frequency	Yield (Annual)
1	10	6%	1	6%
2	10	6%	2	6%
3	10	0%	1	6%
4	10	6%	1	5%
5	9	6%	1	6%

How would you rank the bonds from the shortest to longest duration?

a. 5-2-1-4-3

b. 1-2-3-4-5

c. 5-4-3-1-2

d. 2-4-5-1-3

EXAMPLE 1.12: FRM EXAM 2000—QUESTION 110

Which of the following statements are *true*?

I. The convexity of a 10-year zero-coupon bond is higher than the convexity of a 10-year, 6% bond.
II. The convexity of a 10-year zero-coupon bond is higher than the convexity of a 6% bond with a duration of 10 years.
III. Convexity grows proportionately with the maturity of the bond.
IV. Convexity is always positive for all types of bonds.
V. Convexity is always positive for "straight" bonds.

a. I only

b. I and II only

c. I and V only

d. II, III, and V only

1.3.2 Portfolio Duration and Convexity

Fixed-income portfolios often involve very large numbers of securities. It would be impractical to consider the movements of each security individually. Instead, portfolio managers aggregate the duration and convexity across the portfolio. A manager who believes that rates will increase should shorten the portfolio duration relative to that of the benchmark. Say for instance that the benchmark has a duration of 5 years. The manager shortens the portfolio duration to 1 year only. If rates increase by 2%, the benchmark will lose approximately $5y \times 2\% = 10\%$. The portfolio, however, will only lose $1y \times 2\% = 2\%$, hence "beating" the benchmark by 8%.

Because the Taylor expansion involves a summation, the portfolio duration is easily obtained from the individual components. Say we have N components indexed by i. Defining D_p^* and P_p as the portfolio modified duration and value, the portfolio dollar duration (DD) is

$$D_p^* P_p = \sum_{i=1}^{N} D_i^* x_i P_i \qquad (1.31)$$

where x_i is the number of units of bond i in the portfolio. A similar relationship holds for the portfolio dollar convexity (DC). If yields are the same for all components, this equation also holds for the Macaulay duration.

Because the portfolio's total market value is simply the summation of the component market values,

$$P_p = \sum_{i=1}^{N} x_i P_i \qquad (1.32)$$

we can define the portfolio weight w_i as $w_i = x_i P_i / P_p$, provided that the portfolio market value is nonzero. We can then write the portfolio duration as a weighted average of individual durations

$$D_p^* = \sum_{i=1}^{N} D_i^* w_i \qquad (1.33)$$

Similarly, the portfolio convexity is a weighted average of convexity numbers

$$C_p = \sum_{i=1}^{N} C_i w_i \qquad (1.34)$$

As an example, consider a portfolio invested in three bonds, described in Table 1.3. The portfolio is long a 10-year and 1-year bond, and short a 30-year zero-coupon bond. Its market value is $1,301,600. Summing the duration for each component, the portfolio dollar duration is $2,953,800, which translates into a

TABLE 1.3 Portfolio Dollar Duration and Convexity.

	Bond 1	Bond 2	Bond 3	Portfolio
Maturity (years)	10	1	30	
Coupon	6%	0%	0%	
Yield	6%	6%	6%	
Price P_i	$100.00	$94.26	$16.97	
Modified duration D_i^*	7.44	0.97	29.13	
Convexity C_i	68.78	1.41	862.48	
Number of bonds x_i	10,000	5,000	−10,000	
Dollar amounts $x_i P_i$	$1,000,000	$471,300	−$169,700	$1,301,600
Weight w_i	76.83%	36.21%	−13.04%	100.00%
Dollar duration $D_i^* P_i$	$744.00	$91.43	$494.34	
Portfolio DD: $x_i D_i^* P_i$	$7,440,000	$457,161	−$4,943,361	$2,953,800
Portfolio DC: $x_i C_i P_i$	68,780,000	664,533	−146,362,856	−76,918,323

duration of 2.27 years. The portfolio convexity is $-76,918,323/1,301,600 = -59.10$, which is negative due to the short position in the 30-year zero, which has very high convexity.

Alternatively, assume the portfolio manager is given a benchmark which is the first bond. He or she wants to invest in bonds 2 and 3, keeping the portfolio duration equal to that of the target, or 7.44 years. To achieve the target value and dollar duration, the manager needs to solve a system of two equations in the numbers x_1 and x_2:

Value: $\qquad \qquad \$100 = \qquad x_1 \$94.26 + \qquad x_2 \$16.97$

Dol.Duration: $\quad 7.44 \times \$100 = 0.97 \times x_1 \$94.26 + 29.13 \times x_2 \$16.97$

The solution is $x_1 = 0.817$ and $x_2 = 1.354$, which gives a portfolio value of $100 and modified duration of 7.44 years.[7] The portfolio convexity is 199.25, higher than the index. Such a portfolio consisting of very short and very long maturities is called a barbell portfolio. In contrast, a portfolio with maturities in the same range is called a bullet portfolio. Note that the barbell portfolio has a much greater convexity than the bullet bond because of the payment in 30 years. Such a portfolio would be expected to outperform the bullet portfolio if yields moved by a large amount.

In sum, duration and convexity are key measures of fixed-income portfolios. They summarize the linear and quadratic exposure to movements in yields. This explains why they are essential tools for fixed-income portfolio managers.

---

[7] This can be obtained by first expressing x_2 in the first equation as a function of x_1 and then substituting back into the second equation. This gives $x_2 = (100 - 94.26 x_1)/16.97$, and $744 = 91.43 x_1 + 494.34 x_2 = 91.43 x_1 + 494.34(100 - 94.26 x_1)/16.97 = 91.43 x_1 + 2913.00 - 2745.79 x_1$. Solving, we find $x_1 = (-2169.00)/(-2654.36) = 0.817$ and $x_2 = (100 - 94.26 \times 0.817)/16.97 = 1.354$.

EXAMPLE 1.13: FRM EXAM 2002—QUESTION 57

A bond portfolio has the following composition:

1. Portfolio A: price $90,000, modified duration 2.5, long position in 8 bonds
2. Portfolio B: price $110,000, modified duration 3, short position in 6 bonds
3. Portfolio C: price $120,000, modified duration 3.3, long position in 12 bonds

All interest rates are 10%. If the rates rise by 25 basis points, then the bond portfolio value will

  a. Decrease by $11,430
  b. Decrease by $21,330
  c. Decrease by $12,573
  d. Decrease by $23,463

EXAMPLE 1.14: FRM EXAM 2006—QUESTION 61

Consider the following portfolio of bonds (par amounts are in millions of USD).

Bond	Price	Par amount held	Modified Duration
A	101.43	3	2.36
B	84.89	5	4.13
C	121.87	8	6.27

What is the value of the portfolio's DV01 (dollar value of 1 basis point)?

  a. 8,019
  b. 8,294
  c. 8,584
  d. 8,813

1.4 IMPORTANT FORMULAS

Compounding: $(1 + y)^T = (1 + y^S/2)^{2T} = e^{y^C T}$

Fixed-coupon bond valuation: $P = \sum_{t=1}^{T} \frac{C_t}{(1+y)^t}$

Taylor expansion: $P_1 = P_0 + f'(y_0)\Delta y + \frac{1}{2} f''(y_0)(\Delta y)^2 + \cdots$

Duration as exposure: $\frac{dP}{dy} = -D^* \times P$, $DD = D^* \times P$, $DVBP = DD \times 0.0001$

Conventional duration: $D^* = \frac{D}{(1+y)}$, $D = \sum_{t=1}^{T} \frac{tC_t}{(1+y)^t}/P$

Convexity: $\frac{d^2P}{dy^2} = C \times P$, $C = \sum_{t=1}^{T} \frac{t(t+1)C_t}{(1+y)^{t+2}}/P$

Price change: $\Delta P = -[D^* \times P](\Delta y) + 0.5[C \times P](\Delta y)^2 + \cdots$

Consol: $P = \frac{c}{y}F$, $D = \frac{(1+y)}{y}$

Portfolio duration and convexity: $D_p^* = \sum_{i=1}^{N} D_i^* w_i$, $C_p = \sum_{i=1}^{N} C_i w_i$

1.5 ANSWERS TO CHAPTER EXAMPLES

Example 1.1: FRM Exam 2002—Question 48

a. The EAR is defined by $FV/PV = (1 + EAR)^T$. So $EAR = (FV/PV)^{1/T} - 1$. Here, $T = 1/12$. So, $EAR = (1,000/987)^{12} - 1 = 17.0\%$.

Example 1.2: FRM Exam 2002—Question 51

c. The time T relates the current and future values such that $FV/PV = 2 = (1 + 8\%)^T$. Taking logs of both sides, this gives $T = \ln(2)/\ln(1.08) = 9.006$.

Example 1.3: FRM Exam 2006—Question 75

a. Without doing any computation, the Macaulay duration must be 10 years because this is a zero-coupon bond. With annual compounding, modified duration is $D^* = 10/(1 + 10\%)$, or close to 9 years.

Example 1.4: FRM Exam 2007—Question 115

c. The change in price is given by $\Delta P = -[D^* \times P](\Delta y) + \frac{1}{2}[C \times P](\Delta y)^2$ $= -[8 \times 100](0.0025) + 0.5[150 \times 100](0.0025)^2 = -2.000000 + 0.046875 = -1.953125$.

Example 1.5: FRM Exam 2007—Question 55

a. When yields drop, the duration approximation gives the smallest price increase, so the answer must be either a. or b. Figure 1.2 shows that the full repricing curve for decreases in yields is slightly higher than the duration and convexity approximation. Alternatively, differentiating Equation 1.18 once more give a negative term for the third-order derivative. Combined with δy^3, which is negative, the third-order term must be positive.

Example 1.6: FRM Exam 2003—Question 13

d. As in Table 1.2, we lay out the cash flows and find

Period t	Payment C_t	Yield y	$PV_t = C_t/(1+y)^t$	tPV_t
1	100	5.00	95.24	95.24
2	100	5.00	90.71	181.41
3	1100	5.00	950.22	2850.66
Sum:			1136.16	3127.31

Duration is then 2.75, and modified duration 2.62.

Example 1.7: FRM Exam 2002—Question 118

b. For coupon-paying bonds, Macaulay duration is slightly less than the maturity, which is 1.5 year here. So, b. would be a good guess. Otherwise, we can compute duration exactly.

Example 1.8: Duration and Coupon

c. Going back to the duration equation for the consol, Equation (1.27), we see that it does not depend on the coupon but only on the yield. Hence, the durations must be the same. The price of bond A, however, must be half that of bond B.

Example 1.9: FRM Exam 2004—Question 16

c. Higher duration is associated with physical characteristics that push payments into the future, i.e., longer term, lower coupons, and less frequent coupon payments, as well as lower yields, which increase the relative weight of payments in the future.

Example 1.10: FRM Exam 2001—Question 104

b. With a fixed coupon, the duration goes up to the level of a consol with the same coupon. See Figure 1.7.

Example 1.11: FRM Exam 2000—Question 106

a. The nine-year bond (number 5) has shorter duration because the maturity is shortest, at nine years, among comparable bonds. Next, we have to decide between bonds 1 and 2, which only differ in the payment frequency. The semiannual bond (number 2) has a first payment in six months and has shorter duration than the annual bond. Next, we have to decide between bonds 1 and 4, which only differ in the yield. With lower yield, the cash flows further in the future have a higher weight, so that bond 4 has greater duration. Finally, the zero-coupon bond has the longest duration. So, the order is 5-2-1-4-3.

Example 1.12: FRM Exam 2000—Question 110

c. Because convexity is proportional to the square of time to payment, the convexity of a bond is mainly driven by the cash flows far into the future. Answer I. is correct because the 10-year zero has only one cash flow, whereas the coupon bond has several others that reduce convexity. Answer II. is false because the 6% bond with 10-year duration must have cash flows much further into the future, say in 30 years, which will create greater convexity. Answer III. is false because convexity grows with the square of time. Answer IV. is false because some bonds, for example MBSs or callable bonds, can have negative convexity. Answer V. is correct because convexity must be positive for coupon-paying bonds.

Example 1.13: FRM Exam 2002—Question 57

a. The portfolio dollar duration is $D^* P = \sum x_i D_i^* P_i = +8 \times 2.5 \times \$90,000 - 6 \times 3.0 \times \$110,000 + 12 \times 3.3 \times \$120,000 = \$4,572,000$. The change in portfolio value is then $-(D * P)(\Delta y) = -\$4,572,000 \times 0.0025 = -\$11,430$.

Example 1.14: FRM Exam 2006—Question 61

c. First, the market value of each bond is obtained by multiplying the par amount by the ratio of the market price divided by 100. Next, this is multiplied by D^* to get the dollar duration DD. Summing, this gives $85.841 million. We multiply by 1,000,000 to get dollar amounts and by 0.0001 to get the DV01, which gives $8,584.

Bond	Price	Par	Mkt value	D^*	DD
A	101.43	3	3.043	2.36	7.181
B	84.89	5	4.245	4.13	15.530
C	121.87	8	9.750	6.27	61.130
Sum		8			85.841

APPENDIX: APPLICATIONS OF INFINITE SERIES

When bonds have fixed coupons, the bond valuation problem often can be interpreted in terms of combinations of infinite series. The most important infinite series result is for a sum of terms that increase at a geometric rate:

$$1 + a + a^2 + a^3 + \cdots = \frac{1}{1 - a} \tag{1.35}$$

This can be proved, for instance, by multiplying both sides by $(1 - a)$ and canceling out terms.

Equally important, consider a geometric series with a finite number of terms, say N. We can write this as the difference between two infinite series:

$$1 + a + a^2 + a^3 + \cdots + a^{N-1}$$
$$= (1 + a + a^2 + a^3 + \cdots) - a^N(1 + a + a^2 + a^3 + \cdots) \tag{1.36}$$

such that all terms with order N or higher will cancel each other.

We can then write

$$1 + a + a^2 + a^3 + \cdots + a^{N-1} = \frac{1}{1-a} - a^N \frac{1}{1-a} \tag{1.37}$$

These formulas are essential to value bonds. Consider first a consol with an infinite number of coupon payments with a fixed coupon rate c. If the yield is y and the face value F, the value of the bond is

$$P = cF \left[\frac{1}{(1+y)} + \frac{1}{(1+y)^2} + \frac{1}{(1+y)^3} + \cdots \right]$$
$$= cF \frac{1}{(1+y)}[1 + a^2 + a^3 + \cdots]$$
$$= cF \frac{1}{(1+y)} \left[\frac{1}{1-a} \right]$$
$$= cF \frac{1}{(1+y)} \left[\frac{1}{(1-1/(1+y))} \right]$$
$$= cF \frac{1}{(1+y)} \left[\frac{(1+y)}{y} \right]$$
$$= \frac{c}{y} F$$

Similarly, we can value a bond with a *finite* number of coupons over T periods at which time the principal is repaid. This is really a portfolio with three parts:

1. A long position in a consol with coupon rate c
2. A short position in a consol with coupon rate c that starts in T periods
3. A long position in a zero-coupon bond that pays F in T periods.

Note that the combination of (1) and (2) ensures that we have a finite number of coupons. Hence, the bond price should be:

$$P = \frac{c}{y} F - \frac{1}{(1+y)^T} \frac{c}{y} F + \frac{1}{(1+y)^T} F = \frac{c}{y} F \left[1 - \frac{1}{(1+y)^T} \right] + \frac{1}{(1+y)^T} F \tag{1.38}$$

where again the formula can be adjusted for different compounding methods.

This is useful for a number of purposes. For instance, when $c = y$, it is immediately obvious that the price must be at par, $P = F$. This formula also can be used to find closed-form solutions for duration and convexity.

Fundamentals of Probability

The preceding chapter has laid out the foundations for understanding how bond prices move in relation to yields. More generally, the instrument can be described by a payoff function, which links the price to the underlying risk factor. Next, we have to characterize movements in bond yields, or more generally, any relevant risk factor in financial markets.

This is done with the tools of probability, a mathematical abstraction that describes the distribution of risk factors. Each risk factor is viewed as a random variable whose properties are described by a probability distribution function. These distributions can be processed with the payoff function to create a distribution of the profit and loss profile for the trading portfolio.

This chapter reviews the fundamental tools of probability theory for risk managers. Section 2.1 lays out the foundations, characterizing random variables by their probability density and distribution functions. These functions can be described by their principal moments, mean, variance, skewness, and kurtosis. Distributions with multiple variables are described in Section 2.2. Section 2.3 then turns to functions of random variables. Section 2.4 presents some examples of important distribution functions for risk management, including the uniform, normal, lognormal, Student's, binomial, and Poisson. Finally, Section 2.5 discusses limit distributions, which can be used to characterize the average and tails of independent random variables.

2.1 CHARACTERIZING RANDOM VARIABLES

The classical approach to probability is based on the concept of the random variable (rv). This can be viewed as the outcome from throwing a die, for example. Each realization is generated from a fixed process. If the die is perfectly symmetrical, with six faces, we could say that the probability of observing a face with a six in one throw is $p = 1/6$. Although the event itself is random, we can still make a number of useful statements from a fixed data-generating process.

The same approach can be taken to financial markets, where stock prices, exchange rates, yields, and commodity prices can be viewed as random variables. The assumption of a fixed data-generating process for these variables, however, is more tenuous than for the preceding experiment.

2.1.1 Univariate Distribution Functions

A random variable X is characterized by a distribution function,

$$F(x) = P(X \leq x) \tag{2.1}$$

which is the probability that the realization of the random variable X ends up less than or equal to the given number x. This is also called a cumulative distribution function.

When the variable X takes discrete values, this distribution is obtained by summing the step values less than or equal to x. That is,

$$F(x) = \sum_{x_j \leq x} f(x_j) \tag{2.2}$$

where the function $f(x)$ is called the frequency function or the probability density function (p.d.f.). Here, $f(x)$ is the probability of observing x. This function is characterized by its shape as well as fixed parameters, θ.

When the variable is continuous, the distribution is given by

$$F(x) = \int_{-\infty}^{x} f(u) du \tag{2.3}$$

The density can be obtained from the distribution using

$$f(x) = \frac{dF(x)}{dx} \tag{2.4}$$

Often, the random variable will be described interchangeably by its distribution or its density.

These functions have notable properties. The density $f(u)$ must be positive for all u. As x tends to infinity, the distribution tends to unity as it represents the total probability of any draw for x:

$$\int_{-\infty}^{\infty} f(u) du = 1 \tag{2.5}$$

Figure 2.1 gives an example of a density function $f(x)$, on the top panel, and of a cumulative distribution function $F(x)$ on the bottom panel. $F(x)$ measures the area under the $f(x)$ curve to the left of x, which is represented by the shaded area. Here, this area is 0.24. For small values of x, $F(x)$ is close to zero. Conversely, for large values of x, $F(x)$ is close to unity.

Example: Density Functions

A gambler wants to characterize the probability density function of the outcomes from a pair of dice. Because each has six faces, there are $6^2 = 36$ possible throw combinations. Out of these, there is one occurrence of an outcome of two (each die showing one). So, the frequency of an outcome of two is one. We can have two occurrences of a three (a one and a two and vice versa), and so on.

The gambler compiles the frequency of each value, from 2 to 12, as shown in Table 2.1. From this, he or she can compute the probability of each outcome. For instance, the probability of observing three is equal to 2, the frequency $n(x)$, divided by the total number of outcomes, of 36, which gives 0.0556. We can verify that all the probabilities indeed add up to one, since all occurrences must be accounted for. From the table, we see that the probability of an outcome of 3 or less is 8.33%.

TABLE 2.1   Probability Density Function

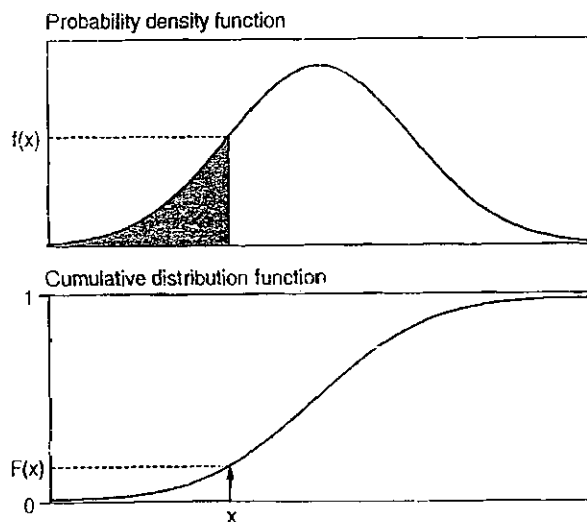| Outcome x_i	Frequency $n(x)$	Probability $f(x)$		Cumulative Probability $F(x)$
2	1	1/36	0.0278	0.0278
3	2	2/36	0.0556	0.0833
4	3	3/36	0.0833	0.1667
5	4	4/36.	0.1111	0.2778
6	5	5/36	0.1389	0.4167
7	6	6/36	0.1667	0.5833
8 .	5	5/36	0.1389	0.7222
9	4	4/36	0.1111	0.8333
10	3	3/36	0.0833	0.9167
11	2	2/36	0.0556	0.9722
12	1	1/36	0.0278	1.0000
Sum	36	1	1.0000	

FIGURE 2.1   Density and Distribution Functions

2.1.2 Moments

A random variable is characterized by its distribution function. Instead of having to report the whole function, it is convenient to summarize it by a few parameters, or moments.

For instance, the expected value for x, or mean, is given by the integral

$$\mu = E(X) = \int_{-\infty}^{+\infty} xf(x)dx \qquad (2.6)$$

which measures the *central tendency*, or *center of gravity* of the population.

The distribution can also be described by its quantile, which is the cutoff point x with an associated probability c:

$$F(x) = \int_{-\infty}^{x} f(u)du = c \qquad (2.7)$$

So, there is a probability of c that the random variable will fall *below* x. Because the total probability adds up to one, there is a probability of $p = 1 - c$ that the random variable will fall *above* x. Define this quantile as $Q(X, c)$. The 50% quantile is known as the median.

In fact, value at risk (VAR) can be interpreted as the cutoff point such that a loss will not happen with probability greater than $p = 95\%$, say. If $f(u)$ is the distribution of profit and losses on the portfolio, VAR is defined from

$$F(x) = \int_{-\infty}^{x} f(u)du = (1 - p) \qquad (2.8)$$

where p is the right-tail probability, and c the usual left-tail probability. VAR can be defined as minus the quantile itself, or alternatively, the deviation between the expected value and the quantile,

$$VAR(c) = E(X) - Q(X, c) \qquad (2.9)$$

Note that VAR is typically reported as a loss, i.e., a positive number, which explains the negative sign. Figure 2.2 shows an example with $c = 5\%$.
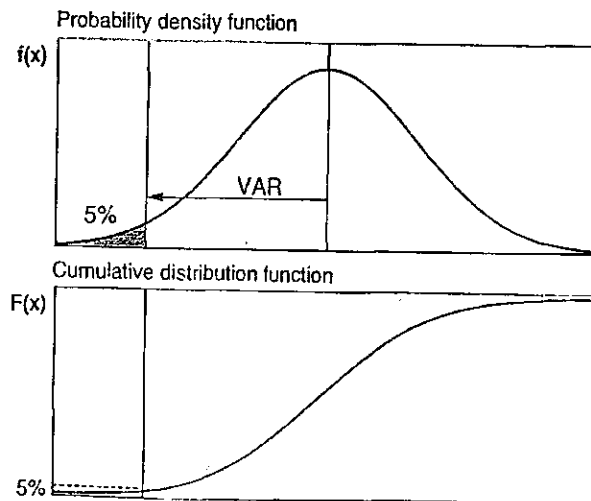

Probability density function

Cumulative distribution function

FIGURE 2.2 VAR as a Quantile
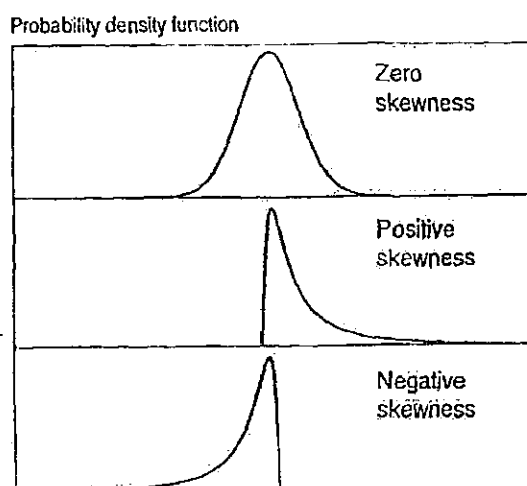
Probability density function

FIGURE 2.3 Effect of Skewness

Another useful moment is the squared dispersion around the mean, or variance

$$\sigma^2 = V(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)dx \qquad (2.10)$$

The standard deviation is more convenient to use as it has the same units as the original variable X

$$SD(X) = \sigma = \sqrt{V(X)} \qquad (2.11)$$

Next, the scaled third moment is the skewness, which describes departures from symmetry. It is defined as

$$\gamma = \left(\int_{-\infty}^{+\infty} [x - E(X)]^3 f(x)dx \right) \Big/ \sigma^3 \qquad (2.12)$$

Negative skewness indicates that the distribution has a long left tail, which indicates a high probability of observing large negative values. If this represents the distribution of profits and losses for a portfolio, this is a dangerous situation. Figure 2.3 displays distributions with various signs for the skewness.

The scaled fourth moment is the kurtosis, which describes the degree of "flatness" of a distribution, or width of its tails. It is defined as

$$\delta = \left(\int_{-\infty}^{+\infty} [x - E(X)]^4 f(x)dx \right) \Big/ \sigma^4 \qquad (2.13)$$

Because of the fourth power, large observations in the tail will have a large weight and hence create large kurtosis. Such a distribution is called leptokurtic, or fat-tailed. This parameter is very important for risk measurement. A kurtosis of 3 is considered average. High kurtosis indicates a higher probability of extreme movements. A distribution with kurtosis lower than 3 is called platykurtic. Figure 2.4 displays distributions with various values for the kurtosis.
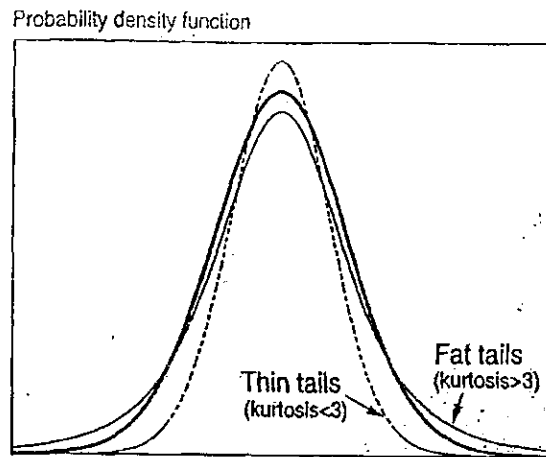
Probability density function

FIGURE 2.4 Effect of Kurtosis

Example: Computing Moments

Our gambler wants to know the expected value of the outcome of throwing two dice. He computes the product of each outcome and associated probability, as shown in Table 2.2. For instance, the first entry is $xf(x) = 2 \times 0.0278 = 0.0556$, and so on. Summing across all events, the mean is $\mu = 7.000$. This is also the median, since the distribution is perfectly symmetrical.

Next, we can use Equation (2.10) to compute the variance. The first term is $(x - \mu)^2 f(x) = (2 - 7)^2 0.0278 = 0.6944$. These terms add up to 5.8333, or, taking the square root, $\sigma = 2.4152$. The skewness terms sum to zero, because for each entry with a positive deviation $(x - \mu)^3$, there is an identical one with a negative sign and with the same probability. Finally, the kurtosis terms $(x - \mu)^4 f(x)$ sum to 80.5. Dividing by $\sigma^4 = 34.0278$, this gives a kurtosis of $\delta = 2.3657$.

TABLE 2.2 Computing Moments of a Distribution

Outcome x_i	Prob. $f(x)$	Mean $xf(x)$	Variance $(x - \mu)^2 f(x)$	Skewness $(x - \mu)^3 f(x)$	Kurtosis $(x - \mu)^4 f(x)$
2	0.0278	0.0556	0.6944	−3.4722	17.3611
3	0.0556	0.1667	0.8889	−3.5556	14.2222
4	0.0833	0.3333	0.7500	−2.2500	6.7500
5	0.1111	0.5556	0.4444	−0.8889	1.7778
6	0.1389	0.8333	0.1389	−0.1389	0.1389
7	0.1667	1.1667	0.0000	0.0000	0.0000
8	0.1389	1.1111	0.1389	0.1389	0.1389
9	0.1111	1.0000	0.4444	0.8889	1.7778
10	0.0833	0.8333	0.7500	2.2500	6.7500
11	0.0556	0.6111	0.8889	3.5556	14.2222
12	0.0278	0.3333	0.6944	3.4722	17.3611
Sum	1.0000	7.0000	$\sigma^2 = 5.8333$	0.0000	80.5000
Denominator				$\sigma^3 = 14.0888$	$\sigma^4 = 34.0278$
		Mean $\mu = 7.00$	StdDev $\sigma = 2.4152$	Skewness $\gamma = 0.0000$	Kurtosis $\delta = 2.3657$

2.2 MULTIVARIATE DISTRIBUTION FUNCTIONS

In practice, portfolio payoffs depend on numerous random variables. To simplify, start with two random variables. This could represent two currencies, or two interest rate factors, or default and credit exposure, to give just a few examples.

2.2.1 Joint Distributions

We can extend Equation (2.1) to

$$F_{12}(x_1, x_2) = P(X_1 \le x_1, X_2 \le x_2) \tag{2.14}$$

which defines a joint bivariate distribution function. In the continuous case, this is also

$$F_{12}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{12}(u_1, u_2) du_1 du_2 \tag{2.15}$$

where $f(u_1, u_2)$ is now the joint density. In general, adding random variables considerably complicates the characterization of the density or distribution functions.

The analysis simplifies considerably if the variables are independent. In this case, the joint density separates out into the product of the densities:

$$f_{12}(u_1 u_2) = f_1(u_1) \times f_2(u_2) \tag{2.16}$$

and the integral reduces to

$$F_{12}(x_1, x_2) = F_1(x_1) \times F_2(x_2) \tag{2.17}$$

This is very convenient because we only need to know the individual densities to reconstruct the joint density. For example, a credit loss can be viewed as a combination of (1) default, which is a random variable with a value of one for default and zero otherwise, and (2) the exposure, which is a random variable representing the amount at risk, for instance the positive market value of a swap. If the two variables are independent, we can construct the distribution of the credit loss easily. In the case of the two dice, the events are indeed independent. As a result, the probability of a joint event is simply the product of probabilities. For instance, the probability of throwing two ones is equal to $1/6 \times 1/6 = 1/36$.

It is also useful to characterize the distribution of x_1 abstracting from x_2. By integrating over all values of x_2, we obtain the marginal density

$$f_1(x_1) = \int_{-\infty}^{\infty} f_{12}(x_1, u_2) du_2 \tag{2.18}$$

and similarly for x_2. We can then define the conditional density as

$$f_{1 \cdot 2}(x_1 \mid x_2) = \frac{f_{12}(x_1, x_2)}{f_2(x_2)} \tag{2.19}$$

Here, we keep x_2 fixed and divide the joint density by the marginal probability of x_2. This normalization is necessary to ensure that the conditional density is a proper density function that integrates to one. This relationship is also known as Bayes' rule.

2.2.2 Copulas

When the two variables are independent, the joint density is simply the product of the marginal densities. It is rarely the case, however, that financial variables are independent. Dependencies can be modeled by a function called the copula, which links, or attaches, marginal distributions into a joint distribution. Formally, the copula is a function of the marginal distributions $F(x)$, plus some parameters, θ, that are specific to this function (and not to the marginals). In the bivariate case, it has two arguments

$$c_{12}[F_1(x_1), F_2(x_2); \theta] \tag{2.20}$$

The link between the joint and marginal distribution is made explicit by *Sklar's theorem*, which states that, for any joint density, there exists a copula that links the marginal densities

$$f_{12}(x_1, x_2) = f_1(x_1) \times f_2(x_2) \times c_{12}[F_1(x_1), F_2(x_2); \theta] \tag{2.21}$$

With independence, the copula function is a constant always equal to one.

Thus the copula contains all the information on the nature of the dependence between the random variables but gives no information on the marginal distributions. Complex dependencies can be modeled with different copulas. Copulas are now used extensively for modeling financial instruments such as collateralized debt obligations (CDOs). As we shall see in a later chapter, CDOs involve movements in many random variables, which are the default events for the companies issuing the debt.

2.2.3 Covariances and Correlations

When dealing with two random variables, the comovement can be described by the covariance

$$\text{Cov}(X_1, X_2) = \sigma_{12} = \int_1 \int_2 [x_1 - E(X_1)][x_2 - E(X_2)] f_{12}(x_1, x_2) dx_1 dx_2 \tag{2.22}$$

It is often useful to scale the covariance into a unitless number, called the correlation coefficient, obtained as

$$\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sigma_1 \sigma_2} \tag{2.23}$$

The correlation coefficient is a measure of linear dependence. One can show that the correlation coefficient always lies in the $[-1, +1]$ interval. A correlation of one

means that the two variables always move in the same direction. A correlation of minus one means that the two variables always move in opposite direction.

If the variables are independent, the joint density separates out and this becomes

$$\text{Cov}(X_1, X_2) = \left\{ \int_1 [x_1 - E(X_1)] f_1(x_1) dx_1 \right\} \left\{ \int_2 [x_2 - E(X_2)] f_2(x_2) dx_2 \right\} = 0$$

by Equation (2.6), since the average deviation from the mean is zero. In this case, the two variables are said to be uncorrelated. Hence independence implies zero correlation (the reverse is not true, however).

Example: Multivariate Functions

Consider two variables, such as the exchange rates for the Canadian dollar and the euro. Table 2.3a describes the joint density function $f_{12}(x_1, x_2)$, assuming two payoffs only for each variable. Note first that the density indeed sums to $0.30 + 0.20 + 0.15 + 0.35 = 1.00$.

TABLE 2.3a   Joint Density Function

x_1 x_2	-5	$+5$
-10	0.30	0.15
$+10$	0.20	0.35

From this, we can compute the marginal density for each variable, along with its mean and standard deviation. For instance, the marginal probability of $x_1 = -5$ is given by $f_1(x_1) = f_{12}(x_1, x_2 = -10) + f_{12}(x_1, x_2 = +10) = 0.30 + 0.20 = 0.50$. The marginal probability of $x_1 = +5$ must be 0.50 as well. Table 2.3b shows that the means and standard deviations are, respectively, $\bar{x}_1 = 0.0, \sigma_1 = 5.0$, and $\bar{x}_2 = 1.0, \sigma_2 = 9.95$.

Finally, Table 2.3c details the computation of the covariance, which gives $\text{Cov} = 15.00$. Dividing by the product of the standard deviations, we get $\rho = \text{Cov}/(\sigma_1 \sigma_2) = 15.00/(5.00 \times 9.95) = 0.30$. The positive correlation indicates that when one variable goes up, the other is more likely to go up than down.

TABLE 2.3b   Marginal Density Functions

	Variable 1				Variable 2		
x_i	Prob. $f_1(x_1)$	Mean $x_1 f_1(x_1)$	Variance $(x_1 - \bar{x}_1)^2 f_1(x_1)$	x_2	Prob. $f_2(x_2)$	Mean $x_2 f_2(x_2)$	Variance $(x_2 - \bar{x}_2)^2 f_2(x_2)$
-5	0.50	-2.5	12.5	-10	0.45	-4.5	54.45
$+5$	0.50	$+2.5$	12.5	$+10$	0.55	$+5.5$	44.55
Sum	1.00	0.0	25.0	Sum	1.00	1.0	99.0
		$\bar{x}_1 = 0.0$	$\sigma_1 = 5.0$			$\bar{x}_2 = 1.0$	$\sigma_2 = 9.95$

	$x_1 = -5$	$x_1 = +5$
$x_2 = -10$	$(-5 - 0)(-10 - 1)0.30 = 16.50$	$(+5 - 0)(-10 - 1)0.15 = -8.25$
$x_2 = +10$	$(-5 - 0)(+10 - 1)0.20 = -9.00$	$(+5 - 0)(+10 - 1)0.35 = 15.75$
Sum		$Cov = 15.00$

EXAMPLE 2.1: FRM EXAM 2000—QUESTION 81

Which one of the following statements about the correlation coefficient is *false?*

a. It always ranges from -1 to $+1$.

b. A correlation coefficient of zero means that two random variables are independent.

c. It is a measure of linear relationship between two random variables.

d. It can be calculated by scaling the covariance between two random variables.

EXAMPLE 2.2: FRM EXAM 2007—QUESTION 93

The joint probability distribution of random variables X and Y is given by $f(x, y) = k \times x \times y$ for $x = 1, 2, 3$, $y = 1, 2, 3$, and k is a positive constant. What is the probability that $X + Y$ will exceed 5?

a. 1/9

b. 1/4

c. 1/36

d. Cannot be determined

2.3 FUNCTIONS OF RANDOM VARIABLES

Risk management is about uncovering the distribution of portfolio values. Consider a security that depends on a unique source of risk, such as a bond. The risk manager could model the change in the bond price as a random variable directly. The problem with this choice is that the distribution of the bond price is not stationary, because the price converges to the face value at expiration.

Instead, the practice is to model the change in yields as a random variable because its distribution is better behaved. The next step is to use the relationship between the bond price and the yield to uncover the distribution of the bond price.

This illustrates a general principle of risk management, which is to model the risk factor first, then to derive the distribution of the instrument from information about the function that links the instrument value to the risk factor. This may not be easy to do, unfortunately, if the relationship is highly nonlinear. In what follows, we first focus on the mean and variance of simple transformations of random variables.

2.3.1 Linear Transformation of Random Variables

Consider a transformation that multiplies the original random variable by a constant and add a fixed amount, $Y = a + bX$. The expectation of Y is

$$E(a + bX) = a + bE(X) \qquad (2.24)$$

and its variance is

$$V(a + bX) = b^2 V(X) \qquad (2.25)$$

Note that adding a constant never affects the variance since the computation involves the *difference* between the variable and its mean. The standard deviation is

$$SD(a + bX) = bSD(X) \qquad (2.26)$$

Example: Currency Position Plus Cash

A dollar-based investor has a portfolio consisting of $1 million in cash plus a position in 1,000 million Japanese yen. The distribution of the dollar/yen exchange rate X has a mean of $E(X) = 1/100 = 0.01$ and volatility of $SD(X) = 0.10/100 = 0.001$.

The portfolio value can be written as $Y = a + bX$, with fixed parameters (in millions) $a = \$1$ and $b = Y1,000$. Therefore, the portfolio expected value is $E(Y) = \$1 + Y1,000 \times 1/100 = \$11$ million, and the standard deviation is $SD(Y) = Y1,000 \times 0.001 = \$1$ million.

2.3.2 Sum of Random Variables

Another useful transformation is the summation of two random variables. A portfolio, for instance, could contain one share of Intel plus one share of Microsoft. The rate of return on each stock behaves as a random variable.

The expectation of the sum $Y = X_1 + X_2$ can be written as

$$E(X_1 + X_2) = E(X_1) + E(X_2) \tag{2.27}$$

and its variance is

$$V(X_1 + X_2) = V(X_1) + V(X_2) + 2\text{Cov}(X_1, X_2) \tag{2.28}$$

When the variables are uncorrelated, the variance of the sum reduces to the sum of variances. Otherwise, we have to account for the cross-product term.

> **KEY CONCEPT**
>
> The expectation of a sum is the sum of expectations. The variance of a sum, however, is only the sum of variances if the variables are uncorrelated.

2.3.3 Portfolios of Random Variables

More generally, consider a linear combination of a number of random variables. This could be a portfolio with fixed weights, for which the rate of return is

$$Y = \sum_{i=1}^{N} w_i X_i \tag{2.29}$$

where N is the number of assets, X_i is the rate of return on asset i, and w_i its weight.

To shorten notation, this can be written in matrix notation, replacing a string of numbers by a single vector:

$$Y = w_1 X_1 + w_2 X_2 + \cdots + w_N X_N = [w_1 w_2 \ldots w_N] \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} = w'X \tag{2.30}$$

where w' represents the transposed vector (i.e., horizontal) of weights and X is the vertical vector containing individual asset returns. The appendix for this chapter provides a brief review of matrix multiplication.

The portfolio expected return is now

$$E(Y) = \mu_p = \sum_{i=1}^{N} w_i \mu_i \tag{2.31}$$

which is a weighted average of the expected returns $\mu_i = E(X_i)$. The variance is

$$V(Y) = \sigma_p^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} w_i w_j \sigma_{ij} = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{N} \sum_{j<i}^{N} w_i w_j \sigma_{ij}$$

(2.32)

Using matrix notation, the variance can be written as

$$\sigma_p^2 = [w_1 \ldots w_N] \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1N} \\ \vdots & & & & \vdots \\ \sigma_{N1} & \sigma_{N2} & \sigma_{N3} & \cdots & \sigma_N \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix}$$

Defining Σ as the covariance matrix, the variance of the portfolio rate of return can be written more compactly as

$$\sigma_p^2 = w' \Sigma w$$

(2.33)

This is a useful expression to describe the risk of the total portfolio.

Example: Computing the Risk of a Portfolio

Consider a portfolio invested in Canadian dollars and euros. The joint density function is given by Table 2.3a. Here, x_1 describes the payoff on the Canadian dollar, with $\mu_1 = 0.00$, $\sigma_1 = 5.00$, and $\sigma_1^2 = 25$. For the euro, $\mu_2 = 1.00$, $\sigma_2 = 9.95$, and $\sigma_2^2 = 99$. The covariance was computed as $\sigma_{12} = 15.00$, with the correlation $\rho = 0.30$. If we have 60% invested in Canadian dollar and 40% in euros, what is the portfolio volatility?

Following Equation (2.33), we write

$$\sigma_p^2 = [0.60\ 0.40] \begin{bmatrix} 25 & 15 \\ 15 & 99 \end{bmatrix} \begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix} = [0.60\ 0.40] \begin{bmatrix} 25 \times 0.60 + 15 \times 0.40 \\ 15 \times 0.60 + 99 \times 0.40 \end{bmatrix}$$

$$\sigma_p^2 = [0.60\ 0.40] \begin{bmatrix} 21.00 \\ 48.60 \end{bmatrix} = 0.60 \times 21.00 + 0.40 \times 48.60 = 32.04$$

Therefore, the portfolio volatility is $\sigma_p = \sqrt{32.04} = 5.66$. Note that this is hardly higher than the volatility of the Canadian dollar alone, even though the risk of the euro is much higher. The portfolio risk has been kept low due to a diversification effect, or low correlation between the two assets.

2.3.4 Product of Random Variables

Some risks result from the product of two random variables. A credit loss, for instance, arises from the product of the occurrence of default and the loss given default.

Using Equation (2.22), the expectation of the product $Y = X_1 X_2$ can be written as

$$E(X_1 X_2) = E(X_1)E(X_2) + \text{Cov}(X_1, X_2) \qquad (2.34)$$

When the variables are independent, this reduces to the product of the means. The variance is more complex to evaluate. With independence, it reduces to:

$$V(X_1 X_2) = E(X_1)^2 V(X_2) + V(X_1)E(X_2)^2 + V(X_1)V(X_2) \qquad (2.35)$$

2.3.5 Distributions of Transformations of RVs

The preceding results focus on the mean and variance of simple transformations only. They do not fully describe the distribution of the transformed variable $Y = g(X)$. This, unfortunately, is usually complicated for all but the simplest transformations $g(\cdot)$ and densities $f(X)$.

Even if there is no closed-form solution for the density, we can describe the cumulative distribution function of Y when $g(X)$ is a one-to-one transformation from X into Y. This implies that the function can be inverted, or that for a given y, we can find x such that $x = g^{-1}(y)$. We can then write

$$P[Y \le y] = P[g(X) \le y] = P[X \le g^{-1}(y)] = F_X(g^{-1}(y)) \qquad (2.36)$$

where $F(\cdot)$ is the cumulative distribution function of X. Here, we assumed the relationship is positive. Otherwise, the right-hand term is changed to $1 - F_X(g^{-1}(y))$.

This allows us to derive the quantile of, say, the bond price from information about the probability distribution of the yield. Suppose we consider a zero-coupon bond, for which the market value V is

$$V = \frac{100}{(1+r)^T} \qquad (2.37)$$

where r is the yield. This equation describes V as a function of r, or $Y = g(X)$. Using $r = 6\%$ and $T = 30$ years, the current price is $V = \$17.41$. The inverse function $X = g^{-1}(Y)$ is

$$r = (100/V)^{1/T} - 1 \qquad (2.38)$$

We wish to estimate the probability that the bond price could fall below a cut-off price $V = \$15$. We invert the price-yield function and compute the associated yield level, $g^{-1}(y) = (100/\$15)^{1/30} - 1 = 6.528\%$. Lower prices are associated with higher yield levels. Using Equation (2.36), the probability is given by
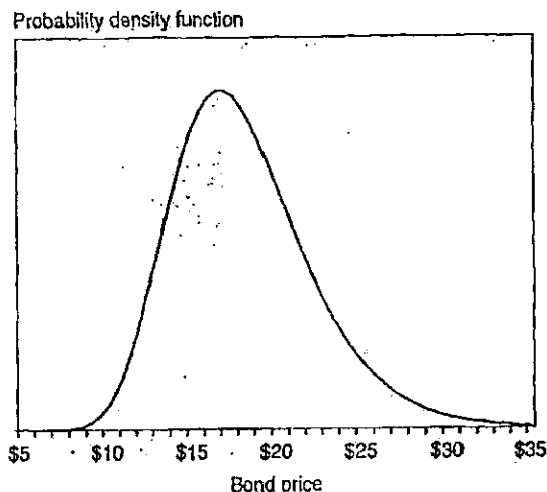
$$P[V \le \$15] = P[r \ge 6.528\%]$$

Probability density function

$5	$10	$15	$20	$25	$30	$35

Bond price

FIGURE 2.5   Density Function for the Bond Price

Assuming the yield change is normal with volatility 0.8%, this gives a probability of 25.5 percent.[1] Even though we do not know the density of the bond price, this method allows us to trace out its cumulative distribution by changing the cutoff price of $15. Taking the derivative, we can recover the density function of the bond price. Figure 2.5 shows that this p.d.f. is skewed to the right.

On the extreme right, if the yield falls to zero, the bond price will go to $100. On the extreme left, if the yield goes to infinity, the bond price will fall to, but not go below, zero. Relative to the current value of $17.41, there is a greater likelihood of large movements up than down.

This method, unfortunately, cannot be easily extended. For general density functions and transformations, risk managers turn to numerical methods, especially when the number of random variables is large. This is why credit risk models, for instance, all describe the distribution of credit losses through simulations.

EXAMPLE 2.3: FRM EXAM 2007—QUESTION 127

Suppose that A and B are random variables, each follows a standard normal distribution, and the covariance between A and B is 0.35. What is the variance of $(3A + 2B)$?

a. 14.47

b. 17.20

c. 9.20

d. 15.10

---

[1] We shall see later that this is obtained from the standard normal variable $z = (6.528 - 6.000)/0.80 = 0.660$. Using standard normal tables, or the NORMSDIST(−0.660) Excel function, this gives 25.5%.

EXAMPLE 2.4: FRM EXAM 2002—QUESTION 70

Given that x and y are random variables, and a, b, c and d are constant, which one of the following definitions is *wrong*.

a. $E(ax + by + c) = aE(x) + bE(y) + c$, if x and y are correlated.

b. $V(ax + by + c) = V(ax + by) + c$, if x and y are correlated.

c. $\text{Cov}(ax + by, cx + dy) = acV(x) + bdV(y) + (ad + bc)\text{Cov}(x, y)$, if x and y are correlated.

d. $V(x - y) = V(x + y) = V(x) + V(y)$, if x and y are uncorrelated.

2.4 IMPORTANT DISTRIBUTION FUNCTIONS

2.4.1 Uniform Distribution

The simplest continuous distribution function is the uniform distribution. This is defined over a range of values for x, $a \leq x \leq b$. The density function is

$$f(x) = \frac{1}{(b-a)}, \quad a \leq x \leq b \tag{2.39}$$

which is constant and indeed integrates to unity. This distribution puts the same weight on each observation within the allowable range, as shown in Figure 2.6. We denote this distribution as $U(a, b)$.

Realization of the uniform random variable
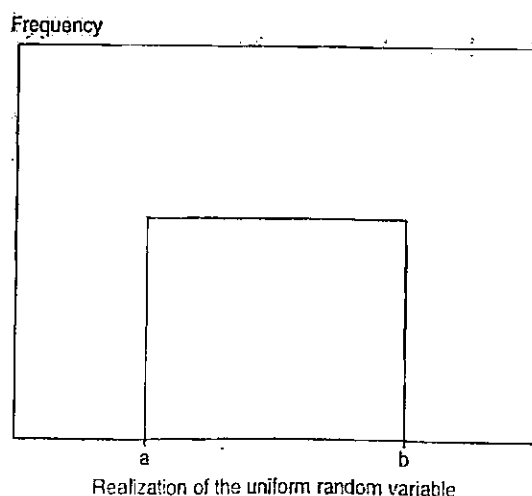
FIGURE 2.6 Uniform Density Function

Its mean and variance are given by

$$E(X) = \frac{a+b}{2}$$ (2.40)

$$V(X) = \frac{(b-a)^2}{12}$$ (2.41)

The uniform distribution $U(0, 1)$ is widely used as a starting distribution for generating random variables from any distribution $F(Y)$ in simulations. We need to have analytical formulas for the p.d.f. $f(Y)$ and its cumulative distribution $F(Y)$. As any cumulative distribution function ranges from zero to unity, we first draw X from $U(0, 1)$ and then compute $y = F^{-1}(x)$. The random variable Y will then have the desired distribution $f(Y)$.

EXAMPLE 2.5: FRM EXAM 2002—QUESTION 119

The random variable X with density function $f(x) = 1/(b-a)$ for $a < x < b$, and 0 otherwise, is said to have a uniform distribution over (a, b). Calculate its mean.

a. $(a + b)/2$
b. $a - b/2$
c. $a + b/4$
d. $a - b/4$

2.4.2 Normal Distribution

Perhaps the most important continuous distribution is the normal distribution, which represents adequately many random processes. This has a bell-like shape with more weight in the center and tails tapering off to zero. The daily rate of return in a stock price, for instance, has a distribution similar to the normal p.d.f.

The normal distribution can be characterized by its first two moments only, the mean μ and variance σ^2. The first parameter represents the location; the second, the dispersion. The normal density function has the following expression:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$ (2.42)

Its mean is $E[X] = \mu$ and variance $V[X] = \sigma^2$. We denote this distribution as $N(\mu, \sigma^2)$. Because the function can be fully specified by these two parameters, it is called a parametric function.

Instead of having to deal with different parameters, it is often more convenient to use a standard normal variable as ϵ, which has been standardized, or

TABLE 2.4 Lower Quantiles of the Standardized Normal Distribution

	Confidence Level (percent)								
	99.99	99.9	99	97.72	97.5	95	90	84.13	50
$c(-\alpha)$	-3.715	-3.090	-2.326	-2.000	-1.960	-1.645	-1.282	-1.000	-0.000

normalized, so that $E(\epsilon) = 0$, $V(\epsilon) = \sigma(\epsilon) = 1$. Figure 2.7 plots the standard normal density.

First, note that the function is symmetrical around the mean. Its mean of zero is the same as its mode (which is also the most likely, or highest, point on this curve) and median (which is such that the area to the left is a 50 percent probability). The skewness of a normal distribution is 0, which indicates that it is symmetrical around the mean. The kurtosis of a normal distribution is 3. Distributions with fatter tails have a greater kurtosis coefficient.

About 95% of the distribution is contained between values of $\epsilon_1 = -2$ and $\epsilon_2 = +2$, and 68% of the distribution falls between values of $\epsilon_1 = -1$ and $\epsilon_2 = +1$. Table 2.4 gives the values that correspond to right-tail probabilities, such that

$$\int_{-\alpha}^{\infty} f(\epsilon)d\epsilon = c \qquad (2.43)$$

For instance, the value of -1.645 is the quantile that corresponds to a 95% probability.[2]

This distribution plays a central role in finance because it represents adequately the behavior of many financial variables. It enters, for instance, the Black–Scholes
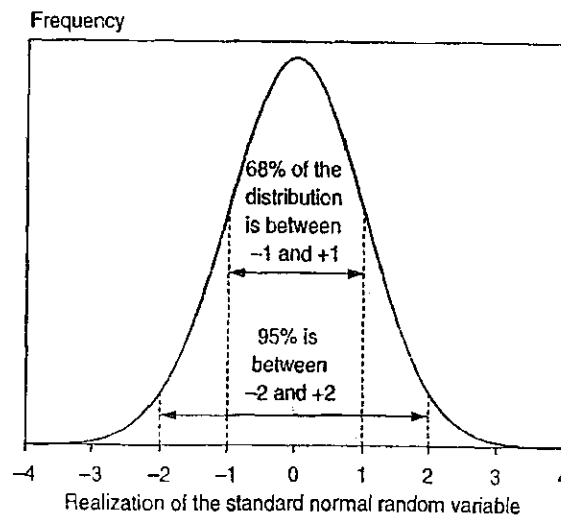
FIGURE 2.7 Normal Density Function

---

[2] More generally, the cumulative distribution can be found from the Excel function NORMDIST(·). For example, we can verify that NORMSDIST(−1.645) yields 0.04999, or a 5% left-tail probability.

option pricing formula where the function $N(\cdot)$ represents the cumulative standardized normal distribution function.

The distribution of any normal variable can then be recovered from that of the standard normal, by defining

$$X = \mu + \epsilon\sigma \tag{2.44}$$

Using Equations (2.24) and (2.25), we can show that X has indeed the desired moments, as $E(X) = \mu + E(\epsilon)\sigma = \mu$ and $V(X) = V(\epsilon)\sigma^2 = \sigma^2$.

Define, for instance, the random variable as the change in the dollar value of a portfolio. The expected value is $E(X) = \mu$. To find the quantile of X at the specified confidence level c, we replace ϵ by $-\alpha$ in Equation (2.44). This gives $Q(X, c) = \mu - \alpha\sigma$. Using Equation (2.9), we can compute VAR as

$$\text{VAR} = E(X) - Q(X, c) = \mu - (\mu - \alpha\sigma) = \alpha\sigma \tag{2.45}$$

For example, a portfolio with a standard deviation of $10 million would have a VAR, or potential downside loss, of $16.45 million at the 95% confidence level.

KEY CONCEPT

With normal distributions, the VAR of a portfolio is obtained from the product of the portfolio standard deviation and a standard normal deviate factor that reflects the confidence level, for instance 1.645 at the 95% level.

An important property of the normal distribution is that it is one of the few distributions that is *stable* under addition. In other words, a linear combination of jointly normally distributed random variables has a normal distribution.[3] This is extremely useful because we only need to know the mean and variance of the portfolio to reconstruct its whole distribution.

KEY CONCEPT

A linear combination of jointly normal variables has a normal distribution.

---

[3] Strictly speaking, this is only true under either of the following conditions: (1) the univariate variables are independently distributed, or (2) the variables are multivariate normally distributed (this invariance property also holds for jointly elliptically distributed variables).

When we have N random variables, the joint normal density can be written as a function of the vector x, of the means μ, and the covariance matrix Σ:

$$f(x_1, \cdots, x_N) = \frac{1}{(2\pi)^{N/2} \mid \Sigma \mid^{1/2}} \exp\left[-\frac{1}{2}(x - \mu)'\Sigma(x - \mu)\right] \qquad (2.46)$$

Using the concept of copulas, this can be separated into N different marginal normal densities and a joint normal copula. For two random variables, Equation (2.21) showed

$$f_{12}(x_1, x_2) = f_1(x_1) \times f_2(x_2) \times c_{12}[F_1(x_1), F_2(x_2); \theta]$$

Here, both f_1 and f_2 are normal marginals. They have parameters μ_1 and σ_1, and μ_2 and σ_2. In addition, c_{12} is the normal copula. Note that its sole parameter is the correlation coefficient ρ_{12}. This additional information is required to construct the covariance matrix Σ and defines the strength of the dependency between the two variables.

EXAMPLE 2.6: FRM EXAM 2005—QUESTION 62

Let Z be a standard normal random variable. An event X is defined to happen if either Z takes a value between -0.5 and $+0.5$ or Z takes any value greater than 1.5. What is the probability of event X happening if $N(0.5) = 0.6915$ and $N(-1.5) = 0.0668$, where $N(.)$ is the cumulative distribution function of a standard normal variable?

a. 0.2583
b. 0.3753
c. 0.4498
d. 0.7583

EXAMPLE 2.7: FRM EXAM 2003—QUESTION 21

Which of the following statements about the normal distribution is *not* accurate?

a. Kurtosis equals 3.
b. Skewness equals 1.
c. The entire distribution can be characterized by two moments, mean and variance.
d. The normal density function has the following expression: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(x - \mu)^2]$

EXAMPLE 2.8: FRM EXAM 2006—QUESTION 11

Which type of distribution produces the lowest probability for a variable to exceed a specified extreme value which is greater than the mean, assuming the distributions all have the same mean and variance?

a. A leptokurtic distribution with a kurtosis of 4

b. A leptokurtic distribution with a kurtosis of 8

c. A normal distribution

d. A platykurtic distribution

2.4.3 Lognormal Distribution

The normal distribution is a good approximation for many financial variables, such as the rate of return on a stock, $r = (P_1 - P_0)/P_0$, where P_0 and P_1 are the stock prices at time 0 and 1.

Strictly speaking, this is inconsistent with reality since a normal variable has infinite tails on both sides. In theory, r could end up below -1, which implies $P_1 < 0$. In reality, due to the limited liability of corporations, stock prices cannot turn negative. In many situations, however, this is an excellent approximation. For instance, with short horizons or small price moves, the probability of having a negative price is so small that it is negligible. If this is not the case, we need to resort to other distributions that prevent prices from going negative. One such distribution is the lognormal.

A random variable X is said to have a lognormal distribution if its logarithm $Y = \ln(X)$ is normally distributed. Define here $X = (P_1/P_0)$. Because the argument X in the logarithm function must be positive, the price P_1 can never go below zero.

The lognormal density function has the following expression

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(\ln(x) - \mu)^2\right], \quad x > 0 \tag{2.47}$$

Note that this is more complex than simply plugging $\ln(x)$ in Equation (2.42), because x also appears in the denominator. Its mean is

$$E[X] = \exp\left[\mu + \frac{1}{2}\sigma^2\right] \tag{2.48}$$

and variance $V[X] = \exp[2\mu + 2\sigma^2] - \exp[2\mu + \sigma^2]$. The parameters were chosen to correspond to those of the normal variable, $E[Y] = E[\ln(X)] = \mu$ and $V[Y] = V[\ln(X)] = \sigma^2$.

Conversely, if we set $E[X] = \exp[r]$, the mean of the associated normal variable is $E[Y] = E[\ln(X)] = (r - \sigma^2/2)$. We will see later that this adjustment is also
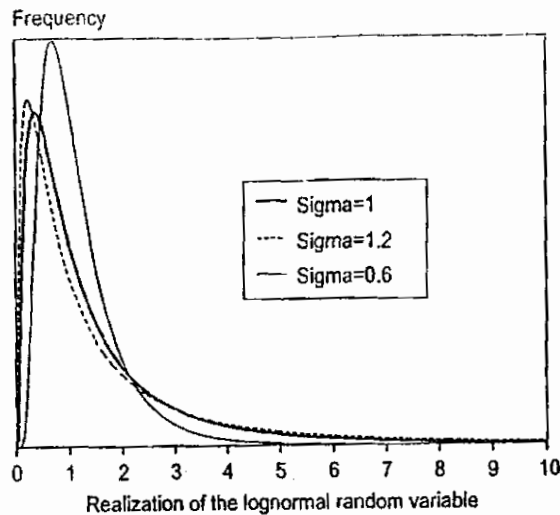
FIGURE 2.8 Lognormal Density Function

used in the Black–Scholes option valuation model, where the formula involves a trend in $(r - \sigma^2/2)$ for the log-price ratio.

Figure 2.8 depicts the lognormal density function with $\mu = 0$, and various values $\sigma = 1.0, 1.2, 0.6$. Note that the distribution is skewed to the right. The tail increases for greater values of σ. This explains why as the variance increases, the mean is pulled up in Equation (2.48).

We also note that the distribution of the bond price in our previous example, Equation (2.37), resembles a lognormal distribution. Using continuous compounding instead of annual compounding, the price function is

$$V = 100 \exp(-rT) \tag{2.49}$$

which implies $\ln(V/100) = -rT$. Thus if r is normally distributed, V has a lognormal distribution.

EXAMPLE 2.9: FRM EXAM 1999—QUESTION 5

Which of the following statements best characterizes the relationship between the normal and lognormal distributions?

a. The lognormal distribution is the logarithm of the normal distribution.
b. If the natural log of the random variable X is lognormally distributed, then X is normally distributed.
c. If X is lognormally distributed, then the natural log of X is normally distributed.
d. The two distributions have nothing to do with one another.

EXAMPLE 2.10: FRM EXAM 2007—QUESTION 21

The skew of a lognormal distribution is always

a. Positive
b. Negative
c. 0
d. 3

EXAMPLE 2.11: FRM EXAM 2002—QUESTION 125

Consider a stock with an initial price of $100. Its price one year from now is given by $S = 100 \times \exp(r)$, where the rate of return r is normally distributed with a mean of 0.1 and a standard deviation of 0.2. With 95% confidence, after rounding, S will be between

a. $67.57 and $147.99
b. $70.80 and $149.20
c. $74.68 and $163.56
d. $102.18 and $119.53

EXAMPLE 2.12: FRM EXAM 2000—QUESTION 128

For a lognormal variable X, we know that $\ln(X)$ has a normal distribution with a mean of zero and a standard deviation of 0.5. What are the expected value and and the variance of X?

a. 1.025 and 0.187
b. 1.126 and 0.217
c. 1.133 and 0.365
d. 1.203 and 0.399

2.4.4 Student's t Distribution

Another important distribution is the Student's t distribution. This arises in hypothesis testing, because it describes the distribution of the ratio of the estimated coefficient to its standard error.

This distribution is characterized by a parameter k known as the degrees of freedom. Its density is

$$f(x) = \frac{\Gamma[(k+1)/2]}{\Gamma(k/2)} \frac{1}{\sqrt{k\pi}} \frac{1}{(1+x^2/k)^{(k+1)/2}} \qquad (2.50)$$

where Γ is the gamma function, defined as $\Gamma(k) = \int_0^\infty x^{k-1}e^{-x}dx$. As k increases, this function converges to the normal p.d.f.

The distribution is symmetrical with mean zero and variance

$$V[X] = \frac{k}{k-2} \qquad (2.51)$$

provided $k > 2$. Its kurtosis is

$$\delta = 3 + \frac{6}{k-4} \qquad (2.52)$$

provided $k > 4$. Its has fatter tails than the normal which often provides a better representation of typical financial variables. Typical estimated values of k are around four to six for stock returns. Figure 2.9 displays the density for $k = 4$ and $k = 50$. The latter is close to the normal. With $k = 4$, however, the p.d.f. has fatter tails. As was done for the normal density, we can also use the Student's t to compute VAR as a function of the volatility

$$\text{VAR} = \alpha_k \sigma \qquad (2.53)$$

where the multiplier now depends on the degrees of freedom k.

As for the multivariate normal distribution, the joint Student distribution can be separated into two components. The marginals have the Student's distribution
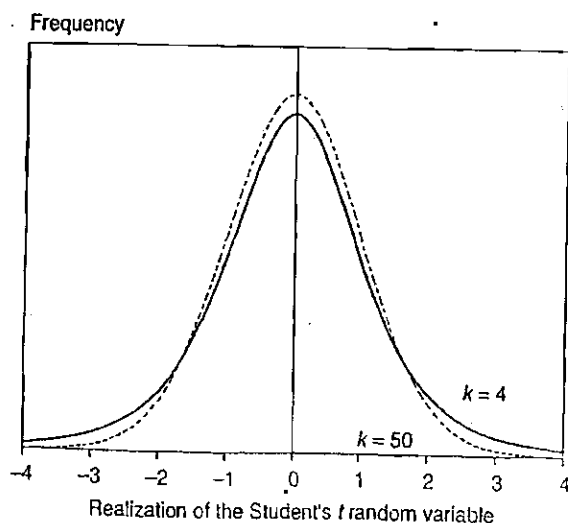
FIGURE 2.9   Student's t Density Function

described in Equation (2.50). In addition, the copula has a specific shape, which is the Student's copula. This copula allows for stronger dependencies in the tails than the normal copula. Marginals and copulas of different types can be used, as best fits the data. For example, one could use normal marginals and a Student's copula. This creates substantial flexibility in the statistical modeling of random variables.

Another distribution derived from the normal is the chi-square distribution, which can be viewed as the sum of independent squared standard normal variables

$$x = \sum_{j=1}^{k} z_j^2 \tag{2.54}$$

where k is also called the degrees of freedom. Its mean is $E[X] = k$ and variance $V[X] = 2k$. For k sufficiently large, $\chi^2(k)$ converges to a normal distribution $N(k, 2k)$. This distribution describes the sample variance.

Finally, another associated distribution is the F distribution, which can be viewed as the ratio of independent chi-square variables divided by their degrees of freedom

$$F(a, b) = \frac{\chi^2(a)/a}{\chi^2(b)/b} \tag{2.55}$$

This distribution appears in joint tests of regression coefficients.

EXAMPLE 2.13: FRM EXAM 2003—QUESTION 18

Which of the following statements is the most accurate about the relationship between a normal distribution and a Student's t-distribution that have the same mean and standard deviation?

a. They have the same skewness and the same kurtosis.

b. The Student's t-distribution has larger skewness and larger kurtosis.

c. The kurtosis of a Student's t-distribution converges to that of the normal distribution as the number of degrees of freedom increases.

d. The normal distribution is a good approximation for the Student's t-distribution when the number of degrees of freedom is small.

2.4.5 Binomial Distribution

Consider now a random variable that can take discrete values between zero and n. This could be, for instance, the number of times VAR is exceeded over the last year, also called the number of exceptions. Thus, the binomial distribution plays an important role for the backtesting of VAR models.

A binomial variable can be viewed as the result of n independent Bernoulli trials, where each trial results in an outcome of $y = 0$ or $y = 1$. This applies, for example, to credit risk. In case of default, we have $y = 1$, otherwise $y = 0$. Each Bernoulli variable has expected value of $E[Y] = p$ and variance $V[Y] = p(1 - p)$.

A random variable is defined to have a binomial distribution if the discrete density function is given by

$$f(x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, \ldots, n \qquad (2.56)$$

where $\binom{n}{x}$ is the number of combinations of n things taken x at a time, or

$$\binom{n}{x} = \frac{n!}{x!(n - x)!} \qquad (2.57)$$

and the parameter p is between zero and one. This distribution also represents the total number of successes in n repeated experiments where each success has a probability of p.

The binomial variable has mean and variance

$$E[X] = pn \qquad (2.58)$$

$$V[X] = p(1 - p)n \qquad (2.59)$$

It is described in Figure 2.10 in the case where $p = 0.25$ and $n = 10$. The probability of observing $X = 0, 1, 2 \ldots$ is 5.6%, 18.8%, 28.1% and so on.

For instance, we want to know what is the probability of observing $x = 0$ exceptions out of a sample of $n = 250$ observations when the true probability is 1%. We should expect to observe 2.5 exceptions on average across many such samples. There will be, however, some samples with no exceptions at all. This
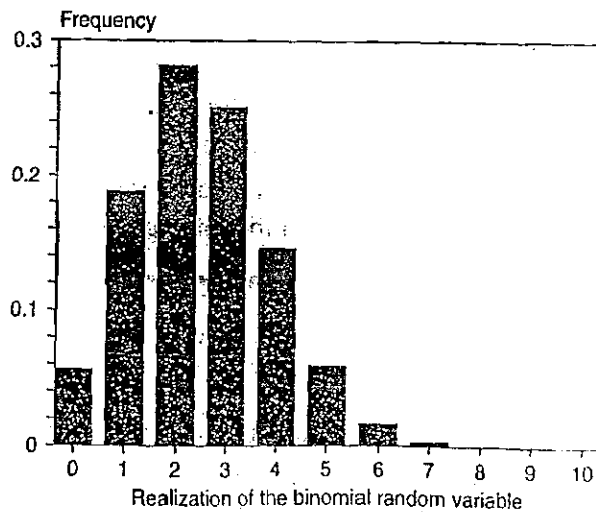
FIGURE 2.10 Binomial Density Function with $p = 0.25, n = 10$

probability is

$$f(X = 0) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} = \frac{250!}{1 \times 250!}0.01^0 0.99^{250} = 0.081$$

So, we would expect to observe 8.1% of samples with zero exceptions, under the null hypothesis. We can repeat this calculation with different values for x. For example, the probability of observing 8 exceptions is $f(X = 8) = 0.02\%$ only. We can use this information to test the null hypothesis. Because this probability is so low, observing 8 exceptions would make us question whether the true probability is 1%.

EXAMPLE 2.14: FRM EXAM 2006—QUESTION 84

On a multiple-choice exam with four choices for each of six questions, what is the probability that a student gets fewer than two questions correct simply by guessing?

a. 0.46%
b. 23.73%
c. 35.60%
d. 53.39%

2.4.6 Poisson Distribution

The Poisson distribution is a discrete distribution, which typically is used to describe the number of events occurring over a fixed period of time, assuming events are independent of each other. It is defined as

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, \ldots \tag{2.60}$$

where λ is a positive number representing the average arrival rate during the period. This distribution, for example, is widely used to represent the frequency, or number of occurrences, of operational losses over a year.

The parameter λ represents the expected value of X and also its variance

$$E[X] = \lambda \tag{2.61}$$

$$V[X] = \lambda \tag{2.62}$$

The Poisson distribution is the limiting case of the binomial distribution as n goes to infinity and p goes to zero, while $np = \lambda$ remains fixed. In addition, when λ

is large the Poisson distribution is well approximated by the normal distribution with mean and variance of λ, through the central limit theorem.

If the number of arrivals follows a Poisson distribution, then the time period between arrivals follows an exponential distribution with mean $1/\lambda$. The latter has density taking the form $f(x) = \lambda e^{-\lambda x}$, for $x \geq 0$. For example, if we expect $\lambda = 12$ losses per year, the average time interval between losses should be 1 year divided by 12, or one month.

EXAMPLE 2.15: FRM EXAM 2004—QUESTION 60

When can you use the normal distribution to approximate the Poisson distribution, assuming you have n independent trials each with a probability of success of p

a. When the mean of the Poisson distribution is very small

b. When the variance of the Poisson distribution is very small

c. When the number of observations is very large and the success rate is close to 1

d. When the number of observations is very large and the success rate is close to 0

2.5 LIMIT DISTRIBUTIONS

2.5.1 Distribution of Averages

The normal distribution is extremely important because of the central limit theorem (CLT), which states that the mean of n independent and identically distributed variables converges to a normal distribution as the number of observations n increases. This very powerful result is valid for any underlying distribution, as long as the realizations are independent. For instance, the distribution of total credit losses converges to a normal distribution as the number of loans increases to a large value, assuming defaults are always independent of each other.

Define $\bar{X}$ as the mean $\frac{1}{n}\sum_{i=1}^{n} X_i$, where each variable has mean μ and standard deviation σ. We have

$$\bar{X} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right) \tag{2.63}$$

Standardizing the variable, we can write

$$\frac{\bar{X} - \mu}{(\sigma/\sqrt{n})} \rightarrow N(0, 1) \tag{2.64}$$

Thus, the normal distribution is the limiting distribution of the average, which explain why it has such a prominent place in statistics.

As an example, consider the binomial variable, which is the sum of independent Bernoulli trials. When n is large, we can use the CLT and approximate the binomial distribution by the normal distribution. Using Equation (2.64) for the sum, we have

$$z = \frac{x - pn}{\sqrt{p(1-p)n}} \rightarrow N(0, 1) \tag{2.65}$$

which is much easier to evaluate than the binomial distribution.

Consider for example the issue of whether the number of exceptions x we observe is compatible with a 99% VAR. For our example, the mean and variance of x are $E[X] = 0.01 \times 250 = 2.5$ and $V[X] = 0.01(1 - 0.01) \times 250 = 2.475$. We observe $x = 8$, which gives $z = (8 - 2.5)/\sqrt{2.475} = 3.50$. We can now compare this number to the standard normal distribution. Say for instance that we decide to reject the hypothesis that VAR is correct if the statistic falls outside a 95% two-tailed confidence band.[4] This interval is $(-1.96, +1.96)$ for the standardized normal distribution. Here, the value of 3.50 is much higher than the cutoff point of $+1.96$. As a result, we would reject the null hypothesis that the true probability of observing an exception is 1% only. In other words, there are simply too many exceptions to be explained by bad luck. It is more likely that the VAR model underestimates risk.

2.5.2 Distribution of Tails

The CLT deals with the mean, or center of the distribution. For risk management purposes, it is also useful to examine the tails of the distribution.

Another powerful theorem is given by extreme value theory (EVT). The EVT theorem says that the limit distribution for values x beyond a cutoff point u belongs to the following family

$$\begin{aligned} F(y) &= 1 - (1 + \xi y)^{-1/\xi}, \quad \xi \neq 0 \\ F(y) &= 1 - exp(-y), \quad\quad \xi = 0 \end{aligned} \tag{2.66}$$

where $y = (x - u)/\beta$. To simplify, we define the loss x as a positive number so that y is also positive. The distribution is characterized by $\beta > 0$, a *scale* parameter, and by ξ, a *shape* parameter that determines the speed at which the tail disappears.

This distribution is called the Generalized Pareto Distribution, because it subsumes other distributions as special cases. For instance, the normal distribution corresponds to $\xi = 0$, in which case the tails disappear at an exponential speed. Typical financial data have $\xi > 0$, which implies *fat tails*. This class of distribution

---

[4] Note that the choice of this confidence level has nothing to do with the VAR confidence level. Here, the 95% level represents the rate at which the decision rule will commit the error of falsely rejecting a correct model.
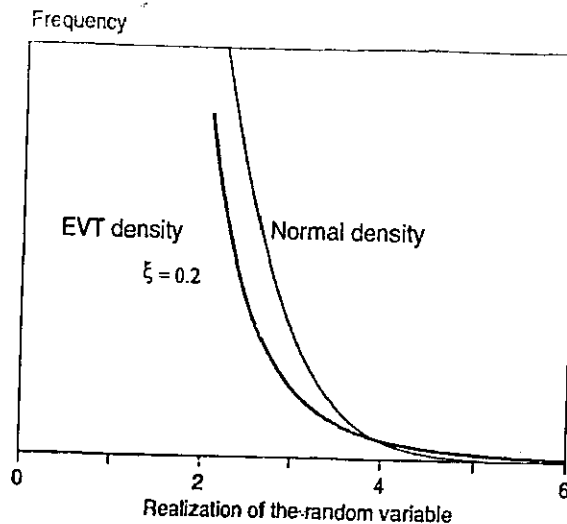
FIGURE 2.11 EVT and Normal Densities

includes the Gumbel, Fréchet, and Weibull families, as $\xi \to 0$, $\xi > 0$, and $\xi < 0$, respectively.

Figure 2.11 illustrates the shape of the density function for U.S. stock market data. The normal density falls off fairly quickly. With $\xi = 0.2$, the EVT density has a fatter tail than the normal density, implying a higher probability of experiencing large losses. This is an important observation for risk management purposes. Note that the EVT density is only defined for the tail, i.e., when the loss x exceeds an arbitrary cutoff point, which is taken as 2 in this case.

EXAMPLE 2.16: FRM EXAM 2007—QUESTION 110

Which of the following statements regarding extreme value theory (EVT) is *incorrect*?

a. In contrast to conventional approaches for estimating VAR, EVT only considers the tail behavior of the distribution.

b. Conventional approaches for estimating VAR that assume that the distribution of returns follows a unique distribution for the entire range of values may fail to properly account for the fat tails of the distribution of returns.

c. EVT attempts to find the optimal point beyond which all values belong to the tail and then models the distribution of the tail separately.

d. By smoothing the tail of the distribution, EVT effectively ignores extreme events and losses that can generally be labeled outliers.

2.6 IMPORTANT FORMULAS

Probability density function: $f(x) = \text{Prob}(X = x)$

(Cumulative) distribution function: $F(x) = \int_{-\infty}^{x} f(u)du$

Mean: $E(X) = \mu = \int x f(x)dx$

Variance: $V(X) = \sigma^2 = \int [x - \mu]^2 f(x)dx$

Skewness: $\gamma = \left(\int [x - \mu]^3 f(x)dx \right) / \sigma^3$

Kurtosis: $\delta = \left(\int [x - \mu]^4 f(x)dx \right) / \sigma^4$

Quantile, VAR: $\text{VAR} = E(X) - Q(X, c) = \alpha\sigma$

Independent joint densities: $f_{12}(x_1 x_2) = f_1(x_1) \times f_2(x_2)$

Marginal densities: $f_1(x_1) = \int f_{12}(x_1, u_2)du_2$,

Conditional densities: $f_{1\cdot2}(x_1 \mid x_2) = \frac{f_{12}(x_1, x_2)}{f_2(x_2)}$

Copula, Sklar's theorem: $f_{12}(x_1, x_2) = f_1(x_1) \times f_2(x_2) \times c_{12}[F_1(x_1), F_2(x_2); \theta]$

Covariance: $\sigma_{12} = \int_1 \int_2 [x_1 - \mu_1][x_2 - \mu_2] f_{12} dx_1 dx_2$

Correlation: $\rho_{12} = \sigma_{12}/(\sigma_1 \sigma_2)$

Linear transformation of random variables: $E(a + bX) = a + bE(X)$,
$\qquad V(a + bX) = b^2 V(X)$, $\sigma(a + bX) = b\sigma(X)$

Sum of random variables: $E(X_1 + X_2) = \mu_1 + \mu_2$,
$\qquad V(X_1 + X_2) = \sigma_1^2 + \sigma_2^2 + 2\sigma_{12}$

Portfolios of random variables: $Y = w'X$, $E(Y) = \mu_p = w'\mu$, $\sigma_p^2 = w'\Sigma w$

Product of random variables: $E(X_1 X_2) = \mu_1\mu_2 + \sigma_{12}$,
$\qquad V(X_1 X_2) = \mu_1^2\sigma_2^2 + \sigma_1^2\mu_2^2 + \sigma_1^2\sigma_2^2$

Uniform distribution: $E(X) = \frac{a+b}{2}$, $V(X) = \frac{(b-a)^2}{12}$

Normal distribution: $E(X) = \mu$, $V(X) = \sigma^2$, $\gamma = 0$, $\delta = 3$

Lognormal distribution: for X if $Y = \ln(X)$ is normal, $E[X] = \exp\left[\mu + \frac{1}{2}\sigma^2\right]$,
$\qquad$ - $V[X] = \exp[2\mu + 2\sigma^2] - \exp[2\mu + \sigma^2]$

Student's t distribution: $V[X] = \frac{k}{k-2}$, $\gamma = 0$, $\delta = 3 + \frac{6}{k-4}$

Binomial distribution: $E[X] = pn$, $V[X] = p(1 - p)n$

Poisson distribution: $E[X] = \lambda$, $V[X] = \lambda$

Distribution of averages (CLT): $\bar{X} \to N\left(\mu, \frac{\sigma^2}{n}\right)$

Distribution of tails (EVT): $y = (x - u)/\beta \to$ Generalized Pareto distribution

2.7 ANSWERS TO CHAPTER EXAMPLES

Example 2.1: FRM Exam 2000—Question 81

b. Correlation is a measure of linear association. Independence implies zero correlation, but the reverse is not always true.

Example 2.2: FRM Exam 2007—Question 83

b. The function $x \times y$ is described in the following table. The sum of the entries is 36. The scaling factor k must be such that the total probability is one. Therefore, we have $k = 1/36$. The table shows one instance where $x + y > 5$, which is $x = 3$, $y = 3$. The probability is $p = 9/36 = 1/4$.

$x \times y$	$x = 1$	2	3
$y = 1$	1	2	3
2	2	4	6
3	3	6	9

Example 2.3: FRM Exam 2007—Question 127

b. The variance is $V(3A + 2B) = 3^2 V(A) + 2^2 V(B) + 2 \times 3 \cdot 2 \, \text{Cov}(A, B) = 9 \cdot 1 + 4 \cdot 1 + 12 \times 0.35 = 17.2$.

Example 2.4: FRM Exam 2002—Question 70

b. Statement a. is correct, as it is a linear operation. Statement c. is correct, as in Equation (2.32). Statement d. is correct, as the covariance term is zero if the variables are uncorrelated. Statement b. is false, as adding a constant c to a variable cannot change the variance. The constant drops out because it is also in the expectation.

Example 2.5: FRM Exam 2002—Question 119

a. The mean is the center of the distribution, which is the average of a and b.

Example 2.6: FRM Exam 2005—Question 62

c. The event is the sum of the probabilities $P(-0.5 < Z < +0.5)$ and $P(Z > +1.5)$. Given the symmetry of the normal distribution, or that $N(d) = 1 - N(-d)$, this gives $P(-0.5 < Z < +0.5) = 2P(0 < Z < +0.5) = 2(P(Z < +0.5) - 0.5) = 2(N(0.5) - 0.5) = 2(0.6915 - 0.5) = 0.3830$ and $P(Z > +1.5) = N(-1.5) = 0.0668$. The sum is 0.4498.

Example 2.7: FRM Exam 2003—Question 21

b. Skewness is 0, kurtosis 3, the entire distribution is described by μ and σ, and the p.d.f. is correct.

Example 2.8: FRM Exam 2006—Question 11

d. A platykurtic distribution has kurtosis less than 3, less than the normal p.d.f. because all other answers have higher kurtosis, this produces the lowest extreme values.

Example 2.9: FRM Exam 1999—Question 5

c. X is said to be lognormally distributed if its logarithm $Y = \ln(X)$ is normally distributed.

Example 2.10: FRM Exam 2007—Question 21

a. A lognormal distribution is skewed to the right. Intuitively, if this represents the distribution of prices, prices can fall at most by 100% but can increase by more than that.

Example 2.11: FRM Exam 2002—Question 125

c. Note that this is a two-tailed confidence band, so that $\alpha = 1.96$. We find the extreme values from $\$100\exp(\mu \pm \alpha\sigma)$. The lower limit is then $V_1 = \$100\exp(0.10 - 1.96 \times 0.2) = \$100\exp(-0.292) = \$74.68$. The upper limit is $V_2 = \$100\exp(0.10 + 1.96 \times 0.2) = \$100\exp(0.492) = \$163.56$.

Example 2.12: FRM Exam 2000—Question 128

c. Using Equation (2.48), we have $E[X] = \exp[\mu + 0.5\sigma^2] = \exp[0 + 0.5 * 0.5^2] = 1.1331$. Assuming there is no error in the answers listed for the variance, it is sufficient to find the correct answer for the expected value.

Example 2.13: FRM Exam 2003—Question 18

c. The two distributions have the same skewness of zero but the Student's t has higher kurtosis. As the number of degrees of freedom increases, the Student converges to the normal, so c. is the correct answer.

Example 2.14: FRM Exam 2006—Question 84

d. We use the density given by Equation (2.56). The number of trials is $n = 6$. The probability of guessing correctly just by chance is $p = 1/4 = 0.25$. The probability of zero lucky guesses is $\binom{6}{0}0.25^0 0.75^6 = 0.75^6 = 0.17798$. The probability of one lucky guess is $\binom{6}{1}0.25^1 0.75^5 = 6 \ 0.25 \ 0.75^5 = 0.35596$. The sum is 0.5339.

    Note that the same analysis can be applied to the distribution of scores on the FRM examination with 140 questions. It would be virtually impossible to have a score of zero, assuming random guesses; this probability is $0.75^{140} = 3.2E - 18$ Also, the expected percentage score under random guesses is $p = 25\%$.

Example 2.15: FRM Exam 2004—Question 60

c. The normal approximation to the Poisson improves when the success rate, λ is very high. Because this is also the mean and variance, answers a. and b. are

wrong. In turn, the binomial density is well approximated by the Poisson density when $np = \lambda$ is large.

Example 2.16: FRM Exam 2007—Question 110

d. EVT only uses information in the tail, so statement a. is correct. Conventional approaches such as delta-normal VAR assume a fixed p.d.f. for the entire distribution, which may understate the extent of fat tails. So, statement b. is correct. The first step in EVT is to choose a cutoff point for the tail, then to estimate the parameters of the tail distribution, so statement c. is correct. Finally, EVT does not ignore extreme events (as long as they are in the sample).

APPENDIX: REVIEW OF MATRIX MULTIPLICATION

This appendix briefly reviews the mathematics of matrix multiplication. Say that we have two matrices, A and B that we wish to multiply to obtain the new matrix $C = AB$. The respective dimensions are $(n \times m)$ for A, that is, n rows and m columns, and $(m \times p)$ for B. The number of columns for A must exactly match (or conform) to the number of rows for B. If so, this will result in a matrix C of dimensions $(n \times p)$.

We can write the matrix A in terms of its individual components a_{ij}, where i denotes the row and j denotes the column:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

As an illustration, take a simple example where the matrices are of dimension (2×3) and (3×2).

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

$$C = AB = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

To multiply the matrices, each row of A is multiplied element-by-element by each column of B. For instance, c_{12} is obtained by taking

$$c_{12} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32}$$

The matrix C is then:

$$C = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}$$

Matrix multiplication can be easily implemented in Excel using the function MMULT. First, we highlight the cells representing the output matrix C, say f1:g2. Then we enter the function, for instance MMULT(a1:c2; d1:e3), where the first range represents the first matrix A, here 2 by 3, and the second range represents the matrix B, here 3 by 2. The final step is to hit the three keys Control-Shift-Return simultaneously.

Fundamentals of Statistics

The preceding chapter was mainly concerned with the theory of probability, including distribution theory. In practice, researchers have to find methods to choose among distributions and to estimate distribution parameters from real data. The subject of sampling brings us now to the theory of statistics. Whereas probability assumes the distributions are known, statistics attempts to make inferences from actual data.

Here, we sample from the distribution of a population, say the change in the exchange rate, to make inferences about the population. The questions are, what is the best distribution for this random variable and what are the best parameters for this distribution? Risk measurement, however, typically deals with large numbers of random variables. So, we also want to characterize the relationships between the risk factors to which the portfolio is exposed. For example, do we observe that movements in the yen/dollar rate are correlated with the dollar/euro rate? Another type of problem is to develop decision rules to test some hypotheses, for instance whether the volatility remains stable over time.

These examples illustrate two important problems in statistical inference, i.e., estimation and tests of hypotheses. With estimation, we wish to estimate the value of an unknown parameter from sample data. With tests of hypotheses, we wish to verify a conjecture about the data.

This chapter reviews the fundamental tools of statistics theory for risk managers. Section 3.1 discusses the sampling of real data and the construction of returns. The problem of parameter estimation is presented in Section 3.2. Section 3.3 then turns to regression analysis, summarizing important results as well as common pitfalls in their interpretation.

3.1 REAL DATA

To start with an example, let us say that we observe movements in the daily yen/dollar exchange rate and wish to characterize the distribution of tomorrow's exchange rate.

The risk manager's job is to assess the range of potential gains and losses on a trader's position. He or she observes a sequence of past spot prices $S_0, S_1, \ldots, S_t$, from which we have to infer the distribution of tomorrow's price, S_{t+1}.

3.1.1 Measuring Returns

The truly random component in tomorrow's price is not its level, but rather its change relative to today's price. We measure the *relative rate of change* in the spot price:

$$r_t = (S_t - S_{t-1})/S_{t-1} \tag{3.1}$$

Alternatively, we could construct the logarithm of the price ratio:

$$R_t = \ln[S_t/S_{t-1}] \tag{3.2}$$

which is equivalent to using continuous instead of discrete compounding. This is also

$$R_t = \ln[1 + (S_t - S_{t-1})/S_{t-1}] = \ln[1 + r_t]$$

Because $\ln(1 + x)$ is close to x if x is small, R_t should be close to r_t provided the return is small. For daily data, there is typically little difference between R_t and r_t.

The return defined so far is the capital appreciation return, which ignores the income payment on the asset. Define the dividend or coupon as D_t. In the case of an exchange rate position, this is the interest payment in the foreign currency over the holding period. The total return on the asset is

$$r_t^{TOT} = (S_t + D_t - S_{t-1})/S_{t-1} \tag{3.3}$$

When the horizon is very short, the income return is typically very small compared to the capital appreciation return.

The next question is whether the sequence of variables r_t can be viewed as independent observations. If so, one could hypothesize, for instance, that the random variables are drawn from a normal distribution $N(\mu, \sigma^2)$. We could then proceed to estimate μ and σ^2 from the data and use this information to create a distribution for tomorrow's spot price change.

Independent observations have the very nice property that their joint distribution is the product of their marginal distribution, which considerably simplifies the analysis. The obvious question is whether this assumption is a workable approximation. In fact, there are good economic reasons to believe that rates of change on financial prices are close to independent.

The hypothesis of efficient markets postulates that current prices convey all relevant information about the asset. If so, any change in the asset price must be due to news, or events which are by definition impossible to forecast (otherwise, it would not be news). This implies that changes in prices are unpredictable and, hence, satisfy our definition of independent random variables.

This hypothesis, also known as the random walk theory, implies that the conditional distribution of returns depends only on current prices, and not on the previous history of prices. If so, technical analysis must be a fruitless exercise. Technical analysts try to forecast price movements from past price patterns.

If in addition the distribution of returns is constant over time, the variables are said to be independently and identically distributed (i.i.d.). So, we could consider that the observations r_t are independent draws from the same distribution $N(\mu, \sigma^2)$.

Later, we will consider deviations from this basic model. Distributions of financial returns typically display fat tails. Also, variances are not constant and display some persistence; expected returns can also slightly vary over time.

3.1.2 Time Aggregation

It is often necessary to translate parameters over a given horizon to another horizon. For example, we may have raw data for daily returns, from which we compute a daily volatility that we want to extend to a monthly volatility.

Returns can be easily related across time when we use the log of the price ratio, because the log of a product is the sum of the logs of the individual terms. The two-day return, for example, can be decomposed as

$$R_{02} = \ln[S_2/S_0] = \ln[(S_2/S_1) \times (S_1/S_0)] = \ln[S_1/S_0] + \ln[S_2/S_1] = R_{01} + R_{12} \tag{3.4}$$

This decomposition is only approximate if we use discrete returns, however.

The expected return and variance are then $E(R_{02}) = E(R_{01}) + E(R_{12})$ and $V(R_{02}) = V(R_{01}) + V(R_{12}) + 2\text{Cov}(R_{01}, R_{12})$. Assuming returns are uncorrelated and have identical distributions across days, we have $E(R_{02}) = 2E(R_{01})$ and $V(R_{02}) = 2V(R_{01})$.

Generalizing over T days, we can relate the moments of the T-day returns R_T to those of the 1-day returns R_1:

$$E(R_T) = E(R_1)T \tag{3.5}$$

$$V(R_T) = V(R_1)T \tag{3.6}$$

Expressed in terms of volatility, this yields the square root of time rule:

$$SD(R_T) = SD(R_1)\sqrt{T} \tag{3.7}$$

KEY CONCEPT

When successive returns are uncorrelated, the volatility increases as the horizon extends following the square root of time.

More generally, the variance can be added up from different values across different periods. For instance, the variance over the next year can be computed as the average monthly variance over the first three months, multiplied by 3, plus the average variance over the last nine months, multiplied by 9. This type of analysis

is routinely used to construct a term structure of implied volatilities, which are derived from option data for different maturities.

It should be emphasized that this holds only if returns have constant parameters across time and are uncorrelated. When there is non-zero correlation across days, the two-day variance is

$$V(R_2) = V(R_1) + V(R_1) + 2\rho V(R_1) = 2V(R_1)(1 + \rho) \qquad (3.8)$$

Because we are considering correlations in the time series of the same variable, ρ is called the autocorrelation coefficient, or the serial autocorrelation coefficient. A positive value for ρ implies that a movement in one direction in one day is likely to be followed by another movement in the same direction the next day. A positive autocorrelation signals the existence of a trend. In this case, Equation (3.8) shows that the two-day variance is greater than the one obtained by the square root of time rule.

A negative value for ρ implies that a movement in one direction in one day is likely to be followed by a movement in the other direction the next day. So, prices tend to revert back to a mean value. A negative autocorrelation signals

EXAMPLE 3.1: FRM EXAM 1999—QUESTION 4

A fundamental assumption of the random walk hypothesis of market returns is that returns from one time period to the next are statistically independent. This assumption implies

 a. Returns from one time period to the next can never be equal.
 b. Returns from one time period to the next are uncorrelated.
 c. Knowledge of the returns from one time period does not help in predicting returns from the next time period.
 d. Both b) and c) are true.

EXAMPLE 3.2: FRM EXAM 2002—QUESTION 3

Consider a stock with daily returns that follow a random walk. The annualized volatility is 34%. Estimate the weekly volatility of this stock assuming that the year has 52 weeks.

 a. 6.80%
 b. 5.83%
 c. 4.85%
 d. 4.71%

EXAMPLE 3.3: FRM EXAM 2002—QUESTION 2

Assume we calculate a one-week VAR for a natural gas position by rescaling the daily VAR using the square-root rule. Let us now assume that we determine the *true* gas price process to be mean-reverting and recalculate the VAR.

Which of the following statements is true?

a. The recalculated VAR will be less than the original VAR.
b. The recalculated VAR will be equal to the original VAR.
c. The recalculated VAR will be greater than the original VAR.
d. There is no necessary relation between the recalculated VAR and the original VAR.

---

mean reversion. In this case, the two-day variance is less than the one obtained by the square root of time rule.

3.1.3 Portfolio Aggregation

Let us now turn to aggregation of returns across assets. Consider, for example, an equity portfolio consisting of investments in N shares. Define the number of each share held as q_i with unit price S_i. The portfolio value at time t is then

$$W_t = \sum_{i=1}^{N} q_i S_{i,t} \tag{3.9}$$

We can write the weight assigned to asset i as

$$w_{i,t} = \frac{q_i S_{i,t}}{W_t} \tag{3.10}$$

which by construction sum to unity. Using weights, however, rules out situations with zero net investment, $W_t = 0$, such as some derivatives positions. But we could have positive and negative weights if short selling is allowed, or weights greater than one if the portfolio can be leveraged.

The next period, the portfolio value is

$$W_{t+1} = \sum_{i=1}^{N} q_i S_{i,t+1} \tag{3.11}$$

assuming that the unit price incorporates any income payment. The gross, or dollar, return is then

$$W_{t+1} - W_t = \sum_{i=1}^{N} q_i (S_{i,t+1} - S_{i,t})$$

(3.12)

and the *rate* of return is

$$\frac{W_{t+1} - W_t}{W_t} = \sum_{i=1}^{N} \frac{q_i S_{i,t}}{W_t} \frac{(S_{i,t+1} - S_{i,t})}{S_{i,t}} = \sum_{i=1}^{N} w_{i,t} \frac{(S_{i,t+1} - S_{i,t})}{S_{i,t}}$$

(3.13)

So, the portfolio rate of return is a linear combination of the asset returns

$$r_{p,t+1} = \sum_{i=1}^{N} w_{i,t}\, r_{i,t+1}$$

(3.14)

The dollar return is then

$$W_{t+1} - W_t = \left[\sum_{i=1}^{N} w_{i,t} r_{i,t+1} \right] W_t$$

(3.15)

and has a normal distribution if the individual returns are also normally distributed.

Alternatively, we could express the individual positions in dollar terms,

$$x_{i,t} = w_{i,t} W_t = q_i S_{i,t}$$

(3.16)

The dollar return is also, using dollar amounts,

$$W_{t+1} - W_t = \left[\sum_{i=1}^{N} x_{i,t} r_{i,t+1} \right]$$

(3.17)

As we have seen in the previous chapter, the variance of the portfolio dollar return is

$$V[W_{t+1} - W_t] = x' \Sigma x$$

(3.18)

Because the portfolio follows a normal distribution, it is fully characterized by it expected return and variance. The portfolio VAR is then

$$VAR = \alpha \sqrt{x' \Sigma x}$$

(3.19)

where α depends on the selected density function.

EXAMPLE 3.4: FRM EXAM 2004—QUESTION 39

Consider a portfolio with 40% invested in asset X and 60% invested in asset Y. The mean and variance of return on X are 0 and 25, respectively. The mean and variance of return on Y are 1 and 121, respectively. The correlation coefficient between X and Y is 0.3. What is the nearest value for portfolio volatility?

a. 9.51
b. 8.60
c. 13.38
d. 7.45

3.2 PARAMETER ESTIMATION

Armed with our i.i.d. sample of T observations, we can start estimating the parameters of interest, such as the sample mean, the variance, and other moments.

3.2.1 Distribution of Estimates

As in the previous chapter, define x_i as the realization of a random sample. The expected return, or mean, $\mu = E(X)$ can be estimated by the sample mean,

$$m = \widehat{\mu} = \frac{1}{T}\sum_{i=1}^{T} x_i \tag{3.20}$$

Intuitively, we assign the same weight of $1/T$ to all observations because they all have the same probability. The variance, $\sigma^2 = E[(X - \mu)^2]$, can be estimated by the sample variance,

$$s^2 = \widehat{\sigma}^2 = \frac{1}{(T-1)}\sum_{i=1}^{T}(x_i - \widehat{\mu})^2 \tag{3.21}$$

Note that we divide by $T - 1$ instead of T. This is because we estimate the variance around an unknown parameter, the mean. So, we have fewer degrees of freedom than otherwise. As a result, we need to adjust s^2 to ensure that its expectation equals the true value. In most situations, however, T is large so that this adjustment is minor.

It is essential to note that these estimated values depend on the particular sample and, hence, have some inherent variability. The sample mean itself is

distributed as

$$m = \widehat{\mu} \sim N(\mu, \sigma^2/T) \tag{3.22}$$

If the population distribution is normal, this exactly describes the distribution of the sample mean. Otherwise, the central limit theorem states that this distribution is only valid asymptotically, i.e., for large samples.

For the distribution of the sample variance $\widehat{\sigma}^2$, one can show that, when X is normal, the following ratio is distributed as a chi-square with $(T - 1)$ degrees of freedom

$$\frac{(T - 1)\widehat{\sigma}^2}{\sigma^2} \sim \chi^2(T - 1) \tag{3.23}$$

If the sample size T is large enough, the chi-square distribution converges to a normal distribution:

$$\widehat{\sigma}^2 \sim N\left(\sigma^2, \sigma^4 \frac{2}{(T - 1)}\right) \tag{3.24}$$

Using the same approximation, the sample standard deviation has a normal distribution with a standard error of

$$se(\widehat{\sigma}) = \sigma\sqrt{\frac{1}{2T}} \tag{3.25}$$

We can use this information for hypothesis testing. For instance, we would like to detect a constant trend in X. Here, the null hypothesis is that $\mu = 0$. To answer the question, we use the distributional assumption in Equation (3.22) and compute a standard normal variable as the ratio of the estimated mean to its standard error, or

$$z = \frac{(m - 0)}{\sigma/\sqrt{T}} \tag{3.26}$$

Because this is now a standard normal variable, we would not expect to observe values far away from zero. Typically, we would set the significance level at 95 percent, which translates into a two-tailed interval for z of $[-1.96, +1.96]$. Roughly, this means that, if the absolute value of z is greater than two, we would reject the hypothesis that m came from a distribution with a mean of zero. We can have some confidence that the true μ is indeed different from zero.

In fact, we do not know the true σ and use the estimated s instead. The distribution is a Student's t with T degrees of freedom:

$$t = \frac{(m - 0)}{s/\sqrt{T}} \tag{3.27}$$

for which the cutoff values can be found from tables. For large values of T, however, this distribution is close to the normal.

We want to characterize movements in the monthly yen/dollar exchange rate from historical data, taken over 1990 to 1999. Returns are defined in terms of continuously compounded changes, as in Equation (3.2). The sample size is $T = 120$, and estimated parameters are $m = -0.28\%$ and $s = 3.55\%$ (per month).

Using Equation (3.22), the standard error of the mean is approximately $se(m) = s/\sqrt{T} = 0.32\%$. For the null of $\mu = 0$, this gives a t-ratio of $t = m/se(m) = -0.28\%/0.32\% = -0.87$. Because this number is less than 2 in absolute value, we cannot reject the hypothesis that the mean is zero at the 95% confidence level. This is a typical result for financial series. The mean is not precisely estimated.

Next, we turn to the precision in the sample standard deviation. By Equation (3.25), its standard error is $se(s) = \sigma\sqrt{\frac{1}{(2T)}} = 0.229\%$. For the null of $\sigma = 0$, this gives a ratio of $z = s/se(s) = 3.55\%/0.229\% = 15.5$, which is very high. So, the volatility is not zero. Therefore, there is much more precision in the measurement of s than in that of m.

Furthermore, we can construct 95% confidence intervals around the estimated values. These are:

$$[m - 1.96 \times se(m), m + 1.96 \times se(m)] = [-0.92\%, +0.35\%]$$
$$[s - 1.96 \times se(s), s + 1.96 \times se(s)] = [3.10\%, 4.00\%]$$

So, we could be reasonably confident that the volatility is between 3% and 4%, but we cannot even be sure that the mean is different from zero.

3.2.2 Choosing Significance Levels for Tests

Hypothesis testing requires the choice of a significance level, which needs careful consideration. Two types of errors can arise, as described in Table 3.1. A *type 1* error involves rejecting a correct model. A *type 2* error involves accepting an incorrect model. For a given test, increasing the significance level will decrease the probability of a type 1 error but increase the probability of a type 2 error. Thus, the choice of the significance level should reflect the cost of each of these errors.

This type of situation arises, for example, when a risk manager or regulator must decide whether to accept a VAR model. The first step is to record the number of exceptions, or losses worse than VAR forecasts constructed at the 99% level

TABLE 3.1 Decision Errors

Decision:	Model	
	Correct	Incorrect
Accept	OK	Type 2 error
Reject	Type 1 error	OK

EXAMPLE 3.6: FRM EXAM 2007—QUESTION 2

Which of the following statements regarding hypothesis testing is *incorrect*?

a. Type II error refers to the failure to reject the null hypothesis when it is actually false.

b. Hypothesis testing is used to make inferences about the parameters of a given population on the basis of statistics computed for a sample that is drawn from that population.

c. All else being equal, the decrease in the chance of making a type I error comes at the cost of increasing the probability of making a type II error.

d. The p-value decision rule is to reject the null hypothesis if the p-value is greater than the significance level.

3.3 REGRESSION ANALYSIS

Regression analysis has particular importance for risk management, because it can be used to explain and forecast financial variables.

3.3.1 Bivariate Regression

In a linear regression, the dependent variable y is projected on a set of N predetermined independent variables, x. In the simplest bivariate case we write

$$y_t = \alpha + \beta x_t + \epsilon_t, \qquad t = 1, \ldots, T \qquad (3.28)$$

where α is called the intercept, or constant, β is called the slope, and ϵ is called the residual, or error term. This could represent a time-series or a cross-section.

The ordinary least squares (OLS) assumptions are

- *The errors are independent of x.*
- *The errors have a normal distribution with zero mean and constant variance, conditional on x.*
- *The errors are independent across observations.*

Based on these assumptions, the usual methodology is to estimate the coefficients by minimizing the sum of squared errors. Beta is estimated by

$$\widehat{\beta} = \frac{[1/(T-1)]\sum_t (x_t - \bar{x})(y_t - \bar{y})}{[1/(T-1)]\sum_t (x_t - \bar{x})^2} \qquad (3.29)$$

where $\bar{x}$ and $\bar{y}$ correspond to the means of x_i and y_t. Alpha is estimated by

$$\widehat{\alpha} = \bar{y} - \widehat{\beta}\bar{x} \tag{3.30}$$

Note that the numerator in Equation (3.29) is also the sample covariance between two series x_i and x_j, which can be written as

$$\widehat{\sigma_{ij}} = \frac{1}{(T-1)}\sum_{t=1}^{T}(x_{t,i} - \widehat{\mu_i})(x_{t,j} - \widehat{\mu_j}) \tag{3.31}$$

To interpret β, we can take the covariance between y and x, which is

$$\text{Cov}(y, x) = \text{Cov}(\alpha + \beta x + \epsilon, x) = \beta \text{Cov}(x, x) = \beta V(x)$$

because ϵ is conditionally independent of x. This shows that the population β is also

$$\beta(y, x) = \frac{\text{Cov}(y, x)}{V(x)} = \frac{\rho(y, x)\sigma(y)\sigma(x)}{\sigma^2(x)} = \rho(y, x)\frac{\sigma(y)}{\sigma(x)} \tag{3.32}$$

The regression fit can be assessed by examining the size of the residuals, obtained by subtracting the fitted values $\widehat{y}_t$ from y_t,

$$\widehat{\epsilon}_t = y_t - \widehat{y}_t = y_t - \widehat{\alpha} - \widehat{\beta}x_t \tag{3.33}$$

and taking the estimated variance as

$$V(\widehat{\epsilon}) = \frac{1}{(T-2)}\sum_{t=1}^{T}\widehat{\epsilon}_t^2 \tag{3.34}$$

We divide by $T - 2$ because the estimator uses two unknown quantities, $\widehat{\alpha}$ and $\widehat{\beta}$. Also note that, because the regression includes an intercept, the average value of $\widehat{\epsilon}$ has to be exactly zero.

The quality of the fit can be assessed using a unitless measure called the regression R-square, also called coefficient of determination. This is defined as

$$R^2 = 1 - \frac{\text{SSE}}{\text{SSY}} = 1 - \frac{\sum_t \widehat{\epsilon}_t^2}{\sum_t (y_t - \bar{y})^2} \tag{3.35}$$

where SSE is the sum of squared errors, and SSY is the sum of squared deviations of y around its mean. If the regression includes a constant, we always have $0 \leq R^2 \leq 1$. In this case, R-square is also the square of the usual correlation coefficient,

$$R^2 = \rho(y, x)^2 \tag{3.36}$$

The R^2 measures the degree to which the size of the errors is smaller than that of the original dependent variables y. To interpret R^2, consider two extreme cases.

On one hand, if the fit is excellent, all the errors will be zero, and the numerator in Equation (3.35) will be zero, which gives $R^2 = 1$. On the other hand, if the fit is poor, SSE will be as large as SSY and the ratio will be one, giving $R^2 = 0$.

Alternatively, we can interpret the R-square by decomposing the variance of $y_t = \alpha + \beta x_t + \epsilon_t$. Because ϵ and x are uncorrelated, this yields

$$V(y) = \beta^2 V(x) + V(\epsilon) \tag{3.37}$$

Dividing by $V(y)$,

$$1 = \frac{\beta^2 V(x)}{V(y)} + \frac{V(\epsilon)}{V(y)} \tag{3.38}$$

Because the R-square is also $R^2 = 1 - V(\epsilon)/V(y)$, it is equal to $= \beta^2 V(x)/V(y)$, which is the contribution in the variation of y due to β and x.

Finally, we can derive the distribution of the estimated coefficients, which is normal and centered around the true values. For the slope coefficient, $\widehat{\beta} \sim N(\beta, V(\widehat{\beta}))$, with variance given by

$$V(\widehat{\beta}) = V(\widehat{\epsilon}) \frac{1}{\sum_t (x_t - \bar{x})^2} \tag{3.39}$$

This can be used to test whether the slope coefficient is significantly different from zero. The associated test statistic

$$t = \widehat{\beta}/\sigma(\widehat{\beta}) \tag{3.40}$$

has a Student's t distribution. Typically, if the absolute value of the statistic is above 2, we would reject the hypothesis that there is no relationship between y and x. This corresponds to a two-tailed significance level of 5%.

3.3.2 Autoregression

A particularly useful application is a regression of a variable on a lagged value of itself, called autoregression

$$y_t = \alpha + \beta_k y_{t-k} + \epsilon_t, \qquad t = 1, \ldots, T \tag{3.41}$$

If the β coefficient is significant, previous movements in the variable can be used to predict future movements. Here, the coefficient β_k is known as the kth-order autocorrelation coefficient.

Consider for instance a first-order autoregression, where the daily change in the yen/dollar rate is regressed on the previous day's value. A positive coefficient $\widehat{\beta}_1$

indicates a trend. A negative coefficient indicates mean reversion. As an example, assume that we find that $\widehat{\beta}_1 = 0.10$, with zero intercept. One day, the yen goes up by 2%. Our best forecast for the next day is another upmove of

$$E[y_{t+1}] = \beta_1 y_t = 0.1 \times 2\% = 0.2\%$$

Autocorrelation changes normal patterns in risk across horizons. When there is no autocorrelation, risk increases with the square root of time. With positive autocorrelation, shocks have a longer-lasting effect and risk increases faster than the square root of time.

3.3.3 Multivariate Regression

More generally, the regression in Equation (3.28) can be written, with N independent variables:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_T \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1N} \\ \vdots & & & & \\ x_{T1} & x_{T2} & x_{T3} & \cdots & x_{TN} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_T \end{bmatrix} \tag{3.42}$$

This can include the case of a constant when the first column of X is a vector of ones, in which case β_1 is the usual α. In matrix notation,

$$y = X\beta + \epsilon \tag{3.43}$$

The estimated coefficients can be written in matrix notation as

$$\widehat{\beta} = (X'X)^{-1}X'y \tag{3.44}$$

and their covariance matrix as

$$V(\widehat{\beta}) = \sigma^2(\epsilon)(X'X)^{-1} \tag{3.45}$$

We can extend the t-statistic to a multivariate environment. Say we want to test whether the last m coefficients are jointly zero. Define $\widehat{\beta}_m$ as these grouped coefficients and $V_m(\widehat{\beta})$ as their covariance matrix. We set up a statistic

$$F = \frac{\widehat{\beta}'_m V_m(\widehat{\beta})^{-1}\widehat{\beta}_m/m}{SSE/(T-N)} \tag{3.46}$$

which has an F-distribution with m and $T - N$ degrees of freedom. As before, we would reject the hypothesis if the value of F is too large compared to critical values from tables.
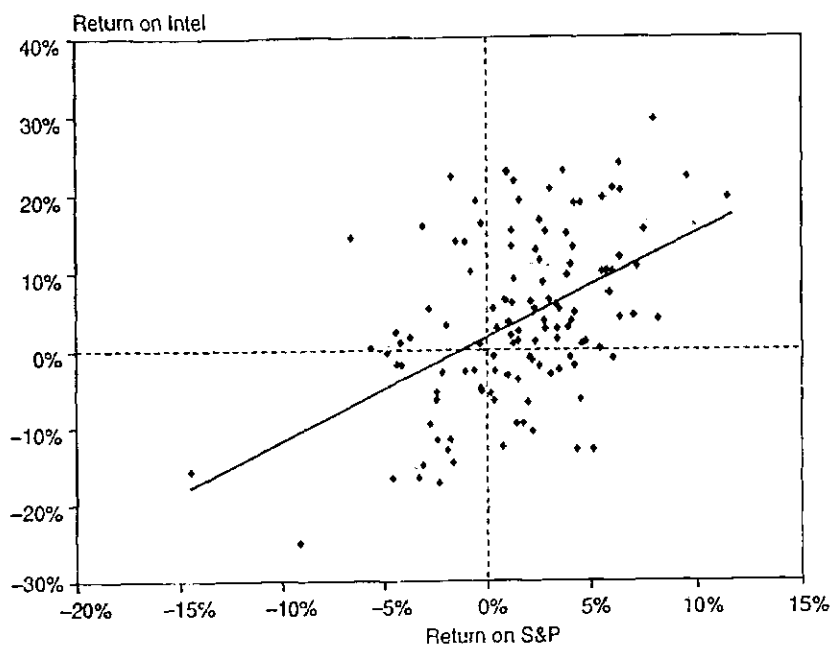
FIGURE 3.1 Intel Return vs. S&P Return

3.3.4 Example

This section gives the example of a regression of a stock return on the market. Such analysis is commonly used to assess whether movements in the stock can be hedged using stock-market index futures.

We consider 10 years of data for Intel and the S&P 500, using total rates of return over a month. Figure 3.1 plots the 120 combination of returns, or (y_t, x_t). Apparently, there is a positive relationship between the two variables, as shown by the straight line that represents the regression fit $(\hat{y}_t, x_t)$.

Table 3.2 displays the regression results. The regression shows a positive relationship between the two variables, with $\hat{\beta} = 1.349$. This is significantly positive, with a standard error of 0.229 and t-statistic of 5.90. The t-statistic is very high, with an associated probability value (p-value) close to zero. Thus, we can be fairly confident of a positive association between the two variables.

This beta coefficient is also called systematic risk, or exposure to general market movements. Typically, technology stocks have greater systematic risk than

TABLE 3.2 Regression Results
$y = \alpha + \beta x$, $y =$ Intel return, $x =$ S&P return

R-square	0.228
Standard error of y	10.94%
Standard error of ϵ	9.62%

Coefficient	Estimate	Standard Error	T-statistic	P-value
Intercept $\hat{\alpha}$	0.0168	0.0094	1.78	0.77
Intercept $\hat{\beta}$	1.349	0.229	5.90	0.00

the average. Indeed, the slope in Intel's regression is greater than unity. To test whether β is significantly different from 1, we can compute a z-score as

$$z = \frac{(\hat{\beta} - 1)}{s(\hat{\beta})} = \frac{(1.349 - 1)}{0.229} = 1.53$$

This is less than the usual cutoff value of 2, so we cannot say for certain that Intel's systematic risk is greater than one.

The R-square of 22.8% can be also interpreted by examining the reduction in dispersion from y to $\hat{\epsilon}$, which is from 10.94% to 9.62%. The R-square can be written as

$$R^{2} = 1 - \frac{9.62\%^{2}}{10.94\%^{2}} = 22.8\%$$

Thus, about 23% of the variance of Intel's returns can be attributed to the market.

EXAMPLE 3.7: FRM EXAM 2004—QUESTION 4

Consider the following linear regression model: $Y = a + b\,X + e$. Suppose $a = 0.05$, $b = 1.2$, SD(Y) = 0.26, SD(e) = 0.1, what is the correlation between X and Y?

a. 0.923
b. 0.852
c. 0.701
d. 0.462

EXAMPLE 3.8: FRM EXAM 2007—QUESTION 22

Consider two stocks, A and B. Assume their annual returns are jointly normally distributed, the marginal distribution of each stock has mean 2% and standard deviation 10%, and the correlation is 0.9. What is the expected annual return of stock A if the annual return of stock B is 3%?

a. 2%
b. 2.9%
c. 4.7%
d. 1.1%

EXAMPLE 3.9: FRM EXAM 2004—QUESTION 23

Which of the following statements about the linear regression of the return of a portfolio over the return of its benchmark presented below are correct?

Portfolio parameter	Value
Beta	1.25
Alpha	0.26
Coefficient of determination	0.66
Standard deviation of error	2.42

I. The correlation is 0.71.

II. 34% of the variation in the portfolio return is explained by variation in the benchmark return.

III. The portfolio is the dependent variable.

IV. For an estimated portfolio return of 12%, the confidence interval at 95% is (7.16% to −16.84%).

    a. II and IV

    b. III and IV

    c. I, II, and III

    d. II, III, and IV

3.3.5 Pitfalls with Regressions

As with any quantitative method, the usefulness of regression analysis depends on the underlying assumptions being fulfilled for the problem at hand. Potential problems of interpretation are now briefly mentioned.

The original OLS setup assumes that the X variables are predetermined (i.e., exogenous or fixed), as in a controlled experiment. In practice, regressions are performed on actual, existing data that do not satisfy these strict conditions. In the previous regression, returns on the S&P are certainly not predetermined.

If the X variables are stochastic, however, most of the OLS results are still valid as long as the X variables are distributed independently of the errors and their distribution does not involve β and σ^2.

Violations of this assumption are serious because they create biases in the slope coefficients. Biases could lead the researcher to come to the wrong conclusion. For instance, we could have measurement errors in the X variables, which causes the measured X to be correlated with ϵ. This so-called errors in the variables problem causes a downward bias, or reduces the estimated slope coefficients from their

true values. Note that errors in the y variables are not an issue, because they are captured by the error component ϵ.

A related problem is that of specification error. Suppose the true model has N variables but we only use a subset N_1. If the omitted variables are correlated with the included variables, the estimated coefficients will be biased. This is a very serious problem because it is difficult to identify. Biases in the coefficients cause problems with estimation.

Another class of problems has to do with potential biases in the standard errors of the coefficients. These errors are especially serious if standard errors are underestimated, creating a sense of false precision in the regression results and perhaps leading to the wrong conclusions. The OLS approach assumes that the errors are independent across observations. This is generally the case for financial time series, but often not in cross-sectional setups. For instance, consider a cross-section of mutual fund returns on some attribute. Mutual fund families often have identical funds, except for the fee structure (e.g., called A for a front load, B for a deferred load). These funds, however, are invested in the same securities and have the same manager. Thus, their returns are certainly not independent. If we run a standard OLS regression with all funds, the standard errors will be too small because we overestimate the number of independent observations. More generally, one has to check that there is no systematic correlation pattern in the residuals. Even with time series, problems can arise with autocorrelation in the errors. Biases in the standard errors cause problems with inference, as one could conclude erroneously that a coefficient is statistically significant.

Problems with efficiency arise when the estimation does not use all available information. For instance, the residuals can have different variances across observations, in which case we have heteroskedasticity. This is the opposite of the constant variance case, or homoskedasticity. Conditional heteroskedasticity occurs when the variance is systematically related to the independent variables. For instance, large values of X could be associated with high error variances. These problems can be identified by diagnostic checks on the residuals. If heteroskedasticity is present, one could construct better standard errors, or try an alternative specification. This is much less of a problem than problems with estimation or inference, however. Inefficient estimates do not necessarily create biases.

Also, regressions may be subject to multicollinearity. This arises when the X variables are highly correlated. Some of the variables may be superfluous, for example using two currencies that are fixed to each other. As a result, the matrix $(X'X)$ in Equation (3.44) will be unstable, and the estimated β unreliable. This problem will show up in large standard errors, however. It can be fixed by discarding some of the variables that are highly correlated with others.

Last, even if all the OLS conditions are satisfied, one has to be extremely careful about using a regression for forecasting. Unlike physical systems, which are inherently stable, financial markets are dynamic and relationships can change quickly. Indeed, financial anomalies, which show up as strongly significant coefficients in historical regressions, have an uncanny ability to disappear as soon as one tries to exploit them.

EXAMPLE 3.10: FRM EXAM 2004—QUESTION 59

Which of the following statements regarding linear regression is *false?*

a. Heteroskedasticity occurs when the variance of residuals is not the same across all observations in the sample.
b. Unconditional heteroskedasticity leads to inefficient estimates, whereas conditional heteroskedasticity can lead to problems with both inference and estimation.
c. Serial correlation occurs when the residual terms are correlated with each other.
d. Multicollinearity occurs when a high correlation exists between or among two or more of the independent variables in a multiple regression.

EXAMPLE 3.11: FRM EXAM 1999—QUESTION 2

Under what circumstances could the explanatory power of regression analysis be overstated?

a. The explanatory variables are not correlated with one another.
b. The variance of the error term decreases as the value of the dependent variable increases.
c. The error term is normally distributed.
d. An important explanatory variable is omitted that influences the explanatory variables included, and the dependent variable.

3.4 IMPORTANT FORMULAS

Discrete returns, log returns: $r_t = (S_t - S_{t-1})/S_{t-1}, \quad R_t = \ln[S_t/S_{t-1}]$

Time aggregation: $E(R_T) = E(R_1)T, \quad V(R_T) = V(R_1)T, \quad SD(R_T) = SD(R_1)\sqrt{T}$

Portfolio rate of return, variance: $r_{p,t+1} = \sum_{i=1}^{N} w_{i,t}\, r_{i,t+1} = w'R,$

$\qquad V[r_{p,t+1}] = w'\Sigma w$

Estimated mean, variance: $m = \widehat{\mu} = \frac{1}{T}\sum_{i=1}^{T} x_i, \quad s^2 = \widehat{\sigma}^2 = \frac{1}{(T-1)}\sum_{i=1}^{T}(x_i - \widehat{\mu})^2$

Distribution of estimated mean, variance, standard deviation: $m = \widehat{\mu} \sim$

$\qquad N(\mu, \sigma^2/T), \quad \frac{(T-1)\widehat{\sigma}^2}{\sigma^2} \sim \chi^2(T-1), \quad \widehat{\sigma}^2 \to N\left(\sigma^2, \sigma^4 \frac{2}{(T-1)}\right), \quad se(\widehat{\sigma}) = \sigma\sqrt{\frac{1}{2T}}$

Bivariate, multivariate regression: $y_t = \alpha + \beta x_t + \epsilon_t, \quad y = X\beta + \epsilon$

Estimated beta: $\widehat{\beta} = (X'X)^{-1}X'y$

Population beta: $\beta(y, x) = \frac{Cov(y,x)}{V(x)} = \frac{\rho(y,x)\sigma(y)\sigma(x)}{\sigma^2(x)} = \rho(y, x)\frac{\sigma(y)}{\sigma(x)}$

Regression R-square: $R^2 = 1 - \frac{SSE}{SSY} = 1 - \frac{\sum_t \hat{\epsilon}_t^2}{\sum_t (y_t - \bar{y})^2}$

Variance decomposition: $V(y) = \beta^2 V(x) + V(\epsilon)$

T-statistic for hypothesis of zero coefficient: $t = \hat{\beta}/\sigma(\hat{\beta})$

3.5 ANSWERS TO CHAPTER EXAMPLES

Example 3.1: FRM Exam 1999—Question 4

d. Efficient markets implies that the distribution of future returns does not depend on past returns. Hence, returns cannot be correlated. It could happen, however, that return distributions are independent, but that, just by chance, two successive returns are equal.

Example 3.2: FRM Exam 2002—Question 3

d. Assuming a random walk, we can use the square root of time rule. The weekly volatility is then $34\% \times 1/\sqrt{52} = 4.71\%$.

Example 3.3: FRM Exam 2002—Question 2

a. With mean reversion, the volatility grows more slowly than the square root of time

Example 3.4: FRM Exam 2004—Question 39

d. The variance of the portfolio is given by $\sigma_p^2 = (0.4)^2 25 + (0.6)^2 121 + 2(0.4)(0.6)0.3 \sqrt{25 \times 121} = 55.48$. Hence, the volatility is 7.45.

Example 3.5: FRM Exam 2007—Question 137

c. The significance level is the probability of committing a type 1 error, or rejecting a correct model. This is also P(reject H0 | H0 is true). On the other hand, the type 2 error rate is P(not reject H0 | H0 is false).

Example 3.6: FRM Exam 2007—Question 2

d. We would reject the null if the observed p-value is *lower* (not greater) than the significance level.

Example 3.7: FRM Exam 2004—Question 4

a. We can find the volatility of X from the variance decomposition, Equation (3.37). This gives $V(x) = [V(y) - V(e)]/\beta^2 = [0.26^2 - 0.10^2]/1.2^2 = 0.04$. Then $SD(X) = 0.2$, and $\rho = \beta SD(X)/SD(Y) = 1.20.2/0.26 = 0.923$.

Example 3.8: FRM Exam 2007—Question 22

b. The information in this question can be used to construct a regression model of A on B. We have $R_A = 2\% + 0.9(10\%/10\%)(R_B - 2\%) + \epsilon$. Next, replacing R_B by 3% gives $\hat{R}_A = 2\% + 0.9(3\% - 2\%) = 2.9\%$.

Example 3.9: FRM Exam 2004—Question 23

b. The correlation is given by $\sqrt{0.66} = 0.81$, so I. is incorrect. Next, 66% of the variation in Y is explained by the benchmark, so answer II. is incorrect. The portfolio return is indeed the dependent variable Y, so answer III. is correct. Finally, to find the 95% two-tailed confidence interval, we use α from a normal distribution, which covers 95% within plus or minus 1.96, close to 2.00. The interval is then $y - 2SD(e)$, $y + 2SD(e)$, or $(7.16 - 16.84)$. So answers III. and IV. are correct.

Example 3.10: FRM Exam 2004—Question 59

b. Heteroskedasticity indeed occurs when the variance of the residuals is not constant, so a. is correct. This leads to inefficient estimates but otherwise does not cause problems with inference and estimation. Statements c. and d. are correct.

Example 3.11: FRM Exam 1999—Question 2

d. If the true regression includes a third variable z that influences both y and x, the error term will not be conditionally independent of x, which violates one of the assumptions of the OLS model. This will artificially increase the explanatory power of the regression. Intuitively, the variable x will appear to explain more of the variation in y simply because it is correlated with z.

Monte Carlo Methods

The two preceding chapters dealt with probability and statistics. The former involves the generation of random variables from known distributions. The second deals with estimation of distribution parameters from actual data. With estimated distributions in hand, we can proceed to the next step, which is the simulation of random variables for the purpose of risk management. Such simulations, called Monte Carlo simulations, are central to financial engineering and risk management. They allow financial engineer to price complex financial instruments. They allow risk managers to build the distribution of portfolios that are too complex to model analytically.

Simulation methods are quite flexible and are becoming easier to implement with technological advances in computing. Their drawbacks should not be underestimated, however. For all their elegance, simulation results depend heavily on the model's assumptions: the shape of the distribution, the parameters, and the pricing functions. Risk managers need to be keenly aware of the effect that errors in these assumptions can have on the results.

This chapter shows how Monte Carlo methods can be used for risk management. Section 4.1 introduces a simple case with just one source of risk. Section 4.2 shows how to apply these methods to construct value at risk (VAR) measures, as well as to price derivatives. Multiple sources of risk are then considered in Section 4.3.

4.1 SIMULATIONS WITH ONE RANDOM VARIABLE

Simulations involve creating artificial random variables with properties similar to those of the risk factors in the portfolio. These include stock prices, exchange rates, bond yields or prices, and commodity prices.

4.1.1 Simulating Markov Processes

In efficient markets, financial prices should display a random walk pattern. More precisely, prices are assumed to follow a Markov process, which is a particular stochastic process independent of its past history; the entire distribution of the future price relies on the current price only. The past history is irrelevant. These

processes are built from the following components, described in order of increasing complexity.

- The Wiener process. This describes a variable Δz, whose change is measured over the interval Δt such that its mean change is zero and variance proportional to Δt:

$$\Delta z \sim N(0, \Delta t) \tag{4.1}$$

If ϵ is a standard normal variable $N(0, 1)$, this can be written as $\Delta z = \epsilon \sqrt{\Delta t}$. In addition, the increments Δz are independent across time.

- The generalized Wiener process. This describes a variable Δx built up from a Wiener process, with in addition a constant trend a per unit time and volatility b:

$$\Delta x = a\Delta t + b\Delta z \tag{4.2}$$

A particular case is the martingale, which is a zero drift stochastic process, $a = 0$, which leads to $E(\Delta x) = 0$. This has the convenient property that the expectation of a future value is the current value

$$E(x_T) = x_0 \tag{4.3}$$

- The Ito process. This describes a generalized Wiener process, whose trend and volatility depend on the *current* value of the underlying variable and time:

$$\Delta x = a(x, t)\Delta t + b(x, t)\Delta z \tag{4.4}$$

This is a Markov process because the distribution depends only on the current value of the random variable x, as well as time. In addition, the innovation in this process has a normal distribution.

4.1.2 The Geometric Brownian Motion

A particular example of Ito process is the geometric Brownian motion (GBM), which is described for the variable S as

$$\Delta S = \mu S \Delta t + \sigma S \Delta z \tag{4.5}$$

The process is geometric because the trend and volatility terms are proportional to the current value of S. This is typically the case for stock prices, for which *rates of returns* appear to be more stationary than raw dollar returns, ΔS. It is also used for currencies. Because $\Delta S/S$ represents the capital appreciation only, abstracting from dividend payments, μ represents the expected total rate of return on the asset minus the rate of income payment, or dividend yield in the case of stocks.

Example: A Stock Price Process

Consider a stock that pays no dividends, has an expected return of 10% per annum, and volatility of 20% per annum. If the current price is $100, what is the process for the change in the stock price over the next week? What if the current price is $10?

The process for the stock price is

$$\Delta S = S(\mu \Delta t + \sigma \sqrt{\Delta t} \times \epsilon)$$

where ϵ is a random drawn from a standard normal distribution. If the interval is one week, or $\Delta t = 1/52 = 0.01923$, the mean is $\mu \Delta t = 0.10 \times 0.01923 = 0.001923$ and $\sigma \sqrt{\Delta t} = 0.20 \times \sqrt{0.01923} = 0.027735$. The process is $\Delta S = \$100(0.001923 + 0.027735 \times \epsilon)$. With an initial stock price at $100, this gives $\Delta S = 0.1923 + 2.7735\epsilon$. With an initial stock price at $10, this gives $\Delta S = 0.01923 + 0.27735\epsilon$. The trend and volatility are scaled down by a factor of 10.

This model is particularly important because it is the underlying process for the Black–Scholes formula. The key feature of this distribution is the fact that the volatility is proportional to S. This ensures that the stock price will stay positive. Indeed, as the stock price falls, its variance decreases, which makes it unlikely to experience a large downmove that would push the price into negative values. As the limit of this model is a normal distribution for $dS/S = d\ln(S)$, S follows a lognormal distribution.

This process implies that, over an interval $T - t = \tau$, the logarithm of the ending price is distributed as

$$\ln(S_T) = \ln(S_t) + (\mu - \sigma^2/2)\tau + \sigma \sqrt{\tau}\, \epsilon \qquad (4.6)$$

where ϵ is a standardized normal variable.

Example: A Stock Price Process (Continued)

Assume the price in one week is given by $S = \$100\exp(R)$, where R has annual expected value of 10% and volatility of 20%. Construct a two-tailed 95% confidence interval for S.

The standard normal deviates that corresponds to a 95% confidence interval are $\alpha_{MIN} = -1.96$ and $\alpha_{MAX} = 1.96$. In other words, we have 2.5% in each tail. The 95% confidence band for R is then $R_{MIN} = \mu \Delta t - 1.96\sigma \sqrt{\Delta t} = 0.001923 - 1.96 \times 0.027735 = -0.0524$ and $R_{MAX} = \mu \Delta t + 1.96\sigma \sqrt{\Delta t} = 0.001923 + 1.96 \times 0.027735 = 0.0563$. This gives $S_{MIN} = \$100\exp(-0.0524) = \$94.89$, and $S_{MAX} = \$100\exp(0.0563) = \$105.79$.

Whether a lognormal distribution is much better than the normal distribution depends on the horizon considered. If the horizon is one day only, the choice of the lognormal versus normal assumption does not really matter. It is highly unlikely that the stock price would drop below zero in one day, given typical volatilities. On the other hand, if the horizon is measured in years, the two assumptions do lead to different results. The lognormal distribution is more realistic as it prevents prices form turning negative.

In simulations, this process is approximated by small steps with a normal distribution with mean and variance given by

$$\frac{\Delta S}{S} \sim N(\mu \Delta t, \sigma^2 \Delta t) \tag{4.7}$$

To simulate the future price path for S, we start from the current price S_t and generate a sequence of independent standard normal variables ϵ, for $i = 1, 2, \ldots, n$. The next price S_{t+1} is built as $S_{t+1} = S_t + S_t(\mu \Delta t + \sigma \epsilon_1 \sqrt{\Delta t})$. The following price S_{t+2} is taken as $S_{t+1} + S_{t+1}(\mu \Delta t + \sigma \epsilon_2 \sqrt{\Delta t})$, and so on until we reach the target horizon, at which point the price $S_{t+n} = S_T$ should have a distribution close to the lognormal.

Table 4.1 illustrates a simulation of a process with a drift (μ) of 0% and volatility (σ) of 20% over the total interval, which is divided into 100 steps.

The initial price is \$100. The local expected return is $\mu \Delta t = 0.0/100 = 0.0$ and the volatility is $0.20 \times \sqrt{1/100} = 0.02$. The second column shows the realization of a uniform $U(0, 1)$ variable. The value for the first step is $u_1 = 0.0430$. The next column transforms this variable into a normal variable with mean 0.0 and volatility of 0.02, which gives -0.0343. The price increment is then obtained by multiplying the random variable by the previous price, which gives $-\$3.433$. This generates a new value of $S_1 = \$100 - \$3.43 = \$96.57$. The process is repeated until the final price of \$125.31 is reached at the 100th step.

This experiment can be repeated as often as needed. Define K as the number of replications, or random trials. Figure 4.1 displays the first three trials. Each leads

TABLE 4.1 Simulating a Price Path

Step i	Random Variable Uniform u_i	Random Variable Normal $\mu \Delta t + \sigma \Delta z$	Price Increment ΔS_i	Price S_{t+i}
0				100.00
1	0.0430	−0.0343	−3.433	96.57
2	0.8338	0.0194	1.872	98.44
3	0.6522	0.0078	0.771	99.21
4	0.9219	0.0284	2.813	102.02
...				
99				124.95
100	0.5563	0.0028	0.354	125.31
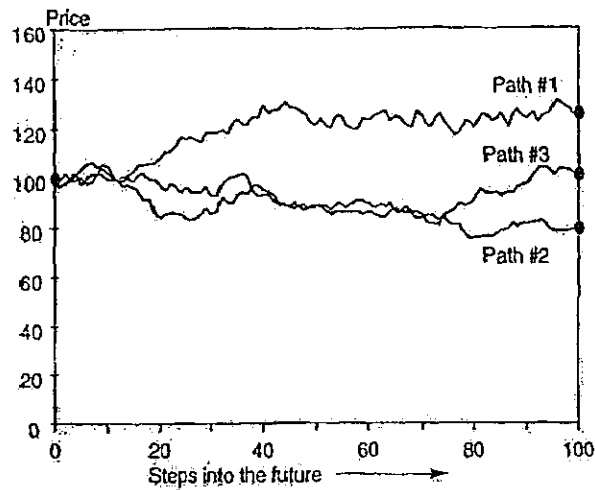
FIGURE 4.1   Simulating Price Paths

to a simulated final value S_T^i. This generates a distribution of simulated prices S_T. With just one step $n = 1$, the distribution must be normal. As the number of steps n grows large, the distribution tends to a lognormal distribution.

While very useful to model stock prices, this model has shortcomings. Price increments are assumed to have a normal distribution. In practice, we observe that price changes for most financial assets typically have fatter tails than the normal distribution. Returns may also experience changing variances.

In addition, as the time interval Δt shrinks, the volatility shrinks as well. This implies that large discontinuities cannot occur over short intervals. In reality, some assets experience discrete jumps, such as commodities, or securities issued by firms that go bankrupt. In such cases, the stochastic process should be changed to accommodate these observations.

EXAMPLE 4.1: FRM EXAM 2003—QUESTION 40

In the geometric Brownian motion process for a variable S,

  I. S is normally distributed.
 II. $d\ln(S)$ is normally distributed.
III. dS/S is normally distributed.
 IV. S is lognormally distributed.
    a. I only
    b. II, III, and IV
    c. IV only
    d. III and IV

EXAMPLE 4.2: FRM EXAM 2002—QUESTION 126

Consider that a stock price S that follows a geometric Brownian motion $dS = aSdt + bSdz$, with b strictly positive. Which of the following statements is *false*?

a. If the drift a is positive, the price one year from now will be above today's price.

b. The instantaneous rate of return on the stock follows a normal distribution.

c. The stock price S follows a lognormal distribution.

d. This model does not impose mean reversion.

4.1.3 Drawing Random Variables

Most spreadsheets or statistical packages have functions that can generate uniform or standard normal random variables. This can be easily extended to distributions that better reflect the data, e.g., with fatter tails or non-zero skewness.

The methodology involves the inverse cumulative probability distribution function (p.d.f.). Take the normal distribution as an example. By definition, the cumulative p.d.f. $N(x)$ is always between 0 and 1. Because we have an analytical formula for this function, it can be easily inverted.

First, we generate a uniform random variable u drawn from $U(0, 1)$. Next, we compute x, such that $u = N(x)$, or $x = N^{-1}(u)$. For example, set $u = 0.0430$, as in the first line of Table 4.1. This gives $x = -1.717.$[1] Because u is less than 0.5, we verify that x is negative. The variable can be transformed into any normal variable by multiplying by the standard deviation and adding the mean. More generally, any distribution function can be generated as long as the cumulative distribution function can be inverted.

4.1.4 Simulating Yields

The GBM process is widely used for stock prices and currencies. Fixed-income products are another matter, however.

Bond prices display long-term reversion to the face value, which represents the repayment of principal at maturity (assuming there is no default). Such process is inconsistent with the GBM process, which displays no such mean reversion. The volatility of bond prices also changes in a predictable fashion, as duration shrinks to zero. Similarly, commodities often display mean reversion.

---

[1] In Excel, a uniform random variable can be generated with the function u_i = RAND(). From this, a standard normal random variable can be computed with NORMSINV(u_i).

These features can be taken into account by modeling bond yields directly in a first step. In the next step, bond prices are constructed from the value of yields and a pricing function. The dynamics of interest rates r_t can be modeled by

$$\Delta r_t = \kappa(\theta - r_t)\Delta t + \sigma r_t{}^{\gamma}\Delta z_t \qquad (4.8)$$

where Δz_t is the usual Wiener process. Here, we assume that $0 \leq \kappa < 1, \theta \geq 0$, $\sigma \geq 0$. Because there is only one stochastic variable for yields, the model is called a one-factor model.

This Markov process has a number of interesting features. First, it displays mean reversion to a long-run value of θ. The parameter κ governs the speed of mean reversion. When the current interest rate is high, i.e., $r_t > \theta$, the model creates a negative drift $\kappa(\theta - r_t)$ toward θ. Conversely, low current rates create a positive drift toward θ.

The second feature is the volatility process. This model includes the Vasicek model when $\gamma = 0$. Changes in yields are normally distributed because Δr is then a linear function of Δz, which is itself normal. The Vasicek model is particularly convenient because it leads to closed-form solutions for many fixed-income products. The problem, however, is that it could potentially lead to negative interest rates when the initial rate starts from a low value. This is because the volatility of the change in rates does not depend on the level, unlike that in the geometric Brownian motion.

Equation (4.8) is more general, however, because it includes a power of the yield in the variance function. With $\gamma = 1$, this is the lognormal model. Ignoring the trend, this gives $\Delta r_t = \sigma r_t \Delta z_t$, or $\Delta r_t / r_t = \sigma \Delta z_t$. This implies that the *rate of change* in the yield dr/r has a fixed variance. Thus, as with the GBM model, smaller yields lead to smaller movements, which makes it unlikely the yield will drop below zero. This model is more appropriate than the normal model when the initial yield is close to zero.

With $\gamma = 0.5$, this is the Cox, Ingersoll, and Ross (CIR) model. Ultimately, the choice of the exponent γ is an empirical issue. Recent research has shown that $\gamma = 0.5$ provides a good fit to the data.

This class of models is known as equilibrium models. They start with some assumptions about economic variables and imply a process for the short-term interest rate r. These models generate a predicted term structure, whose shape depends on the model parameters and the initial short rate. The problem with these models, however, is that they are not flexible enough to provide a good fit to today's term structure. This can be viewed as unsatisfactory, especially by practitioners who argue they cannot rely on a model that cannot be trusted to price today's bonds.

In contrast, no-arbitrage models are designed to be consistent with today's term structure. In this class of models, the term structure is an input into the parameter estimation. The earliest model of this type was the Ho and Lee model:

$$\Delta r_t = \theta(t)\Delta t + \sigma \Delta z_t \qquad (4.9)$$

where $\theta(t)$ is a function of time chosen so that the model fits the initial term structure. This was extended to incorporate mean reversion in the Hull and White model:

$$\Delta r_t = [\theta(t) - ar_t]\Delta t + \sigma \Delta z_t \qquad (4.10)$$

Finally, the Heath, Jarrow, and Morton model goes one step further and assumes that the volatility is a function of time.

The downside of these no-arbitrage models is that they do not impose any consistency between parameters estimated over different dates. The function $\theta(t)$ could be totally different from one day to the next, which is illogical. No-arbitrage models are also more sensitive to outliers, or data errors in bond prices used to fit the term structure.

4.1.5 Binomial Trees

Simulations are very useful to mimic the uncertainty in risk factors, especially with numerous risk factors. In some situations, however, it is also useful to describe the uncertainty in prices with discrete trees. When the price can take one of two steps, the tree is said to be binomial.

The binomial model can be viewed as a discrete equivalent to the geometric Brownian motion. As before, we subdivide the horizon T into n intervals, $\Delta t = T/n$. At each "node," the price is assumed to go either up with probability p, or down with probability $1 - p$.

The parameters u, d, p are chosen so that, for a small time interval, the expected return and variance equal those of the continuous process. One could choose

$$u = e^{\sigma\sqrt{\Delta t}}, \quad d = (1/u), \quad p = \frac{e^{\mu\Delta t} - d}{u - d} \qquad (4.11)$$

This matches the mean, for example,

$$E\left[\frac{S_1}{S_0}\right] = pu + (1 - p)d = \frac{e^{\mu\Delta t} - d}{u - d}u + \frac{u - e^{\mu\Delta t}}{u - d}d$$

$$= \frac{e^{\mu\Delta t}(u - d) - du + ud}{u - d} = e^{\mu\Delta t}$$

Table 4.2 shows how a binomial tree is constructed. As the number of steps increases, the discrete distribution of S_T converges to the lognormal distribution. This model will be used in a later chapter to price options.

TABLE 4.2 Binomial Tree

Step			
0	1	2	3

$$
\begin{array}{ccccccc}
 & & & & & & u^3S \\
 & & & & & \nearrow & \\
 & & & & u^2S & & \\
 & & & \nearrow & & \searrow & \\
 & & uS & & & & u^2dS \\
 & \nearrow & & \searrow & & \nearrow & \\
S & & & & udS & & \\
 & \searrow & & \nearrow & & \searrow & \\
 & & dS & & & & d^2uS \\
 & & & \searrow & & \nearrow & \\
 & & & & d^2S & & \\
 & & & & & \searrow & \\
 & & & & & & d^3S
\end{array}
$$
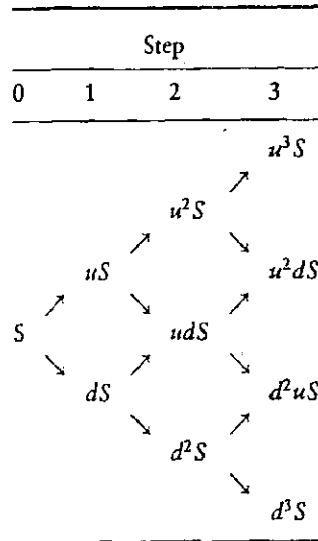
EXAMPLE 4.3: FRM EXAM 1999—QUESTION 25

The Vasicek model defines a risk-neutral process for r which is $dr = a(b-r)dt + \sigma dz$, where a, b, and σ are constant, and r represents the rate of interest. From this equation we can conclude that the model is a

a. Monte Carlo-type model
b. Single-factor term-structure model
c. Two-factor term-structure model
d. Decision tree model

EXAMPLE 4.4: FRM EXAM 1999—QUESTION 26

The term $a(b-r)$ in the previous question represents which term?

a. Gamma
b. Stochastic
c. Reversion
d. Vega

EXAMPLE 4.5: FRM EXAM 2000—QUESTION 118

Which group of term-structure models do the Ho-Lee, Hull-White, and Heath, Jarrow, and Morton models belong to?

a. No-arbitrage models
b. Two-factor models
c. Lognormal models
d. Deterministic models

EXAMPLE 4.6: FRM EXAM 2000—QUESTION 119

A plausible stochastic process for the short-term rate is often considered to be one where the rate is pulled back to some long-run average level. Which one of the following term-structure models does *not* include this characteristic?

a. The Vasicek model
b. The Ho-Lee model
c. The Hull-White model
d. The Cox-Ingersoll-Ross model

4.2 IMPLEMENTING SIMULATIONS

4.2.1 Simulation for VAR

Implementing Monte Carlo (MC) methods for risk management follows these steps:

1. Choose a stochastic process for the risk factor price S (i.e., its distribution and parameters, starting from the current value S_t).
2. Generate pseudo-random variables representing the risk factor at the target horizon, S_T.
3. Calculate the value of the portfolio at the horizon, $F_T(S_T)$.
4. Repeat steps 2 and 3 as many times as necessary. Call K the number of replications.

These steps create a distribution of values, $F_T^1, \ldots, F_T^K$, which can be sorted to derive the VAR. We measure the cth quantile $Q(F_T, c)$ and the average value

Ave(F_T). If VAR is defined as the deviation from the expected value on the target date, we have

$$\text{VAR}(c) = \text{Ave}(F_T) - Q(F_T, c) \tag{4.12}$$

4.2.2 Simulation for Derivatives

Readers familiar with derivatives pricing will have recognized that this method is similar to the Monte Carlo method for valuing derivatives. In that case, we simply focus on the expected value on the target date discounted into the present:

$$F_t = e^{-r(T-t)} \text{Ave}(F_T) \tag{4.13}$$

Thus, derivatives valuation focuses on the discounted center of the distribution, while VAR focuses on the quantile on the target date.

Monte Carlo simulations have been long used to price derivatives. As will be seen in a later chapter, pricing derivatives can be done by assuming that the underlying asset grows at the risk-free rate r (assuming no income payment). For instance, pricing an option on a stock with expected return of 20% is done assuming that (1) the stock grows at the risk-free rate of 10% and (2) we discount at the same risk-free rate. This is called the risk-neutral approach.

In contrast, risk measurement deals with actual distributions, sometimes called physical distributions. For measuring VAR, the risk manager must simulate asset growth using the actual expected return μ of 20%. Therefore, risk management uses physical distributions, whereas pricing methods use risk-neutral distributions.

It should be noted that simulation methods are not applicable to all types of options. These methods assume that the value of the derivative instrument at expiration can be priced solely as a function of the end-of-period price S_T, and perhaps of its sample path. This is the case, for instance, with an Asian option, where the payoff is a function of the price *averaged* over the sample path. Such an option is said to be path-dependent.

Simulation methods, however, are inadequate to price American options, because such options can be exercised early. The optimal exercise decision, however, is complex to model because it should take into account *future* values of the option. This cannot be done with regular simulation methods, which only consider present and past information. Instead, valuing American options requires a backward recursion, for example with binomial trees. This method examines whether the option should be exercised or not, starting from the end and working backward in time until the starting time.

4.2.3 Accuracy

Finally, we should mention the effect of sampling variability. Unless K is extremely large, the empirical distribution of S_T will only be an approximation of the true

distribution. There will be some natural variation in statistics measured from Monte Carlo simulations. Since Monte Carlo simulations involve *independent* draws, one can show that the standard error of statistics is inversely related to the square root of K. Thus more simulations will increase precision, but at a slow rate. For example, accuracy is increased by a factor of ten going from $K = 10$ to $K = 1,000$, but then requires going from $K = 1,000$ to $K = 100,000$ for the same factor of 10.

This accuracy issue is worse for risk management than for pricing, because the quantiles are estimated less precisely than the average. For VAR measures, the precision is also a function of the selected confidence level. Higher confidence levels generate fewer observations in the left tail and hence less-precise VAR measures. A 99% VAR using 1,000 replications should be expected to have only 10 observations in the left tail, which is not a large number. The VAR estimate is derived from the tenth and eleventh sorted number. In contrast, a 95% VAR is measured from the fiftieth and fifty-first sorted numbers, which is more precise. In addition, the precision of the estimated quantile depends on the shape of the distribution. Relative to a symmetric distribution, a short option position has negative skewness, or a long left tail. The observations in the left tail therefore will be more dispersed, making is more difficult to estimate VAR precisely.

Various methods are available to speed up convergence:

- **Antithetic Variable Technique.** This technique uses twice the same sequence of random draws from t to T. It takes the original sequence and changes the sign of all their values. This creates twice the number of points in the final distribution of F_T without running twice the number of simulations.

- **Control Variate Technique.** This technique is used to price options with trees when a similar option has an analytical solution. Say that f_E is a European option with an analytical solution. Going through the tree yields the values of an American and European option, F_A and F_E. We then assume that the error in F_A is the same as that in F_E, which is known. The adjusted value is $F_A - (F_E - f_E)$.

- **Quasi-Random Sequences.** These techniques, also called Quasi Monte Carlo (QMC), create draws that are not independent but instead are designed to fill the sample space more uniformly. Simulations have shown that QMC methods converge faster than Monte Carlo. In other words, for a fixed number of replications K, QMC values will be on average closer to the true value.

The advantage of traditional MC, however, is that it also provides a standard error, which is on the order of $1/\sqrt{K}$ because draws are independent. So, we have an idea of how far the estimate might be from the true value, which is useful to decide on the number of replications. In contrast, QMC methods give no measure of precision.

Monte Carlo Methods

EXAMPLE 4.7: FRM EXAM 2005—QUESTION 67

Which of the following statements about Monte Carlo simulation is *false*?

a. Monte Carlo simulation can be used with a lognormal distribution.
b. Monte Carlo simulation can generate distributions for portfolios that contain only linear positions.
c. One drawback of Monte Carlo simulation is that it is computationally very intensive.
d. Assuming the underlying process is normal, the standard error resulting from Monte Carlo simulation is inversely related to the square root of the number of trials.

EXAMPLE 4.8: FRM EXAM 2007—QUESTION 66

A risk manager has been requested to provide some indication of accuracy of a Monte Carlo simulation. Using 1,000 replications of a normally distributed variable S, the relative error in the one-day 99% VAR is 5%. Under these conditions,

a. Using 1,000 replications of a long option position on S should create a larger relative error.
b. Using 10,000 replications should create a larger relative error.
c. Using another set of 1,000 replications will create an exact measure of 5.0% for relative error.
d. Using 1,000 replications of a short option position on S should create a larger relative error.

EXAMPLE 4.9: SAMPLING VARIATION

The measurement error in VAR, due to sampling variation, should be greater with

a. More observations and a high confidence level (e.g., 99%)
b. Fewer observations and a high confidence level
c. More observations and a low confidence level (e.g., 95%)
d. Fewer observations and a low confidence level

4.3 MULTIPLE SOURCES OF RISK

We now turn to the more general case of simulations with many sources of financial risk. Define N as the number of risk factors. If the factors S_j are independent, the randomization can be performed independently for each variable. For the GBM model,

$$\Delta S_{j,t} = S_{j,t-1}\mu_j \Delta t + S_{j,t-1}\sigma_j \epsilon_{j,t}\sqrt{\Delta t} \tag{4.14}$$

where the standard normal variables ϵ are independent across time and factor $j = 1, \ldots, N$.

In general, however, risk factors are correlated. The simulation can be adapted by, first, drawing a set of independent variables η, and, second, transforming them into correlated variables ϵ. As an example, with two factors only, we write

$$\epsilon_1 = \eta_1$$
$$\epsilon_2 = \rho\eta_1 + (1 - \rho^2)^{1/2}\eta_2 \tag{4.15}$$

Here, ρ is the correlation coefficient between the variables ϵ. Because the η's have unit variance and are uncorrelated, we verify that the variance of ϵ_2 is one, as required

$$V(\epsilon_2) = \rho^2 V(\eta_1) + [(1 - \rho^2)^{1/2}]^2 V(\eta_2) = \rho^2 + (1 - \rho^2) = 1$$

Furthermore, the correlation between ϵ_1 and ϵ_2 is given by

$$\text{Cov}(\epsilon_1, \epsilon_2) = \text{Cov}(\eta_1, \rho\eta_1 + (1 - \rho^2)^{1/2}\eta_2) = \rho\text{Cov}(\eta_1, \eta_1) = \rho$$

Defining ϵ as the *vector* of values, we verified that the covariance matrix of ϵ is

$$V(\epsilon) = \begin{bmatrix} \sigma^2(\epsilon_1) & \text{Cov}(\epsilon_1, \epsilon_2) \\ \text{Cov}(\epsilon_1, \epsilon_2) & \sigma^2(\epsilon_2) \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = R$$

Note that this covariance matrix, which is the expectation of squared deviations from the mean, can also be written as

$$V(\epsilon) = E[(\epsilon - E(\epsilon)) \times (\epsilon - E(\epsilon))'] = E(\epsilon \times \epsilon')$$

because the expectation of ϵ is 0. To generalize this approach to many more risk factors, however, we need a systematic way to derive the transformation in Equation (4.15).

4.3.1 The Cholesky Factorization

We would like to generate N joint values of ϵ that display the correlation structure $V(\epsilon) = E(\epsilon\epsilon') = R$. Because the matrix R is a symmetric real matrix, it can be decomposed into its so-called Cholesky factors:

$$R = TT' \qquad (4.16)$$

where T is a lower triangular matrix with zeros on the upper right corners (above the diagonal). This is known as the Cholesky factorization.

As in the previous section, we first generate a vector of independent η. Thus, their covariance matrix is $V(\eta) = I$, where I is the identity matrix with zeros everywhere except for the diagonal.

We then construct the transformed variable $\epsilon = T\eta$. The covariance matrix is now $V(\epsilon) = E(\epsilon\epsilon') = E((T\eta)(T\eta)') = E(T\eta\eta'T') = TE(\eta\eta')T' = TV(\eta)T' = TIT' = TT' = R$. This transformation therefore generates ϵ variables with the desired correlations.

To illustrate, let us go back to our two-variable case. The correlation matrix can be decomposed into its Cholesky factors:

$$\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 \end{bmatrix}$$

To find the entries a_{11}, a_{21}, a_{22}, we solve each of the three equations

$$a_{11}^2 = 1$$

$$a_{11}a_{21} = \rho$$

$$a_{21}^2 + a_{22}^2 = 1$$

This gives $a_{11} = 1$, $a_{21} = \rho$, and $a_{22} = (1 - \rho^2)^{1/2}$. The Cholesky factorization is then

$$\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & (1-\rho^2)^{1/2} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ 0 & (1-\rho^2)^{1/2} \end{bmatrix}$$

Note that this conforms to Equation (4.15):

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & (1-\rho^2)^{1/2} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

In practice, this decomposition yields a number of useful insights. The decomposition will fail if the number of independent factors implied in the

correlation matrix is less than N. For instance, if $\rho = 1$, the two assets are perfectly correlated, meaning that we have twice the same factor. This could be for instance the case of two currencies fixed to each other. The decomposition gives $a_{11} = 1, a_{21} = 1, a_{22} = 0$. The new variables are then $\epsilon_1 = \eta_1$ and $\epsilon_2 = \eta_1$. In this case, the second variable η_2 is totally superfluous and simulations can be performed with one variable only.

4.3.2 The Curse of Dimensionality

Modern risk management is about measuring the risk of large portfolios, typically exposed to a large number of risk factors. The problem is that the number of computations increases geometrically with the number of factors N. The covariance matrix, for instance, has dimensions $N(N+1)/2$. A portfolio with 500 variables requires a matrix with 125,250 entries.

In practice, the risk manager should simplify the number of risk factors, discarding those that do not contribute significantly to the risk of the portfolio. Simulations based on the full set of variables would be inordinately time-consuming. The art of simulation is to design parsimonious experiments that represent the breadth of movements in risk factors.

This can be done by an economic analysis of the risk factors and portfolio strategies, as done in Part Three of this handbook. Alternatively, the risk manager can perform a statistical decomposition of the covariance matrix. A widely-used method for this is the principal-component analysis (PCA), which finds linear combinations of the risk factors that have maximal explanatory power. This type of analysis, which is as much an art as it is a science, can be used to reduce the dimensionality of the risk space.

EXAMPLE 4.10: FRM EXAM 2007—QUESTION 28

Let N be an $n \times 1$ vector of independent draws from a standard normal distribution, and let V be a covariance matrix of market time-series data. Then, if L is a diagonal matrix of the eigenvalues of V, E is a matrix of the eigenvectors of V, and $C'C$ is the Cholesky factorization of V, which of the following would generate a normally distributed random vector with mean zero and covariance matrix V to be used in a Monte Carlo simulation?

a. $NC'CN$

b. NC'

c. $E'LE$

d. Cannot be determined from data given

EXAMPLE 4.11: FRM EXAM 2006—QUESTION 82

Consider a stock that pays no dividends, has a volatility of 25% pa and an expected return of 13% pa. The current stock price is $S_0 = \$30$. This implies the model $S_{t+1} = S_t(1 + 0.13\Delta t + 0.25\sqrt{\Delta t}\,\epsilon)$, where ϵ is a standard normal random variable. To implement this simulation, you generate a path of the stock price by starting at $t = 0$, generating a sample for ϵ, updating the stock price according to the model, incrementing t by 1 and repeating this process until the end of the horizon is reached. Which of the following strategies for generating a sample for ϵ will implement this simulation properly?

a. Generate a sample for ϵ by using the inverse of the standard normal cumulative distribution of a sample value drawn from a uniform distribution between 0 and 1.

b. Generate a sample for ϵ by sampling from a normal distribution with mean 0.13 and standard deviation 0.25.

c. Generate a sample for ϵ by using the inverse of the standard normal cumulative distribution of a sample value drawn from a uniform distribution between 0 and 1. Use Cholesky decomposition to correlate this sample with the sample from the previous time interval.

d. Generate a sample for ϵ by sampling from a normal distribution with mean 0.13 and standard deviation 0.25. Use Cholesky decomposition to correlate this sample with the sample from the previous time interval.

EXAMPLE 4.12: FRM EXAM 2006—QUESTION 83

Continuing with the previous question, you have implemented the simulation process discussed above using a time interval $\Delta t = 0.001$, and you are analyzing the following stock price path generated by your implementation.

t	S_{t-1}	ϵ	ΔS
0	30.00	0.0930	0.03
1	30.03	0.8493	0.21
2	30.23	0.9617	0.23
3	30.47	0.2460	0.06
4	30.53	0.4769	0.12
5	30.65	0.7141	0.18

Given this sample, which of the following simulation steps most likely contains an error.

a. Calculation to update the stock price

b. Generation of random sample value for ϵ

c. Calculation of the change in stock price during each period

d. None of the above

4.4 IMPORTANT FORMULAS

The Wiener process: $\Delta z \sim N(0, \Delta t)$

The generalized Wiener process: $\Delta x = a\Delta t + b\Delta z$

The Ito process: $\Delta x = a(x, t)\Delta t + b(x, t)\Delta z$

The geometric Brownian motion: $\Delta S = \mu S \Delta t + \sigma S \Delta z$

One-factor equilibrium model for yields: $\Delta r_t = \kappa(\theta - r_t)\Delta t + \sigma r_t^{\gamma} \Delta z_t$

Vasicek model, $\gamma = 0$

Lognormal model, $\gamma = 1$

CIR model, $\gamma = 0.5$

No-arbitrage models:

Ho and Lee model, $\Delta r_t = \theta(t)\Delta t + \sigma \Delta z_t$

Hull and White model, $\Delta r_t = [\theta(t) - ar_t]\Delta t + \sigma \Delta z_t$

Heath, Jarrow, and Morton model, $\Delta r_t = [\theta(t) - ar_t]\Delta t + \sigma(t)\Delta z_t$

Binomial trees: $u = e^{\sigma\sqrt{\Delta t}}$, $d = (1/u)$, $p = \frac{e^{\mu\Delta t} - d}{u - d}$

Cholesky factorization: $R = TT'$, $\epsilon = T\eta$

4.5 ANSWERS TO CHAPTER EXAMPLES

Example 4.1: FRM Exam 2003—Question 40

b. Both dS/S or $d\ln(S)$ are normally distributed. As a result, S is lognormally distributed. The only incorrect answer is I.

Example 4.2: FRM Exam 2002—Question 126

a. All the statements are correct except a., which is too strong. The expected price is higher than today's price but certainly not the price in all states of the world.

Example 4.3: FRM Exam 1999—Question 25

b. This model postulates only one source of risk in the fixed-income market. This is a single-factor term-structure model.

Example 4.4: FRM Exam 1999—Question 26

c. This represents the expected return with mean reversion.

Example 4.5: FRM Exam 2000—Question 118

a. These are no-arbitrage models of the term structure, implemented as either one-factor or two-factor models.

Example 4.6: FRM Exam 2000—Question 119

b. Both the Vasicek and CIR models are one-factor equilibrium models with mean reversion. The Hull-White model is a no-arbitrage model with mean reversion. The Ho and Lee model is an early no-arbitrage model without mean-reversion.

Example 4.7: FRM Exam 2005—Question 67

b. MC simulations do account for options. The first step is to simulate the process of the risk factor. The second step prices the option, which properly accounts for non-linearity.

Example 4.8: FRM Exam 2007—Question 66

d. Short option positions have long left tails, which makes it more difficult to estimate a left-tailed quantile precisely. Accuracy with independent draws increases with the square root of K. Thus increasing the number of replications should shrink the standard error, so answer b. is incorrect.

Example 4.9: Sampling Variation

b. Sampling variability (or imprecision) increases with (1) fewer observations and (2) greater confidence levels. To show (1), we can refer to the formula for the precision of the sample mean, which varies inversely with the square root of the number of data points. A similar reasoning applies to (2). A greater confidence level involves fewer observations in the left tails, from which VAR is computed.

Example 4.10: FRM Exam 2007—Question 28

b. In the notation of the text, N is the vector of i.i.d. random variables η and $C'C = TT'$. The transformed variable is $T\eta$, or $C'N$, or its transpose.

Example 4.11: FRM Exam 2006—Question 82

a. The variable ϵ should have a standard normal distribution, i.e., with mean zero and unit standard deviation. Answer b. is incorrect because ϵ is transformed afterward to the desired mean and standard deviation. The Cholesky decomposition is not applied here because the sequence of random variables has no serial correlation.

Example 4.12: FRM Exam 2006—Question 83

b. The random variable ϵ should have a standard normal distribution, which means that it should have negative as well as positive values, which should average close to zero. This is not the case here. This is probably a uniform variable instead.
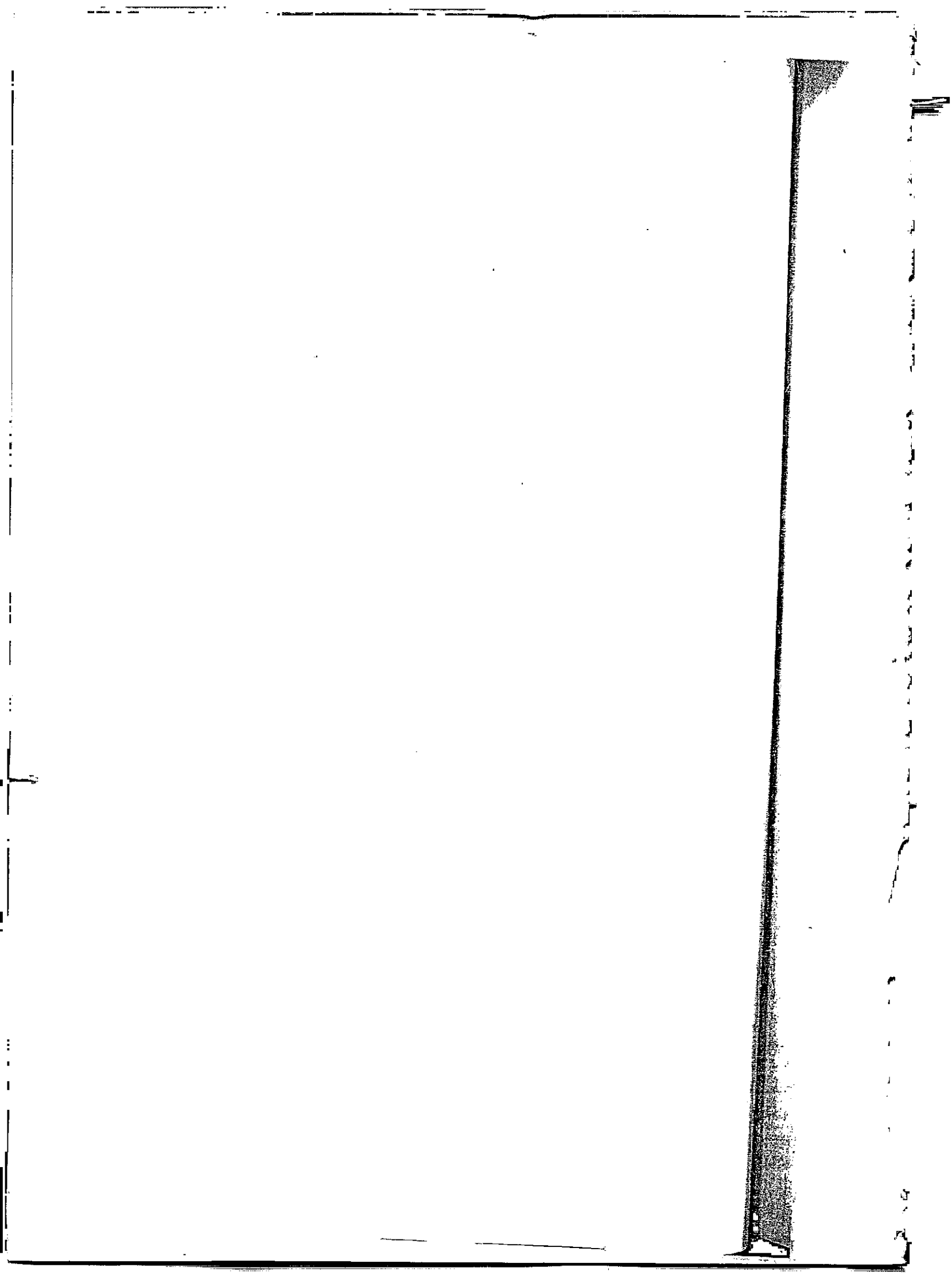
PART
TWO

Capital Markets
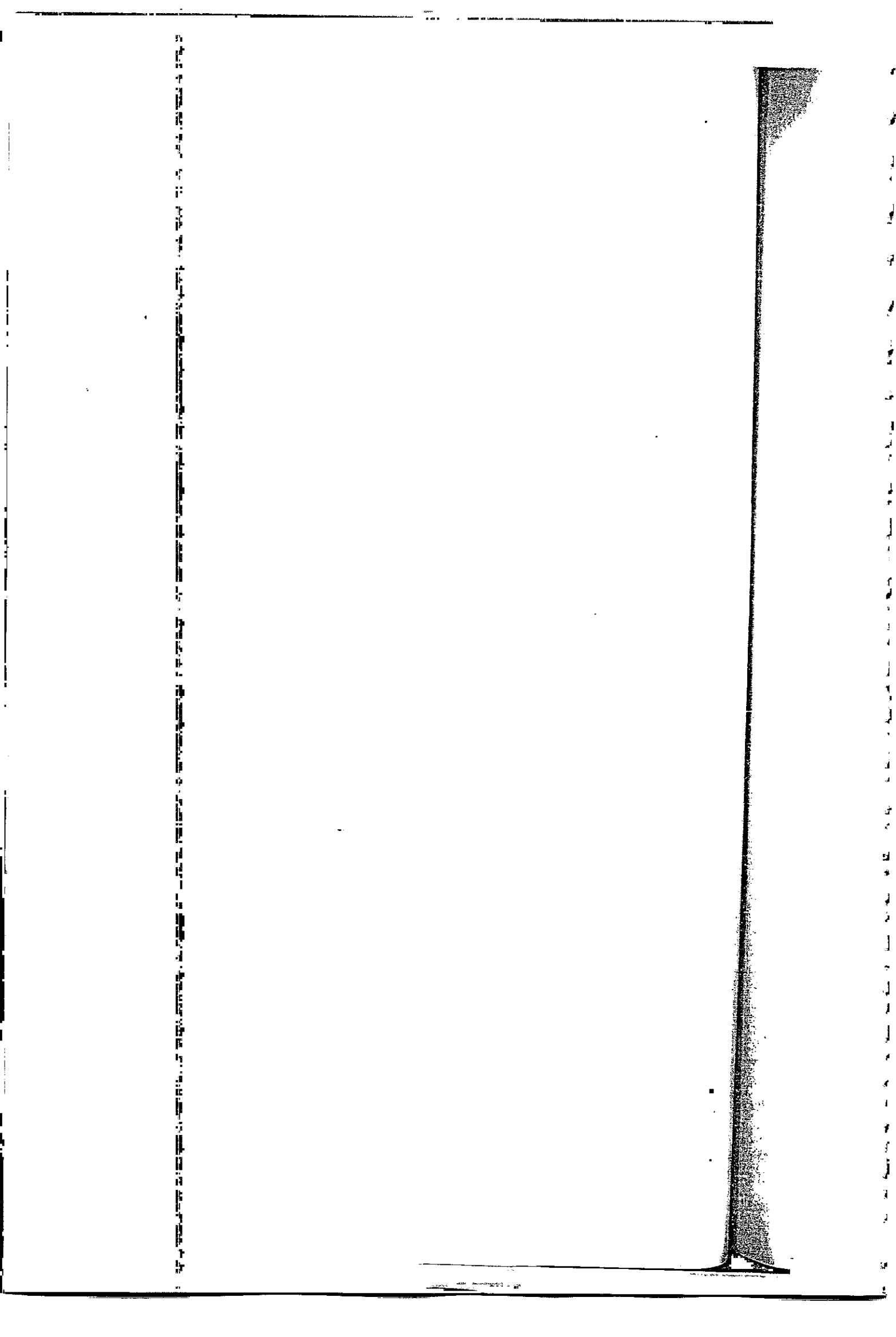
Introduction to Derivatives

This chapter provides an overview of derivative instruments. Derivatives are financial contracts traded in private over-the-counter (OTC) markets, or on organized exchanges. As the term implies, derivatives derive their value from some underlying index, typically the price of an asset. Depending on the type of relationship, they can be broadly classified into two categories: linear and nonlinear instruments.

To the first category belong forward contracts, futures, and swaps. Their value is a linear function of the underlying index. These are *obligations* to exchange payments according to a specified schedule. Forward contracts are relatively simple to evaluate and price. So are futures, which are traded on exchanges. Swaps are more complex but generally can be reduced to portfolios of forward contracts. To the second category belong options, which are traded both OTC and on exchanges. Their value is a nonlinear function of the underlying index. These will be covered in the next chapter.

This chapter describes the general characteristics as well as the pricing of linear derivatives. Pricing is the first step toward risk measurement. The second step consists of combining the valuation formula with the distribution of underlying risk factors to derive the distribution of contract values. This will be done later, in the market risk section.

Section 5.1 provides an overview of the size of the derivatives markets. Section 5.2 then presents the valuation and pricing of forwards. Sections 5.3 and 5.4 introduce futures and swap contracts, respectively.

5.1 OVERVIEW OF DERIVATIVES MARKETS

A derivative instrument can be generally defined as a private contract whose value derives from some underlying asset price, reference rate or index—such as a stock, bond, currency, or a commodity. In addition, the contract must also specify a principal, or notional amount, which is defined in terms of currency, shares, bushels, or some other unit. Movements in the value of the derivative depend on the notional and the underlying price or index.

In contrast with securities, such as stocks and bonds, which are issued to raise capital, derivatives are contracts, or private agreements between two parties. Thus

111

the sum of gains and losses on derivatives contracts must be zero. For any gain made by one party, the other party must have suffered a loss of equal magnitude.

. At the broadest level, derivatives markets can be classified by the underlying instrument, as well as by the type of trading. Table 5.1 describes the size and growth of the global derivatives markets. As of 2007, the total notional amounts add up to $677 trillion, of which $596 trillion is on OTC markets and $81 trillion on organized exchanges. These markets have grown exponentially, from $56 trillion in 1995.

The table shows that interest rate contracts are the most widespread type of derivatives, especially swaps. On the OTC market, currency contracts are also widely used, especially outright forwards and forex swaps, which are a combination of spot and short-term forward transactions. Among exchange-traded instruments, interest rate futures and options are the most common.

The magnitude of the notional amount of $677 trillion is difficult to grasp. This number is several times the world gross domestic product (GDP), which amounted to approximately $48 trillion in 2006. It is also greater than the total

TABLE 5.1 Global Derivatives Markets, 1995–2007 (Billions of U.S. Dollars)

	Notional Amounts	
	March 1995	Dec. 2007
OTC Instruments	47,530	596,004
Interest rate contracts	26,645	393,138
Forwards (FRAs)	4,597	26,599
Swaps	18,283	309,588
Options	3,548	56,951
Foreign exchange contracts	13,095	56,238
Forwards and forex swaps	8,699	29,144
Swaps	1,957	14,347
Options	2,379	12,748
Equity-linked contracts	579	8,509
Forwards and swaps	52	2,233
Options	527	6,276
Commodity contracts	318	9,000
Credit default swaps	0	57,894
Others	6,893	71,225
Exchange-Traded Instruments	8,838	80,578
Interest rate contracts	8,380	71,052
Futures	5,757	26,770
Options	2,623	44,282
Foreign exchange contracts	88	292
Futures	33	159
Options	55	133
Stock-index contracts	370	9,234
Futures	128	1,132
Options	242	8,102
Total	55,910	676,582

Source: Bank for International Settlements

outstanding value of stocks, which was $54 trillion and of debt securities, which was $69 trillion at that time.

Notional amounts give an indication of equivalent positions in cash markets. For example, a long futures contract on a stock index with a notional of $1 million is equivalent to a cash position in the stock market of the same magnitude.

Notional amounts, however, do not give much information about the risks of the positions. The current (positive) market value of OTC derivatives contracts, for instance, is estimated at $15 trillion. This is less than 3% of the notional. More generally, the risk of these derivatives is best measured by the potential change in mark-to-market values over the horizon. In other words, by a value-at-risk measure.

5.2 FORWARD CONTRACTS

5.2.1 Definition

The most common transactions in financial instruments are spot transactions, that is, for physical delivery as soon as practical (perhaps in two business days or in a week). Historically, grain farmers went to a centralized location to meet buyers for their product. As markets developed, the farmers realized that it would be beneficial to trade for delivery at some future date. This allowed them to hedge out price fluctuations for the sale of their anticipated production.

This gave rise to forward contracts, which are private agreements to exchange a given asset against cash (or sometimes another asset) at a fixed point in the future. The terms of the contract are the quantity (number of units or shares), date, and price at which the exchange will be done.

A position which implies buying the asset is said to be long. A position to sell is said to be short. Any gain to one party must be a loss to the other.

These instruments represent contractual obligations, as the exchange must occur whatever happens to the intervening price, unless default occurs. Unlike an option contract, there is no choice in taking delivery or not.

To avoid the possibility of losses, the farmer could enter a forward sale of grain for dollars. By so doing, he locks up a price now for delivery in the future. We then say that the farmer is hedged against movements in the price.

We use the notations

$$t = \text{current time}$$
$$T = \text{time of delivery}$$
$$\tau = T - t = \text{time to maturity}$$
$$S_t = \text{current spot price of the asset in dollars}$$
$$F_t(T) = \text{current forward price of the asset for delivery at } T$$
$$\text{(also written as } F_t \text{ or } F \text{ to avoid clutter)}$$
$$V_t = \text{current value of contract}$$
$$r = \text{current domestic risk-free rate for delivery at } T$$
$$n = \text{quantity, or number of units in contract}$$

The face amount, or principal value of the contract is defined as the amount nF to pay at maturity, like a bond. This is also called the notional amount. We will assume that interest rates are continuously compounded so that the present value of a dollar paid at expiration is $PV(\$1) = e^{-r\tau}$.

Say that the initial forward price is $F_t = \$100$. A speculator agrees to buy $n = 500$ units for F_t at T. At expiration, the payoff on the forward contract is determined as follows:

1. The speculator pays $nF = \$50,000$ in cash and receives 500 units of the underlying.
2. The speculator could then sell the underlying at the prevailing spot price S_T, for a profit of $n(S_T - F)$. For example, if the spot price is at $S_T = \$120$, the profit is $500 \times (\$120 - \$100) = \$10,000$. This is also the mark-to-market value of the contract at expiration.

In summary, the value of the forward contract at expiration, for one unit of the underlying asset is

$$V_T = S_T - F \tag{5.1}$$

Here, the value of the contract at expiration is derived from the purchase and physical delivery of the underlying asset. There is a payment of cash in exchange for the actual asset.

Another mode of settlement is cash settlement. This involves simply measuring the market value of the asset upon maturity, S_T, and agreeing for the "long" to receive $nV_T = n(S_T - F)$. This amount can be positive or negative, involving a profit or loss.

Figures 5.1 and 5.2 present the payoff patterns on long and short positions in a forward contract, respectively. It is important to note that the payoffs are *linear*
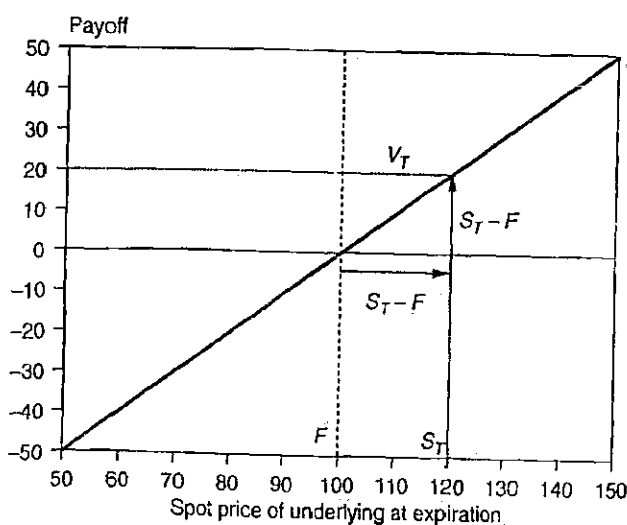
FIGURE 5.1   Payoff of Profits on Long Forward Contract
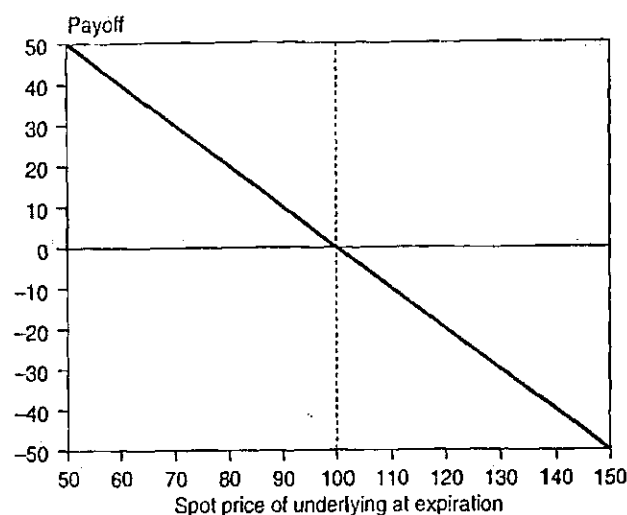
Payoff

FIGURE 5.2 Payoff of Profits on Short Forward Contract

in the underlying spot price. Also, the positions in the two figures are symmetrical around the horizontal axis. For a given spot price, the sum of the profit or loss for the long and the short is zero, because these are private contracts.

5.2.2 Valuing Forward Contracts

When evaluating forward contracts, two important questions arise. First, how is the current forward price F_t determined? Second, what is the current value V_t of an outstanding forward contract?

Initially, we assume that the underlying asset pays no income. This will be generalized in the next section. We also assume no transaction costs, that is, zero bid–ask spread on spot and forward quotations as well as the ability to lend and borrow at the same risk-free rate.

Generally, forward contracts are established so that their initial value is zero. This is achieved by setting the forward price F_t appropriately by a no-arbitrage relationship between the cash and forward markets. No-arbitrage is a situation where positions with the same payoffs have the same price. This rules out situations where arbitrage profits can exist. Arbitrage is a zero-risk, zero-net investment strategy that still generates profits.

Consider these strategies:

- Buy one share/unit of the underlying asset at the spot price S_t and hold to time T.
- Enter a forward contract to buy one share/unit of same underlying asset at the forward price F_t. In order to have sufficient funds at maturity to pay F_t, we invest the present value of F_t in an interest-bearing account. This is the present value $F_t e^{-r\tau}$. The forward price F_t is set so that the initial cost of the forward contract, V_t, is zero.

The two portfolios are economically equivalent because they will be identical at maturity. Each will contain one share of the asset. Hence their up-front cost must be the same. To avoid arbitrage, we must have:

$$S_t = F_t e^{-r\tau} \qquad (5.2)$$

This equation defines the fair forward price F_t such that the initial value of the contract is zero. More generally, the term multiplying F_t is the present value factor for maturity τ, or PV($1). For instance, assuming $S_t = \$100$, $r = 5\%$, $\tau = 1$, we have $F_t = S_t e^{r\tau} = \$100 \times \exp(0.05 \times 1) = \$105.13$.

We see that the forward rate is higher than the spot rate. This reflects the fact that there is no down payment to enter the forward contract, unlike for the cash position. As a result, the forward price must be higher than the spot price to reflect the time value of money.

Abstracting from transaction costs, any deviation creates an arbitrage opportunity. This can be taken advantage of by buying the cheap asset and selling the expensive one. Assume for instance that $F = \$110$. We determined that the fair value is $S_t e^{r\tau} = \$105.13$, based on the cash price. We apply the principle of buying low at $105.13 and selling high at $110. We can lock in a sure profit by:

1. Buying now the asset spot at $100
2. Selling now the asset forward at $110

This can be done by borrowing the $100 to buy the asset now. At expiration, we will owe principal plus interest, or $105.13 but receive $110, for a profit of $4.87. This would be a blatant arbitrage opportunity, or "money machine."

Now consider a mispricing where $F = \$102$. We apply the principle of buying low at $102 and selling high at $105.13. We can lock in a sure profit by:

1. Short-selling now the asset spot at $100
2. Buying now the asset forward at $102

From the short sale, we invest the cash, which will grow to $105.13. At expiration, we will have to deliver the stock but this will be acquired through the forward purchase. We pay $102 for this and are left with a profit of $3.13.

This transaction involves the short-sale of the asset, which is more involved than an outright purchase. When purchasing, we pay $100 and receive one share of the asset. When short-selling, we borrow one share of the asset and promise to give it back at a future date; in the meantime, we sell it at $100.[1]

---

[1] In practice, we may not get full access to the proceeds of the sale when it involves individual stocks. The broker will typically only allow us to withdraw 50% of the cash. The rest is kept as a performance bond should the transaction lose money.

5.2.3 Valuing an Off-Market Forward Contract

We can use the same reasoning to evaluate an outstanding forward contract, with a locked-in delivery price of K. In general, such a contract will have non zero value because K differs from the prevailing forward rate. Such a contract is said to be **off-market**.

Consider these strategies:

- Buy one share/unit of the underlying asset at the spot price S_t and hold it until time T.
- Enter a forward contract to buy one share/unit of the same underlying asset at the price K; in order to have sufficient funds at maturity to pay K, we invest the present value of K in an interest-bearing account. This present value is also $Ke^{-r\tau}$. In addition, we have to pay the market value of the forward contract, or V_t.

The up-front cost of the two portfolios must be identical. Hence, we must have $V_t + Ke^{-r\tau} = S_t$, or

$$V_t = S_t - Ke^{-r\tau} \tag{5.3}$$

which defines the market value of an outstanding long position.[2] This gains value when the underlying S increases in value. A short position would have the reverse sign. Later, we will extend this relationship to the measurement of risk by considering the distribution of the underlying risk factors, S_t and r.

For instance, assume we still hold the previous forward contract with $F_t = \$105.13$ and after one month the spot price moves to $S_t = \$110$. The fixed rate is $K = \$105.13$ throughout the life of the contract. The interest has not changed at $r = 5\%$, but the maturity is now shorter by one month, $\tau = 11/12$. The new value of the contract is $V_t = S_t - Ke^{-r\tau} = \$110 - \$105.13\exp(-0.05 \times 11/12) = \$110 - \$100.42 = \$9.58$. The contract is now more valuable than before because the spot price has moved up.

5.2.4 Valuing Forward Contracts with Income Payments

We previously considered a situation where the asset produces no income payment. In practice, the asset may be

- A stock that pays a regular dividend
- A bond that pays a regular coupon
- A stock index that pays a dividend stream approximated by a continuous yield
- A foreign currency that pays a foreign-currency denominated interest rate

Whichever income is paid on the asset, we can usefully classify the payment into discrete, that is, fixed dollar amounts at regular points in time, or on a continuous

---

[2] Note that V_t is not the same as the forward price F_t. The former is the value of the contract; the latter refers to a specification of the contract.

basis, that is, accrued in proportion to the time the asset is held. We must assume that the income payment is fixed or is certain. More generally, a storage cost is equivalent to a negative dividend.

We use these definitions:

$$D = \text{discrete (dollar) dividend or coupon payment}$$
$$r_t^*(T) = \text{foreign risk-free rate for delivery at } T$$
$$q_t(T) = \text{dividend yield}$$

Whether the payment is a dividend or a foreign interest rate, the principle is the same. We can afford to invest less in the asset up-front to get one unit at expiration. This is because the income payment can be reinvested into the asset. Alternatively, we can borrow against the value of the income payment to increase our holding of the asset.

It is also important to note that all prices (S, F) are measured in the domestic currency. For example S could be expressed in terms of the U.S. dollar price of the euro, in which case r is the U.S. interest rate and r^* is the euro interest rate. Conversely, if S is the Japanese yen price of the U.S. dollar, r will represent the Japanese interest rate, and r^* the U.S. interest rate.

Continuing our example, consider a stock priced at $100 that pays a dividend of $D = \$1$ in three months. The present value of this payment discounted over three months is $De^{-r\tau} = \$1 \exp(-0.05 \times 3/12) = \$0.99$. We only need to put up $S_t - PV(D) = \$100.00 - 0.99 = \$99.01$ to get one share in one year. Put differently, we buy 0.9901 fractional shares now and borrow against the (sure) dividend payment of $1 to buy an additional 0.0099 fractional share, for a total of one share.

The pricing formula in Equation (5.2) is extended to

$$F_t e^{-r\tau} = S_t - PV(D) \tag{5.4}$$

where PV(D) is the present value of the dividend/coupon payments. If there is more than one payment, PV(D) represents the sum of the present values of each individual payment, discounted at the appropriate risk-free rate. With storage costs, we need to *add* the present value of storage costs PV(C) to the right side of Equation (5.4).

The approach is similar for an asset that pays a continuous income, defined per unit time instead of discrete amounts. Holding a foreign currency, for instance, should be done through an interest-bearing account paying interest that accrues with time. Over the horizon τ, we can afford to invest less up-front, $S_t e^{-r^*\tau}$ in order to receive one unit at maturity. The right-hand side of Equation (5.4) is now

$$F_t e^{-r\tau} = S_t e^{-r^*\tau} \tag{5.5}$$

Hence, the forward price should be

$$F_t = S_t e^{-r^*\tau}/e^{-r\tau} \tag{5.6}$$

If instead interest rates are annually compounded, this gives

$$F_t = S_t(1 + r)^\tau/(1 + r^*)^\tau \tag{5.7}$$

Equation (5.6) can be also written in terms of the forward premium or discount, which is

$$\frac{(F_t - S_t)}{S_t} = e^{-r^*\tau}/e^{-r\tau} = \exp(r - r^*)\tau \approx (r - r^*)\tau \tag{5.8}$$

If $r^* < r$, we have $F_t > S_t$ and the asset trades at a forward premium. Conversely, if $r^* > r$, $F_t < S_t$ and the asset trades at a forward discount. Thus, the forward price is higher or lower than the spot price, depending on whether the yield on the asset is lower than or higher than the domestic risk-free interest rate.

Equation (5.6) is also known as interest rate parity when dealing with currencies. Also note that both the spot and forward prices must be expressed in dollars per unit of the foreign currency when the domestic currency interest rate is r. This is the case, for example, for the dollar euro or dollar/pound exchange rate. If, on the other hand, the exchange rate is expressed in foreign currency per dollar, then r must be the rate on the foreign currency. For the yen/dollar rate, for example, S is in yen per dollar, r is the yen interest rate, and r^* is the dollar interest rate.

> **KEY CONCEPT**
>
> The forward price differs from the spot price to reflect the time value of money and the income yield on the underlying asset. It is higher than the spot price if the yield on the asset is lower than the domestic risk-free interest rate, and vice versa.

With income payments, the value of an outstanding forward contract is

$$V_t = S_t e^{-r^*\tau} - K e^{-r\tau} \tag{5.9}$$

If F_t is the new, current forward price, we can also write

$$V_t = F_t e^{-r\tau} - K e^{-r\tau} = (F_t - K)e^{-r\tau} \tag{5.10}$$

This provides a useful alternative formula for the valuation of a forward contract. The intuition here is that we could liquidate the outstanding forward contract by entering a reverse position at the current forward rate. The payoff at expiration is $(F - K)$, which, discounted back to the present, gives Equation (5.10).

> **KEY CONCEPT**
>
> The current value of an outstanding forward contract can be found by entering an offsetting forward position and discounting the net cash flow at expiration.

EXAMPLE 5.1: FRM EXAM 2005—QUESTION 2

What is the no-arbitrage price of a forward contract if the time to expiration is three months, the underlying asset is worth $1,000, the continuously compounded annualized risk-free rate is 6%, and storage costs are expressed in terms of a continuous annualized yield of 3%?

a. USD 1,008
b. USD 972
c. USD 1,023
d. USD 1,039

EXAMPLE 5.2: FRM EXAM 2005—QUESTION 16

Suppose that U.S. interest rates rise from 3% to 4% this year. The spot exchange rate quotes at 112.5 JPY/USD and the forward rate for a one-year contract is at 110.5. What is the Japanese interest rate?

a. 1.81%
b. 2.15%
c. 3.84%
d. 5.88%

EXAMPLE 5.3: FRM EXAM 2002—QUESTION 56

Consider a forward contract on a stock market index. Identify the false statement. Everything else being constant,

a. The forward price depends directly upon the level of the stock market index.
b. The forward price will fall if underlying stocks increase the level of dividend payments over the life of the contract.
c. The forward price will rise if time to maturity is increased.
d. The forward price will fall if the interest rate is raised.

EXAMPLE 5.4: FRM EXAM 2007—QUESTION 119

A three-month futures contract on an equity index is currently priced at USD 1,000. The underlying index stocks are valued at USD 990 and pay dividends at a continuously compounded rate of 2%. The current continuously compounded risk-free rate is 4%. The potential arbitrage profit per contract, given this set of data, is closest to

     a. USD 10.00
     b. USD 7.50
     c. USD 5.00
     d. USD 1.50

5.3 FUTURES CONTRACTS

5.3.1 Definitions of Futures

Forward contracts allow users to take positions that are economically equivalent to those in the underlying cash markets. Unlike cash markets, however, they do not involve substantial up-front payments. Thus, forward contracts can be interpreted as having *leverage*. Leverage is efficient as it makes our money work harder.

Leverage creates credit risk for the counterparty, however. For a cash trade, there is no leverage. When a speculator buys a stock at the price of $100, the counterparty receives the cash and has no credit risk. Instead, when a speculator enters a forward contract to buy an asset at the price of $105, there is no up-front payment. In effect, the speculator borrows from the counterparty to invest in the asset. There is a risk of default should the value of the contract to the speculator fall sufficiently. In response, futures contracts have been structured so as to minimize credit risk for all counterparties. Otherwise, from a market risk standpoint, futures contracts are basically identical to forward contracts.

Futures contracts are standardized, negotiable, and exchange-traded contracts to buy or sell an underlying asset. They differ from forward contracts as follows.

- Trading on organized exchanges. In contrast to forwards, which are OTC contracts tailored to customers' needs, futures are traded on organized exchanges (either with a physical location or electronic).
- Standardization. Futures contracts are offered with a limited choice of expiration dates. They trade in fixed contract sizes. This standardization ensures an active secondary market for many futures contracts, which can be easily traded, purchased, or resold. In other words, most futures contracts have good liquidity. The trade-off is that futures are less precisely suited to the need of some hedgers, which creates basis risk (to be defined later).

- Clearinghouse. Futures contracts are also standardized in terms of the counterparty. After each transaction is confirmed, the clearinghouse basically interposes itself between the buyer and the seller, ensuring the performance of the contract. Thus, unlike forward contracts, counterparties do not have to worry about the credit risk of the other side of the trade. Instead, the credit risk is that of the clearinghouse (or the broker), which is generally excellent.

- Marking-to-market. As the clearinghouse now has to deal with the credit risk of the two original counterparties, it has to monitor credit risk closely. This is achieved by daily marking-to-market, which involves settlement of the gains and losses on the contract every day. This will avoid the accumulation of large losses over time, potentially leading to an expensive default.

- Margins. Although daily settlement accounts for past losses, it does not provide a buffer against future losses. This is the goal of margins, which represent upfront posting of collateral that can be seized should the other party default. If the equity in the account falls below the maintenance margin, the customer is required to provide additional funds to cover the initial margin. The level of margin depends on the instrument and the type of position; in general, less volatile instruments or hedged positions require lower margins.

Example: Margins for a Futures Contract

Consider a futures contract on 1,000 units of an asset worth $100. A long futures position is economically equivalent to holding $100,000 worth of the asset directly. To enter the futures position, a speculator has to post only $5,000 in margin, for example. This amount is placed in an equity account with the broker. The next day, the futures price moves down by $3, leading to a loss of $3,000 for the speculator. The profit or loss is added to the equity account, bringing it down to $5,000−$3,000 = $2,000. The speculator would then receive a margin call from the broker, asking to have an additional $3,000 of capital posted to the account. If he or she fails to meet the margin call, the broker has the right to liquidate the position.

Since futures trading is centralized on an exchange, it is easy to collect and report aggregate trading data. Volume is the number of contracts traded during the day, which is a flow item. Open interest represents the outstanding number of contracts at the close of the day, which is a stock item.

5.3.2 Valuing Futures Contracts

Valuation principles for futures contracts are very similar to those for forward contracts. The main difference between the two types of contracts is that any profit or loss accrues *during* the life of the futures contract instead of all at once, at expiration.

When interest rates are assumed constant or deterministic, forward and futures prices must be equal. With stochastic interest rates, there may be a small difference, depending on the correlation between the value of the asset and interest rates.

If the correlation is zero, then it makes no difference whether payments are received earlier or later. The futures price must be the same as the forward price. In contrast, consider a contract whose price is positively correlated with the interest rate. If the value of the contract goes up, it is more likely that interest rates will go up as well. This implies that profits can be withdrawn and reinvested at a higher rate. Relative to forward contracts, this marking-to-market feature is beneficial to a long futures position. As a result, the futures price must be higher in equilibrium.

In practice, this effect is only observable for interest-rate futures contracts, whose value is *negatively* correlated with interest rates. Because this feature is unattractive for the long position, the futures price must be *lower* than the forward price. Chapter 8 will explain how to compute the adjustment, called the convexity effect.

EXAMPLE 5.5: FRM EXAM 2004—QUESTION 38

An investor enters into a short position in a gold futures contract at USD 294.20. Each futures contract controls 100 troy ounces. The initial margin is USD 3,200, and the maintenance margin is USD 2,900. At the end of the first day, the futures price drops to USD 286.6. Which of the following is the amount of the variation margin at the end of the first day?

a. 0
b. USD 34
c. USD 334
d. USD 760

EXAMPLE 5.6: FRM EXAM 2004—QUESTION 66

Which one of the following statements is *incorrect* regarding the margining of exchange-traded futures contracts?

a. Day trades and spread transactions require lower margin levels.
b. If an investor fails to deposit variation margin in a timely manner the positions may be liquidated by the carrying broker.
c. Initial margin is the amount of money that must be deposited when a futures contract is opened.
d. A margin call will be issued only if the investor's margin account balance becomes negative.

5.4 SWAP CONTRACTS

Swap contracts are OTC agreements to exchange a *series* of cash flows according to prespecified terms. The underlying asset can be an interest rate, an exchange rate, an equity, a commodity price, or any other index. Typically, swaps are established for longer periods than forwards and futures.

For example, a 10-year currency swap could involve an agreement to exchange every year 5 million dollars against 3 million pounds over the next 10 years, in addition to a principal amount of 100 million dollars against 50 million pounds at expiration. The principal is also called notional principal.

Another example is that of a five-year interest rate swap in which one party pays 8% of the principal amount of 100 million dollars in exchange for receiving an interest payment indexed to a floating interest rate. In this case, since both payments are the same amount in the same currency, there is no need to exchange principal at maturity.

Swaps can be viewed as a portfolio of forward contracts. They can be priced using valuation formulas for forwards. Our currency swap, for instance, can be viewed as a combination of 10 forward contracts with various face values, maturity dates, and rates of exchange. We will give detailed examples in later chapters.

5.5 IMPORTANT FORMULAS

Forward price, no income on the asset: $F_t e^{-r\tau} = S_t$

Forward price, income on the asset:

discrete dividend, $F_t e^{-r\tau} = S_t - PV(D)$,

continuous dividend, $F_t e^{-r\tau} = S_t e^{-r^*\tau}$

Forward premium or discount: $\frac{(F_t - S_t)}{S_t} \approx (r - r^*)\tau$

Valuation of outstanding forward contract:

$V_t = S_t e^{-r^*\tau} - K e^{-r\tau} = F_t e^{-r\tau} - K e^{-r\tau} = (F_t - K)e^{-r\tau}$

5.6 ANSWERS TO CHAPTER EXAMPLES

Example 5.1: FRM Exam 2005—Question 2

c. The asset has a carrying cost, which is the equivalent to a negative dividend yield. Using Equation (5.6), with $S = 1000, r = 0.06, r^* = -0.03$, and $T = 3/12$, we have $F = 1000 \times \exp(+0.03 \times 1/12)/\exp(-0.06 \times 1/12) = 1,022.8$.

Example 5.2: FRM Exam 2005—Question 16

b. As is the convention in the currency markets, the exchange rate is defined as the yen price of the dollar, which is the foreign currency. The foreign currency interest rate is the latest U.S. dollar rate, or 4%. Assuming discrete compounding, the

pricing formula for forward contracts is $F(JPY/USD)/(1 + rT) = S(JPY/USD)/(1 + r^*T)$. Therefore, $(1 + rT) = (F/S)(1 + r^*T) = (110.5/112.5)(1.04) = 1.0215$, and $r = 2.15\%$. Using continuous compounding gives a similar result. Another approach would consider the forward discount on the dollar, which is $(F - S)/S = -1.8\%$. Thus, the dollar is 1.8% cheaper forward than spot, which must mean that the Japanese interest rate must be approximately 1.8% lower than the U.S. interest rate.

Example 5.3: FRM Exam 2002—Question 56

d. Defining the dividend yield as q, the forward price depends on the cash price according to $F \exp(-rT) = S \exp(-qT)$. This can also be written as $F = S \exp[+(r - q)T]$. Generally, $r > q$. Statement a. is correct: F depends directly on S. Statement b. is also correct, as higher q decreases the term between brackets and hence F. Statement c. is correct because the term $r - q$ is positive, leading to a larger term in brackets as the time to maturity T increases. Statement d. is false, as increasing r makes the forward contract more attractive, or increases F.

Example 5.4: FRM Exam 2007—Question 119

c. The fair value of the futures contract is given by $F = S \exp(-r^*T)/\exp(-rT) = 990 \exp(-0.02 \times 3/12) \exp(-0.04 \times 3/12) = 994.96$. Hence, the actual futures price is too high by $(1,000 - 995) = 5$.

Example 5.5: FRM Exam 2004—Question 38

a. This is a tricky question. Because the investor is short and the price fell, the position creates a profit and there is no variation margin.

On the other hand, for the long the loss is $760, which would bring the equity to $3,200 - 760 = 2,440$. Because this is below the maintenance margin of 2,900, an additional payment of $760 is required to bring back the equity to the initial margin.

Example 5.6: FRM Exam 2004—Question 66

d. All the statements are correct, except d. If the margin account balance falls below the maintenance margin (not zero), a margin call will be issued.
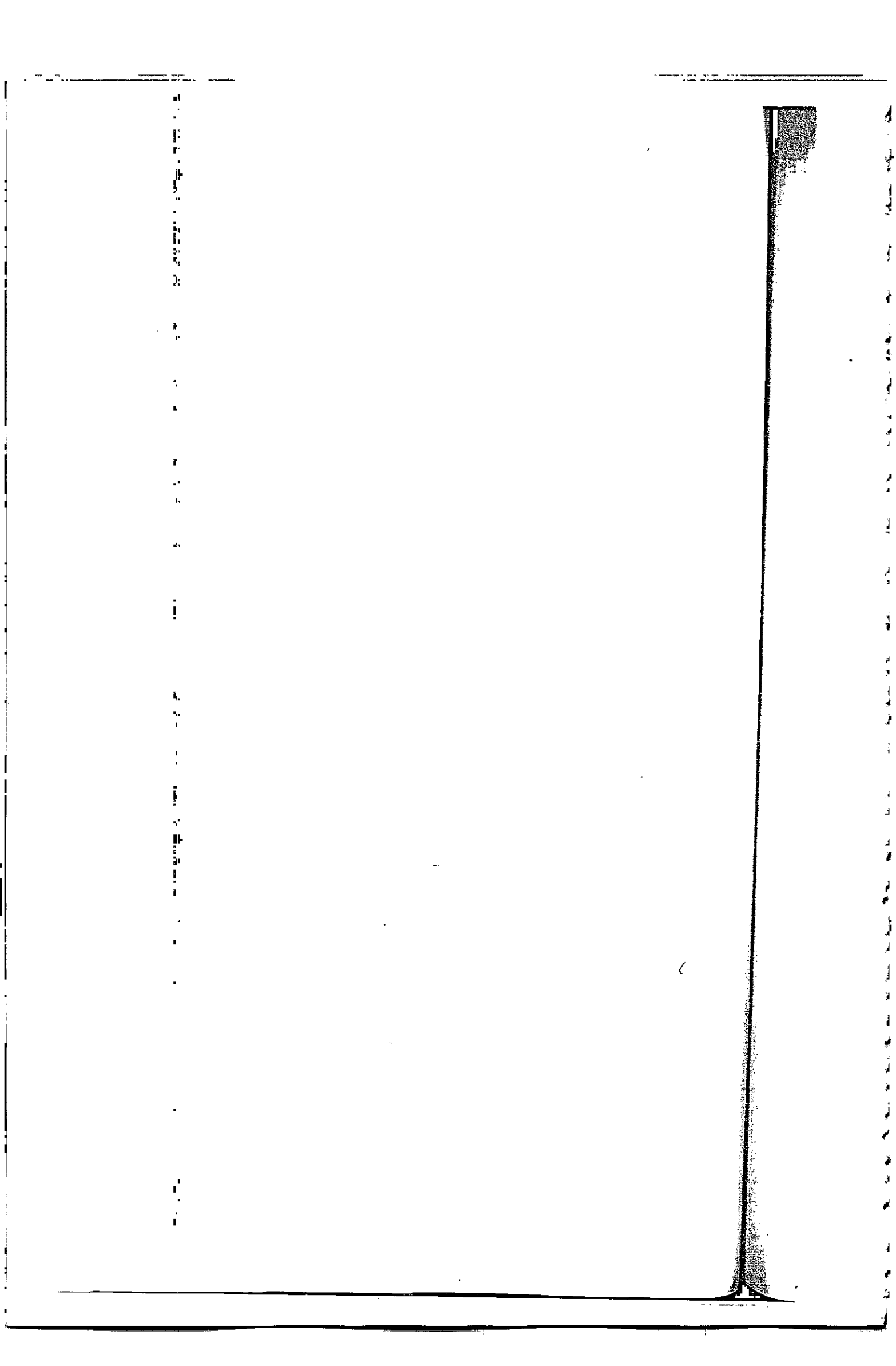
Options

This chapter now turns to nonlinear derivatives, or options. As described in Table 5.1, options account for a large part of the derivatives markets. On organized exchanges, options represent $52 trillion in derivatives outstanding. Over-the-counter options add up to about $76 trillion in notional amounts.

Although the concept behind these instruments are not new, option markets have blossomed since the early 1970s, because of a breakthrough in pricing options, the Black–Scholes formula, and advances in computing power. We start with plain, vanilla options, calls and puts. These are the basic building blocks of many financial instruments. They are also more common than complicated, exotic options.

The purpose of this chapter is to present a compact overview of important concepts for options, including their pricing. We will cover option sensitivities (the "Greeks") in a future chapter. Section 6.1 presents the payoff functions on basic options and combinations thereof. We then discuss option premiums in Section 6.2. The Black–Scholes pricing approach is presented in Section 6.3. Next, Section 6.4 briefly summarizes more complex options. Finally, Section 6.5 shows how to value options using a numerical, binomial tree model.

6.1 OPTION PAYOFFS

6.1.1 Basic Options

Options are instruments that give their holder the *right* to buy or sell an asset at a specified price until a specified expiration date. The specified delivery price is known as the delivery price, or exercise price, or strike price, and is denoted by K.

Options to buy are **call options**. Options to sell are **put options**. As options confer a right to the purchaser of the option, but not an obligation, they will be exercised only if they generate profits. In contrast, forwards involve an obligation to either buy or sell and can generate profits or losses. Like forward contracts, options can be purchased or sold. In the latter case, the seller is said to write the option.

Depending on the timing of exercise, options can be classified into European or American options. European options can be exercised at maturity only. American options can be exercised at any time, before or at maturity. Because American

127

options include the right to exercise at maturity, they must be at least as valuable as European options. In practice, however, the value of this early exercise feature is small, as an investor can generally receive better value by reselling the option on the open market instead of exercising it.

We use these notations, in addition to those in the previous chapter:

$$K = \text{exercise price}$$
$$c = \text{value of European call option}$$
$$C = \text{value of American call option}$$
$$p = \text{value of European put option}$$
$$P = \text{value of American put option}$$

To illustrate, take an option on an asset that currently trades at $85 with a delivery price of $100 in one year. If the spot price stays at $85 at expiration, the holder of the call will not exercise the option, because the option is not profitable with a stock price less than $100. In contrast, if the price goes to $120, the holder will exercise the right to buy at $100, will acquire the stock now worth $120, and will enjoy a "paper" profit of $20. This profit can be realized by selling the stock. For put options, a profit accrues if the spot price falls below the exercise price $K = \$100$.

Thus, the payoff profile of a long position in the call option at expiration is

$$C_T = \text{Max}(S_T - K, 0) \tag{6.1}$$

The payoff profile of a long position in a put option is

$$P_T = \text{Max}(K - S_T, 0) \tag{6.2}$$

If the current asset price S_t is close to the strike price K, the option is said to be at-the-money. If the current asset price S_t is such that the option could be exercised now at a profit, the option is said to be in-the-money. If the remaining situation, the option is said to be out-of-the-money. A call will be in-the-money if $S_t > K$. A put will be in-the-money if $S_t < K$.

As in the case of forward contracts, the payoff at expiration can be cash settled. Instead of actually buying the asset, the contract could simply pay $20 if the price of the asset is $120.

Because buying options can generate only profits (at worst zero) at expiration, an option contract must be a valuable asset (or at worst have zero value). This means that a payment is needed to acquire the contract. This up-front payment, which is much like an insurance premium, is called the option "premium." This premium cannot be negative. An option becomes more expensive as it moves in-the-money.

Thus, the payoffs on options must take into account this cost (for long positions) or benefit (for short positions). To compute the total payoff, we should translate all option payoffs by the *future* value of the premium, that is, ce^{rt}, for European call options.
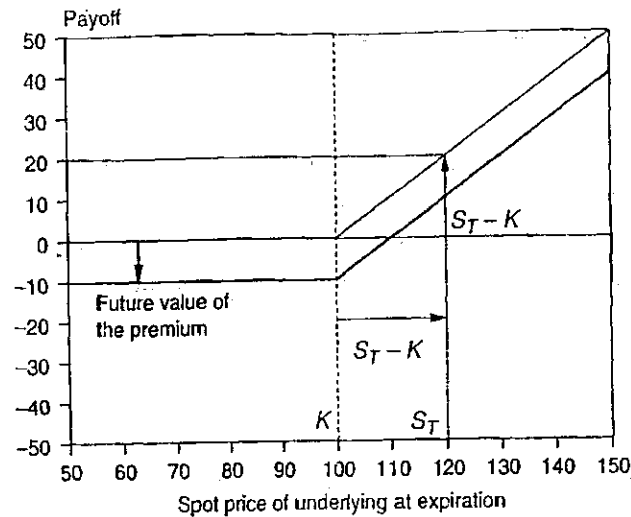
FIGURE 6.1   Profit Payoffs on Long Call

Figure 6.1 displays the total profit payoff on a call option as a function of the asset price at expiration. Assuming that $S_T = \$120$, the proceeds from exercise are $\$120 - \$100 = \$20$, from which we have to subtract the future value of the premium, say $\$10$. In the graphs that follow, we always take into account the cost of the option.

Figure 6.2 summarizes the payoff patterns on long and short positions in a call and a put contract. Unlike those of forwards, these payoffs are nonlinear in the underlying spot price. Sometimes they are referred to as the "hockey stick" diagrams. This is because forwards are obligations, whereas options are rights. Note that the positions for the same contract are symmetrical around the horizontal axis. For a given spot price, the sum of the profit or loss for the long and for the short is zero.

In the market-risk section (Part Three) of this handbook, we will combine these payoffs with the distribution of the risk factors. Even so, it is immediately
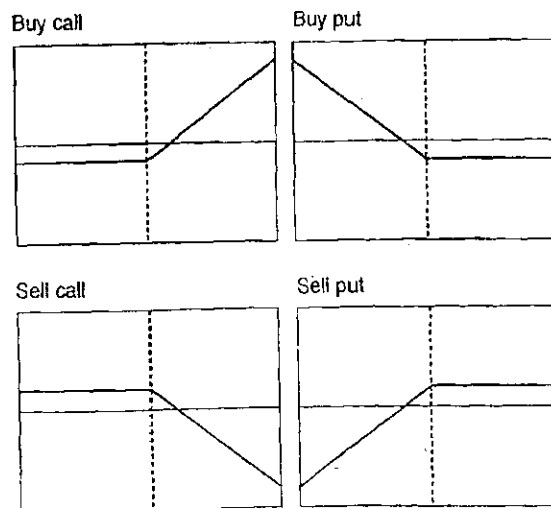
FIGURE 6.2   Profit Payoffs on Long and Short Calls and Puts

obvious that long option positions have limited downside risk, which is the loss of the premium. Short call option positions have unlimited downside risk because there is no upper limit on S. The worst loss on short put positions occurs if S goes to zero.

So far, we have covered options on cash instruments. Options can also be struck on futures. When exercising a call, the investor becomes long the futures contract. Conversely, exercising a put creates a short position in the futures contract. Because positions in futures are equivalent to leveraged positions in the underlying cash instrument, options on cash instruments and on futures are equivalent.

6.1.2 Put–Call Parity

These option payoffs can be used as the basic building blocks for more complex positions. A long position in the underlying asset can be decomposed into a long call plus a short put, as shown in Figure 6.3.

The figure shows that the long call provides the equivalent of the upside while the short put generates the same downside risk as holding the asset. This link creates a relationship between the value of the call and that of the put, also known as put–call parity. The relationship is illustrated in Table 6.1, which examines the payoff at initiation and at expiration under the two possible states of the world. We only consider European options with the same maturity and exercise price. Also, we assume that there is no income payment on the underlying asset.

The portfolio consists of a long position in the call, a short position in the put, and an investment to ensure that we will be able to pay the exercise price at maturity. Long positions are represented by negative values, as they are outflows.

The table shows that the final payoffs to portfolio (1) add up to S_T in the two states of the world, which is the same as a long position in the asset itself. Hence, to avoid arbitrage, the initial payoff must be equal to the current cost of the asset,
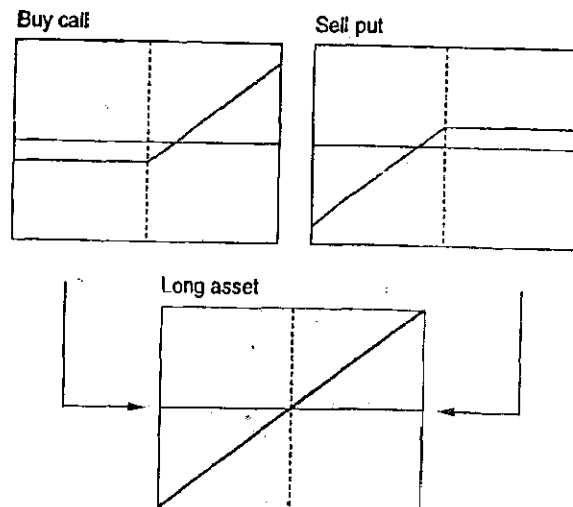
FIGURE 6.3 Decomposing a Long Position in the Asset

TABLE 6.1 Put–Call Parity

Portfolio	Position	Initial Payoff	Final Payoff $S_T < K$	$S_T \geq K$
(1)	Buy call	$-c$	0	$S_T - K$
	Sell put	$+p$	$-(K - S_T)$	0
	Invest	$-Ke^{-r\tau}$	K	K
	Total	$-c + p - Ke^{-r\tau}$	S_T	S_T
(2)	Buy asset	$-S$	S_T	S_T

which is $S_t = S$. So, we must have $-c + p - Ke^{-r\tau} = -S$. More generally, with income paid at the rate of r^*, put-call parity can be written as

$$c - p = Se^{-r^*\tau} - Ke^{-r\tau} = (F - K)e^{-r\tau} \tag{6.3}$$

Because $c \geq 0$ and $p \geq 0$, this relationship can be also used to determine lower bounds for European calls and puts. Note that the relationship does not hold exactly for American options since there is a likelihood of early exercise, which could lead to mismatched payoffs.

Finally, this relationship can be used to determine the implied dividend yield from market prices. We observe c, p, S, and r and can solve for y or r^*. This yield is used for determining the forward rate in dividend swaps, which are contracts where the payoff is indexed to the actual dividends paid over the horizon, minus the implied dividends.

KEY CONCEPT

A long position in an asset is equivalent to a long position in a European call with a short position in an otherwise identical put, combined with a risk-free position.

EXAMPLE 6.1: FRM EXAM 2007—QUESTION 84

According to put–call parity, buying a put option on a stock is equivalent to

a. Buying a call option and buying the stock with funds borrowed at the risk-free rate
b. Selling a call option and buying the stock with funds borrowed at the risk-free rate
c. Buying a call option, selling the stock, and investing the proceeds at the risk-free rate
d. Selling a call option, selling the stock, and investing the proceeds at the risk-free rate

EXAMPLE 6.2: FRM EXAM 2005—QUESTION 72

A one-year European put option on a non-dividend-paying stock with strike at EUR 25 currently trades at EUR 3.19. The current stock price is EUR 23 and its annual volatility is 30%. The annual risk-free interest rate is 5%. What is the price of a European call option on the same stock with the same parameters as that of the above put option? Assume continuous compounding.

a. EUR 1.19

b. EUR 3.97

c. EUR 2.41

d. Cannot be determined with the data provided

EXAMPLE 6.3: FRM EXAM 2002—QUESTION 25

The price of a non-dividend-paying stock is $20. A six-month European call option with a strike price of $18 sells for $4. A European put option on the same stock, with the same strike price and maturity, sells for $1.47. The continuously compounded risk-free interest rate is 6% per annum. Are these three securities (the stock and the two options) consistently priced?

a. No, there is an arbitrage opportunity worth $2.00.

b. No, there is an arbitrage opportunity worth $2.53.

c. No, there is an arbitrage opportunity worth $14.00.

d. Yes.

EXAMPLE 6.4: FRM EXAM 2006—QUESTION 74

Jeff is an arbitrage trader, who wants to calculate the implied dividend yield on a stock while looking at the over-the-counter price of a five-year European put and call on that stock. He has the following data: $S = \$85$, $K = \$90$, $r = 5\%$, $c = \$10$, $p = \$15$. What is the continuous implied dividend yield of that stock?

a. 2.48%

b. 4.69%

c. 5.34%

d. 7.71%

6.1.3 Combination of Options

Options can be combined in different ways, either with each other or with the underlying asset. Consider first combinations of the underlying asset and an option. A long position in the stock can be accompanied by a short sale of a call to collect the option premium. This operation, called a covered call, is described in Figure 6.4. Likewise, a long position in the stock can be accompanied by a purchase of a put to protect the downside. This operation is called a protective put.

Options can also be combined with an underlying position to limit the range of potential gains and losses. Suppose an investor is long a stock, currently trading at $10. The investor can buy a put with a low strike price (e.g., $7), partially financed by the sale of a call with a high strike (e.g., $12). Ignoring the net premium, the highest potential gain is $2 and the worst loss is $3. Such strategy is called a collar. If the strike prices were the same, this would be equivalent to a short stock position, which creates a net payoff of exactly zero.

We can also combine a call and a put with the same or different strike prices and maturities. When the strike prices of the call and the put, and their maturities, are the same, the combination is referred to as a straddle. Figure 6.5 shows how to construct a long straddle, i.e., buying a call and a put with the same maturity and strike price. This position is expected to benefit from a large price move, whether up or down. The reverse position is a short straddle. When the strike prices are different, the combination is referred to as a strangle. Since strangles are out-of-the-money, they are cheaper to buy than straddles.

Thus far, we have concentrated on positions involving two classes of options. One can, however, establish positions with one class of options, called spreads. Calendar, or horizontal spreads correspond to different maturities. Vertical spreads correspond to different strike prices. The names of the spreads are derived from the manner in which they are listed in newspapers: time is listed horizontally and strike prices are listed vertically. Diagonal spreads move across maturities and strike prices.
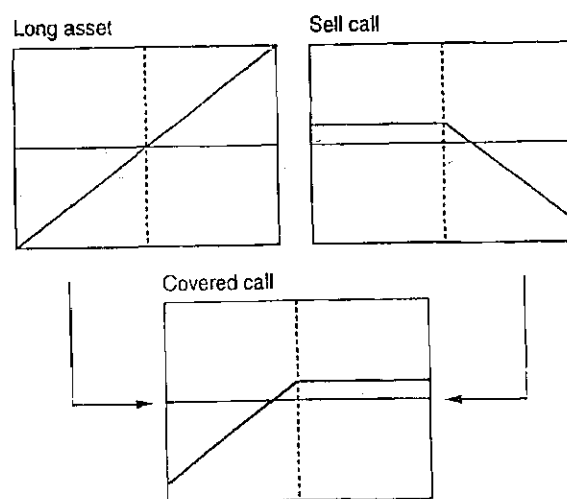
FIGURE 6.4 Creating a Covered Call
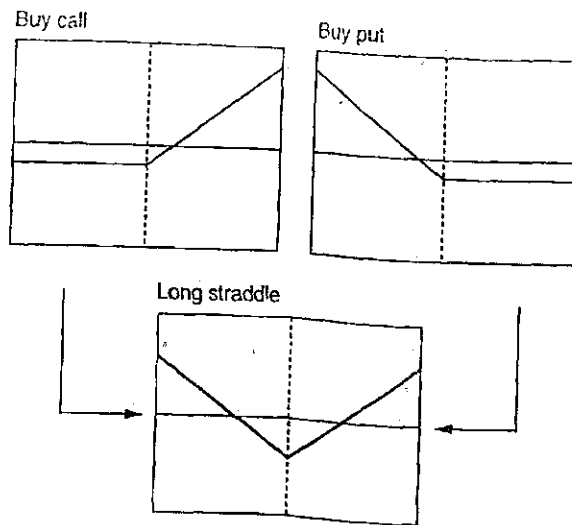
FIGURE 6.5 Creating a Long Straddle

For instance, a bull spread is positioned to take advantage of an increase in the price of the underlying asset. Conversely, a bear spread represents a bet on a falling price. Figure 6.6 shows how to construct a bull(ish) vertical spread with two calls with the same maturity. This could also be constructed with puts, however. Here, the spread is formed by buying a call option with a low exercise price K_1 and selling another call with a higher exercise price K_2. Note that the cost of the first call $c(S, K_1)$ must exceed the cost of the second call $c(S, K_2)$, because the first option is more in-the-money than the second. Hence, the sum of the two premiums represents a net cost. At expiration, when $S_T > K_2$, the payoff is $\text{Max}(S_T - K_1, 0) - \text{Max}(S_T - K_2, 0) = (S_T - K_1) - (S_T - K_2) = K_2 - K_1$, which is positive. Thus, this position is expected to benefit from an upmove, while incurring only limited downside risk.

Spreads involving more than two positions are referred to as butterfly or sandwich spreads. A butterfly spread involves three types of options with the same maturity: for example, a long call at a strike price K_1, two short calls at a
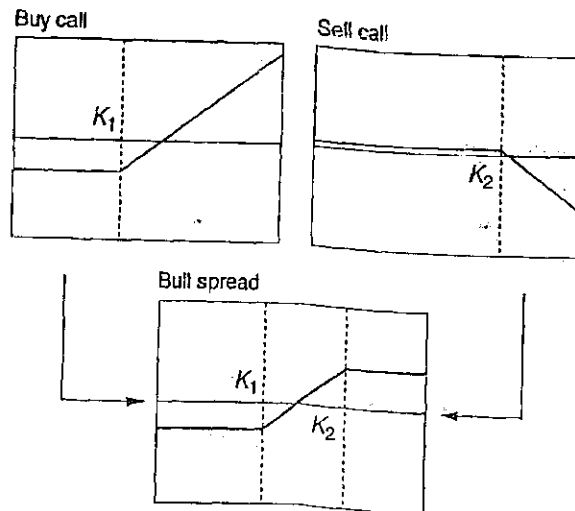
FIGURE 6.6 Creating a Bull Spread
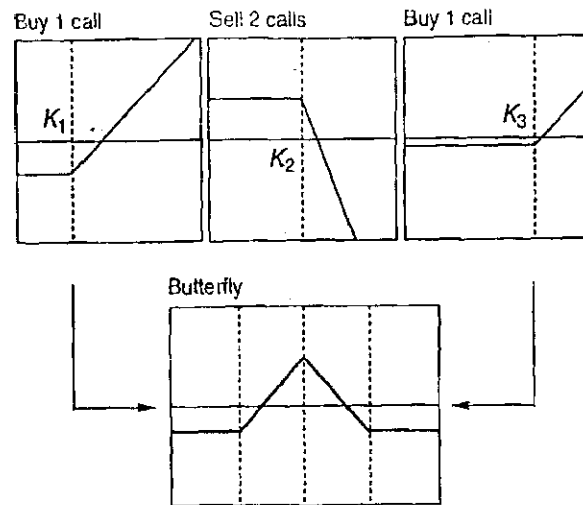
FIGURE 6.7 Creating a Butterfly Spread

higher strike price K_2, and a long call position at a higher strike price K_3, with the same spacing. Figure 6.7 shows that this position is expected to benefit when the underlying asset price stays stable, close to K_2. The double position in the middle is called the body, and the others the wings. A sandwich spread is the opposite of a butterfly spread.

EXAMPLE 6.5: FRM EXAM 2001—QUESTION 90

Which of the following is the riskiest form of speculation using option contracts?

a. Setting up a spread using call options
b. Buying put options
c. Writing naked call options
d. Writing naked put options

EXAMPLE 6.6: FRM EXAM 2007—QUESTION 103

An investor sells a June 2008 call of ABC Limited with a strike price of USD 45 for USD 3 and buys a June 2008 call of ABC Limited with a strike price of USD 40 for USD 5. What is the name of this strategy and the maximum profit and loss the investor could incur?

a. Bear spread, maximum loss USD 2, maximum profit USD 3
b. Bull spread, maximum loss Unlimited, maximum profit USD 3
c. Bear spread, maximum loss USD 2, maximum profit unlimited
d. Bull spread, maximum loss USD 2, maximum profit USD 3

EXAMPLE 6.7: FRM EXAM 2006—QUESTION 45

A portfolio manager wants to hedge his bond portfolio against changes in interest rates. He intends to buy a put option with a strike price below the portfolio's current price in order to protect against rising interest rates. He also wants to sell a call option with a strike price above the portfolio's current price in order to reduce the cost of buying the put option. What strategy is the manager using?

a. Bear spread
b. Strangle
c. Collar
d. Straddle

EXAMPLE 6.8: FRM EXAM 2002—QUESTION 42

Consider a bearish option strategy of buying one $50 strike put for $7, selling two $42 strike puts for $4 each, and buying one $37 put for $2. All options have the same maturity. Calculate the final profit (P/L) per share of the strategy if the underlying is trading at $33 at expiration.

a. $1 per share
b. $2 per share
c. $3 per share
d. $4 per share

EXAMPLE 6.9: FRM EXAM 2003—QUESTION 72

Which of the following regarding option strategies is/are not correct?

I. A long strangle involves buying a call and a put with equal strike prices.
II. A short bull spread involves selling a call at lower strike price and buying another call at higher strike price.
III. Vertical spreads are formed by options with different maturities.
IV. A long butterfly spread is formed by buying two options at two different strike prices and selling another two options at the same strike price.

a. I only
b. I and III only
c. I and II only
d. III and IV only

6.2 OPTION PREMIUMS

6.2.1 General Relationships

So far, we have examined the payoffs at expiration only. Also important is the instantaneous relationship between the option value and the current price S, which is displayed in Figures 6.8 and 6.9.

For a call, a higher price S increases the current value of the option, but in a nonlinear, convex fashion. For a put, lower values for S increase the value of the option, also in a convex fashion. As time goes by, the curved line approaches the hockey stick line.

交流QQ: 792644267
电话: 13167301953
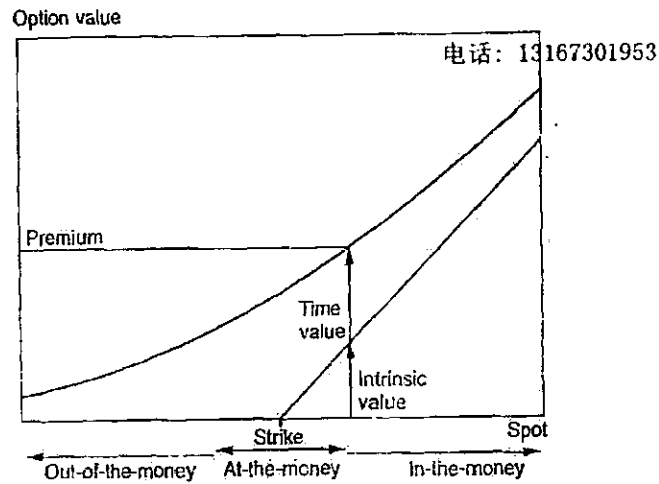
FIGURE 6.8 Relationship between Call Value and Spot Price

MSN:1sqing2022019

FRM资料专业提供
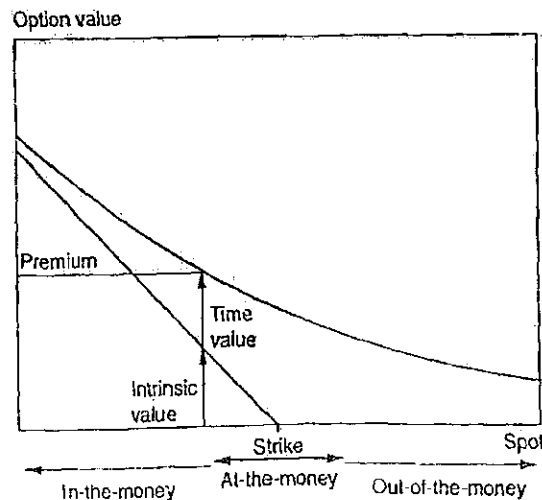
FIGURE 6.9 Relationship between Put Value and Spot Price

Figures 6.8 and 6.9 decompose the current premium into

- An intrinsic value, which basically consists of the value of the option if exercised today, or Max($S_t - K$, 0) for a call, and Max($K - S_t$, 0) for a put
- A time value, which consists of the remainder, reflecting the possibility that the option will create further gains in the future

Consider for example a one-year call with strike $K = \$100$. The current price is $S = \$120$ and interest rate $r = 5\%$. The asset pays no dividend. Say the call premium is $\$26.17$. This can be decomposed into an intrinsic value of $\$120 - \$100 = \$20$ and time value of $\$6.17$. The time value increases with the volatility of the underlying asset. It also generally increases with the maturity of the option.

As shown in Figures 6.8 and 6.9, options can be classified into:

- At-the-money, when the current spot price is close to the strike price
- In-the-money, when the intrinsic value is large
- Out-of-the-money, when the spot price is much below the strike price for calls and conversely for puts (out-of-the-money options have zero intrinsic value)

We can also identify some general bounds for European options that should always be satisfied; otherwise there would be an arbitrage opportunity (a money machine). For simplicity, assume there is no dividend. We know that a European option is worth less than an American option. First, the current value of a call must be less than, or equal to, the asset price:

$$c_t \leq C_t \leq S_t \qquad (6.4)$$

This is because, in the limit, an option with zero exercise price is equivalent to holding the stock. We are sure to exercise the option.

Second, the value of a European call must be greater than, or equal to, the price of the asset minus the present value of the strike price:

$$c_t \geq S_t - Ke^{-r\tau} \qquad (6.5)$$

To prove this, we could simply use put-call parity, or Equation (6.3) with $r^* = 0$, imposing the condition that $p \geq 0$. Note that, since $e^{-r\tau} < 1$, we must have $S_t - Ke^{-r\tau} > S_t - K$ before expiration. Thus, $S_t - Ke^{-r\tau}$ is a more informative lower bound than $S_t - K$. As an example, continue with our call option. The lower bound is $S_t - Ke^{-r\tau} = \$120 - \$100\exp(-5\% \times 1) = \$24.88$. This is more informative than $S - K = \$20$.

We can also describe upper and lower bounds for put options. The value of a put cannot be worth more than K

$$p_t \leq P_t \leq K \qquad (6.6)$$

which is the upper bound if the price falls to zero. Using put–call parity, we can show that the value of a European put must satisfy the following lower bound

$$p_t \geq Ke^{-r\tau} - S_t \tag{6.7}$$

6.2.2 Early Exercise of Options

These relationships can be used to assess the value of early exercise for American options. The basic trade-off arises between the value of the American option dead, that is, exercised, or alive, that is, nonexercised. Thus, the choice is between exercising the option and selling it on the open market.

Consider an American call on a non-dividend-paying stock. By exercising early, the holder gets exactly $S_t - K$. The value of the option alive, however, must be worth more than that of the equivalent European call. From Equation (6.5), this must satisfy $c_t \geq S_t - Ke^{-r\tau}$, which is strictly greater than $S_t - K$. Hence, an American call on a non-dividend-paying stock *should never* be exercised early.

In our example, the lower bound on the European call is $24.88. If we exercise the American call, we only get $S - K = \$120 - \$100 = \$20$. Because this is less than the minimum value of the European call, the American call should not be exercised. As a result, the value of the American feature is zero and we always have $c_t = C_t$.

The only reason to exercise early a call is to capture a dividend payment. Intuitively, a high income payment makes holding the asset more attractive than holding the option. Thus, American options on income-paying assets may be exercised early. Note that this applies also to options on futures, since the implied income stream on the underlying is the risk-free rate.

KEY CONCEPT

An American call option on a non-dividend-paying stock (or asset with no income) should never be exercised early. If the asset pays income, early exer cise may occur, with a probability that increases with the size of the income payment.

For an American put, we must have

$$P_t \geq K - S_t \tag{6.8}$$

because it could be exercised now. Unlike the relationship for calls, this lower bound $K - S_t$ is strictly greater than the lower bound for European puts $Ke^{-r\tau} - S_t$. So, we could have early exercise.

To decide whether to exercise early, the holder of the option has to balance the benefit of exercising, which is to receive K now instead of later, against the loss of killing the time value of the option. Because it is better to receive money now than later, it may be worth exercising the put option early.

Thus, American puts on nonincome paying assets *could* be exercised early, unlike calls. The probability of early exercise decreases for lower interest rates and with higher income payments on the asset. In each case, it becomes less attractive to sell the asset.

KEY CONCEPT

An American put option on a non-dividend-paying stock (or asset with no income) may be exercised early. If the asset pays income, the possibility of early exercise decreases with the size of the income payments.

EXAMPLE 6.10: FRM EXAM 2002—QUESTION 50

Given strictly positive interest rates, the best way to close out a long American call option position early (option written on a stock that pays no dividends) would be to

a. Exercise the call
b. Sell the call
c. Deliver the call
d. Do none of the above

EXAMPLE 6.11: FRM EXAM 2005—QUESTION 15

You have been asked to verify the pricing of a two-year European call option with a strike price of USD 45. You know that the initial stock price is USD 50, and the continuous risk-free rate is 3%. To verify the possible price range of this call, you consider using price bounds. What is the difference between the upper and lower bounds for that European call?

a. 0.00
b. 7.62
c. 42.38
d. 45.00

6.3 VALUING OPTIONS

6.3.1 Pricing by Replication

We now turn to the pricing of options. The philosophy of pricing models consists of replicating the payoff on the instrument. To avoid arbitrage, the price of the instrument must then equal the price of the replicating portfolio.

Consider a call option on a stock whose price is represented by a binomial process. The initial price of $S_0 = \$100$ can only move up or down, to two values (hence the name "binomial"), $S_1 = \$150$ or $S_2 = \$50$. The option is a call with $K = \$100$, and therefore can only take values of $c_1 = \$50$ or $c_2 = \$0$. We assume that the rate of interest is $r = 25\%$, so that a dollar invested now grows to $\$1.25$ at maturity.

$$
\begin{array}{l}
 S_1 = \$150 \quad c_1 = \$50 \\[4pt]
S_0 = \$100 \quad \nearrow \\[4pt]
 \searrow \\[4pt]
 S_2 = \$50 \quad c_2 = \$0
\end{array}
$$
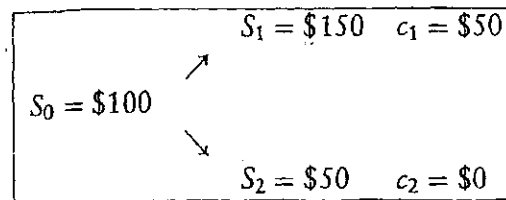
The key idea of derivatives pricing is that of replication. In other words, we replicate the payoff on the option by a suitable portfolio of the underlying asset plus a position, long or short, in a risk-free bill. This is feasible in this simple setup because we have two states of the world and two instruments, the stock and the bond. To prevent arbitrage, the current value of the derivative must be the same as that of the portfolio.

The portfolio consists of n shares and a risk-free investment currently valued at B (a negative value implies borrowing). We set $c_1 = nS_1 + B$, or $\$50 = n\$150 + B$ and $c_2 = nS_2 + B$, or $\$0 = n\$50 + B$ and solve the 2 by 2 system, which gives $n = 0.5$ and $B = -\$25$. At time $t = 0$, the value of the loan is $B_0 = \$25/1.25 = \$20$. The current value of the portfolio is $nS_0 + B_0 = 0.5 \times \$100 - \$20 = \$30$. Hence, the current value of the option must be $c_0 = \$30$. This derivation shows the essence of option pricing methods.

Note that we did not need the actual probability of an upmove. Define this as p. To see how this can be derived, we write the current value of the stock as the discounted expected payoff assuming investors were risk-neutral:

$$S_0 = [p \times S_1 + (1 - p) \times S_2]/(1 + r) \tag{6.9}$$

where the term between brackets is the expectation of the future spot price, given by the probability times its value for each state. Solving for $100 = [p \times 150 + (1 - p) \times 50]/1.25$, we find a risk-neutral probability of $p = 0.75$. We now value the option in the same fashion:

$$c_0 = [p \times c_1 + (1 - p) \times c_2]/(1 + r) \tag{6.10}$$

which gives

$$c_0 = [0.75 \times \$50 + 0.25 \times \$0]/1.25 = \$30$$

This simple example illustrates a very important concept, which is that of risk-neutral pricing.

6.3.2 Black–Scholes Valuation

The Black–Scholes (BS) model is an application of these ideas that provides an elegant closed-form solution to the pricing of European calls. The derivation of the model is based on four assumptions:

Black–Scholes Model Assumptions

- *The price of the underlying asset moves in a continuous fashion.*
- *Interest rates are known and constant.*
- *The variance of underlying asset returns is constant.*
- *Capital markets are perfect* (i.e., short-sales are allowed, there are no transaction costs or taxes, and markets operate continuously).

The most important assumption behind the model is that prices are continuous. This rules out discontinuities in the sample path, such as jumps, which cannot be hedged in this model.

The statistical process for the asset price is modeled by a geometric Brownian motion: Over a very short time interval, dt, the logarithmic return has a normal distribution with mean $= \mu dt$ and variance $= \sigma^2 dt$. The total return can be modeled as

$$dS/S = \mu dt + \sigma dz \qquad (6.11)$$

where the first term represents the drift component, and the second is the stochastic component, with dz distributed normally with mean zero and variance dt.

This process implies that the logarithm of the ending price is distributed as

$$\ln(S_T) = \ln(S_0) + (\mu - \sigma^2/2)\tau + \sigma\sqrt{\tau}\,\epsilon \qquad (6.12)$$

where ϵ is a $N(0, 1)$ random variable. Hence, the price is lognormally distributed.

Based on these assumptions, Black and Scholes (1972) derived a closed-form formula for European options on a non-dividend-paying stock, called the Black–Scholes model. The key point of the analysis is that a position in the option can be replicated by a "delta" position in the underlying asset. Hence, a portfolio combining the asset and the option in appropriate proportions is "locally"

risk-free, that is, for small movements in prices. To avoid arbitrage, this portfolio must return the risk-free rate.

As a result, we can directly compute the present value of the derivative as the discounted expected payoff

$$f_t = E_{RN}[e^{-r\tau} F(S_T)] \tag{6.13}$$

where the underlying asset is assumed to grow at the risk-free rate, and the discounting is also done at the risk-free rate. Here, the subscript RN refers to the fact that the analysis assumes risk neutrality. In a risk-neutral world, the expected return on all securities must be the risk-free rate of interest, r. The reason is that risk-neutral investors do not require a risk premium to induce them to take risks. The BS model value can be computed assuming that all payoffs grow at the risk-free rate and are discounted at the same risk-free rate.

This risk-neutral valuation approach is without a doubt the most important tool in derivatives pricing. Before the Black–Scholes breakthrough, Samuelson had derived a very similar model in 1965, but with the asset growing at the rate μ and discounting as some other rate μ^*.[1] Because μ and μ^* are unknown, the Samuelson model was not practical. The risk-neutral valuation is merely an artificial method to obtain the correct solution, however. It does not imply that investors are risk-neutral.

Furthermore, this approach has limited uses for risk management. The BS model can be used to derive the risk-neutral probability of exercising the option. For risk management, however, what matters is the actual probability of exercise, also called physical probability. This can differ from the RN probability.

We now turn to the formulation of the BS model. In the case of a European call, the final payoff is $F(S_T) = Max(S_T - K, 0)$. Initially, we assume no dividend payment on the asset. The current value of the call is given by:

$$c = SN(d_1) - Ke^{-r\tau} N(d_2) \tag{6.14}$$

where $N(d)$ is the cumulative distribution function for the standard normal distribution:

$$N(d) = \int_{-\infty}^{d} \Phi(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-\frac{1}{2}x^2} dx$$

with Φ defined as the standard normal density function. $N(d)$ is also the area to the left of a standard normal variable with value equal to d, as shown in Figure 6.10. Note that, since the normal density is symmetrical, $N(d) = 1 - N(-d)$, or the area to the left of d is the same as the area to the right of $-d$.

---

[1] Samuelson, Paul (1965), Rational Theory of Warrant Price, *Industrial Management Review* 6, 13–39.

Probability density function $\Phi(d)$

$N(d_1)$

Delta
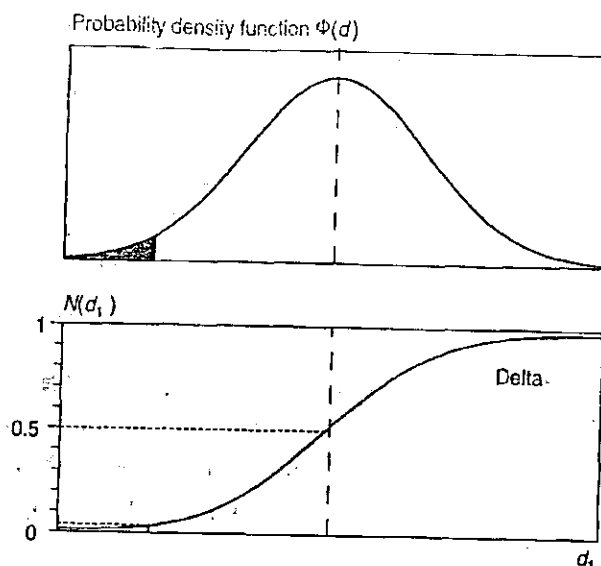
FIGURE 6.10 Cumulative Distribution Function

The values of d_1 and d_2 are:

$$d_1 = \frac{\ln(S/Ke^{-r\tau})}{\sigma\sqrt{\tau}} + \frac{\sigma\sqrt{\tau}}{2}, \quad d_2 = d_1 - \sigma\sqrt{\tau}$$

By put-call parity, the European put option value is

$$p = S[N(d_1) - 1] - Ke^{-r\tau}[N(d_2) - 1] \tag{6.15}$$

Example: Computing the Black–Scholes Value

Consider an at-the-money call on a stock worth $S = \$100$, with a strike price of $K = \$100$ and maturity of six months. The stock has annual volatility of $\sigma = 20\%$ and pays no dividend. The risk-free rate is $r = 5\%$.

First, we compute the present value factor, which is $e^{-r\tau} = \exp(-0.05 \times 6/12) = 0.9753$. We then compute the value of $d_1 = \ln[S/Ke^{-r\tau}]/\sigma\sqrt{\tau} + \sigma\sqrt{\tau}/2 = 0.2475$ and $d_2 = d_1 - \sigma\sqrt{\tau} = 0.1061$. Using standard normal tables or the NORMSDIST Excel function, we find $N(d_1) = 0.5977$ and $N(d_2) = 0.5422$. Note that both values are greater than 0.5 since d_1 and d_2 are both positive. The option is at-the-money. As S is close to K, d_1 is close to zero and $N(d_1)$ close to 0.5.

The value of the call is $c = SN(d_1) - Ke^{-r\tau}N(d_2) = \$6.89$.

The value of the call can also be viewed as an equivalent position of $N(d_1) = 59.77\%$ in the stock and some borrowing: $c = \$59.77 - \$52.88 = \$6.89$. Thus this is a leveraged position in the stock.

The value of the put is $\$4.42$. Buying the call and selling the put costs $\$6.89 - \$4.42 = \$2.47$. This indeed equals $S - Ke^{-r\tau} = \$100 - \$97.53 = \$2.47$, which confirms put–call parity.

We should note that Equation (6.14) can be reinterpreted in view of the discounting formula in a risk-neutral world, Equation (6.13)

$$c = E_{RN}[e^{-r\tau}\text{Max}(S_T - K, 0)] = e^{-r\tau}[\int_K^\infty Sf(S)dS] - Ke^{-r\tau}[\int_K^\infty f(S)dS]$$

$$(6.16)$$

We see that the integral term multiplying K is the risk-neutral probability of exercising the call, or that the option will end up in-the-money $S > K$. Matching this up with (6.14), this gives

$$\text{Risk} - \text{neutral probability of exercise} = \left[\int_K^\infty f(S)dS\right] = N(d_2) \qquad (6.17)$$

6.3.3 Extensions

Merton (1973) expanded the BS model to the case of a stock paying a continuous dividend yield q. Garman and Kohlhagen (1983) extended the formula to foreign currencies, reinterpreting the yield as the foreign rate of interest $q = r^*$, in what is called the **Garman-Kohlhagen** model.

The Merton model then replaces all occurrences of S by $Se^{-r^*\tau}$. The call is worth

$$c = Se^{-r^*\tau}N(d_1) - Ke^{-r\tau}N(d_2) \qquad (6.18)$$

It is interesting to take the limit of Equation (6.14) as the option moves more in-the-money, that is, when the spot price S is much greater than K. In this case, d_1 and d_2 become very large and the functions $N(d_1)$ and $N(d_2)$ tend to unity. The value of the call then tends to

$$c(S >> K) = Se^{-r^*\tau} - Ke^{-r\tau} \qquad (6.19)$$

which is the valuation formula for a forward contract. A call that is deep in-the-money is equivalent to a long forward contract, because we are almost certain to exercise.

The Black model (1976) applies the same formula to options on futures. The only conceptual difference lies in the income payment to the underlying instrument. With an option on cash, the income is the dividend or interest on the cash instrument. In contrast, with a futures contract, the economically equivalent stream of income is the riskless interest rate. The intuition is that a futures can be viewed as equivalent to a position in the underlying asset with the investor setting aside an amount of cash equivalent to the present value of F.

6.3.4 Market versus Model Prices

In practice, the BS model is widely used to price options. All of the parameters are observable, except for the volatility. If we observe a market price, however, we can solve for the volatility parameter that sets the model price equal to the market price. This is called the implied standard deviation (ISD).

If the model were correct, the ISD should be constant across strike prices. In fact, this is not what we observe. Plots of the ISD against the strike price generally display what is called a volatility smile pattern, meaning that ISDs increase for low and high values of K. This effect has been observed in a variety of markets, such as foreign currency options. For equity index options, the effect is more asymmetrical, with very high ISDs for low strike prices. Because of the negative slope, this is called a volatility skew. Before the stock market crash of October 1987, this effect was minor. Since then, it has become more pronounced.

A related concept is the term structure of volatility, which refers to the observation that the ISD differs across maturities. This arises because the option market incorporates many types of events over the life of the option. For instance, the realized volatility of a stock tends to increase on the day of an earnings announcement.

These observations can be usefully summarized by an implied volatility surface, which is a three-dimensional plot of ISD across maturities and strike prices. Option traders typically observe the shape of this volatility surface to identify sectors where the ISDs seem out of line. To predict returns on options, however, traders also need to forecast the evolution of the implied volatility surface.

Figure 6.11 gives an example of the evolution of a volatility skew. The initial curve has an ISD of 18% for ATM options, with a strike price of $K = 100$. The
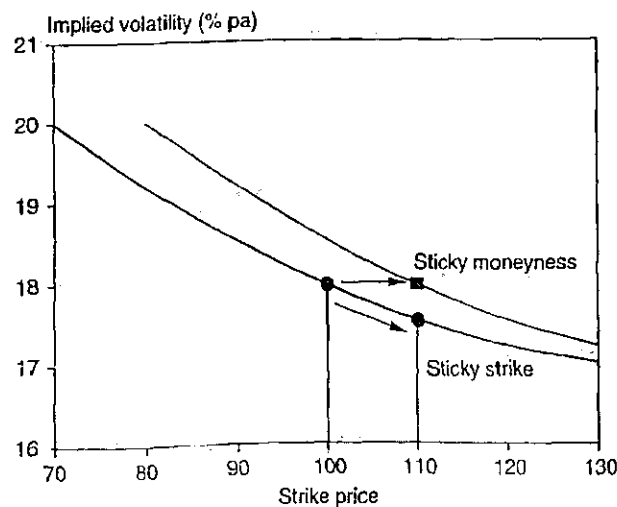
FIGURE 6.11 Evolution of Volatility Skew

question is what the curve will look like if the spot price moves from $S = 100$ to $S = 110$. Traders typically use heuristic approaches to the extrapolation of the curve over the investment horizon.

In a first scenario, called sticky strike, the curve does not change and the ISD drops from 18 to 17.5. This assumes that there is no structural change in the volatility curve and that the price movement is largely temporary. In a second scenario, called sticky moneyness, the curve shifts to the right and the ISD stays at 18 (the moneyness is not changed and is still 100%). This assumes a permanent shift in the volatility curve. Based on these assumptions, the trader can then examine the return on different option trading strategies and choose the most appropriate one. More generally, this demonstrates that the implied volatility is a major risk factor when trading options.

- Finally, the ISD of a portfolio of assets can be related to the ISD of its components through the implied correlation. Normally, the portfolio variance is related to the individual volatilities using

$$\sigma_p^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{N} \sum_{j<i}^{N} w_i w_j \rho_{ij} \sigma_i \sigma_j \qquad (6.24)$$

Assume now that there is a constant correlation ρ across all pairs of assets that maintains the portfolio variance

$$\sigma_p^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{N} \sum_{j<i}^{N} w_i w_j (\rho) \sigma_i \sigma_j \qquad (6.25)$$

This correlation is a weighted average of the pairwise correlations ρ_{ij}. With option ISDs measured for the portfolio and all the constituents, we can use Equation (6.25) to infer the portfolio implied correlation. This implied correlation is a summary measure of diversification benefits across the portfolio. All else equal, an increasing correlation increases the total portfolio risk.

EXAMPLE 6.12: FRM EXAM 2005—QUESTION 40

Which of the following statements is *wrong*?

a. The Black–Scholes formula holds only in a risk-neutral world.

b. The futures price of a stock depends on the risk-free rate.

c. An American put option is generally priced higher than a similar European put option.

d. Binomial trees can be used to price American options.

EXAMPLE 6.13: FRM EXAM 2001—QUESTION 91

Using the Black–Scholes model, calculate the value of a European call option given the following information: Spot rate = 100; Strike price = 110; Risk-free rate = 10%; Time to expiry = 0.5 years; $N(d_1) = 0.457185$; $N(d_2) = 0.374163$.

a. $10.90
b. $9.51
c. $6.57
d. $4.92

EXAMPLE 6.14: PROBABILITY OF EXERCISE

In the Black–Scholes expression for a European call option the term used to compute option probability of exercise is

a. d_1
b. d_2
c. $N(d_1)$
d. $N(d_2)$

6.4 OTHER OPTION CONTRACTS

The options described so far are standard, plain-vanilla options. Many other types of options, however, have been developed.

Binary options, also called digital options pay a fixed amount, say Q, if the asset price ends up above the strike price

$$c_T = Q \times I(S_T - K) \tag{6.26}$$

where $I(x)$ is an indicator variable that takes the value of 1 if $x \geq 0$ and 0 otherwise. The payoff function is illustrated in Figure 6.12 when $K = \$100$.

Because the probability of ending in-the-money in a risk-neutral world is $N(d_2)$, the initial value of this option is simply

$$c = Qe^{-r\tau} N(d_2) \tag{6.27}$$
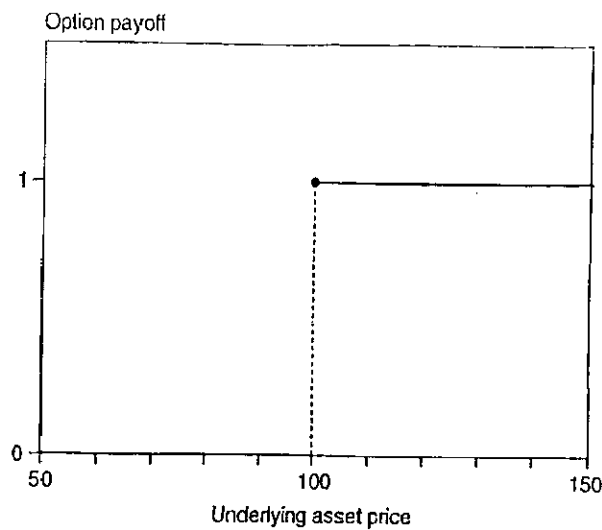
FIGURE 6.12 Payoff on a Binary Option

These options involve a sharp discontinuity around the strike price. Just below K, their value is zero. Just above, the value is the notional Q. Due to this discontinuity, these options are very difficult to hedge.

Another important class of options are barrier options. Barrier options are options where the payoff depends on the value of the asset hitting a barrier during a certain period of time. A knock-out option disappears if the price hits a certain barrier. A knock-in option comes into existence when the price hits a certain barrier.

An example of a knock-out option is the **down-and-out call**. This disappears if S hits a specified level H during its life. In this case, the knock-out price H must be lower than the initial price S_0. The option that appears at H is the down-and-in call. With identical parameters, the two options are perfectly complementary. When one disappears, the other appears. As a result, these two options must add up to a regular call option. Similarly, an up-and-out call ceases to exist when S reaches $H > S_0$. The complementary option is the up-and-in call.

Figure 6.13 compares price paths for the four possible combinations of calls. In all figures, the dark line describes the relevant price path, during which the option is alive; the grey line describes the remaining path.

The graphs illustrates that the down-and-out call and down-and-in call add up to the regular price path of a regular European call option. Thus at initiation, the value of these two options must add up to

$$c = c_{DO} + c_{DI} \tag{6.28}$$

Because all these values are positive (or at worst zero), the value of each premium c_{DO} and c_{DI} must be no greater than that of c. A similar reasoning applies to the two options in the right panels. Sometimes the option offers a rebate if it is knocked out.

Similar combinations exist for put options. An up-and-out put ceases to exist when S reaches $H > S_0$. A down-and-out put ceases to exist when S reaches $H < S_0$. The only difference with Figure 6.13 is that the option is exercised at maturity if $S < K$.
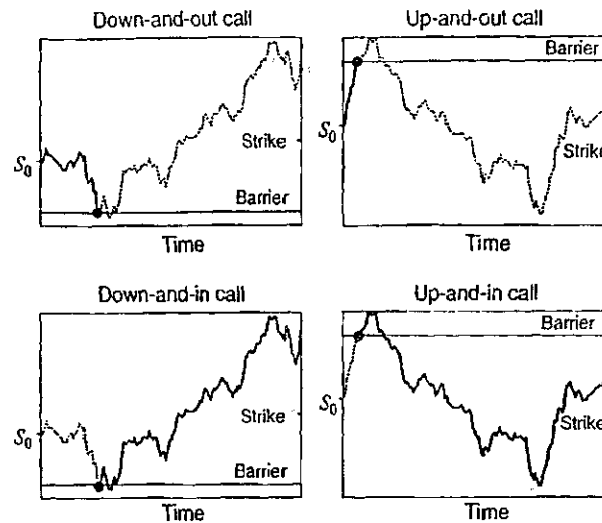
FIGURE 6.13 Paths for Knock-Out and Knock-In Call Options

Barrier options are attractive because they are "cheaper" than the equivalent European option. This, of course, reflects the fact that they are less likely to be exercised than other options.

In addition, these options are difficult to hedge because a discontinuity arises as the spot price get closer to the barrier. Just above the barrier, the option has positive value. For a very small movement in the asset price, going below the barrier, this value disappears.

Another widely used class of options are Asian options. Asian options, or average rate options, generate payoffs that depend on the average value of the underlying spot price during the life of the option, instead of the ending value. Define this as $S_{AVE}(t, T)$. The final payoff for a call is

$$c_T = \text{Max}(S_{AVE}(t, T) - K, 0) \tag{6.29}$$

Because an average is less variable than the final value at the end of the same period, such options are "cheaper" than regular options due to lower volatility. In fact, the price of the option can be treated like that of an ordinary option with the volatility set equal to $\sigma/\sqrt{3}$ and an adjustment to the dividend yield.[3] As a result of the averaging process, such options are easier to hedge than ordinary options.

Chooser options allow the holder to choose whether the option is a call or a put. At that point in time, the value of the option is

$$f_t = \text{Max}(c_t, p_t) \tag{6.30}$$

Thus, it is a package of two options, a regular call plus an option to convert to a put. As a result, these options are more expensive than plain-vanilla options.

---

[3] This is only strictly true when the averaging is a geometric average. In practice, average options involve an arithmetic average, for which there is no analytic solution; the lower volatility adjustment is just an approximation.

· Finally, lookback options have payoffs that depend on the extreme values of S over the option's life. Define S_{MAX} as the maximum and S_{MIN} as the minimum. A fixed strike lookback call option pays $Max(S_{MAX} - K, 0)$. A floating strike lookback call option pays $Max(S_T - S_{MIN}, 0)$. These options are even more expensive than regular options.

EXAMPLE 6.15: FRM EXAM 2003—QUESTION 34

Which of the following options is strongly path-dependent?

a. An Asian option
b. A binary option
c. An American option
d. A European call option

EXAMPLE 6.16: FRM EXAM 2006—QUESTION 59

All else being equal, which of the following options would cost more than plain-vanilla options that are currently at-the-money?

I. Lookback options
II. Barrier options
III. Asian options
IV. Chooser option
    a. I only
    b. I and IV
    c. II and III
    d. I, III, and IV

EXAMPLE 6.17: FRM EXAM 2002—QUESTION 19

Of the following options, which one does *not* benefit from an increase in the stock price when the current stock price is $100 and the barrier has not yet been crossed:

a. A down-and-out call with out barrier at $90 and strike at $110
b. A down-and-in call with in barrier at $90 and strike at $110
c. An up-and-in put with barrier at $110 and strike at $100
d. An up-and-in call with barrier at $110 and strike at $100

6.5 VALUING OPTIONS BY NUMERICAL METHODS

Some options have analytical solutions, such as the Black–Scholes models for European vanilla options. For more general options, however, we need to use numerical methods.

The basic valuation formula for derivatives is Equation (6.13), which states that the current value is the discounted present value of expected cash flows, where all assets grow at the risk-free rate and are discounted at the same risk-free rate.

We can use the Monte Carlo simulation methods presented in Chapter 4 to generate sample paths, final option values, and discount them into the present. Such simulation methods can be used for European or even path-dependent options, such as Asian options.

Table 6.2 gives an example. Suppose we need to price a European call with parameters $S = 100, K = 100, T = 1, r = 5\%, r^* = 0, \sigma = 20\%$. We set up the simulation with, for instance, $n = 100$ steps over the horizon of one year. For each step, the trend is $r/n = 0.05/100$; the volatility is $\sigma/\sqrt{(n)} = 0.20/\sqrt{100}$. Each replication starts from a price of \$100 until the horizon. For instance, the first replication gives a final price of $S_T = \$114.06$. The option is in-the-money and is worth $c_T = \$14.06$. We then discount this number into the present and get \$13.37. In the second replication, $S_T = \$75.83$ and the option expires worthless $c_T = 0$. Averaging across the K replications gives an average of \$10.33 in this case. The result is close to the actual Black–Scholes model price of \$10.45, obtained with Equation (6.14). The simulation, however, is much more general. The payoff at expiration could be a complicated function of the final price or even its intermediate values.

Simulation methods, however, cannot account for the possibility of early exercise. Instead, binomial trees must be used to value American options. As explained previously, the method consists of chopping up the time horizon into n intervals Δt and setting up the tree so that the characteristics of price movements fit the lognormal distribution.

At each node, the initial price S can go up to uS with probability p or down to dS with probability $(1 - p)$. The parameters u, d, p are chosen so that, for a small time interval, the expected return and variance equal those of the continuous

TABLE 6.2   Example of Simulation for a European Call Option

Replication	Final Payoff		Discounted Value
	S_T	c_T	
1	114.06	14.06	13.37
2	75.83	0.00	0.00
3	108.76	8.76	8.33
....			
Average			10.33

process. One could choose, for instance,

$$u = e^{\sigma\sqrt{\Delta t}}, \quad d = (1/u), \quad p = \frac{e^{\mu\Delta t} - d}{u - d} \tag{6.31}$$

Since this is a risk-neutral process, the total expected return must be equal to the risk-free rate r. Allowing for an income payment of r^*, this gives $\mu = r - r^*$.

The tree is built starting from the current time to maturity, from the left to the right. Next, the derivative is valued by starting at the end of the tree, and working backward to the initial time, from the right to the left.

Consider first a European call option. At time T (maturity) and node j, the call option is worth $Max(S_{Tj} - K, 0)$. At time $T - 1$ and node j, the call option is the discounted expected value of the option at time T and nodes j and $j + 1$:

$$c_{T-1,j} = e^{-r\Delta t}[pc_{T,j+1} + (1 - p)c_{T,j}] \tag{6.32}$$

We then work backward through the tree until the current time.

For American options, the procedure is slightly different. At each point in time, the holder compares the value of the option *alive* and *dead* (i.e., exercised). The American call option value at node $T - 1, j$ is

$$C_{T-1,j} = Max[(S_{T-1,j} - K), c_{T-1,j}] \tag{6.33}$$

Example: Computing an American Option Value

Consider an at-the-money call on a foreign currency with a spot price of $100, a strike price of $K = \$100$, and a maturity of six months. The annualized volatility is $\sigma = 20\%$. The domestic interest rate is $r = 5\%$; the foreign rate is $r^* = 8\%$. Note that we require an income payment for the American feature to be valuable. If $r^* = 0$, we know that the American option is worth the same as a European option, which can be priced with the Black–Scholes model. There would be no point in using a numerical method.

First, we divide the period into four intervals, for instance, so that $\Delta t = 0.50/4 = 0.125$. The discounting factor over one interval is $e^{-r\Delta t} = 0.9938$. We then compute:

$$u = e^{\sigma\sqrt{\Delta t}} = e^{0.20\sqrt{0.125}} = 1.0733$$

$$d = (1/u) = 0.9317$$

$$a = e^{(r-r^*)\Delta t} = e^{(-0.03)0.125} = 0.9963$$

$$p = \frac{a - d}{u - d} = (0.9963 - 0.9317)/(1.0733 - 0.9317) = 0.4559$$

The procedure for pricing the option is detailed in Table 6.3. First, we lay out the tree for the spot price, starting with $S = 100$ at time $t = 0$, then $uS = 107.33$ and $dS = 93.17$ at time $t = 1$, and so on.

This allows us to value the European call. We start from the end, at time $t = 4$, and set the call price to $c = S - K = 132.69 - 100.00 = 32.69$ for the highest spot price, 15.19 for the next price and so on, down to $c = 0$ if the spot price is below $K = 100.00$. At the previous step and highest node, the value of the call is

$$c = 0.9938[0.4559 \times 32.69 + (1 - 0.4559) \times 15.19] = 23.02$$

Continuing through the tree to time 0 yields a European call value of \$4.43. The Black–Scholes formula gives an exact value of \$4.76. Note how close the binomial approximation is, with just four steps. A finer partition would quickly improve the approximation.

Next, we examine the American call. At time $t = 4$, the values are the same as above since the call expires. At time $t = 3$ and node $j = 4$, the option holder can either keep the call, in which case the value is still \$23.02, or exercise. When exercised, the option payoff is $S - K = 123.63 - 100.00 = 23.63$. Since this is greater than the value of the option alive, the holder should optimally exercise the option. We replace the European option value by \$23.63 at that node. Continuing through the tree in the same fashion, we find a starting value of \$4.74. The value of the American call is slightly greater than the European call price, as expected.

TABLE 6.3 Computation of American Option Value

	0	1	2	3	4
Spot Price S_t	→	→	→	→	→
					132.69
				123.63	115.19
			115.19	107.33	100.00
		107.33	100.00	93.17	86.81
	100.00	93.17	86.81	80.89	75.36
European Call c_t	←	←	←	←	←
					32.69
				23.02	15.19
			14.15	6.88	0.00
		8.10	3.12	0.00	0.00
	4.43	1.41	0.00	0.00	0.00
Exercised Call $S_t - K$					32.69
				23.63	15.19
			15.19	7.33	0.00
		7.33	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00
American Call C_t	←	←	←	←	←
					32.69
				23.63	15.19
			15.19	7.33	0.00
		8.68	3.32	0.00	0.00
	4.74	1.50	0.00	0.00	0.00

EXAMPLE 6.18: FRM EXAM 2006—QUESTION 86

Which of the following statements about American options is *incorrect*?

a. American options can be exercised at any time until maturity.

b. American options are always worth at least as much as European options.

c. American options can easily be valued with Monte Carlo simulation.

d. American options can be valued with binomial trees.

6.6 IMPORTANT FORMULAS

Payoff on a long call and put: $C_T = Max(S_T - K, 0)$, $P_T = Max(K - S_T, 0)$

Put–call parity: $c - p = Se^{-r^*\tau} - Ke^{-r\tau} = (F - K)e^{-r\tau}$

Bounds on call value (no dividends): $c_t \le C_t \le S_t$, $c_t \ge S_t - Ke^{-r\tau}$

Bounds on put value (no dividends): $p_t \le P_t \le K$, $p_t \ge Ke^{-r\tau} - S_t$

Geometric Brownian motion: $\ln(S_T) = \ln(S_0) + (\mu - \sigma^2/2)\tau + \sigma\sqrt{\tau}\,\epsilon$

Risk-neutral discounting formula: $f_t = E_{RN}[e^{-r\tau}F(S_T)]$

Black–Scholes call option pricing: $c = SN(d_1) - Ke^{-r\tau}N(d_2)$,
$$d_1 = \frac{\ln(S/Ke^{-r\tau})}{\sigma\sqrt{\tau}} + \frac{\sigma\sqrt{\tau}}{2}, \quad d_2 = d_1 - \sigma\sqrt{\tau}$$

Black–Scholes put option pricing: $p = S[N(d_1) - 1] - Ke^{-r\tau}[N(d_2) - 1]$

Black–Scholes pricing with dividend, Garman-Kohlhagen model:
$$c = Se^{-r^*\tau}N(d_1) - Ke^{-r\tau}N(d_2)$$

Black model, option on futures: $c = [FN(d_1) - KN(d_2)]e^{-r\tau}$

Margrabe model: Replace S by S_A, K by S_B, set $\sigma^2_{AB} = \sigma^2_A + \sigma^2_B - 2\rho_{AB}\sigma_A\sigma_B$

Pricing of binary option: $c = Qe^{-r\tau}N(d_2)$

Asian option: $c_T = Max(S_{AVE}(t, T) - K, 0)$

Binomial process: $u = e^{\sigma\sqrt{\Delta t}}$, $d = (1/u)$, $p = \frac{e^{\mu\Delta t} - d}{u - d}$

6.7 ANSWERS TO CHAPTER EXAMPLES

Example 6.1: FRM Exam 2007—Question 84

c. Buying a put creates a gain if the stock price falls, which is similar to selling the stock on the downside. On the upside, the loss is capped by buying a call.

Example 6.2: FRM Exam 2005—Question 72

c. By put–call parity, $c = p + Se^{-r^*\tau} - Ke^{-r\tau} = 3.19 + 23 - 25\exp(-0.05 \times 1) = 3.19 - 0.78 = 2.409$. Note that the volatility information is not useful.

Example 6.3: FRM Exam 2002—Question 25

d. Put–call parity applies to these European options. With no dividend, the relationship is $c - p = S - K\exp(-r\tau)$. The first term is $c - p = \$4 - \$1.47 = \$2.53$. The second term is $S - K\,exp(-r\tau) = \$20 - \$18\exp[-6\%(6/12)] = \$2.53$. Because the two numbers are the same, there is no arbitrage opportunity.

Example 6.4: FRM Exam 2006—Question 74

c. By put–call parity, $c - p = Se^{-r^*\tau} - Ke^{-r\tau}$. Therefore, $Se^{-r^*\tau} = (c - p + Ke^{-r\tau}) = (10 - 15 + 90\exp(0.05 \times 5) = 65.09$. The dividend yield is then $y = -(1/T)\ln(65.09/85) = 5.337\%$.

Example 6.5: FRM Exam 2001—Question 90

c. Long positions in options can lose at worst the premium, so b. is wrong. Spreads involve long and short positions in options and have limited downside loss, so a. is wrong. Writing options exposes the seller to very large losses. In the case of puts, the worst loss is the strike price K, if the asset price goes to zero. In the case of calls, however, the worst loss is in theory unlimited because there is a small probability of a huge increase in S. Between c. and d., c. is the better answer.

Example 6.6: FRM Exam 2007—Question 103

d. This position is graphed in Figure 6.6. It benefits from an increase in the price between 40 and 45, so it is a bull spread. The worst loss occurs below $K_1 = 40$, when none of the options is exercised and the net lost premium is $5 - 3 = 2$. The maximum profit occurs above $K_2 = 45$, when the two options are exercised, for a net profit of $5 minus the lost premium, which gives $3.

Example 6.7: FRM Exam 2006—Question 45

c. The manager is long a portfolio, which is protected by buying a put with low strike price and selling a call with higher strike price. This locks in a range of profits and losses and is a collar. If the strike prices were the same, the hedge would be perfect.

Example 6.8: FRM Exam 2002—Question 42

b. Because the final price is below the lowest of the three strike prices, all the puts will be exercised. The final payoff is $(\$50 - \$33) - 2(\$42 - \$33) + (\$37 - \$33) = \$17 - \$18 + \$4 = \$3$. From this, we have to deduct the up-front cost, which is $-\$7 + 2(\$4) - \$2 = -\$1$. The total profit is then, ignoring the time value of money, $\$3 - \$1 = \$2$ per share.

Example 6.9: FRM Exam 2003—Question 72

b. A strangle involves two different strike prices, so I. is incorrect. A long bull spread involves buying a call and selling a call with $K_1 < K_2$; the short position is inverted, so that II. is correct. Options with different maturities are called horizontal spreads, so answer III. is incorrect. A long butterfly spread indeed involves options with three strike prices, so IV. is correct. Hence, I. and III. are incorrect, and answer b. is the (correct) solution.

Example 6.10: FRM Exam 2002—Question 50

b. When there is no dividend, there is never any reason to exercise an American call early. Instead, the option should be sold to another party.

Example 6.11: FRM Exam 2005—Question 15

d. The upper bound is $S = 50$. The lower bound is $c \geq S - Ke^{-rt} = 50 - 45\exp(-0.03 \times 2) = 42.379$.

Example 6.12: FRM Exam 2005—Question 40

a. The BS formula relies on a method called risk-neutral valuation but does apply to the real world. Otherwise, it would be useless.

Example 6.13: FRM Exam 2001—Question 91

c. We use Equation (6.14) assuming there is no income payment on the asset. This gives $c = SN(d_1) - K\exp(-r\tau)N(d_2) = 100 \times 0.457185 - 110\exp(-0.1 \times 0.5) \times 0.374163 = \$6.568$.

Example 6.14: Probability of Exercise

d. This is the term multiplying the present value of the strike price, by Equation (6.17).

Example 6.15: FRM Exam 2003—Question 34

a. The payoff of an Asian option depends on the average value of S and therefore is path-dependent.

Example 6.16: FRM Exam 2006—Question 59

b. Lookback options use the maximum stock price over the period, which must be more than the value at the end. Hence, they must be more valuable than regular European options. Chooser options involve an additional choice during the life of the option, and as a result are more valuable than regular options. Asian options

involve the average, which is less volatile than the final price, so must be less expensive than regular options. Finally, barrier options can be structured so that the sum of two barrier options is equal to a regular option. Because each premium is positive, a barrier option must be less valuable than regular options.

Example 6.17: FRM Exam 2002—Question 19

b. A down-and-out call where the barrier has not been touched is still alive and hence benefits from an increase in S, so a. is incorrect. A down-and-in call only comes alive when the barrier is touched, so an increase in S brings it away from the barrier. This is not favorable, so b. is correct. An up-and-in put would benefit from an increase in S as this brings it closer to the barrier of $110, so c. is not correct. Finally, an up-and-in call would also benefit if S gets closer to the barrier.

Example 6.18: FRM Exam 2006—Question 86

c. This statement is incorrect because Monte Carlo simulations are strictly backward-looking, and cannot take into account optimal future exercise, which a binomial tree can do.
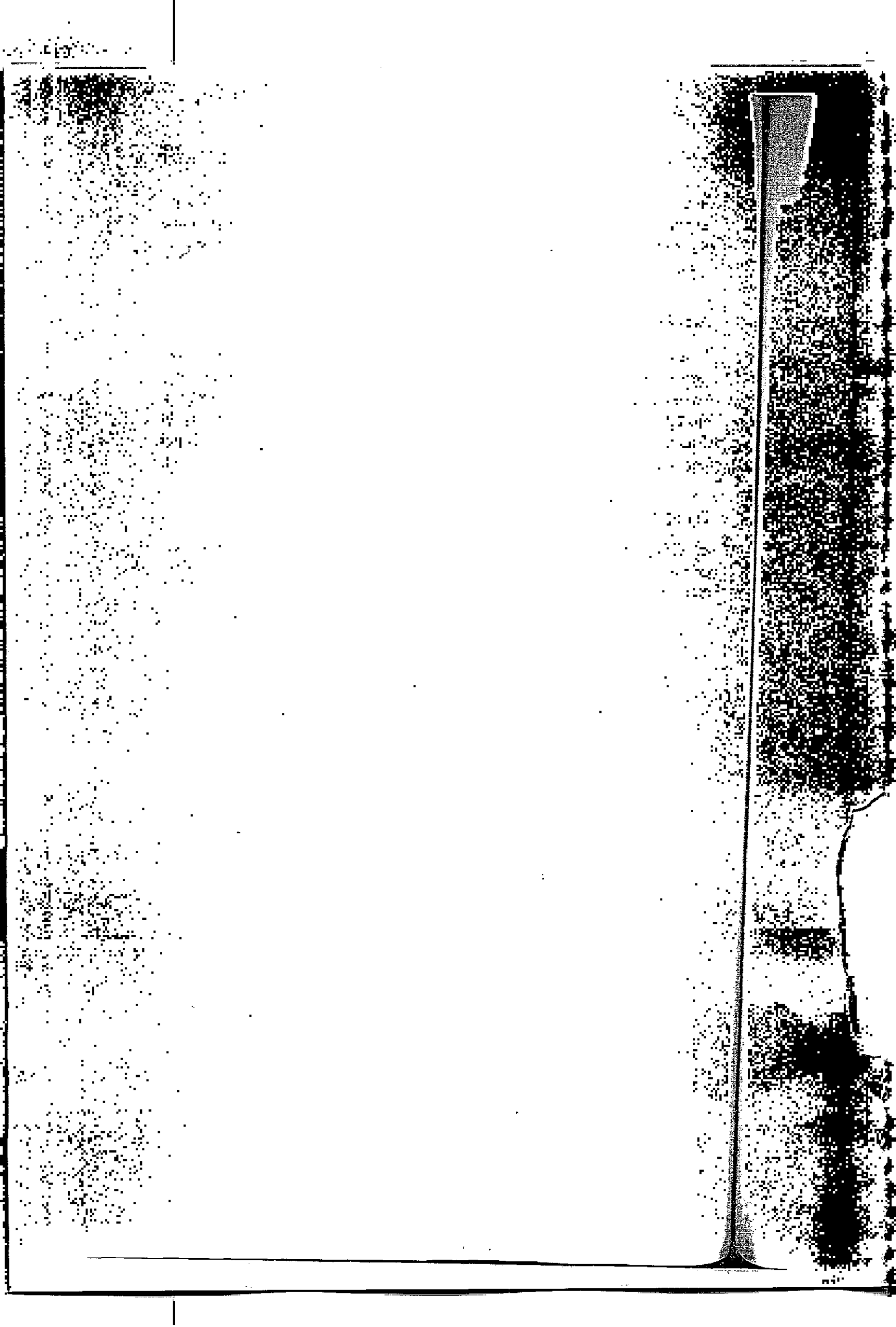
Fixed-Income Securities

The next two chapters provide an overview of fixed-income markets, securities, and their derivatives. Originally, fixed-income securities referred to bonds that promise to make fixed coupon payments. Over time, this narrow definition has evolved to include any security that obligates the borrower to make specific payments to the bondholder on specified dates. Thus, a bond is a security that is issued in connection with a borrowing arrangement. In exchange for receiving cash, the borrower becomes obligated to make a series of payments to the bondholder.

Fixed-income derivatives are instruments whose value derives from some bond price, interest rate, or other bond market variable. Due to their complexity, these instruments are analyzed in the next chapter.

Section 7.1 provides an overview of the different segments of the bond market. Section 7.2 then introduces the various types of fixed-income securities. Section 7.3 reviews the basic tools for analyzing fixed-income securities, including the determination of cash flows, the measurement of duration, and the term structure of interest rates and forward rates. Because of their importance, mortgage-backed securities (MBSs) are analyzed in great detail. MBSs are an example of securitization, which is the process by which assets are pooled and securities representing interests in the pool are issued. This topic is covered in Section 7.4. Section 7.5 then discusses MBSs and collateralized mortgage obligations (CMOs). These new structures illustrate the creativity of financial engineering.

7.1 OVERVIEW OF DEBT MARKETS

Fixed-Income markets are truly global. To help sort the various categories of the bond markets, Table 7.1 provides a broad classification of bonds by borrower and currency types. Bonds issued by resident entities and denominated in the domestic currency are called domestic bonds. In contrast, foreign bonds are those floated by a foreign issuer in the domestic currency and subject to domestic country regulations (e.g., by the government of Sweden in dollars in the United States). Eurobonds are mainly placed outside the country of the currency in which they are denominated and are sold by an international syndicate of financial institutions (e.g., a dollar-denominated bond issued by IBM and marketed in London).[1]

---

[1] These should not be confused with bonds denominated in the euro, which can be of any type.

TABLE 7.1 Classification of Bond Markets

	By resident	By non-resident
In domestic currency	Domestic Bond	Foreign Bond
In foreign currency	Eurobond	Eurobond

Foreign bonds and Eurobonds constitute the international bond market. Finally, global bonds are placed at the same time in the Eurobond and one or more domestic markets with securities fungible between these markets.

The domestic bond market can be further decomposed into these categories:

- Government bonds, issued by central governments, or also called sovereign bonds (e.g., by the United States in dollars)
- Government agency and guaranteed bonds, issued by agencies or guaranteed by the central government (e.g., by Fannie Mae, a U.S. government agency), which are public financial institutions
- State and local bonds, issued by local governments, other than the central government, also known as municipal bonds (e.g., by the state of California)
- Bonds issued by private financial institutions, including banks, insurance companies, or issuers of asset-backed securities (e.g., by Citibank in the U.S. market)
- Corporate bonds, issued by private nonfinancial corporations, including industrials and utilities (e.g., by IBM in the U.S. market)

Table 7.2 breaks down the world debt securities market, which was worth $80 trillion at the end of 2007. This includes the bond markets, traditionally defined as fixed-income securities with remaining maturities beyond one year, and the shorter-term money markets, with maturities below one year. The table includes all publicly tradable debt securities sorted by country of issuer and issuer type.

The table shows that U.S. entities have issued a total of $24.4 trillion in domestic debt and $4.8 trillion in international debt, for a total amount of $29.2 trillion, by far the biggest debt securities market. Next comes the Eurozone market, with a size of $20.0 trillion, and the Japanese market, with $9.0 trillion.

As Table 7.2 shows, the largest sector is for domestic government debt. This sector includes sovereign debt issued by emerging countries in their own currencies, e.g., Mexican peso-denominated debt issued by the Mexican government. Few of these markets have long-term issues, because of their history of high inflation, which renders long-term bonds very risky. In Mexico, for instance, the market consists mainly of Cetes, which are peso-denominated, short-term Treasury bills.

Among international debt, the emerging market sector also includes debt denominated in United States dollars, such as Brady bonds, which are sovereign bonds issued in exchange for bank loans, and the Tesebonos, which are dollar-denominated bills issued by the Mexican government. Brady bonds are hybrid

TABLE 7.2 Global Debt Securities Markets - 2007 (Billions of U.S. dollars)

Country of Issuer	Domestic	Type			Int'l	Total
		Gov't	Financials	Corporates		
United States	24,430	6,592	14,907	2,931	4,770	29,200
Japan	8,856	7,145	983	728	165	9,021
Germany	2,631	1,393	1,048	190	2,193	4,824
Italy	3,044	1,772	934	338	929	3,973
France	2,849	1,405	1,123	321	1,414	4,263
United Kingdom	1,359	903	433	23	3,160	4,519
Canada	1,144	734	277	133	433	1,577
Spain	1,641	496	600	545	1,224	2,865
Netherlands	993	397	519	77	1,633	2,626
South Korea	1,077	466	380	231	106	1,183
Belgium	550	388	113	49	207	757
China	1,687	1,137	446	104	21	1,708
Denmark	586	84	469	33	111	697
Australia	695	117	534	44	483	1,178
Brazil	953	694	251	8	82	1,035
Sweden	389	145	207	37	248	637
Switzerland	243	116	111	16	26	269
Austria	252	120	102	30	352	604
Eurozone	11,960	5,971	4,439	1,550	7,952	19,912
Subtotal	53,379	24,104	23,437	5,838	17,557	70,936
Others	3,794	2,662	650	483	5,157	8,951
Total	57,173	26,766	24,087	6,321	22,714	79,887

Source: Bank for International Settlements

securities whose principal is collateralized by U.S. Treasury zero-coupon bonds. As a result, there is no risk of default on the principal, unlike on coupon payments.

The domestic financial market is also important, especially for mortgage-backed securities. Mortgage-backed securities (MBSs), are securities issued in conjunction with mortgage loans, which are loans secured by the collateral of a specific real estate property. Payments on MBSs are repackaged cash flows supported by mortgage payments made by property owners. MBSs can be issued by government agencies as well as by private financial corporations. More generally, asset-backed securities (ABSs) are securities whose cash flows are supported by assets such as credit card receivables or car loan payments.

Finally, the remainder of the domestic market represents bonds raised by private, nonfinancial corporations. This sector, by far the largest in the United States, is growing rather quickly as corporations recognize that bond issuances are a lower-cost source of funds than bank debt. The advent of the common currency, the euro, is also leading to a growing, more liquid and efficient, corporate bond market in Europe.

7.2 FIXED-INCOME SECURITIES

7.2.1 Instrument Types

Bonds pay interest on a regular basis, semiannual for U.S. Treasury and corporate bonds, annual for others such as Eurobonds, or quarterly for others. The most common types of bonds are

- Fixed-coupon bonds, which pay a fixed percentage of the principal every period and the principal as a balloon, one-time payment at maturity
- Zero-coupon bonds, which pay no coupons but only the principal; their return is derived from price appreciation only
- Annuities, which pay a constant amount over time which includes interest plus amortization, or gradual repayment, of the principal
- Perpetual bonds or consols, which have no set redemption date and whose value derives from interest payments only
- Floating-coupon bonds, which pay interest equal to a reference rate plus a margin, reset on a regular basis; these are usually called floating-rate notes (FRNs)
- Structured notes, which have more complex coupon patterns to satisfy the investor's needs
- Inflation-protected notes, whose principal is indexed to the Consumer Price Index (CPI), hence providing protection against an increasing rate of inflation[2]

There are many variations on these themes. For instance, step-up bonds have fixed coupons that start at a low rate and increase over time.

It is useful to consider floating-rate notes in more detail. Take for instance a 10-year $100 million FRN paying semiannually six-month LIBOR in arrears.[3] Here, LIBOR is the London Interbank Offer Rate, a benchmark cost of borrowing for highly-rated (AA) credits. Every semester, on the reset date, the value of six-month LIBOR is recorded. Say LIBOR is initially at 6%. At the next coupon date, the payment will be $(\frac{1}{2}) \times \$100 \times 6\% = \$3$ million. Simultaneously, we record a new value for LIBOR, say 8%. The next payment will then increase to $4 million, and so on. At maturity, the issuer pays the last coupon plus the principal. Like a cork at the end of a fishing line, the coupon payment "floats" with the current interest rate.

---

[2] In the United States, these government bonds are called Treasury inflation-protected securities (TIPS). The coupon payment is fixed in real terms, say 3%. If after six months, the cumulative inflation is 2%, the principal value of the bond increases from $100 to $100 × (1 + 2%) = $102. The first semiannual coupon payment is then (3%/2) × $102 = $1.53.

[3] Note that the index could be defined differently. The floating payment could be tied to a Treasury rate, or LIBOR with a different maturity—say 3-month LIBOR. The pricing of the FRN will depend on the index. Also, the coupon will typically be set to LIBOR plus some spread that depends on the creditworthiness of the issuer.

APPLICATION: LIBOR AND OTHER BENCHMARK INTEREST RATES

LIBOR, the London Interbank Offer Rate, is a reference rate based on interest rates at which banks borrow unsecured funds from each other in the London interbank market.

LIBOR is published daily by the British Bankers' Association (BBA) around 11:45A.M., London Time, and is computed from an average of the distribution of rates provided by reporting banks. LIBID, the London Interbank Bid Rate, represents the average deposit rate.

LIBOR is calculated for 10 different currencies and various maturities, from overnight to one year. LIBOR rates are the most widely used reference rates for short-term futures contracts, such as the Eurodollar futures.

For the euro, however, EURIBOR, or Euro Interbank Offer Rate, is most often used. It is sponsored by the European Banking Federations (EBF) and published by Reuters at 11A.M., Central European Time (CET). In addition, EONIA (Euro Overnight Index Average) is an overnight unsecured lending rate that is published every day before 7P.M., CET. The same panel of banks contributes to EURIBOR and EONIA. The equivalent for sterling is SONIA (Sterling Overnight Index Average).

Among structured notes, we should mention inverse floaters, also known as reverse floaters, which have coupon payments that vary inversely with the level of interest rates. A typical formula for the coupon is $c = 12\% - $ LIBOR, if positive, payable semiannually. Assume the principal is \$100 million. If LIBOR starts at 6%, the first coupon will be $(1/2) \times \$100 \times (12\% - 6\%) = \$3$ million. If after six months LIBOR moves to 8%, the second coupon will be $(1/2) \times \$100 \times (12\% - 8\%) = \$2$ million. The coupon will go to zero if LIBOR moves above 12%. Conversely, the coupon will increase if LIBOR drops. Hence, inverse floaters do best in a falling interest rate environment.

Bonds can also be issued with option features. The most important are

- Callable bonds, where the issuer has the right to "call" back the bond at fixed prices on fixed dates, the purpose being to call back the bond when the cost of issuing new debt is lower than the current coupon paid on the bond
- Puttable bonds, where the investor has the right to "put" the bond back to the issuer at fixed prices on fixed dates, the purpose being to dispose of the bond should its price deteriorate
- Convertible bonds, where the bond can be converted into the common stock of the issuing company at a fixed price on a fixed date, the purpose being to partake in the good fortunes of the company (these will be covered in the chapter on equities)

The key to analyzing these bonds is to identify and price the option feature. For instance, a callable bond can be decomposed into a long position in a straight bond minus a call option on the bond price. The call feature is unfavorable for investors who require a lower price to purchase the bond, thereby increasing its yield. Conversely, a put feature will make the bond more attractive, increasing its price and lowering its yield. Similarly, the convertible feature allows companies to issue bonds at a lower yield than otherwise.

EXAMPLE 7.1: FRM EXAM 2003—QUESTION 85

With any other factors remaining unchanged, which of the following statements regarding bonds is *not* valid?

a. The price of a callable bond increases when interest rates increase.
b. Issuance of a callable bond is equivalent to a short position in a straight bond plus a long call option on the bond price.
c. The put feature in a puttable bond lowers its yield compared with the yield of an equivalent straight bond.
d. The price of an inverse floater decreases as interest rates increase.

7.2.2 Methods of Quotation

Most bonds are quoted on a clean price basis, that is, without accounting for the accrued income from the last coupon. For U.S. bonds, this clean price is expressed as a percent of the face value of the bond with fractions in thirty-seconds, for instance as 104-12, which means $104 + 12/32$, for the 6.25% May 2030 Treasury bond. Transactions are expressed in number of units, e.g., $20 million face value.

Actual payments, however, must account for the accrual of interest. This is factored into the gross price, also known as the dirty price, which is equal to the clean price plus accrued interest. In the U.S. Treasury market, accrued interest (AI) is computed on an *actual/actual* basis:

$$AI = \text{Coupon} \times \frac{\text{Actual number of days since last coupon}}{\text{Actual number of days between last and next coupon}} \qquad (7.1)$$

The fraction involves the actual number of days in both the numerator and denominator. For instance, say the 6.25% of May 2030 paid the last coupon on November 15 and will pay the next coupon on May 15. The denominator is, counting the number of days in each month, $15 + 31 + 31 + 29 + 31 + 30 + 15 = 182$. If the trade settles on April 26, there are $15 + 31 + 31 + 29 + 31 + 26 = 163$

days into the period. The accrued is computed from the $3.125 coupon as

$$\$3.125 \times \frac{163}{182} = \$2.798763$$

The total, gross price for this transaction is:

$$(\$20,000,000/100) \times [(104 + 12/32) + 2.798763] = \$21,434,753$$

Different markets have different day count conventions. A 30/360 convention, for example, considers that all months count for 30 days exactly. The computation of the accrued interest is tedious but must be performed precisely to settle the trades.

We should note that the accrued interest in the LIBOR market is based on *actual/360*. For instance, the interest accrued on a 6% $1 million loan over 92 days is

$$\$1,000,000 \times 0.06 \times \frac{92}{360} = \$15,333.33$$

Another notable pricing convention is the discount basis for Treasury bills. These bills are quoted in terms of an annualized discount rate (DR) to the face value, defined as

$$DR = (\text{Face} - P)/\text{Face} \times (360/t) \qquad (7.2)$$

where P is the price and t is the actual number of days. The dollar price can be recovered from

$$P = \text{Face} \times [1 - DR \times (t/360)] \qquad (7.3)$$

For instance, a bill quoted at a 5.19% discount with 91 days to maturity could be purchased for

$$\$100 \times [1 - 5.19\% \times (91/360)] = \$98.6881$$

This price can be transformed into a conventional yield to maturity, using

$$F/P = (1 + y \times t/365) \qquad (7.4)$$

which gives 5.33% in this case. Note that the yield is greater than the discount rate because it is a rate of return based on the initial price. Because the price is lower than the face value, the yield must be greater than the discount rate.

FIGURE 7.1   Time Profile of Cash Flows

This approach, however, ignores the shape of the term structure of interest rates. Short maturities, for example, could have much lower rates in which case it is inappropriate to use the same yield. We should really be discounting each cash flow at the zero-coupon rate that corresponds to each time period. Define R_t as the spot interest rate for maturity t and this risk class (i.e., same currency and credit risk). The fair value of the bond is then:

$$\hat{P} = \sum_{t=1}^{T} \frac{C_t}{(1 + R_t)^t} \qquad (7.7)$$

We can then check whether the market price is greater or lower. If the term structure is flat, the two approaches will be identical.

Alternatively, to assess whether a bond is rich or cheap, we can add a fixed amount SS, called the static spread to the spot rates so that the NPV equals the current price:

$$P = \sum_{t=1}^{T} \frac{C_t}{(1 + R_t + SS)^t} \qquad (7.8)$$

All else equal, a bond with a large static spread is preferable to another with a lower spread. It means the bond is cheaper, or has a higher expected rate of return.

It is simpler, but less accurate to compute a yield spread, YS, using yield to maturity, such that

$$P = \sum_{t=1}^{T} \frac{C_t}{(1 + y_T + YS)^t} \qquad (7.9)$$

Table 7.3 gives an example of a 7% coupon, two-year bond. The term structure environment, consisting of spot rates and par yields, is described on the left side.

TABLE 7.3 Bond Price and Term Structure

| Maturity (Year) i | Term Structure | | 7% Bond PVCF Discounted at | | |
	Spot Rate R_i	Par Yield y_i	Spot $SS = 0$	Yield+YS $\Delta y = 0.2386$	Spot + SS $SSs = 0.2482$
1	4.0000	4.0000	6.7308	6.5926	6.7147
2	6.0000	5.9412	95.2296	94.9074	94.7853
Sum			101.9604	101.5000	101.5000
Price			101.5000	101.5000	101.5000

The right side lays out the present value of the cash flows (PVCF). Discounting the two cash flows at the spot rates gives a fair value of $\hat{P} = \$101.9604$. In fact, the bond is selling at a price of $P = \$101.5000$. So, the bond is cheap.

We can convert the difference in prices to annual yields. The yield to maturity on this bond is 6.1798%, which implies a yield spread of YS $= 6.1798 - 5.9412 = 0.2386\%$. Using the static spread approach, we find that adding SS $= 0.2482\%$ to the spot rates gives the current price. The second measure is more accurate than the first.

Cash flows with different credit risks need to be discounted with different rates. For example, the principal on Brady bonds is collateralized by U.S. Treasury securities and carries no default risk, in contrast to the coupons. As a result, it has become common to separate the discounting of the principal from that of the coupons. Valuation is done in two steps. First, the principal is discounted into a present value using the appropriate Treasury yield. The present value of the principal is subtracted from the market value. Next, the coupons are discounted at what is called the stripped yield, which accounts for the credit risk of the issuer.

7.3.3 Duration

Armed with a cash flow profile, we can proceed to compute duration. As we have seen in Chapter 1, duration is a measure of the exposure, or sensitivity, of the bond price to movements in yields. When cash flows are fixed, duration is measured as the weighted maturity of each payment, where the weights are proportional to the present value of the cash flows. Using the same notations as in Equation (7.5), recall that Macaulay duration is

$$D = \sum_{t=1}^{T} t \times w_t = \sum_{t=1}^{T} t \times \frac{C_t/(1+y)^t}{\sum C_t/(1+y)^t} \qquad (7.10)$$

KEY CONCEPT

Duration can be viewed as the weighted average time to wait for each payment, but only when the bond's cash flows are predetermined.

More generally, duration is a measure of interest-rate exposure:

$$\frac{dP}{dy} = -\frac{D}{(1+y)}P = -D^*P \qquad (7.11)$$

where D^* is modified duration. The second term D^*P is also known as the dollar duration. Sometimes this sensitivity is expressed in dollar value of a basis point (also known as DV01), defined as

$$\frac{dP}{0.01\%} = \text{DVBP} \qquad (7.12)$$

For fixed cash flows, duration can be computed using Equation (7.10). Otherwise, we can infer duration from an economic analysis of the security. Consider a floating-rate note (FRN) with no credit risk. Just before the reset date, we know that the coupon will be set to the prevailing interest rate. The FRN is then similar to cash, or a money market instrument, which has no interest rate risk and hence is selling at par with zero duration. Just after the reset date, the investor is locked into a fixed coupon over the accrual period. The FRN is then economically equivalent to a zero-coupon bond with maturity equal to the time to the next reset date.

KEY CONCEPT

The duration of a floating-rate note is the time to wait until the next reset period, at which time the FRN should be at par.

EXAMPLE 7.2: CALLABLE BOND DURATION

A 10-year zero-coupon bond is callable annually at par (its face value) starting at the beginning of year six. Assume a flat yield curve of 10%. What is the bond duration?

a. 5 years

b. 7.5 years

c. 10 years

d. Cannot be determined based on the data given

EXAMPLE 7.3: DURATION OF FLOATERS

A money markets desk holds a floating-rate note with an eight-year maturity. The interest rate is floating at three-month LIBOR rate, reset quarterly. The next reset is in one week. What is the approximate duration of the floating-rate note?

    a. 8 years

    b. 4 years

    c. 3 months

    d. 1 week

7.4 SPOT AND FORWARD RATES

In addition to the cash flows, we also need detailed information on the term structure of interest rates to value fixed-income securities and their derivatives. This information is provided by spot rates, which are zero-coupon investment rates that start at the current time. From spot rates, we can infer forward rates, which are rates that start at a future date. Both are essential building blocks for the pricing of bonds.

Consider for instance a one-year rate that starts in one year. This forward rate is defined as $F_{1,2}$ and can be inferred from the one-year and two-year spot rates, R_1 and R_2. The forward rate is the break-even future rate that equalizes the return on investments of different maturities. An investor has the choice to lock in a two-year investment at the two-year rate, or to invest for a term of one year and roll over at the one-to-two year forward rate. The two portfolios will have the same payoff when the future rate $F_{1,2}$ is such that

$$(1 + R_2)^2 = (1 + R_1)(1 + F_{1,2}) \tag{7.13}$$

For instance, if $R_1 = 4.00\%$ and $R_2 = 4.62\%$, we have $F_{1,2} = 5.24\%$.

More generally, the T-period spot rate can be written as a geometric average of the spot and consecutive one-year forward rates

$$(1 + R_T)^T = (1 + R_1)(1 + F_{1,2}) \ldots (1 + F_{T-1,T}) \tag{7.14}$$

where $F_{i,i+1}$ is the forward rate of interest prevailing now (at time t) over a horizon of i to $i + 1$. This sequence is shown in Figure 7.2. Table 7.4 displays a sequence of spot rates, forward rates, and par yields, using annual compounding. The last column is the discount function, which is simply the current price of a dollar paid at t.

Alternatively, one could infer a series of forward rates for various maturities, all starting in one year

$$(1 + R_3)^3 = (1 + R_1)(1 + F_{1,3})^2, \ldots, (1 + R_T)^T = (1 + R_1)(1 + F_{1,T})^{T-1} \tag{7.15}$$

This defines a term structure in one year, $F_{1,2}, F_{1,3}, \ldots, F_{1,T}$.
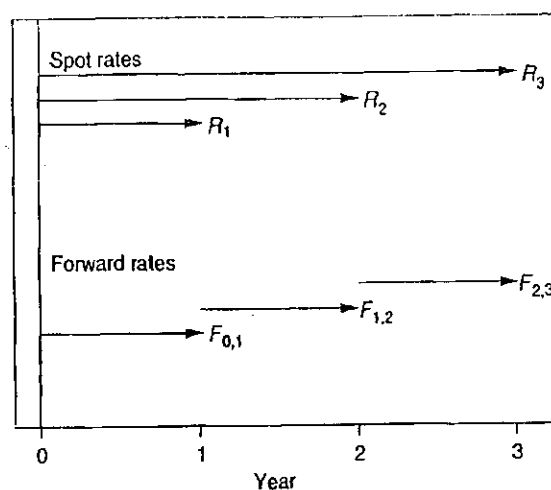
FIGURE 7.2 Spot and Forward Rates

The forward rate can be interpreted as a measure of the slope of the term structure. We can, for instance, expand both sides of Equation (7.13). After neglecting cross-product terms, we have

$$F_{1,2} \approx R_2 + (R_2 - R_1) \tag{7.16}$$

Thus, with an upward-sloping term structure, R_2 is above R_1, and $F_{1,2}$ will also be above R_2.

We can also show that in this situation, the spot rate curve is above the par yield curve. Consider a bond with two payments. The two-year par yield y_2 is implicitly defined from:

$$P = \frac{cF}{(1+y_2)} + \frac{(cF+F)}{(1+y_2)^2} = \frac{cF}{(1+R_1)} + \frac{(cF+F)}{(1+R_2)^2}$$

where P is set to par $P = F$. The par yield can be viewed as a weighted average of spot rates. In an upward-sloping environment, par yield curves involve coupons

TABLE 7.4 Spot Rates, Forward Rates, and Par Yields

Maturity (Year) i	Spot Rate R_i	Forward Rate $F_{i-1,i}$	Par Yield y_i	Discount Function $D(t_i)$
1	4.000	4.000	4.000	0.9615
2	4.618	5.240	4.604	0.9136
3	5.192	6.350	5.153	0.8591
4	5.716	7.303	5.640	0.8006
5	6.112	7.712	6.000	0.7433
6	6.396	7.830	6.254	0.6893
7	6.621	7.980	6.451	0.6383
8	6.808	8.130	6.611	0.5903
9	6.970	8.270	6.745	0.5452
10	7.112	8.400	6.860	0.5030

Interest rate
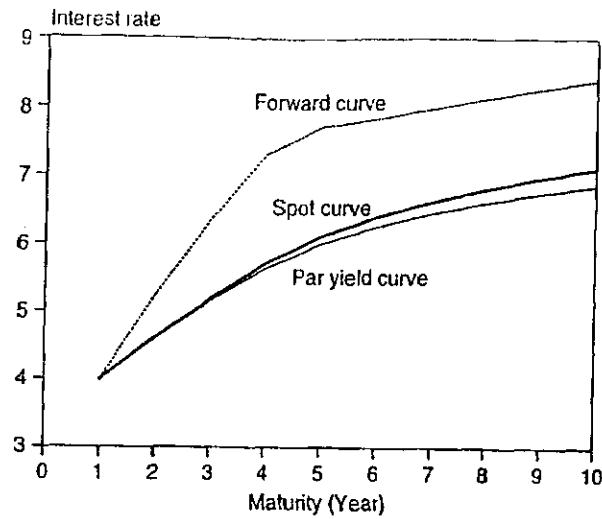
FIGURE 7.3   Upward-Sloping Term Structure
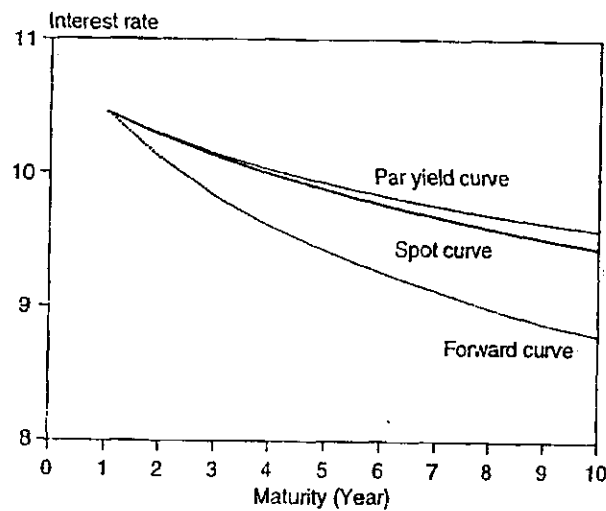
FIGURE 7.4   Downward-Sloping Term Structure

that are discounted at shorter and thus lower rates than the final payment. As a result, the par yield curve lies below the spot rate curve.[4] When the spot rate curve is flat, the spot curve is identical to the par yield curve and to the forward curve. In general, the curves differ. Figure 7.3 displays the case of an upward-sloping term structure. It shows the yield curve is below the spot curve while the forward curve is above the spot curve. With a downward-sloping term structure, as shown in Figure 7.4, the yield curve is above the spot curve, which is above the forward curve.

---

[4]For a formal proof, consider a two-period par bond with a face value of $1 and coupon of y_2. We can write the price of this bond as $1 = y_2/(1 + R_1) + (1 + y_2)/(1 + R_2)^2$. After simplification, this gives $y_2 = R_2(2 + R_2)/(2 + F_{1,2})$. In an upward-sloping environment, $F_{1,2} > R_2$ and thus $y_2 < R_2$.

Note that, because interest rates have to be positive, forward rates have to be positive, otherwise there would be an arbitrage opportunity.[5]

Forward rates allow us to project future cash flows that depend on future rates. The $F_{1,2}$ forward rate, for example, can be taken as the market's expectation of the second coupon payment on an FRN with annual payments and resets. We will also show later that positions in forward rates can be taken easily with derivative instruments.

As a result, the forward rate can be viewed as an expectation of the future spot rate. According to the expectations hypothesis

$$F_{1,2}^t = E(R_1^{t+1}) \tag{7.17}$$

This assumes that there is no risk premium embedded in forward rates. An upward-sloping term structure implies that short-term rates will rise in the future. In Figure 7.3, the forward curve traces out the path of future one-year spot rates.

If this hypothesis is correct, then it does not matter which maturity should be selected for investment purposes. Longer maturities benefit from higher coupons but will suffer a capital loss, due to the increase in rates, that will offset this benefit exactly.

KEY CONCEPT

In an upward-sloping term structure environment, the forward curve is above the spot curve, which is above the par yield curve. According to the expectations hypothesis, this implies a forecast for rising interest rates.

EXAMPLE 7.4: FRM EXAM 2007—QUESTION 32

The price of a three-year zero-coupon government bond is 85.16. The price of a similar four-year bond is 79.81. What is the one-year implied forward rate from year 3 to year 4?

a. 5.4%
b. 5.5%
c. 5.8%
d. 6.7%

---

[5] We abstract from transaction costs and assume we can invest and borrow at the same rate. For instance, $R_1 = 11.00\%$ and $R_2 = 4.62\%$ gives $F_{1,2} = -1.4\%$. This means that $(1 + R_1) = 1.11$ is greater than $(1 + R_2)^2 = 1.094534$. To take advantage of this discrepancy, we borrow $1 million for two years and invest it for one year. After the first year, the proceeds are kept in cash, or under the proverbial mattress, for the second period. The investment gives $1,110,000 and we have to pay back $1,094,534 only. This would create a profit of $15,466 out of thin air, which is highly unlikely in practice.

EXAMPLE 7.5: FRM EXAM 1999—QUESTION 1

Suppose that the yield curve is upward sloping. Which of the following statements is *true*?

a. The forward rate yield curve is above the zero-coupon yield curve, which is above the coupon-bearing bond yield curve.
b. The forward rate yield curve is above the coupon-bearing bond yield curve, which is above the zero-coupon yield curve.
c. The coupon-bearing bond yield curve is above the zero-coupon yield curve, which is above the forward rate yield curve.
d. The coupon-bearing bond yield curve is above the forward rate yield curve, which is above the zero-coupon yield curve.

EXAMPLE 7.6: FRM EXAM 2004—QUESTION 61

According to the pure expectations hypothesis, which of the following statements is *correct* concerning the expectations of market participants in an upward-sloping yield curve environment?

a. Interest rates will increase and the yield curve will flatten.
b. Interest rates will increase and the yield curve will steepen.
c. Interest rates will decrease and the yield curve will flatten.
d. Interest rates will decrease and the yield curve will steepen.

7.5 PREPAYMENT

7.5.1 Describing Prepayment Speed

So far, we considered fixed-income securities with fixed cash flows. In practice, many instruments have uncertain cash flows. Consider an investment in a traditional fixed-rate mortgage. The homeowner has the possibility of making early payments of principal. For the borrower, this represents a long position in an option. For the lender, this is a short position.

· In some cases, these prepayments are random, for instance when the homeowner sells the home due to changing job or family conditions. In other cases, these prepayments are more predictable. When interest rates fall, prepayments increase as homeowners can refinance at a lower cost. This also applies to callable

bonds, where the borrower has the option to call back its bonds at a fixed prices at fixed points in time. Generally, these factors affect mortgage refinancing patterns:

- *Spread between the mortgage rate and current rates*: Increases in the spread increase prepayments. Like a callable bond, there is a greater benefit to refinancing if it achieves a significant cost saving.
- *Age of the loan*: Prepayment rates are generally low just after the mortgage loan has been issued and gradually increase over time until they reach a stable, or "seasoned," level. This effect is known as seasoning.
- *Refinancing incentives*: The smaller the costs of refinancing, the more likely home-owners will refinance often.
- *Previous path of interest rates*: Refinancing is more likely to occur if rates have been high in the past but recently dropped. In this scenario, past prepayments have been low but should rise sharply. In contrast, if rates are low but have been so for a while, most of the principal will already have been prepaid. This path dependence is usually referred to as burnout.
- *Level of mortgage rates*: Lower rates increase affordability and turnover.
- *Economic activity*: An economic environment, where more workers change job location creates greater job turnover, which is more likely to lead to prepayments.
- *Seasonal effects*: There is typically more home-buying in the spring, leading to increased prepayments in early fall.

The prepayment rate is summarized into what is called the conditional prepayment rate (CPR), which is expressed in annual terms. This number can be translated into a monthly number, known as the single monthly mortality (SMM) rate using the adjustment:

$$(1 - SMM)^{12} = (1 - CPR) \qquad (7.18)$$

For instance, if CPR = 6% annually, the monthly proportion of principal paid early will be SMM $= 1 - (1 - 0.06)^{1/12} = 0.005143$, or 0.514% monthly. For a loan with a beginning monthly balance (BMB) of BMB = \$50,525 and a scheduled principal payment of SP = \$67, the prepayment will be $0.005143 \times (\$50,525 - \$67) = \$260$.

To price the mortgage, the portfolio manager should describe the schedule of projected prepayments during the remaining life of the bond. This depends on many factors, including the age of the loan.

Prepayments can be described using an industry standard, known as the Public Securities Association (PSA) prepayment model. The PSA model assumes a CPR of 0.2% for the first month, going up by 0.2% per month for the next 30 months, until 6% thereafter. Formally, this is:

$$CPR = Min[6\% \times (t/30), 6\%] \qquad (7.19)$$
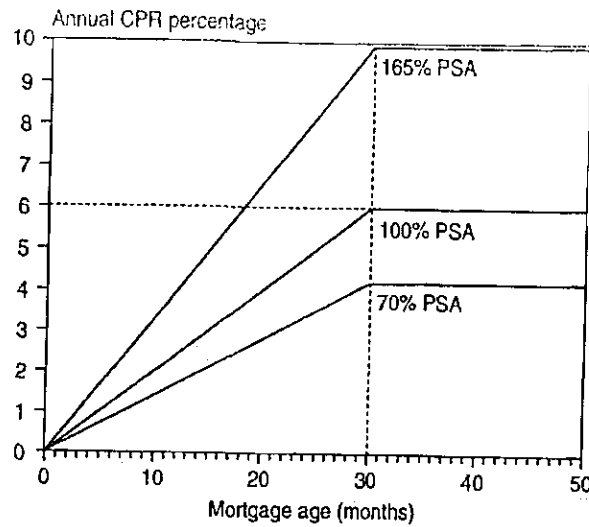
FIGURE 7.5   Prepayment Pattern

This pattern is described in Figure 7.5 as the 100% PSA speed. By convention, prepayment patterns are expressed as a percentage of the PSA speed, for example 165% for a faster pattern and 70% PSA for a slower pattern.

Example: Computing the CPR

Consider a mortgage issued 20 months ago with a speed of 150% PSA. What are the CPR and SMM?

The PSA speed is $Min[6\% \times (20/30), 6\%] = 0.04$. Applying the 150 factor, we have $CPR = 150\% \times 0.04 = 0.06$. This implies $SMM = 0.514\%$.

The next step is to project cash flows based on the prepayment speed pattern. Figure 7.6 displays cash-flow patterns for a 30-year loan with a face amount of $1 million and 6% interest rate. The horizontal, "No prepayment" line, describes the fixed annuity payment of $6,000 without any prepayment. The "100% PSA" line describes an increasing pattern of cash flows, peaking in 30 months and decreasing thereafter. This point corresponds to the stabilization of the CPR at 6%. This pattern is more marked for the "165% PSA" line, which assumes a faster prepayment speed.

Early prepayments create less payments later, which explains why the 100% PSA line is initially higher than the 0% line, then lower as the principal has been paid off more quickly.

7.5.2 Prepayment Risk

Like other fixed-income instruments, mortgages are subject to market risk, due to fluctuations in interest rates, and to credit risk, due to homeowner default. They are also, however, subject to prepayment risk, which is the risk that the principal will be repaid early.
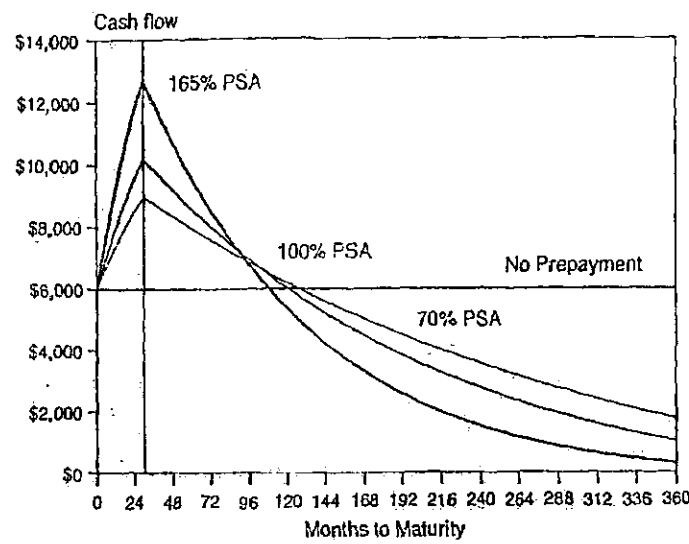
Cash flow

FIGURE 7.6 Cash Flows for Various PSA Speeds

Consider for instance an 8% mortgage, which is illustrated in Figure 7.7. If rates drop to 6%, homeowners will rationally prepay early to refinance the loan. Because the average life of the loan is shortened, this is called contraction risk. Conversely, if rates increase to 10%, homeowners will be less likely to refinance early, and prepayments will slow down. Because the average life of the loan is extended, this is called extension risk.

As shown in Figure 7.7, this feature creates "negative convexity" at point A. This reflects the fact that the investor in a mortgage is short an option. At point B, interest rates are very high and it is unlikely that the homeowner will refinance early. The option is nearly worthless and the mortgage behaves like a regular bond, with positive convexity.
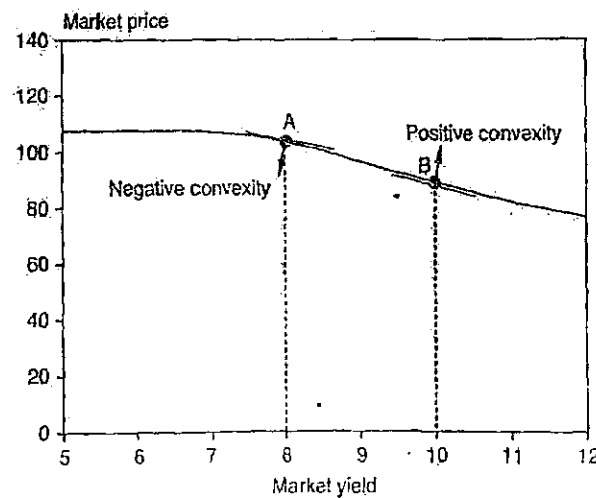
FIGURE 7.7 Negative Convexity of Mortgages

component. The approach most commonly used is the option-adjusted spread (OAS).

Starting from the static spread, the OAS method involves running simulations of various interest rate scenarios and prepayments to establish the option cost. The OAS is then

$$OAS = \text{Static spread} - \text{Option cost} \tag{7.22}$$

which represents the net richness or cheapness of the instrument. Within the same risk class, a security trading at a high OAS is preferable to others.

The OAS is more stable over time than the spread, because the latter is affected by the option component. This explains why during market rallies, i.e., when long-term Treasury yields fall, yield spreads on current coupon mortgages often widen. These mortgages are more likely to be prepaid early, which makes them less attractive. Their option cost increases, pushing up the yield spread.

EXAMPLE 7.7: FRM EXAM 1999—QUESTION 51

Suppose the annual prepayment rate CPR for a mortgage-backed security is 6%. What is the corresponding single-monthly mortality rate SMM?

a. 0.514%

b. 0.334%

c. 0.5%

d. 1.355%

EXAMPLE 7.8: FRM EXAM 2000—QUESTION 3

How would you describe the typical price behavior of a low premium mortgage pass-through security?

a. It is similar to a U.S. Treasury bond.

b. It is similar to a plain-vanilla corporate bond.

c. When interest rates fall, its price increase would exceed that of a comparable duration U.S. Treasury bond.

d. When interest rates fall, its price increase would lag that of a comparable duration U.S. Treasury bond.

EXAMPLE 7.9: FRM EXAM 2003—QUESTION 52

What bond type does the following price-yield curve represent and at which yield level is convexity equal to zero?
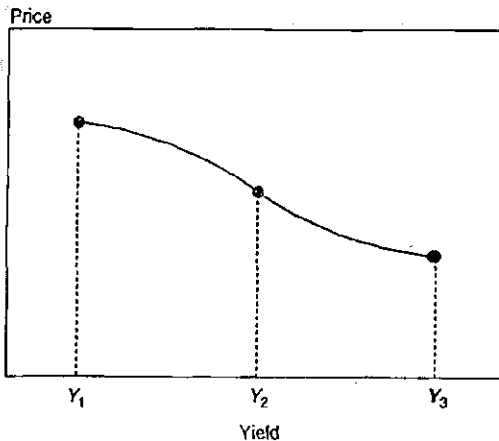
a. Puttable bond with convexity close to zero at Y_2.

b. Puttable bond with convexity close to zero at Y_1 and Y_3.

c. Callable bond with convexity close to zero at Y_2.

d. Callable bond with convexity close to zero at Y_1 and Y_3.

EXAMPLE 7.10: FRM EXAM 2006—QUESTION 93

You are analyzing two comparable (same credit rating, maturity, liquidity, rate) U.S. callable corporate bonds. The following data is available for the nominal spread over the U.S. Treasury yield curve and Z spread and option-adjusted spread (OAS) relative to the U.S. Treasury spot curve.

	X	Y
Nominal spread	145	130
Z spread	120	115
OAS	100	105

The nominal spread on the comparable option-free bonds in the market is 100 basis points. Which of the following statements is correct?

a. X only is undervalued.

b. Y only is undervalued.

c. X and Y both are undervalued.

d. Neither X nor Y is undervalued.

7.6 SECURITIZATION

7.6.1 Principles of Securitization

One problem with mortgage loans is that they are not tradable. In the past, they were originated and held by financial institutions such as savings and loans. This arrangement, however, concentrates risk in an industry that may not be able to hedge it efficiently. Also, it limits the amount of capital that can flow into mortgages. Mortgage-backed securities (MBSs) were created to solve this problem. MBSs are tradable securities that represent claims on pools of mortgage loans.[6]

This is an example of securitization, which is the process by which assets are pooled and securities representing interests in the pool are issued. These assets are created by an originator, or issuer.

The first step of the process is to create a new legal entity, called a special-purpose vehicle (SPV), or special-purpose entity (SPE). The originator then pools a group of assets and sells them to the SPV. In the next step, the SPV issues tradable claims, or securities, that are backed by the financial assets. Figure 7.8 describes a basic securitization structure.

A major advantage of this structure is that it shields the ABS investor from the credit risk of the originator. This requires, however, a clean sale of the assets to the SPV. Otherwise, the creditors of the originators might try to seize the SPV's assets in a bankruptcy proceeding. Other advantages are that pooling offers ready-made diversification across many assets.

The growth of securitization is being fueled by the disintermediation of banks as main providers of capital to everyone. When banks act as financial intermediaries, they raise funds (recorded as liabilities on the balance sheet) that are used for making loans (recorded as assets). With securitization, both assets and liabilities are removed from the balance sheet, requiring less equity capital to operate. Securitization provides regulatory capital relief if it enables the originator to hold proportionately less equity capital than otherwise. For instance, if the capital requirements for mortgages are too high, the bank will benefit from spinning off mortgage loans into securities because its required capital will drop sharply.

For the originator, securitization creates an additional source of funding. Securitization can also be used to manage the bank's risk profile. If the securitized assets have the same risk as the rest of the bank's assets, the relative risk of the bank is not changed, even though its size shrinks. In contrast, if the collateral is much riskier than the rest of the assets, the bank will have lowered its risk profile with securitization.

All sorts of assets can be included in ABSs, including mortgage loans, auto loans, student loans, credit card receivables, accounts receivables, and debt obligations. These assets are called collateral. In general, collecting payments on the collateral requires ongoing servicing activities. This is done by the servicing agent. Usually, the originator also performs the servicing, in exchange for a servicing fee.

---

[6] The MBS market was developed largely by Salomon Brothers in the early 1980s. This is described in a very entertaining book by Michael Lewis (1989), *Liar's Poker*, New York: Norton.
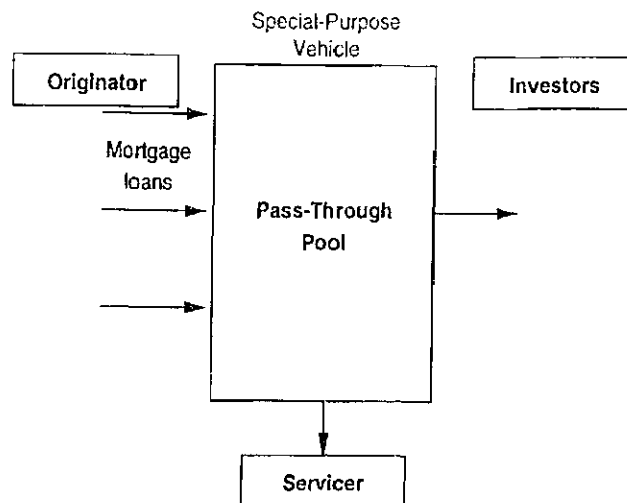
FIGURE 7.8 Securitization

The cash flows from the assets, minus the servicing fees, flow through the SPV to securities holders. When the securitization is structured as a pass-through, there is one class of bonds, and all investor receive the same proportional interests in the cash flows. When the SPV issues different classes of securities, the bonds are called tranches.[7] In addition, derivative instruments can be created to exchange claims on the ABS tranches, as we shall see in Chapter 22.

So far, we examined off–balance sheet securitizations. Another group is on–balance sheet securitizations, called covered bonds or Pfandbriefe in Germany. In these structures, the bank originates the loans and issues securities secured by these loans, which are kept on its books. Such structures are similar to secured corporate bonds, but have stronger legal protection in many European civil-law countries. Another difference is that investors have recourse against the bank in the case of defaults on the mortgages. Effectively, the bank provides a guarantee against credit risk.

In the case of MBS securitizations, the collateral consists of residential or commercial mortgage loans. These are called RMBS and CMBS, respectively. Their basic cash-flow patterns start from an annuity, where the homeowner makes a monthly fixed payment that covers principal and interest. As a result, the net present value of these cash flows is subject to interest rate risk, prepayment risk, and default risk.

In practice, however, most MBSs have third-party guarantees against credit risk. For instance, MBSs issued by Fannie Mae, an agency that is sponsored by the U.S. government, carry a guarantee of full interest and principal payment, even if the original borrower defaults. In this case, the government-sponsored enterprise (GSE) is the mortgage insurer. Such mortgage pass-throughs are sometimes called

---

[7] This is the French word for *slice*, as in a cake.

participation certificates. In contrast, private-label MBSs are exposed to credit risk, and may receive a credit rating.[8]

7.6.2 Issues with Securitization

The financial crisis that started in 2007 has highlighted serious deficiencies with the securitization process. Securitization allows banks to move assets off their balance sheets, freeing up capital and spreading the risk among many different investors. In theory, this provides real benefits.

In practice, failures in the originate-to-distribute model contributed to the credit crisis. The model failed in a number of key places, including underwriting, credit rating, and investor due diligence.

In the traditional banking model, banks or savings institutions underwrite mortgage loans and keep them on their balance sheet. This creates an incentive to screen loans carefully and to monitor their quality closely. In contrast, when a loan is securitized, the originator has less incentives to worry about the quality of the loans because its revenues depend on the volume of issuance. After all, another party bears the losses. This has led to a large increase in poor quality loans, in particular within the subprime category, which were packaged in securities and went bad rather quickly. This is a moral hazard problem, where an institution behaves differently than if it was fully exposed to the risk of the activity.

An additional problem arises when a bank decides to securitize a loan because of negative information about the borrower, rather than for diversification or capital relief reasons. This creates a systematic bias toward lower quality loans among securitized loans. To protect themselves, investors in asset-backed securities can require the issuing bank to retain some of the risk of the loans but this is not always sufficient. This is an adverse selection problem, due to asymmetric information. Indeed, Berndt and Gupta (2008) find that borrowers whose loans are sold in the secondary market have stock prices that underperform other firms.[9]

In addition, most of these loans ended up in complex securitization structures. Investors had relied blindly on credit ratings that turned out to be inaccurate and had not performed sufficient analysis of the risk of the underlying assets.

Finally, because it was so easy to securitize assets, the total amount of lending in the residential sector has gone up, creating additional demand for housing and perhaps pushing housing prices even further away from their fundamental value.

The securitization process also came back to haunt originating banks, many of which had created separate entities that invested in asset-backed securities. One such example are structured investment vehicles (SIVs), which are basically virtual banks, investing in asset-backed securities and funding themselves using

---

[8] Within this class, "Alt-A" loans contain nonstandard features but have borrowers of "A" credit-worthiness. At the low end of the credit scale, securities that are backed by subprime mortgage loans are classified as "home equity ABS" rather than MBS.

[9] An alternative explanation is that borrowers whose loans are sold are not subject to the discipline of bank monitoring, and undertake suboptimal investment. See Berndt, A. and A. Gupta (2008), Moral Hazard and Adverse Selection in the Originate-to-Distribute Model of Bank Credit, Working Paper, Carnegie Mellon University.
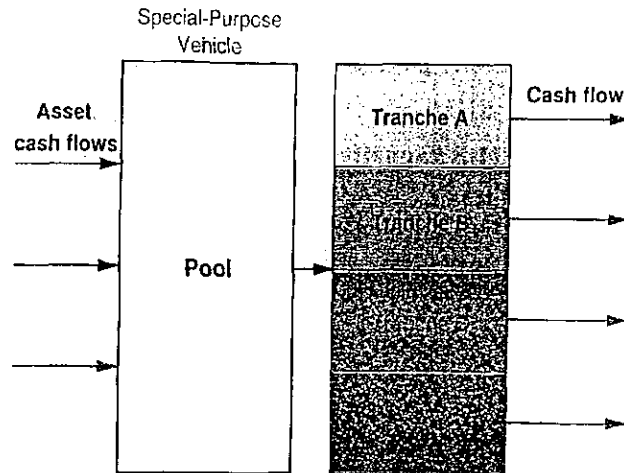
FIGURE 7.9 Tranching

short-term debt. During the credit crisis, investors in this debt became worried about the solvability of these SIVs and refused to roll over their investments. This forced many banks to absorb the assets of the failing SIVs they had sponsored back on their balance sheet, requiring them to raise additional equity capital in a particularly unfavorable environment. Thus securitization failed to provide the expected capital relief.

7.6.3 Tranching

In the case of MBSs, however, a major remaining risk is prepayment risk, which induces negative convexity. This feature is unattractive to investors who want fixed-income securities with predictable payments, so that they can match their liabilities.

In response, the industry has developed new classes of securities based on MBSs with more appealing characteristics. These are the collateralized mortgage obligations (CMOs), which are new securities that redirect the cash flows of an MBS pool to various segments.

Figure 7.9 illustrates this tranching process. The cash flows from the MBS pool go into the SPV, which issues different claims, or tranches with various characteristics. These are structured so that the cash flow from the first tranche, for instance, is more predictable than the original cash flows. The uncertainty is then pushed into the other tranches.

Starting from an MBS pool, financial engineering creates securities that are better tailored to investors' needs. It is important to realize, however, that the cash flows and risks are fully preserved. They are only redistributed across tranches. Whatever transformation is brought about, the resulting package must obey basic laws of conservation for the underlying securities and package of resulting securities.[10]

---

[10]As Lavoisier, the French chemist who was executed during the French revolution said, *Rien ne se perd, rien ne se crée (nothing is lost, nothing is created).*

At every single point in time, we must have the same cash flows going into and coming out of the SPV. As a result, we must have the same market value and the same risk profile. In particular, the weighted duration and convexity of the portfolio of tranches must add up to the original duration and convexity. If Tranche A has less convexity than the underlying securities, the other tranches must have more convexity.

Similar structures apply to collateralized bond obligations (CBOs), collateralized loan obligations (CLOs), and collateralized debt obligations (CDOs), which are a set of tradable bonds backed by bonds, loans, or debt (bonds and loans), respectively. These structures rearrange credit risk and will be explained in more detail in a later chapter.

KEY CONCEPT

Tranching rearranges the total cash flows, total value, and total risk of the underlying securities At all times, the total cash flows, value, and risk of the tranches must equal those of the collateral. If some tranches are less risky than the collateral, others must be more risky.

EXAMPLE 7.11: FRM EXAM 2000—QUESTION 13

A CLO is generally

a. A set of loans that can be traded individually in the market
b. A pass-through
c. A set of bonds backed by a loan portfolio
d. None of the above

EXAMPLE 7.12: FRM EXAM 2004—QUESTION 57

When evaluating asset-backed securitization issues, which of the following would be *least* important during the investor's analysis process?

a. The liability concentration levels of the asset originator
b. The structure of the underlying securitization transaction
c. The quality of the loan servicer for the underlying assets in the transaction
d. The quality of the underlying assets within the securitization structure

7.6.4 Tranching: Inverse Floaters

To illustrate the concept of tranching, we consider a simple example. with a two-tranche structure. The collateral consists of a regular five-year, 6% coupon $100 million note. This can be split up into a floater, that pays LIBOR on a notional of $50 million, and an inverse floater, that pays 12% − LIBOR on a notional of $50 million. Because the coupon C_{IF} on the inverse floater cannot go below zero, this imposes another condition on the floater coupon C_F. The exact formulas are:

$$\text{Coupon}_F = \text{Min}(\text{LIBOR}, 12\%) \quad \text{Coupon}_{IF} = \text{Max}(12\% - \text{LIBOR}, 0)$$

We verify that the outgoing cash flows exactly add up to the incoming flows. For each coupon payment, we have, in millions

$$\$50 \times \text{LIBOR} + \$50 \times (12\% - \text{LIBOR}) = \$100 \times 6\% = \$6$$

so this is a perfect match. At maturity, the total payments of twice $50 million add up to $100 million, so this matches as well.

We can also decompose the risk of the original structure into that of the two components. Assume a flat-term structure and say the duration of the original five-year note is $D = 4.5$ years. The portfolio dollar duration is:

$$\$50,000,000 \times D_F + \$50,000,000 \times D_{IF} = \$100,000,000 \times D$$

Just before a reset, the duration of the floater is close to zero $D_F = 0$. Hence, the duration of the inverse floater must be $D_{IF} = (\$100,000,000/\$50,000,000) \times D = 2 \times D$, or nine years, which is twice that of the original note. Note that the duration of the inverse floater is much greater than its maturity. This illustrates the point that duration is an interest rate sensitivity measure. When cash flows are uncertain, duration is not necessarily related to maturity. Intuitively, the first tranche, the floater, has zero risk so that all of the risk must be absorbed into the second tranche. The total risk of the portfolio is conserved.

This analysis can be easily extended to inverse floaters with greater leverage. Suppose the coupon is tied to twice LIBOR, for example 18% − 2 × LIBOR. The principal must be allocated in the amount x, in millions, for the floater and $100 - x$ for the inverse floater so that the coupon payment is preserved. We set

$$x \times \text{LIBOR} + (100 - x) \times (18\% - 2 \times \text{LIBOR}) = \$6$$

$$[x - 2(100 - x)] \times \text{LIBOR} + (100 - x) \times 18\% = \$6$$

Because LIBOR will change over time, this can only be satisfied if the term between brackets is always zero. This implies $3x - 200 = 0$, or $x = \$66.67$ million. Thus, two-thirds of the notional must be allocated to the floater, and one-third to the inverse floater. The inverse floater now has three times the duration of the original note.

EXAMPLE 7.13: FRM EXAM 1999—QUESTION 79

Suppose that the coupon and the modified duration of a 10-year bond priced to par are 6.0% and 7.5, respectively. What is the approximate modified duration of a 10-year inverse floater priced to par with a coupon of 18% − 2 × LIBOR?

a. 7.5
b. 15.0
c. 22.5
d. 0.0

EXAMPLE 7.14: FRM EXAM 2004—QUESTION 69

With LIBOR at 4%, a manager wants to increase the duration of his portfolio. Which of the following securities should he acquire to increase the duration of his portfolio the most?

a. A 10-year reverse floater that pays 8%− LIBOR, payable annually
b. A 10-year reverse floater that pays 12% − 2×LIBOR, payable annually
c. A 10-year floater that pays LIBOR, payable annually
d. A 10-year fixed rate bond carrying a coupon of 4% payable annually

EXAMPLE 7.15: FRM EXAM 2003—QUESTION 91

Which of the following statements most accurately reflects characteristics of a reverse floater (with no options attached)?

a. A portfolio of reverse floaters carries a marginally higher duration risk than a portfolio of similar maturity normal floaters.
b. A holder of a reverse floater can synthetically convert his position into a fixed rate bond by receiving floating and paying fixed on an interest rate swap.
c. A reverse floater hedges against rising benchmark yields.
d. A reverse floater's price changes by as much as that in a similar maturity fixed rate bond for a given change in yield.

7.6.5 Tranching: CMOs

When the collateral consists of mortgages, CMOs can be defined by prioritizing the payment of principal into different tranches. This is defined as sequential-pay tranches. Tranche A, for instance, will receive the principal payment on the whole underlying mortgages first. This creates more certainty in the cash flows accruing to Tranche A, which makes it more appealing to some investors. Of course, this is to the detriment of others. After principal payments to Tranche A are exhausted, Tranche B then receives all principal payments on the underlying MBS, and so on for other tranches.

Prepayment risk can be minimized further with a planned amortization class (PAC). All prepayment risk is then transferred to other bonds in the CMO structure, called support bonds. PAC bond[5] offer a fixed redemption schedule as long as prepayments on the collateral stay within a specific PSA range, say 100 to 250 PSA, called the PAC collar. When the structure is set up, the principal payment is set at the minimum payment of these two extreme values for every month of its life. Over time, this ensures a more stable pattern of payments.

Another popular construction is the IO/PO structure. This strips the MBS into two components. The interest-only (IO) tranche receives only the interest payments on the underlying MBS. The principal-only (PO) tranche then receives only the principal payments. As before, the market value of the IO and PO must exactly add to that of the MBS. Figure 7.10 describes the price behavior of the IO and PO. Note that the vertical addition of the two components always equals the value of the MBS.

To analyze the PO, it is useful to note that the sum of all principal payments is constant (because we have no default risk). Only the timing is uncertain. In contrast, the sum of all interest payments depends on the timing of principal payments. Later principal payments create greater total interest payments.

If interest rates fall, principal payments will come early, which reflects contraction risk. Because the principal is paid earlier and the discount rate decreases, the
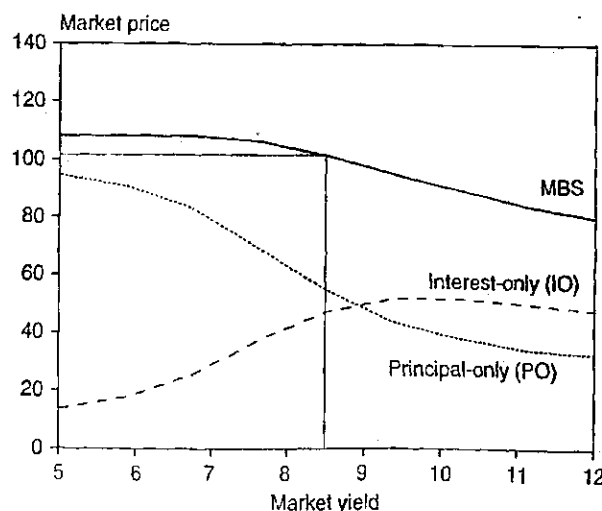
FIGURE 7.10   Creating an IO and PO from an MBS

PO should appreciate sharply in value. On the other hand, the faster prepayments mean less interest payments over the life of the MBS, which is unfavorable to the IO. The IO should depreciate.

Conversely, if interest rates rise, prepayments will slow down, which reflects extension risk. Because the principal is paid later and the discount rate increases, the PO should lose value. On the other hand, the slower prepayments mean more interest payments over the life of the MBS, which is favorable to the IO. The IO appreciates in value, up to the point where the higher discount rate effect dominates. Thus, IOs are bullish securities with negative duration, as shown in Figure 7.10.

EXAMPLE 7.16: FRM EXAM 2006—QUESTION 43

Which of the following mortgage-backed securities has a negative duration?

a. Interest-only strips (IO)
b. Inverse floater
c. Mortgage pass-through
d. Principal-only strips (PO)

EXAMPLE 7.17: FRM EXAM 2004—QUESTION 45

As the CRO of a firm specializing in MBSs, you have been asked to explain how interest-only (IO) strips and principal-only (PO) strips would react if interest rates change. Which of the following is *true*?

a. When interest rates fall, both PO and IO strips will increase in value.
b. When interest rates fall, POs will increase in value, IOs decrease in value.
c. When interest rates rise, POs will increase in value, IOs decrease in value.
d. When interest rates rise, both PO and IO strips will increase in value.

7.7 IMPORTANT FORMULAS

Quotation of Treasury bill as discount rate: $DR = (Face - P)/Face \times (360/t)$

Pricing using spot rate: $\hat{P} = \sum_{t=1}^{T} \frac{C_t}{(1+R_t)^t}$

Spot and forward rate: $(1 + R_2)^2 = (1 + R_1)(1 + F_{1,2})$, $\quad F_{1,2} \approx R_2 + (R_2 - R_1)$

Conditional prepayment rate (CPR), single monthly mortality (SMM) rate:

$$(1 - SMM)^{12} = (1 - CPR)$$

Public Securities Association (PSA) model: $CPR = Min[6\% \times (t/30), 6\%]$

7.8 ANSWERS TO CHAPTER EXAMPLES

Example 7.1: FRM Exam 2003—Question 95

a. Answer b. is valid because a short position in a callable bond is the same as a short position in a straight bond plus a long position in a call. (The issuer can call the bond back.) Answer c. is valid because a put is favorable for the investor, so lowers the yield. Answer d. is valid because an inverse floater has high duration.

Example 7.2: Callable Bond Duration

c. Because this is a zero-coupon bond, it will always trade below par, and the call should never be exercised. Hence, its duration is the maturity, 10 years.

Example 7.3: Duration of Floaters

d. Duration is not related to maturity when coupons are not fixed over the life of the investment. We know that at the next reset, the coupon on the FRN will be set at the prevailing rate. Hence, the market value of the note will be equal to par at that time. The duration or price risk is only related to the time to the next reset, which is one week here.

Example 7.4: FRM Exam 2007—Question 32

d. The forward rate can be inferred from $P_4 = P_3/(1 + F_{3,4})$, or $(1 + R_4)^4 = (1 + R_3)^3(1 + F_{3,4})$. Solving, this gives $F_{3,4} = (85.16/79.81) - 1 = 0.067$.

Example 7.5: FRM Exam 1999—Question 1

a. See Figures 7.3 and 7.4. The coupon yield curve is an average of the spot, zero-coupon curve, hence has to lie below the spot curve when it is upward-sloping. The forward curve can be interpreted as the spot curve plus the slope of the spot curve. If the latter is upward-sloping, the forward curve has to be above the spot curve.

Example 7.6: FRM Exam 2004—Question 61

a. An upward-sloping term structure implies forward rates higher than spot rates, or that short-term rates will increase. Because short-term rates increase more than long-term rates, this implies a flattening of the yield curve.

Example 7.7: FRM Exam 1999—Question 51

a. Using $(1 - 6\%) = (1 - SMM)^{12}$, we find $SMM = 0.51\%$.

Example 7.8: FRM Exam 2000—Question 3

d. MBSs are unlike regular bonds, Treasuries, or corporates, because of their negative convexity. When rates fall, homeowners prepay early, which means that the price appreciation is less than that of comparable duration regular bonds.

Example 7.9: FRM Exam 2003—Question 52

c. This has to be a callable bond because the price is capped if rates fall, reflecting the fact that the borrower would call back the bond. At Y_1, convexity is negative, at Y_2, close to zero.

Example 7.10: FRM Exam 2006—Question 93

b. The nominal spreads and Z spreads do not take into account the call option, Instead, comparisons should focus on the OAS, which is higher for Y, and also higher than the 100bp for option-free bonds.

Example 7.11: FRM Exam 2000—Question 13

c. Like a CMO, a CLO represents a set of tradable securities backed by some collateral, in this case a loan portfolio.

Example 7.12: FRM Exam 2004—Question 57

a. Bankruptcy by the originator would not affect the SPV, so the financial condition of the originator is the least important factor. All of the other factors would be important in evaluating the securitization.

Example 7.13: FRM Exam 1999—Question 79

c. Following the same reasoning as above, we must divide the fixed-rate bonds into 2/3 FRN and 1/3 inverse floater. This will ensure that the inverse floater payment is related to twice LIBOR. As a result, the duration of the inverse floater must be 3 times that of the bond.

Example 7.14: FRM Exam 2004—Question 69

b. The duration of a floater is about zero. The duration of a 10-year regular bond is about nine years. The first reverse floater has a duration of about $2 \times 9 = 18$ years, the second, $3 \times 9 = 27$ years.

Example 7.15: FRM Exam 2003—Question 91

b. The duration of a reverse floater is higher than that of a FRN, which is close to zero, or even than that of a fixed-date bond with the same maturity. So, answers a. and d. are wrong. It loses money when yields rise, so c. is wrong. A reverse floater is equivalent as a long position in a fixed-rate bond plus a receive-fixed/pay-floating swap. Hence, b. is correct.

Example 7.16: FRM Exam 2006—Question 43

a. IOs increase in value as interest rates increase because in this scenario, there will be less prepayment of mortgages. Less early payment means more total interest payments, which increases the value of the IO.

Example 7.17: FRM Exam 2004—Question 45

b. POs have positive duration, IOs negative. Hence, they react in opposite directions to falls in interest rates.

Fixed-Income Derivatives

This chapter turns to the analysis of fixed-income derivatives. These are instruments whose value derives from a bond price, interest rate, or other bond market variable. As discussed in Chapter 5, fixed-income derivatives account for the largest proportion of the global derivatives markets. Understanding fixed-income derivatives is also important because many fixed-income securities have derivative-like characteristics.

This chapter focuses on the use of fixed-income derivatives, as well as their pricing. Pricing involves finding the fair market value of the contract. For risk management purposes, however, we also need to assess the range of possible movements in contract values. This will be further examined in the chapters on market risk (Chapters 10–15) and in Chapter 21 on credit exposure.

This chapter presents the most important interest rate derivatives and discusses fundamentals of pricing. Section 8.1 discusses interest rate forward contracts, also known as forward rate agreements. Section 8.2 then turns to the discussion of interest rate futures, covering Eurodollar and Treasury bond futures. Although these products are dollar-based, similar products exist on other capital markets. Swaps are analyzed in Section 8.3. Swaps are very important instruments due to their widespread use. Finally, interest rate options are covered in Section 8.4, including caps and floors, swaptions, and exchange-traded options.

8.1 FORWARD CONTRACTS

Forward Rate Agreements (FRAs) are over-the-counter financial contracts that allow counterparties to lock in an interest rate starting at a future time. The buyer of an FRA locks in a borrowing rate, the seller locks in a lending rate. In other words, the "long" benefits from an increase in rates and the "short" benefits from a fall in rates.

As an example, consider an FRA that settles in one month on three-month LIBOR. Such an FRA is called 1×4. The first number corresponds to the first settlement date, the second to the time to final maturity. Call τ the period to which LIBOR applies, three months in this case. On the settlement date, in one month, the payment to the long involves the net value of the difference between the spot
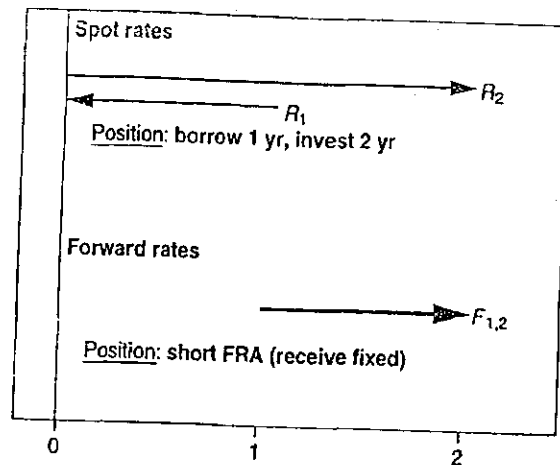
FIGURE 8.1 Decomposition of a Short FRA Position

rate S_T (the prevailing three-month LIBOR rate) and of the locked-in forward rate F. The payoff is $S_T - F$, as with other forward contracts, present valued to the first settlement date. This gives

$$V_T = (S_T - F) \times \tau \times \text{Notional} \times \text{PV}(\$1) \tag{8.1}$$

where $\text{PV}(\$1) = \$1/(1 + S_T\tau)$. The amount is settled in cash.

Figure 8.1 shows that a *short* position in an FRA is equivalent to borrowing short-term to finance a long-term investment. In both cases, there is no up-front investment. The duration is equal to the difference between the durations of the two legs, and can be inferred from the derivative of Equation (8.1). The duration of a short FRA is τ. Its dollar duration is $DD = \tau \times \text{Notional} \times \text{PV}(\$1)$.

Example: Using an FRA

A company will receive $100 million in six months to be invested for a six-month period. The Treasurer is afraid rates will fall, in which case the investment return will be lower. The company needs to take a position that will offset this loss by generating a gain when rates fall. Because a short FRA gains when rates fall, the Treasurer needs to *sell* a 6 × 12 FRA on $100 million at the rate of, say, $F = 5\%$. This locks in an investment rate of 5% starting in six months.

When the FRA expires in six months, assume that the prevailing six-month spot rate is $S_T = 3\%$. This will lower the investment return on the cash received, which is the scenario the Treasurer feared. Using Equation (8.1), the FRA has a payoff of $V_T = -(3\% - 5\%) \times (6/12) \times \$100 \text{ million} = \$1,000,000$, which multiplied by the 4% present value factor gives $980,392$. In effect, this payment offsets the lower return that the company received on a floating investment, guaranteeing a return equal to the forward rate. This contract is also equivalent to borrowing the present value of $100 million for six months and investing the proceeds for 12 months.

8.2 FUTURES

Whereas FRAs are over-the-counter contracts, futures are traded on organized exchanges. We will cover the most important types of futures contracts, Eurodollar and T-bond futures.

8.2.1 Eurodollar Futures

Eurodollar futures are futures contracts tied to a forward LIBOR rate. Since their creation on the Chicago Mercantile Exchange, Eurodollar futures have spread to equivalent contracts such as Euribor futures (denominated in euros),[1] Euroyen futures (denominated in Japanese yen), and so on. These contracts are akin to FRAs involving three-month forward rates starting on a wide range of dates, up to 10 years into the future.

The formula for calculating the value of one contract is

$$P_t = 10,000 \times [100 - 0.25(100 - FQ_t)] = 10,000 \times [100 - 0.25 F_t] \qquad (8.2)$$

where FQ_t is the quoted Eurodollar futures price. This is quoted as 100.00 minus the interest rate F_t, expressed in percent, that is, $FQ_t = 100 - F_t$. The 0.25 factor represents the three-month maturity, or 0.25 years. For instance, if the market quotes $FQ_t = 94.47$, we have $F_t = 100 - 94.47 = 5.53$, and the contract value is $P = 10,000[100 - 0.25 \times 5.53] = \$986,175$. At expiration, the contract value settles to

$$P_T = 10,000 \times [100 - 0.25 S_T] \qquad (8.3)$$

where S_T is the three-month Eurodollar spot rate prevailing at T. Payments are cash settled.

As a result, F_t can be viewed as a three-month forward rate that starts at the maturity of the futures contract. The formula for the contract price may look complicated but in fact is structured so that an increase in the interest rate leads to a decrease in the price of the contract, as is usual for fixed-income instruments. Also, because the change in the price is related to the interest rate by a factor of 0.25, this contract has a constant duration of three months. It is useful to remember that the DV01 is $\$10,000 \times 0.25 \times 0.01 = \$25$.

Chapter 5 has explained that the pricing of forwards is similar to those of futures, except when the value of the futures contract is strongly correlated with the reinvestment rate. This is the case with Eurodollar futures.

Interest rate futures contracts are designed to move like a bond, that is, to lose value when interest rates increase. The correlation is negative. This implies that when interest rates rise, the futures contract loses value and in addition funds have to be provided precisely when the borrowing cost or reinvestment rate is higher.

---

[1] Euribor futures are based on the European Bankers Federations' Euribor Offered Rate (EBF Euribor).

Example: Using Eurodollar Futures

As in the previous section, the Treasurer wants to hedge a future investment of $100 million in six months for a six-month period. The company needs to take a position that will offset the earnings loss by generating a gain when rates fall. Because a long Eurodollar futures position gains when rates fall, the Treasurer should *buy* Eurodollar futures.

If the futures contract trades at $FQ_t = 95.00$, the dollar value of one contract is $P = 10,000 \times [100 - 0.25(100 - 95)] = \$987,500$. The Treasurer needs to buy a suitable number of contracts that will provide the best hedge against the loss of earnings. The computation of this number will be detailed in a future chapter.

Conversely when rates drop, the contract gains value and the profits can be withdrawn but are now reinvested at a lower rate. Relative to forward contracts, this marking-to-market feature is *disadvantageous* to long futures positions. This has to be offset by a *lower* futures contract value. Given that the value is negatively related to the futures rate, by $P_t = 10,000 \times [100 - 0.25 \times F_t]$, this implies a *higher* Eurodollar futures rate F_t.

The difference is called the convexity adjustment and can be described as[2]

$$\text{Futures Rate} = \text{Forward Rate} + (1/2)\sigma^2 t_1 t_2 \qquad (8.4)$$

where σ is the volatility of the change in the short-term rate, t_1 is the time to maturity of the futures contract, and t_2 is the maturity of the rate underlying the futures contract.

Example: Convexity Adjustment

Consider a 10-year Eurodollar contract, for which $t_1 = 10$, $t_2 = 10.25$. The maturity of the futures contract itself is 10 years and that of the underlying rate is 10 years plus three months.

Typically, $\sigma = 1\%$, so that the adjustment is $(1/2)0.01^2 \times 10 \times 10.25 = 0.51\%$. So, if the forward price is 6%, the equivalent futures rate would be 6.51%. Note that the effect is significant for long maturities only. Changing t_1 to one year and t_2 to 1.25, for instance, reduces the adjustment to 0.006%, which is negligible.

8.2.2 T-Bond Futures

T-bond futures are futures contracts tied to a pool of Treasury bonds that consists of all bonds with a remaining maturity greater than 15 years (and noncallable within 15 years). Similar contracts exist on shorter rates, including 2-, 5-, and

---

[2] This formula is derived from the Ho-Lee model. See for instance Hull (2000), *Options, Futures, and Other Derivatives*, Upper Saddle River, NJ: Prentice-Hall.

10-year Treasury notes. Government bond futures also exist in other markets, including Canada, the United Kingdom, the Eurozone, and Japan.

Futures contracts are quoted like T-bonds, for example 97-02, in percent plus thirty seconds, with a notional of $100,000. Thus the price of the contract is $P = \$100,000 \times (97 + 2/32)/100 = \$97,062.50$. The next day, if yields go up and the quoted price falls to 96-0, the new value is $96,000, and the loss on the long position is $P_2 - P_1 = -\$1,062.50$.

It is important to note that the T-bond futures contract is settled by physical delivery. To ensure interchangeability between the deliverable bonds, the futures contract uses a conversion factor (CF) for delivery. This factor multiplies the futures price for payment to the short. The goal of the CF is to attempt to equalize the net cost of delivering the various eligible bonds.

The conversion factor is needed because bonds trade at widely different prices. High coupon bonds trade at a premium, low coupon bonds at a discount. Without this adjustment, the party with the short position (the "short") would always deliver the same, cheap bond and there would be little exchangeability between bonds. Exchangeability is an important feature, however, as it minimizes the possibility of market squeezes. A squeeze occurs when holders of the short position cannot acquire or borrow the securities required for delivery under the terms of the contract.

So, the "short" buys a bond, delivers it, and receives the quoted futures price times a conversion factor that is specific to the delivered bond (plus accrued interest). The short should rationally pick the bond that minimizes the net cost,

$$\text{Cost} = \text{Price} - \text{Futures Quote} \times \text{CF} \tag{8.5}$$

The bond with the lowest net cost is called cheapest to deliver (CTD).

In practice, the CF is set by the exchange at initiation of the contract for each bond. It is computed by discounting the bond's cash flows at a notional 6% rate, assuming a flat term structure. Take for instance the 7 5/8% of 2025. The CF is computed as

$$CF = \frac{(7.625\%/2)}{(1 + 6\%/2)^1} + \cdots + \frac{(1 + 7.625\%/2)}{(1 + 6\%/2)^T} \tag{8.6}$$

which gives CF = 1.1717. High coupon bonds have higher CFs. Also, because the coupon is greater than 6%, the CF is greater than 1.

The net cost calculations are illustrated in Table 8.1 for three bonds. The net cost for the first bond in the table is $\$104.375 - 110.8438 \times 0.9116 = \$3.330$. For the 6% coupon bond, the CF is exactly unity. The net cost for the third bond in the table is $1.874. Because this is the lowest entry, this bond is the CTD for this group. Note how the CF adjustment brings the cost of all bonds much closer to each other than their original prices.

The adjustment is not perfect when current yields are far from 6%, or when the term structure is not flat, or when bonds do not trade at their theoretical prices. Assume for instance that we operate in an environment where yields are flat at

TABLE 8.1 Calculation of CTD

Bond	Price	Futures	CF	Cost
5 1/4% Nov 2028	104.3750	110.8438	0.9116	3.330
6% Feb 2026	112.9063	110.8438	1.0000	2.063
7 5/8% Feb 2025	131.7500	110.8438	1.1717	1.874

5% and all bonds are priced at par. Discounting at 6% will create CF factors that are lower than 1; the longer the maturity of the bond, the greater the difference. The net cost $P - F \times CF$ will then be greater for longer-term bonds. This tends to favor short-term bonds for delivery. When the term structure is upward sloping, the opposite occurs, and there is a tendency for long-term bonds to be delivered.

As a first approximation, this CTD bond drives the characteristics of the futures contract. As before, the equilibrium futures price is given by

$$F_t e^{-r\tau} = S_t - PV(D) \qquad (8.7)$$

where S_t is the gross price of the CTD and $PV(D)$ is the present value of the coupon payments. This has to be further divided by the conversion factor for this bond. The duration of the futures contract is also given by that of the CTD. In fact, this relationship is only approximate because the short has an *option* to deliver the cheapest of a group of bonds. The value of this delivery option should depress the futures price because the party who is long the futures is also short the option. As a result, he requires a lower acquisition price. Unfortunately, this complex option is not easy to evaluate.

EXAMPLE 8.4: FRM EXAM 2005—QUESTION 49

John H., a portfolio manager, is shorting a U.S. Treasury bond futures contract and has decided to deliver. The quoted futures price is USD 95.5.
Among the four deliverable bonds, which is the cheapest-to-deliver?

Bond	A	B	C	D
Quote	125.69	90.31	87.6	128.56
Conversion Factor	1.1979	0.8109	0.8352	1.2249

a. Bond A

b. Bond B

c. Bond C

d. Bond D

EXAMPLE 8.5: FRM EXAM 2007—QUESTION 80

Consider an FRA (forward rate agreement) with the same maturity and compounding frequency as a Eurodollar futures contract. The FRA has a LIBOR underlying. Which of the following statements are true about the relationship between the forward rate and the futures rate?

a. The forward rate is normally higher than the futures rate.
b. They have no fixed relationship.
c. The forward rate is normally lower than the futures rate.
d. They should be exactly the same.

8.3 SWAPS

Swaps are agreements by two parties to exchange cash flows in the future according to a prearranged formula. Interest rate swaps have payments tied to an interest rate. The most common type of swap is the fixed-for-floating swap, where one party commits to pay a fixed percentage of notional against a receipt that is indexed to a floating rate, typically LIBOR. The risk is that of a change in the level of rates.

Other types of swaps are basis swaps, where both payments are indexed to a floating rate. For instance, the swap can involve exchanging payments tied to three-month LIBOR against a three-month Treasury bill rate. The risk is that of a change in the spread between the reference rates.

8.3.1 Instruments

Consider two counterparties, A and B, that can raise funds either at fixed or floating rates, $100 million over 10 years. A wants to raise floating, and B wants to raise fixed.

Table 8.2a displays capital costs. Company A has an absolute advantage in the two markets as it can raise funds at rates systematically lower than B. Company

TABLE 8.2a Cost of Capital Comparison

Company	Fixed	Floating
A	10.00%	LIBOR + 0.30%
B	11.20%	LIBOR + 1.00%

TABLE 8.2b Swap to Company A

Operation	Fixed	Floating
Issue debt	Pay 10.00%	
Enter swap	Receive 10.00%	Pay LIBOR + 0.05%
Net		Pay LIBOR + 0.05%
Direct cost		Pay LIBOR + 0.30%
Savings		0.25%

TABLE 8.2c Swap to Company B

Operation	Floating	Fixed
Issue debt	Pay LIBOR + 1.00%	
Enter swap	Receive LIBOR + 0.05%	Pay 10.00%
Net		Pay 10.95%
Direct cost		Pay 11.20%
Savings		0.25%

A, however, has a comparative advantage in raising fixed as the cost is 1.2% lower than for B. In contrast, the cost of raising floating is only 0.70% lower than for B. Conversely, Company B must have a comparative advantage in raising floating.

This provides a rationale for a swap that will be to the mutual advantage of both parties. If both companies directly issue funds in their final desired market, the total cost will be LIBOR + 0.30% (for A) and 11.20% (for B), for a total of LIBOR + 11.50%. In contrast, the total cost of raising capital where each has a comparative advantage is 10.00% (for A) and LIBOR + 1.00% (for B), for a total of LIBOR + 11.00%. The gain to both parties from entering a swap is 11.50% − 11.00% = 0.50%. For instance, the swap described in Tables 8.2b and 8.2c splits the benefit equally between the two parties.

Company A issues fixed debt at 10.00%, and then enters a swap whereby it promises to pay LIBOR + 0.05% in exchange for receiving 10.00% fixed payments. Its net, effective funding cost is therefore LIBOR + 0.05%, which is less than the direct cost by 25bp.

Similarly, Company B issues floating debt at LIBOR + 1.00%, and then enters a swap whereby it receives LIBOR + 0.05% in exchange for paying 10.00% fixed. Its net, effective funding cost is therefore 11.00%−0.05% = 10.95%, which is less than the direct cost by 25bp. Both parties benefit from the swap.

In terms of actual cash flows, swap payments are typically *netted* against each other. For instance, if the first LIBOR rate is at 9% assuming annual payments, company A would be owed 10% × $100 = $1 million, and would have to pay LIBOR + 0.05%, or 9.05% × $100 = $0.905 million. This gives a net receipt of

$95,000. There is no need to exchange principals since both involve the same amount.

8.3.2 Quotations

Swaps can be quoted in terms of spreads relative to the yield of similar-maturity Treasury notes. For instance, a dealer may quote 10-year swap spreads as 31/34bp against LIBOR. If the current note yield is 6.72, this means that the dealer is willing to pay 6.72 + 0.31 = 7.03% against receiving LIBOR, or that the dealer is willing to receive 6.72 + 0.34 = 7.06% against paying LIBOR. Of course, the dealer makes a profit from the spread, which is rather small, at 3bp only. Equivalently, the outright quote is 7.03/7.06 for the swap.

Note that the swap should trade at a positive credit spread to Treasuries. This is because the other leg is quoted in relation to LIBOR, which also has credit risk. More precisely, swap rates correspond to the credit risk of AA-rated counterparties.

Table 5.1 has shown that the interest rate swap market is by far the largest derivative market in terms of notional. Because the market is very liquid, market quotations for the fixed rate leg have become benchmark interest rates. Thus, swap rates form the basis for the swap curve, which is also called the par curve, because it is equivalent to yields on bonds selling at par. Because the floating-rate leg is indexed to LIBOR, which carries credit risk, the swap curve is normally higher than the par curve for government bonds in the same currency.

8.3.3 Pricing

We now discuss the pricing of interest rate swaps. Consider, for instance, a three-year $100 million swap, where we receive a fixed coupon of 5.50% against LIBOR. Payments are annual and we ignore credit spreads. We can price the swap using either of two approaches, taking the difference between two bond prices or valuing a sequence of forward contracts. This is illustrated in Figure 8.2.

The top part of the figure shows that this swap is equivalent to a long position in a fixed-rate, 5.5% three-year bond and a short position in a three-year floating-rate note (FRN). If B_F is the value of the fixed-rate bond and B_f is the value of the FRN, the value of the swap is $V = B_F - B_f$.

The value of the FRN should be close to par. Just before a reset, B_f will behave exactly like a cash investment, as the coupon for the next period will be set to the prevailing interest rate. Therefore, its market value should be close to the face value. Just after a reset, the FRN will behave like a bond with a six-month maturity. But overall, fluctuations in the market value of B_f should be small.

Consider now the swap value. If at initiation the swap coupon is set to the prevailing par yield, B_F is equal to the face value, $B_F = 100$. Because $B_f = 100$ just before the reset on the floating leg, the value of the swap is zero, $V = B_F - B_f = 0$. This is like a forward contract at initiation.
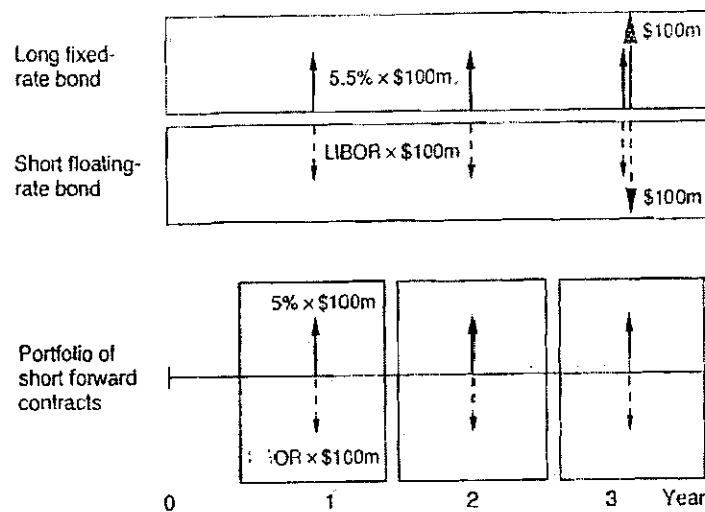
FIGURE 8.2 Alternative Decompositions for Swap Cash Flows

After the swap is consummated, its value will be affected by interest rates. If rates fall, the swap will move in the money, since it receives higher coupons than prevailing market yields. B_F will increase, whereas B_f will barely change.

Thus, the duration of a receive-fixed swap is similar to that of a fixed-rate bond, including the fixed coupons and principal at maturity. This is because the duration of the floating leg is close to zero. The fact that the principals are not exchanged does not mean that the duration computation should not include the principal. Duration should be viewed as an interest rate sensitivity.

KEY CONCEPT

A position in a receive-fixed swap is equivalent to a long position in a bond with similar coupon characteristics and maturity offset by a short position in a floating-rate note. Its duration is close to that of the fixed-rate note.

We now value the three-year swap using term-structure data from the preceding chapter. The time is just before a reset, so $B_f = \$100$ million. We compute B_F (in millions) as

$$B_F = \frac{\$5.5}{(1 + 4.000\%)} + \frac{\$5.5}{(1 + 4.618\%)^2} + \frac{\$105.5}{(1 + 5.192\%)^3} = \$100.95$$

The outstanding value of the swap is therefore $V = \$100.95 - \$100 = \$0.95$ million.

Alternatively, the swap can be valued as a sequence of forward contracts, as shown in the bottom part of Figure 8.2. Recall from Chapter 5 that the value of a unit position in a long forward contract is given by

$$V_i = (F_i - K)\exp(-r_i\tau_i) \tag{8.8}$$

where F_i is the current forward rate, K the prespecified rate, and r_i the spot rate for time τ_i. Extending this to multiple maturities, and to discrete compounding using R_i, the swap can be valued as

$$V = \sum_i n_i(F_i - K)/(1 + R_i)^{\tau_i} \tag{8.9}$$

where n_i is the notional amount for maturity i.

A long forward rate agreement benefits if rates go up. Indeed Equation (8.8) shows that the value increases if F_i goes up. In the case of our swap, we *receive* a fixed rate K. So, the position loses money if rates go up, as we could have received a higher rate. Hence, the sign on Equation (8.9) must be reversed.

Using the forward rates listed in Table 7.4, we find

$$V = -\frac{\$100(4.000\% - 5.50\%)}{(1 + 4.000\%)} - \frac{\$100(5.240\% - 5.50\%)}{(1 + 4.618\%)^2}$$

$$- \frac{\$100(6.350\% - 5.50\%)}{(1 + 5.192\%)^3}$$

$$V = +1.4423 + 0.2376 - 0.7302 = \$0.95 \text{ million}$$

This is identical to the previous result, as it should be. The swap is in-the-money primarily because of the first payment, which pays a rate of 5.5% whereas the forward rate is only 4.00%.

Thus, interest rate swaps can be priced and hedged using a sequence of forward rates, such as those implicit in Eurodollar contracts. In practice, the practice of daily marking-to-market futures induces a slight convexity bias in futures rates, which have to be adjusted downward to get forward rates.

Figure 8.3 compares a sequence of quarterly forward rates with the five-year swap rate prevailing at the same time. Because short-term forward rates are less than the swap rate, the near payments are in-the-money. In contrast, the more distant payments are out-of-the-money. The current market value of this swap is zero, which implies that all the near-term positive values must be offset by distant negative values.
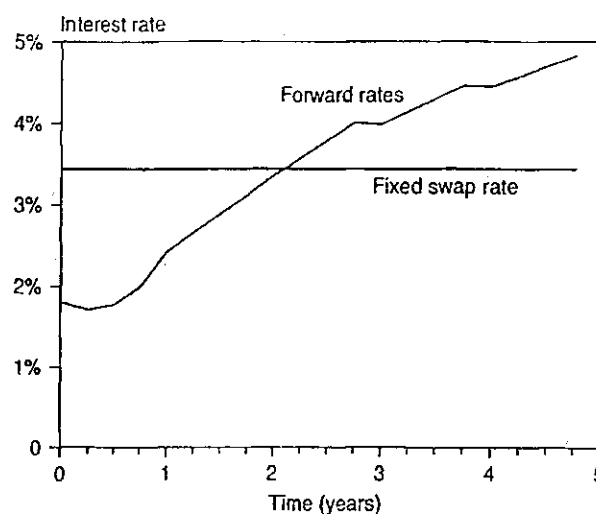
FIGURE 8.3 Sequence of Forward Rates and Swap Rate

EXAMPLE 8.6: FRM EXAM 2005—QUESTION 51

You are given the following information about an interest rate swap: two-year term, semiannual payment, fixed rate = 6%, floating rate = LIBOR + 50 basis points, notional USD 10 million. Calculate the net coupon exchange for the first period if LIBOR is 5% at the beginning of the period and 5.5% at the end of the period.

  a. Fixed-rate payer pays USD 0
  b. Fixed-rate payer pays USD 25,000
  c. Fixed-rate payer pays USD 50,000
  d. Fixed-rate payer receives USD 25,000

EXAMPLE 8.7: FRM EXAM 2000—QUESTION 55

Bank One enters into a five-year swap contract with Mervin Co. to pay LIBOR in return for a fixed 8% rate on a principal of $100 million. Two years from now, the market rate on three-year swaps at LIBOR is 7%. At this time Mervin Co. declares bankruptcy and defaults on its swap obligation. Assume that the net payment is made only at the end of each year for the swap contract period. What is the market value of the loss incurred by Bank One as a result of the default?

  a. $1.927 million
  b. $2.245 million
  c. $2.624 million
  d. $3.011 million

8.4 OPTIONS

There is a large variety of fixed-income options. We will briefly describe here caps and floors, swaptions, and exchange-traded options. In addition to these standalone instruments, fixed-income options are embedded in many securities. For instance, a callable bond can be viewed as a regular bond plus a short position in an option.

When considering fixed-income options, the underlying can be a yield or a price. Due to the negative price–yield relationship, a call option on a bond can also be viewed as a put option on the underlying yield.

8.4.1 Caps and Floors

A cap is a call option on interest rates with unit value

$$C_T = \text{Max}[i_T - K, 0] \tag{8.10}$$

where $K = i_C$ is the cap rate and i_T is the rate prevailing at maturity.

In practice, caps are purchased jointly with the issuance of floating-rate notes that pay LIBOR plus a spread on a periodic basis for the term of the note. By purchasing the cap, the issuer ensures that the cost of capital will not exceed the capped rate. Such caps are really a combination of individual options, called caplets.

The payment on each caplet is determined by C_T, the notional, and an accrual factor. Payments are made in arrears, that is, at the end of the period. For instance, take a one-year cap on a notional of \$1 million and a six-month LIBOR cap rate of 5%. The agreement period is from January 15 to the next January with a reset on July 15. Suppose that on July 15, LIBOR is at 5.5%. On the following January, the payment is

$$\$1 \text{ million} \times (0.055 - 0.05)(184/360) = \$2,555.56$$

using *Actual*/360 interest accrual. If the cap is used to hedge a FRN, this would help to offset the higher coupon payment, which is now 5.5%.

A floor is a put option on interest rates with value

$$P_T = \text{Max}[K - i_T, 0] \tag{8.11}$$

where $K = i_F$ is the floor rate. A collar is a combination of buying a cap and selling a floor. This combination decreases the net cost of purchasing the cap protection. Figure 8.4 shows an example of price path, with a cap rate of 3.5% and a floor rate of 2%. There are three instances where the cap is exercised, leading to a receipt of payment. There is one instance where the rate is below the floor, requiring a payment.
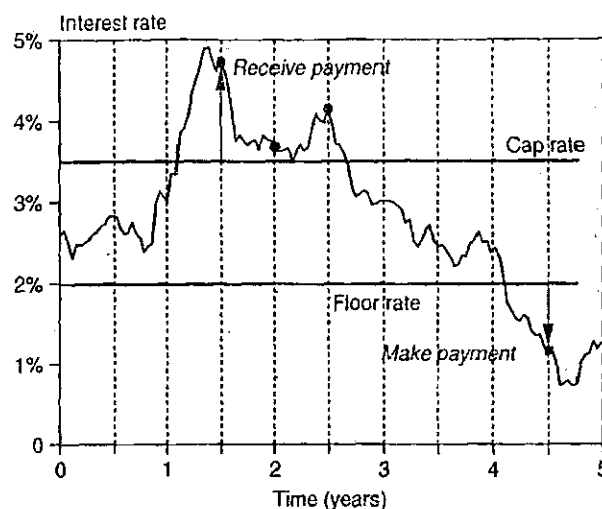
FIGURE 8.4 Exercise of Cap and Floor

When the cap and floor rates converge to the same value $K = i_C = i_F$, the overall debt cost becomes fixed instead of floating. The collar is then the same as a pay-fixed swap, which is the equivalent of put–call parity,

$$\text{Long Cap}(i_C = K) + \text{Short Floor}(i_F = K) = \text{Long Pay-Fixed Swap} \qquad (8.12)$$

Caps are typically priced using a variant of the Black model, assuming that interest rate changes are lognormal. The value of the cap is set equal to a portfolio of N caplets, which are European-style individual options on different interest rates with regularly spaced maturities

$$c = \sum_{j=1}^{N} c_j \qquad (8.13)$$

For each caplet, the unit price is

$$c_j = [FN(d_1) - KN(d_2)]\text{PV}(\$1) \qquad (8.14)$$

where F is the current forward rate for the period t_j to t_{j+1}, K is the cap rate, and PV($\$1$) is the discount factor to time t_{j+1}. To obtain a dollar amount, we must adjust for the notional amount as well as the length of the accrual period.

The volatility entering the pricing model, σ, is that of the forward rate between now and the expiration of the option contract, that is, at t_j. Generally, volatilities are quoted as one number for all caplets within a cap, which is called flat volatilities.

$$\sigma_j = \sigma$$

Alternatively, volatilities can be quoted separately for each forward rate in the caplet, which is called spot volatilities.

Example: Computing the Value of a Cap

Consider the previous cap on $1 million at the capped rate of 5%. Assume a flat term structure at 5.5% and a volatility of 20% pa. The reset is on July 15, in 181 days. The accrual period is 184 days.

Since the term structure is flat, the six-month forward rate starting in six months is also 5.5%. First, we compute the present value factor, which is $PV(\$1) = 1/(1 + 0.055 \times 365/360) = 0.9472$, and the volatility, which is $\sigma\sqrt{\tau} = 0.20\sqrt{181/360} = 0.1418$.

We then compute the value of $d_1 = \ln[F/K]/\sigma\sqrt{\tau} + \sigma\sqrt{\tau}/2 = \ln[0.055/0.05]/0.1418 + 0.1418/2 = 0.7430$ and $d_2 = d_1 - \sigma\sqrt{\tau} = 0.7430 - 0.1418 = 0.6012$. We find $N(d_1) = 0.7713$ and $N(d_2) = 0.7261$. The unit value of the call is $c = [F N(d_1) - K N(d_2)]PV(\$1) = 0.5789\%$. Finally, the total price of the call is $1 million $\times 0.5789\% \times (184/360) = \$2,959$.

Figure 8.3 can be taken as an illustration of the sequence of forward rates. If the cap rate is the same as the prevailing swap rate, the cap is said to be *at-the-money*. In the figure, the near caplets are out-of-the-money because $F_i < K$. The distant caplets, however, are in-the-money.

EXAMPLE 8.8: FRM EXAM 2002—QUESTION 22

An interest rate cap runs for 12 months based on three-month Libor with a strike price of 4%. Which of the following is generally *true*?

   a. The cap consists of three caplet options with maturities of three months, the first one starting today based on three-month LIBOR set in advance and paid in arrears.

   b. The cap consists of four caplets starting today, based on LIBOR set in advance and paid in arrears.

   c. The implied volatility of each caplet will be identical no matter how the yield curve moves.

   d. Rate caps have only a single option based on the maturity of the structure.

EXAMPLE 8.9: FRM EXAM 2004—QUESTION 10

The payoff to a swap where the investor receives fixed and pays floating can be replicated by all of the following *except*

   a. A short position in a portfolio of FRAs

   b. A long position in a fixed rate bond and a short position in a floating rate bond

   c. A short position in an interest rate cap and a long position in a floor

   d. A long position in a floating rate note and a short position in a floor

---

EXAMPLE 8.10: FRM EXAM 2003—QUESTION 27

A portfolio management firm manages the fixed-rate corporate bond portfolio owned by a defined-benefit pension fund. The duration of the bond portfolio is five years; the duration of the pension fund's liabilities is seven years. Assume that the fund sponsor strongly believes that rates will decline over the next six months and is concerned about the duration mismatch between portfolio assets and pension liabilities. Which of the following strategies would be the best way to eliminate the duration mismatch?

   a. Enter into a swap transaction in which the firm pays fixed and receives floating.

   b. Enter into a swap transaction in which the firm receives fixed and pays floating.

   c. Purchase an interest rate cap expiring in six months

   d. Sell Eurodollar futures contracts.

8.4.2 Swaptions

Swaptions are OTC options that give the buyer the right to enter a swap at a fixed point in time at specified terms, including a fixed coupon rate.

These contracts take many forms. A European swaption is exercisable on a single date at some point in the future. On that date, the owner has the right to enter a swap with a specific rate and term. Consider for example a "1Y × 5Y" swaption. This gives the owner the right to enter in one year a long or short position in a five-year swap.

A fixed-term American swaption is exercisable on any date during the exercise period. In our example, this would be during the next year. If, for instance, exercise occurs after six months, the swap would terminate in five years and six months from now. So, the termination date of the swap depends on the exercise date. In contrast, a contingent American swaption has a prespecified termination date, for instance exactly six years from now. Finally, a Bermudan option gives the holder the right to exercise on a specific set of dates during the life of the option.

As an example, consider a company that, in one year, will issue five-year floating-rate debt. The company wishes to have the option to swap the floating payments into fixed payments. The company can purchase a swaption that will give it the right to create a five-year pay-fixed swap at the rate of 8%. If the prevailing swap rate in one year is higher than 8%, the company will exercise the swaption, otherwise not. The value of the option at expiration will be

$$P_T = \text{Max}[V(i_T) - V(K), 0] \tag{8.15}$$

where $V(i)$ is the value of a swap to pay a fixed rate i, i_T is the prevailing swap rate for the swap maturity, and K is the locked-in swap rate. This contract is called a European 6/1 put swaption, or one into five-year payer option.

Such a swap is equivalent to an option on a bond. As this swaption creates a profit if rates rise, it is akin to a one-year put option on a six-year bond. A

TABLE 8.3 Summary of Terminology for OTC Swaps and Options

Product	Buy (long)	Sell (short)
Fixed/Floating Swap	Pay fixed wReceive floating	Pay floating Receive fixed
Cap	Pay premium Receive Max$(i - i_C, 0)$	Receive premium Pay Max$(i - i_C, 0)$
Floor	Pay premium Receive Max$(i_F - i, 0)$	Receive premium Pay Max$(i_F - i, 0)$
Put Swaption (payer option)	Pay premium Option to pay fixed and receive floating	Receive premium If exercised, receive fixed and pay floating
Call Swaption (receiver option)	Pay premium Option to pay floating and receive fixed	Receive premium If exercised, receive floating and pay fixed

put option benefits when the bond value falls, which happens when rates rise. Conversely, a swaption that gives the right to receive fixed is akin to a call option on a bond. Table 8.3 summarizes the terminology for swaps, caps and floors, and swaptions.

Swaptions can be used for a variety of purposes. Consider an investor in a mortgage-backed security (MBS). If long-term rates fall, prepayment will increase, leading to a shortfall in the price appreciation in the bond. This risk can be hedged by buying receiver swaptions. If rates fall, the buyer will exercise the option, which creates a profit to offset the loss on the MBS. Alternatively, this risk can also be hedged by issuing callable debt. This creates a long position in an option that generates a profit if rates fall. As an example, Fannie Mae, a government-sponsored enterprise that invests heavily in mortgages, uses these techniques to hedge its prepayment risk.

Finally, swaptions are typically priced using a variant of the Black model, assuming that interest rates are lognormal. The value of the swaption is then equal to a portfolio of options on different interest rates, all with the same maturity. In practice, swaptions are traded in terms of volatilities instead of option premiums. The applicable forward rate starts at the same time as the option, with a term equal to that of the option.

EXAMPLE 8.11: FRM EXAM 2003—QUESTION 56

As your company's risk manager, you are looking for protection against adverse interest rate changes in five years. Using Black's model for options on futures to price a European swap option (swaption) which gives the option holder the right to cancel a seven-year swap after five years, which of the following would you use in the model?

a. The two-year forward par swap rate starting in five years time
b. The five-year forward par swap rate starting in two years time
c. The two-year par swap rate
d. The five-year par swap rate

8.4.3 Exchange-Traded Options

Among exchange-traded fixed-income options, we describe options on Eurodollar futures and on T-bond futures.

Options on Eurodollar futures give the owner the right to enter a long or short position in Eurodollar futures at a fixed price. The payoff on a put option, for example, is

$$P_T = \text{Notional} \times \text{Max}[K - FQ_T, 0] \times (90/360) \qquad (8.16)$$

where K is the strike price and FQ_T the prevailing futures price quote at maturity. In addition to the cash payoff, the option holder enters a position in the underlying futures. Since this is a put, it creates a short position after exercise, with the counterparty taking the opposing position. Note that, since futures are settled daily, the value of the contract is zero.

Since the futures price can also be written as $FQ_T = 100 - i_T$ and the strike price as $K = 100 - i_C$, the payoff is also

$$P_T = \text{Notional} \times \text{Max}[i_T - i_C, 0] \times (90/360) \qquad (8.17)$$

which is equivalent to that of a cap on rates. Thus, a put on Eurodollar futures is equivalent to a caplet on LIBOR.

In practice, there are minor differences in the contracts. Options on Eurodollar futures are American style instead of European style. Also, payments are made at the expiration date of Eurodollar futures options instead of in arrears.

Options on T-bond futures give the owner the right to enter a long or short position in futures at a fixed price. The payoff on a call option, for example, is

$$C_T = \text{Notional} \times \text{Max}[F_T - K, 0] \qquad (8.18)$$

An investor who thinks that rates will fall, or that the bond market will rally, could buy a call on T-bond futures. In this manner, he or she will participate in the upside, without downside risk.

EXAMPLE 8.12: FRM EXAM 2007—QUESTION 95

To hedge against future, unanticipated, and significant increases in borrowing rates, which of the following alternatives offers the greatest flexibility for the borrower?

a. Interest rate collar
b. Fixed for floating swap
c. Call swaption
d. Interest rate floor

8.5 IMPORTANT FORMULAS

Long FRA $1 \times 4 =$ Invest for one period, borrow for four

Payment on FRA: $V_T = (S_T - F) \times \tau \times$ Notional $\times$ PV($1)

Valuation of Eurodollar contract:
$$P_t = 10,000 \times [100 - 0.25(100 - FQ_t)] = 10,000 \times [100 - 0.25F_t]$$

Eurodollar contract risk: DV01 = $25

Futures convexity adjustment: Futures Rate $=$ Forward Rate $+ \left(\frac{1}{2}\right)\sigma^2 t_1 t_2$ (negative relationship between contract value and rates)

T-bond futures net delivery cost: Cost $=$ Price $-$ Futures Quote $\times$ CF

T-bond futures conversion factor: CF $=$ NPV of bond at 6%

Valuation of interest rate swap: $V = B_F$(fixed $-$ rate) $- B_f$(floating $-$ rate)
Long receive-fixed $=$ long fixed-coupon bond $+$ short FRN

Valuation of interest rate swap as forward contracts:
$$V = \sum_i n_i(F_i - K)/(1 + R_i)^{r_i}$$

Interest-rate cap: $C_T = \text{Max}[i_T - K, 0]$

Interest-rate floor: $P_T = \text{Max}[K - i_T, 0]$

Collar: Long cap plus short floor

Cap valuation: $c = \sum_{j=1}^{N} c_j, \quad c_j = [FN(d_1) - KN(d_2)]\text{PV}(\$1)$

Put swaption ($1Y \times 5Y$): (right to pay fixed, starting in one year for five years)
$$P_T = \text{Max}[V(i_T) - V(K), 0]$$

Call option on Eurodollar futures $=$ cap on rates

8.6 ANSWERS TO CHAPTER EXAMPLES

Example 8.1: FRM Exam 2002—Question 27

b. An FRA defined as $t_1 \times t_2$ involves a forward rate starting at time t_1 and ending at time t_2. The buyer of this FRA locks in a borrowing rate for months 3 to 5. This is equivalent to borrowing for five months and reinvesting the funds for the first two months.

Example 8.2: FRM Exam 2005—Question 57

d. The market-implied forward rate is given by $\exp(-R_2 \times 2) = \exp(-R_1 \times 1 - F_{1,2} \times 1)$, or $F_{1,2} = 2 \times 3.50 - 1 \times 3.25 = 3.75\%$. Given that this is exactly equal to the quoted rate, the value must be zero. If instead this rate was 3.50%, for example, the value would be $V = \$1,000,000 \times (3.75\% - 3.50\%) \times (2 - 1)\exp(-3.50\% \times 2) = 2,331$.

Example 8.3: FRM Exam 2001—Question 70

b. The seller of an FRA agrees to receive fixed. Since rates are now higher than the contract rate, this contract must show a loss for the seller. The loss

is $\$10,000,000 \times (6.85\% - 6.35\%) \times (90/360) = \$12,500$ when paid in arrears, i.e., in nine months. On the settlement date, i.e., brought forward by three months, the loss is $\$12,500/(1 + 6.85\% \times 0.25) = \$12,290$.

Example 8.4: FRM Exam 2005—Question 49

c. Compute the ratio of the price to the CF. This gives, respectively, $125.69/1.1979 = 11.29$, then 12.87, 8.09, and 11.58. Hence, bond C is the cheapest-to-deliver.

Example 8.5: FRM Exam 2007—Question 80

c. Equation (8.4) shows that the futures rate exceeds the forward rate.

Example 8.6: FRM Exam 2005—Question 51

b. The floating leg uses LIBOR at the beginning of the period, plus 50bp, or 5.5%. The payment is given by $\$10,000,000 \times (0.06 - 0.055) \times 0.5 = \$25,000$.

Example 8.7: FRM Exam 2000—Question 55

c. Using Equation (8.9) for three remaining periods, we have the discounted value of the net interest payment, or $(8\% - 7\%)\$100m = \$1m$, discounted at 7%, which is $\$934,579 + \$873,439 + \$816,298 = \$2,624,316$.

Example 8.8: FRM Exam 2002—Question 22

a. Interest rate caps involve multiple options, or caplets. The first one has terms that are set in three months. It locks in $\text{Max}[R(t + 3) - 4\%, 0]$. Payment occurs in arrears in six months. The second one is a function of $\text{Max}[R(t + 6) - 4\%, 0]$. The third is a function of $\text{Max}[R(t + 9) - 4\%, 0]$ and is paid at $t + 12$. The sequence then stops because the cap has a term of 12 months only. This means there are three caplets.

Example 8.9: FRM Exam 2004—Question 10

d. A receive-fixed swap position is equivalent to being long a fixed-rate bond, or being short a portfolio of FRAs (which gain if rates go down), or selling a cap and buying a floor with the same strike price (which gains if rates go up). A short position in a floor does not generate a gain if rates drop. It is asymmetric anyway.

Example 8.10: FRM Exam 2003—Question 27

b. The manager should increase the duration of assets, or buy coupon-paying bonds. This can be achieved by entering a receive-fixed swap, so b. is correct and a. is wrong. Buying a cap will not provide protection if rates drop. Selling Eurodollar futures will lose money if rates drop.

Example 8.11: FRM Exam 2003—Question 56

a. The forward rate should start at the beginning of the option in five years, with a maturity equal to the duration of the option, or two years.

Example 8.12: FRM Exam 2007—Question 95

c. A swaption gives the borrower the flexibility to lock in a low rate. On the other hand, a regular swap does not offer flexibility as an option. A collar fixes a range of rates, but not much flexibility. A floor involves protection if rates go down, not up. (Note that buying a cap would have been another good choice.)

Equity, Currency, and Commodity Markets

Having covered fixed-income instruments, we now turn to equity, currency, and commodity markets. Equities, or common stocks, represent ownership shares in a corporation. Due to the uncertainty in their cash flows, as well as in the appropriate discount rate, equities are much more difficult to value than fixed-income securities. They are also less amenable to the quantitative analysis that is used in fixed-income markets. Equity derivatives, however, can be priced reasonably precisely in relation to underlying stock prices.

Next, the foreign currency markets include spot, forward, options, futures, and swap markets. The foreign exchange markets are by far the largest financial markets in the world, with daily turnover estimated at $1,880 billion in 2004.

Commodity markets consist of agricultural products, metals, energy, and other products. Commodities differ from financial assets as their holding provides an implied benefit known as convenience yield but also incurs storage costs.

Section 9.1 introduces equity markets and presents valuation methods. Section 9.2 briefly discusses convertible bonds and warrants. These differ from the usual equity options in that exercising them creates new shares. Section 9.3 then provides an overview of important equity derivatives, including stock index futures, stock options, stock index options, and equity swaps. Section 9.4 presents a brief introduction to currency markets. Contracts such as futures, forwards, and options have been developed in previous chapters and do not require special treatment. In contrast, currency swaps are analyzed in some detail in Section 9.5 due to their unique features and importance. Finally, Section 9.6 discusses commodity markets.

9.1 EQUITIES

9.1.1 Overview

Common stocks, also called equities, are securities that represent ownership in a corporation. Bonds are *senior* to equities, that is, have a prior claim on the firm's assets in case of bankruptcy. Hence equities represent residual claims to what is left of the value of the firm after bonds, loans, and other contractual obligations have been paid off.

217

TABLE 9.1 Global Equity Markets—2007
(Billions of U.S. Dollars)

United States	19,922
Japan	4,543
Eurozone	10,047
United Kingdom	3,852
Other Europe	2,867
Other Pacific	4,949
Canada	2,187
Developed	48,366
Emerging	12,508
World	60,874

Source: World Federation of Exchanges

Another important feature of common stocks is their limited liability, which means that the most shareholders can lose is their original investment. This is unlike owners of unincorporated businesses, whose creditors have a claim on the personal assets of the owner should the business turn bad.

Table 9.1 describes the global equity markets. The total market value of common stocks was worth approximately $61 trillion at the end of 2007. The United States accounts for the largest share, followed by Japan, the Eurozone, and the United Kingdom. In 2008, global stocks fell by 42%, which implies a loss of market value of about $26 trillion.

Preferred stocks differ from common stock because they promise to pay a specific stream of dividends. So, they behave like a perpetual bond, or consol. Unlike bonds, however, failure to pay these dividends does not result in default. Instead, the corporation must withhold dividends to common stock holders until the preferred dividends have been paid out. In other words, preferred stocks are junior to bonds, but senior to common stocks.

With cumulative preferred dividends, all current and previously postponed dividends must be paid before any dividends on common stock shares can be paid. Preferred stocks usually have no voting rights.

Unlike interest payments, preferred stocks dividends are not tax-deductible expenses. Preferred stocks, however, have an offsetting tax advantage. Corporations that receive preferred dividends only pay taxes on 30% of the amount received, which lowers their income tax burden. As a result, most preferred stocks are held by corporations. The market capitalization of preferred stocks is much lower than that of common stocks, as seen from the IBM example below. Trading volumes are also much lower.

Example: IBM Preferred Stock

IBM issued 11.25 million preferred shares in June 1993. These are traded as 45 million "depositary" shares, each representing one-fourth of the preferred, under the ticker "IBM-A" on the NYSE. Dividends accrue at the rate of $7.50 per annum, or $1.875 per depositary share.

As of April 2001, the depositary shares were trading at $25.4, within a narrow 52-week trading range of [$25.00, $26.25]. Using the valuation formula for a consol, the shares trade at an implied yield of 7.38%. The total market capitalization of the IBM-A shares amounts to approximately $260 million. In comparison, the market value of the common stock is $214,602 million, which is more than 800 times larger.

9.1.2 Valuation

Common stocks are extremely difficult to value. Like any other asset, their value derives from their future benefits, that is, from their stream of future cash flows (i.e., dividend payments) or future stock price.

We have seen that valuing Treasury bonds is relatively straightforward, as the stream of cash flows, coupon, and principal payments, can be easily laid out and discounted into the present.

This is an entirely different affair for common stocks. Consider for illustration a "simple" case where a firm pays out a dividend D over the next year that grows at the constant rate of g. We ignore the final stock value and discount at the constant rate of r, such that $r > g$. The firm's value, P, can be assessed using the net present value formula, like a bond

$$
\begin{aligned}
P &= \sum_{t=1}^{\infty} C_t/(1+r)^t \\
&= \sum_{t=1}^{\infty} D(1+g)^{(t-1)}/(1+r)^t \\
&= [D/(1+r)] \sum_{t=0}^{\infty} [(1+g)/(1+r)]^t \\
&= [D/(1+r)] \times \left[\frac{1}{1-(1+g)/(1+r)} \right] \\
&= [D/(1+r)] \times [(1+r)/(r-g)]
\end{aligned}
$$

This is also the so-called "Gordon-growth" model,

$$
P = \frac{D}{r-g} \tag{9.1}
$$

as long as the discount rate exceeds the growth rate of dividends, $r > g$.

The problem with equities is that the growth rate of dividends is uncertain and that, in addition, it is not clear what the required discount rate should be. To make things even harder, some companies simply do not pay any dividend and instead create value from the appreciation of their share price.

Still, this valuation formula indicates that large variations in equity prices can arise from small changes in the discount rate or in the growth rate of dividends, explaining the large volatility of equities. More generally, the risk and expected return of the equity depends on the underlying business fundamentals as well as on the amount of leverage, or debt in the capital structure.

For financial intermediaries for which the value of underlying assets can be measured precisely, we can value the equity from the underlying assets and the cost of borrowing. This situation, however, is more akin to the pricing of a derivative from the price of the underlying than pricing the asset directly.

EXAMPLE 9.1: LEVERAGE AND RETURN ON EQUITY

A hedge fund leverages its $100 million of investor capital by a factor of three and invests it into a portfolio of junk bonds yielding 14%. If its borrowing costs are 8%, what is the yield on investor capital?

a. 14%
b. 18%
c. 26%
d. 42%

9.2 CONVERTIBLE BONDS AND WARRANTS

9.2.1 Definitions

We now turn to convertible bonds and warrants. While these instruments have option-like features, they differ from regular options. When a call option is exercised, for instance, the "long" purchases an outstanding share from the "short." There is no net creation of shares. In contrast, the exercise of convertible bonds, of warrants, (and of executive stock options) entails the creation of new shares, as the option is sold by the corporation itself. Because the number of shares goes up, the existing shares are said to be diluted by the creation of new shares.

Warrants are long-term call options issued by a corporation on its own stock. They are typically created at the time of a bond issue, but they trade separately from the bond to which they were originally attached. When a warrant is exercised, it results in a cash inflow to the firm which issues more shares.

Convertible bonds are bonds issued by a corporation that can be converted into equity at certain times using a predetermined exchange ratio. They are equivalent to a regular bond plus a warrant. This allows the company to issue debt with a lower coupon than otherwise.

For example, a bond with a conversion ratio of 10 allows its holder to convert one bond with par value of $1,000 into 10 shares of the common stock. The conversion price, which is really the strike price of the option, is $1,000/10 = $100. The corporation will typically issue the convertible deep out-of-the-money, for example when the stock price is at $50. When the stock price moves, for instance to $120, the bond can be converted into stock for an immediate option profit of ($120 - $100) × 10 = $200.

Conversion value: Stock price times conversion ratio
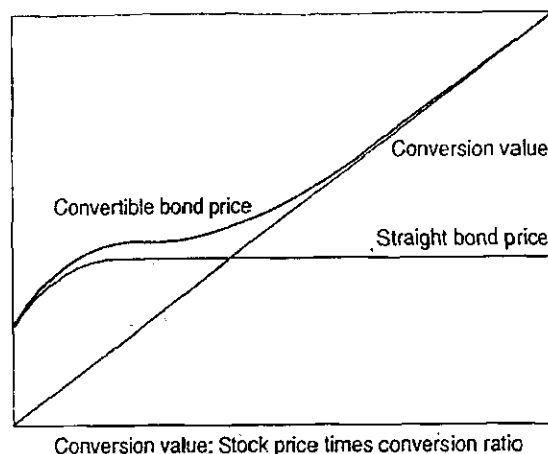
FIGURE 9.1   Convertible Bond Price and Conversion Value

Figure 9.1 describes the relationship between the value of the convertible bond and the conversion value, defined as the current stock price times the conversion ratio. The convertible bond value must be greater than the price of an otherwise identical straight bond and the conversion value.

For high values of the stock price, the firm is unlikely to default and the straight bond price is constant, reflecting the discounting of cash flows at the risk-free rate. In this situation, it is almost certain the option will be exercised and the convertible value is close to the conversion value. For low values of the stock price, the firm is likely to default and the straight bond price drops, reflecting the likely loss upon default. In this situation, it is almost certain the option will not be exercised, and the convertible value is close to the straight bond value. In the intermediate region, the convertible value depends on both the conversion and straight bond values. The convertible is also sensitive to interest rate risk.

Example: A Convertible Bond

Consider an 8% annual coupon, 10-year convertible bond with a face value of $1,000. The yield on similar maturity straight debt issued by the company is currently 8.50%, which gives a current value of straight debt of $967. The bond can be converted into common stock at a ratio of 10-to-1.

Assume first that the stock price is $50. The conversion value is then $500, much less than the straight debt value of $967. This corresponds to the left area of Figure 9.1. If the convertible trades at $972, its promised yield is 8.42%. This is close to the yield of straight debt, as the option has little value.

Assume next that the stock price is $150. The conversion value is then $1,500, much higher than the straight debt value of $967. This corresponds to the right area of Figure 9.1. If the convertible trades at $1,505, its promised yield is 2.29%. In this case, the conversion option is in-the-money, which explains why the yield is so low.

9.2.2 Valuation

Warrants can be valued by adapting standard option pricing models to the dilution effect of new shares. Consider a company with N outstanding shares and M outstanding warrants, each allowing the holder to purchase γ shares at the fixed price of K. At origination, the value of the firm includes the warrant, or

$$V_0 = NS_0 + MW_0 \tag{9.2}$$

where S_0 is the initial stock price just before issuing the warrant, and W_0 is the up-front value of the warrant.

After dilution, the total value of the firm includes the value of the firm before exercise (including the original value of the warrants) plus the proceeds from exercise, i.e., $V_T + M\gamma K$. The number of shares then increases to $N + \gamma M$. The total payoff to the warrant holder is

$$W_T = \gamma \text{Max}(S_T - K, 0) = \gamma(S_T - K) = \gamma\left(\frac{V_T + M\gamma K}{N + \gamma M} - K\right) \tag{9.3}$$

which must be positive. After simplification, this is also

$$W_T = \gamma\left(\frac{V_T - NK}{N + \gamma M}\right) = \frac{\gamma}{N + \gamma M}(V_T - NK) = \frac{\gamma N}{N + \gamma M}\left(\frac{V_T}{N} - K\right) \tag{9.4}$$

which is equivalent to $n = \gamma N/(N + \gamma M)$ options on the stock price. The warrant can be valued by standard option models with the asset value equal to the stock price plus the warrant proceeds, multiplied by the factor n,

$$W_0 = n \times c\left(S_0 + \frac{M}{N}W_0, K, \tau, \sigma, r, d\right) \tag{9.5}$$

with the usual parameters. Here, the unit asset value is $\frac{V_0}{N} = S_0 + \frac{M}{N}W_0$. This must be solved iteratively since W_0 appears on both sides of the equation. If, however, M is small relative to the current float, or number of outstanding shares N, the formula reduces to a simple call option in the amount γ

$$W_0 = \gamma\, c(S_0, K, \tau, \sigma, r, d) \tag{9.6}$$

Example: Pricing a Convertible Bond

Consider a zero-coupon, 10-year convertible bond with face value of $1,000. The yield on similar maturity straight debt issued by the company is currently 8.158%, using continuous compounding, which gives a straight debt value of $442.29.

The bond can be converted into common stock at a ratio of 10-to-1 at expiration only. This gives a strike price of $K = $100. The current stock price is $60. The stock pays no dividend and has annual volatility of 30%. The risk-free rate is 5%, also continuously compounded.

Ignoring dilution effects, the Black–Scholes model gives an option value of $216.79. So, the theoretical value for the convertible bond is $P = \$442.29 + \$216.79 = \$659.08$. If the market price is lower than $659, the convertible is said to be cheap. This, of course, assumes that the pricing model and input assumptions are correct.

One complication is that most convertibles are also callable at the discretion of the firm. Convertible securities can be called for several reasons. First, an issue can be called to force conversion into common stock when the stock price is high enough. Bondholders have typically a month during which they can still convert, in which case this is a forced conversion. This call feature gives the corporation more control over conversion. It also allows the company to raise equity capital by forcing the bondholders to pay the exercise price.

Second, the call may be exercised when the option value is worthless and the firm can refinance its debt at a lower coupon. This is similar to the call of a nonconvertible bond, except that the convertible must be *busted*, which occurs when the stock price is much lower than the conversion price.

In terms of risk factors, a long position in a convertible bond is exposed to increasing interest rates and credit spreads, like regular corporate bonds, but also to factors that decrease the value of the embedded call, such as a decreasing stock price and decreasing implied volatility.

EXAMPLE 9.2: FRM EXAM 2001—QUESTION 119

A corporate bond with face value of $100 is convertible at $40 and the corporation has called it for redemption at $106. The bond is currently selling at $115 and the stock's current market price is $45. Which of the following would a bondholder most likely do?

a. Sell the bond

b. Convert the bond into common stock

c. Allow the corporation to call the bond at 106

d. None of the above

EXAMPLE 9.3: FRM EXAM 2001—QUESTION 117

What is the main reason why convertible bonds are generally issued with a call?

a. To make their analysis less easy for investors

b. To protect against unwanted takeover bids

c. To reduce duration

d. To force conversion if in-the-money

9.3 EQUITY DERIVATIVES

Equity derivatives can be traded on over-the-counter markets as well as organized exchanges. We only consider the most popular instruments.

9.3.1 Stock Index Futures

Stock index futures are actively traded all over the world. In fact, the turnover corresponding to the notional amount is often greater than the total amount of trading in physical stocks in the same market. The success of these contracts can be explained by their versatility for risk management. Stock index futures allow investors to manage efficiently their exposure to broad stock market movements. Speculators can take easily directional bets with futures, on the upside or downside. Hedgers also find that futures provide a cost-efficient method to protect against price risk.

Perhaps the most active contract is the S&P 500 futures contract on the Chicago Mercantile Exchange (CME). The contract notional is defined as $250 times the index level. Table 9.2 displays quotations as of December 31, 1999.

The table shows that most of the volume was concentrated in the "near" contract, that is, March in this case. Translating the trading volume in number of contracts into a dollar equivalent, we find $250 × 1484.2 × 34,897, which gives $13 billion. So, these markets are very liquid. As a comparison, the average daily volume was $35 billion in 2001. This was close to the trading volume of $42 billion for stocks on the New York Stock Exchange (NYSE).

We can also compute the daily profit on a long position, which would have been $250 × (+3.40), or $850 on that day. In relative terms, this daily move was +3.4/1480.8, which is only 0.23%. The typical daily standard deviation is about 1%, which gives a typical profit or loss of $3,710.50.

These contracts are cash settled. They do not involve delivery of the underlying stocks at expiration. In terms of valuation, the futures contract is priced according to the usual cash-and-carry relationship,

$$F_t e^{-r\tau} = S_t e^{-y\tau} \tag{9.7}$$

where y is now the dividend yield defined per unit time. For instance, the yield on the S&P was $y = 0.94$ percent per annum on that day.

Here, we assume that the dividend yield is known in advance and paid on a continuous basis. In general, this is not necessarily the case but can be viewed as

TABLE 9.2 Sample S&P Futures Quotations

Maturity	Open	Settle	Change	Volume	Open Interest
March	1480.80	1484.20	+3.40	34,897	356,791
June	1498.00	1503.10	+3.60	410	8,431

a good approximation. With a large number of firms in the index, dividends will be spread reasonably evenly over the quarter.

To check if the futures contract was fairly valued, we need the spot price, $S = 1469.25$, the short-term interest rate, $r = 5.3\%$, and the number of days to maturity, which was 76 (to March 16). Note that rates are not continuously compounded. The present value factor is $PV(\$1) = 1/(1 + r\tau) = 1/(1 + 5.3\%(76/365)) = 0.9891$. Similarly, the present value of the dividend stream is $1/(1 + y\tau) = 1/(1 + 0.94\%(76/365)) = 0.9980$. The fair price is then

$$F = [S/(1 + y\tau)] (1 + r\tau) = [1469.25 \times 0.9980]/0.9891 = 1482.6$$

This is rather close to the settlement value of $F = 1484.2$. The discrepancy is probably because the quotes were not measured simultaneously. Because the yield is less than the interest rate, the forward price is greater than the spot price.

Figure 9.2 displays the convergence of futures and cash prices for the December 1999 S&P 500 futures contract traded on the CME. The futures price is always above the spot price. The correlation between the two prices is very high, reflecting the cash-and-carry relationship in Equation (9.7).

Because financial institutions engage in stock index arbitrage, we would expect the cash-and-carry relationship to hold very well, One notable exception was during the market crash of October 19, 1987. The market lost more than 20% in a single day. Throughout the day, however, futures prices were more up-to-date than cash prices because of execution delays in cash markets. As a result, the S&P stock index futures value was very cheap compared with the underlying cash market. Arbitrage, however, was made difficult due to chaotic market conditions.
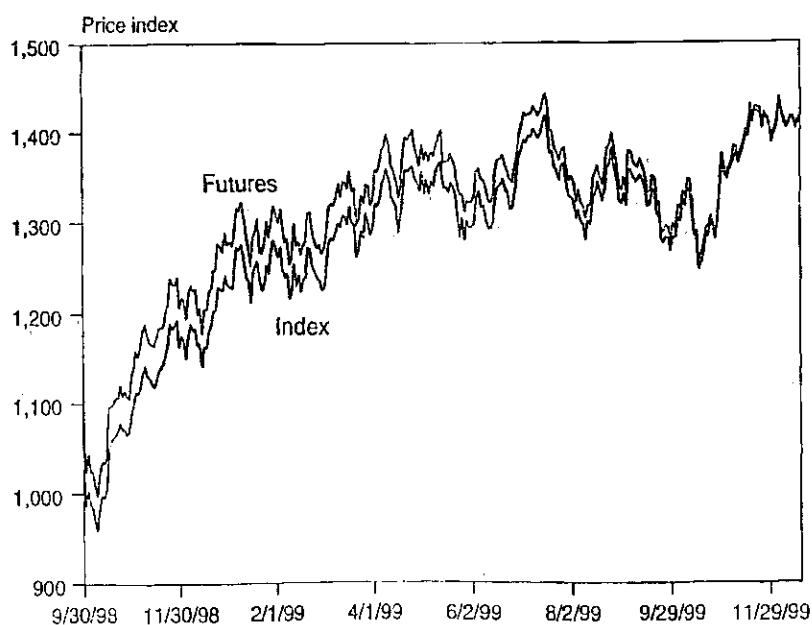
FIGURE 9.2 Futures and Cash Prices for S&P 500 Futures

EXAMPLE 9.4: FRM EXAM 2000—QUESTION 12

Suppose the price for a six-month S&P index futures contract is 552.3. If the risk-free interest rate is 7.5% per year and the dividend yield on the stock index is 4.2% per year, and the market is complete and there is no arbitrage, what is the price of the index today?

- a. 543.26
- b. 552.11
- c. 555.78
- d. 560.02

9.3.2 Single Stock Futures

In late 2000, the United States passed legislation authorizing trading in single stock futures, which are futures contracts on individual stocks. Such contracts were already trading in Europe and elsewhere. In the United States, electronic trading started in November 2002 and now takes place on "OneChicago," a joint venture of Chicago exchanges.

Each contract gives the obligation to buy or sell 100 shares of the underlying stock. Settlement usually involves physical delivery, that is, the exchange of the underlying stock. Relative to trading in the underlying stocks, single stock futures have many advantages. Positions can be established more efficiently due to their low margin requirements, which are generally 20% of the cash value. In contrast, margin for stocks are higher. Also, short selling eliminates the costs and inefficiencies associated with the stock loan process. Other than physical settlement, these contracts trade like stock index futures.

9.3.3 Equity Options

Options can be traded on individual stocks, on stock indices, or on stock index futures. In the United States, stock options trade, for example, on the Chicago Board Options Exchange (CBOE). Each option gives the right to buy or sell a round lot of 100 shares. Settlement involves physical delivery.

Traded options are typically American-style, so their valuation should include the possibility of early exercise. In practice, however, their values do not differ much from those of European options, which can be priced by the Black–Scholes model. When the stock pays no dividend, the values are the same. For more precision, we can use numerical models such as binomial trees to take into account dividend payments.

The most active *index* options in the United States are options on the S&P 100 and S&P 500 index traded on the CBOE. The former are American-style, while the latter are European-style. These options are cash settled, as it would be too

complicated to deliver a basket of 100 or 500 underlying stocks. Each contract is for $100 times the value of the index. European options on stock indices can be priced using the Black–Scholes formula, using y as the dividend yield on the index as we have done in the previous section for stock index futures.

Finally, options on S&P 500 stock index futures are also popular. These give the right to enter a long or short futures position at a fixed price. Exercise is cash settled.

9.3.4 Equity Swaps

An equity swap is an agreement to exchange cash flows tied to the return on a stock market index in exchange for a fixed or floating rate of interest. An example is a swap that provides the return on the S&P 500 index every six months in exchange for payment of LIBOR plus a spread. The swap will be typically priced so as to have zero value at initiation. Equity swaps can be valued as portfolios of forward contracts, as in the case of interest rate swaps. We will later see how to price currency swaps. The same method can be used for equity swaps.

These swaps are used by investment managers to acquire exposure to, for example, an emerging stock market, without having to invest in the market itself. In some cases, these swaps can also be used to skirt restrictions on foreign investments.

9.3.5 Variance Swaps

A variance swap is a forward contract on the variance. The payoff is computed as

$$V_T = (\sigma^2_{t_0,T} - K_V)N \tag{9.8}$$

where N is the notional amount, σ^2 is the realized variance over the life of the contract, usually measured as

$$\sigma^2 = \frac{252}{\tau} \sum_{i=1}^{\tau} [\ln(S_i/S_{i-1})]^2 \tag{9.9}$$

and K_V is the strike price, or forward price. Variance swaps can be written on any underlying asset, but are most common for equities or equity indices. They allow trades based on direct views on variance. Long positions are bets on high volatility.

For example, suppose a dealer quotes a one-year contract on the S&P 500 index, with $K_V = (15\%)^2$ and notional of $N = \$100,000/$(one volatility point)^2. If at expiration the realized volatility is 17%, the payoff to the long position is $[\$100,000/(1^2)][(17)^2 - (15)^2] = \$100,000(289 - 225) = \$6,400,000$. Therefore, the payoff is a quadratic function of the volatility. In theory, it is

unlimited.[1] Like any forward contract, K_V is determined so that the initial value of the contract is zero. In fact, the widely quoted VIX index is the fair strike price for a variance swap on the S&P 500 index, quoted as volatility.

The market value of an outstanding variance swap with $\tau = T - t$ days remaining to maturity is

$$V_t = Ne^{-r\tau}[w(\sigma^2_{t_0,t} - K_V) + (1 - w)(K_t - K_V)] \qquad (9.10)$$

where $\sigma^2_{t_0,t}$ is the elapsed variance between the initial time t_0 and the current time, t, w is the fraction of days elapsed since t_0, and K_t is the current forward price. We can verify that at the initial time, $w = 0$, and V_0 is simply proportional to $K_t - K_V$, which is zero if the contract starts at-the-market. At expiration, this converges to Equation (9.8).

Such contracts also allow correlation trading. Consider for example an index of two stocks. A variance swap is available for each constituent stock as well as for the index. The realized variance of the index depends on the two variances as well as the correlation coefficient. All else equal, a higher correlation translates into a higher portfolio variance. A long correlation trade would buy a variance swap on the index and short variance swaps on the components.[2] If the correlation increases, the long position should gain more than the short positions, thereby generating a gain.

9.4 CURRENCY MARKETS

9.4.1 Overview

The forex, or currency markets have enormous trading activity, with daily turnover estimated at $3,210 billion in 2007. Their size and growth is described in Table 9.3. This trading activity dwarfs that of bond or stock markets. In comparison, the daily trading volume on the New York Stock Exchange (NYSE) is approximately $80 billion. Even though the largest share of these transactions is between dealers, or with other financial institutions, the volume of trading with other, nonfinancial institutions is still quite large, at $549 billion daily.

Spot transactions are exchanges of two currencies for settlement as soon as it is practical, typically in two business days. They account for about 35% of trading volume. Other transactions are outright forward contracts and forex swaps. Outright forward contracts are agreements to exchange two currencies at a future date, and account for about 12% of the total market. Forex swaps involve two transactions, an exchange of currencies on a given date and a reversal at a later date, and account for 53% of the total market. Note that forex swaps are typically

---

[1] In practice, most contracts are capped to a maximum value for the variance equal to $m^2 K_V$. Volatility swaps are also available but are much less common. This is because variance swaps can be hedged relatively easily, using a combination of options. This is not the case for volatility swaps.
[2] Note that keeping the position variance-neutral requires a greater notional amount for the index swaps than for the sum of component swaps.

TABLE 9.3 Average Daily Trading Volume in Currency Markets (Billions of U.S. Dollars)

Year	Spot	Forwards & Forex Swaps	Total
1989	350	240	590
1992	416	404	820
1995	517	673	1,190
1998	592	898	1,490
2001	399	811	1,210
2004	656	1,224	1,880
2007	1,005	2,076	3,210
Of which, between:			
Dealers			1,374
Financials			1,287
Others			549

Source: Bank for International Settlements surveys.

of a short-term nature and should not be confused with long-term currency swaps, which involve a stream of payments over longer horizons.

In addition to these contracts, the market also includes OTC forex options ($212 billion daily) and exchange-traded derivatives ($72 billion daily). The most active currency futures are traded on the Chicago Mercantile Exchange (CME) and settled by physical delivery. The CME also trades options on currency futures.

As we have seen before, currency forwards, futures, and options can be priced according to standard valuation models, specifying the income payment to be a continuous flow defined by the foreign interest rate, r^*.

Currencies are generally quoted in European terms, that is, in units of the foreign currency per dollar. The yen, for example, is quoted as 120 yen per U.S. dollar. Two notable exceptions are the British pound (sterling) and the euro, which are quoted in American terms, that is, in dollars per unit of the foreign currency, The pound, for example, is quoted as 1.6 dollar per pound.

9.4.2 Currency Products

Thus, currency markets offer the full range of financial instruments. Because of their importance, currency swaps will be examined in more detail in the following section.

One type of instrument, however, is specific to the currency markets. This is the quanto, or quantity-adjusted derivative. The payoff is defined by variables associated with one currency but is paid in another currency.

As an example, suppose a U.S. investor considers buying Japanese stocks, represented by the Nikkei 225 index. Normally, buying the stocks involves taking a position in stocks and in the Japanese yen currency. We could have a situation where the Nikkei increases in value but this gain is wiped out by a fall in the value of the yen. To avoid this currency risk, the investor could buy a quanto forward contract on the index with expiry in one year. The contract has a notional in

dollars, N^{USD} and a forward price F^Q in yen. The payout at expiration depends on the yen value of the index, S_T, and is given by

$$V_T = N^{USD}(S_T - F^Q) \tag{9.11}$$

The forward price on the quanto depends on the usual forward price in the foreign currency but also the covariance between movements in the yen stock price and in the dollar price of the yen

$$F^Q = F \exp(\rho \sigma_{ST} \sigma_{FX} \tau) \tag{9.12}$$

As an example, suppose we consider a one-year contract. The Nikkei index is at $S_0 = 10,000$, the dividend yield is at 2%, and the interest rate at 1%. The usual forward price in yen is $F = S \exp(-[y - r]\tau) = 10,000 \exp(-[0.01 - 0.02] \times 1) = 10,100$. Next, we estimate that the volatility of the Nikkei is $\sigma_{ST} = 20\%$, the volatility of the yen is $\sigma_{FX} = 10\%$, and their correlation 0.3. Based on this information, the quanto forward price should be $F^Q = 10,100 \exp(0.3 \times 0.2 \times 0.1 \times 1) = 10,161$.

Now assume that after one year, the Nikkei has gone up to 11,000 but that the yen has depreciated from \$(1/100) to \$(1/120). Normally, this fall would have wiped out the gain on the index. An initial investment of $V_0 = 10,000/100 = \$100$ would have turned into $V_1 = (11,000 + 2\% \times 10,000)/120 = \$93.33$, which is a loss of \$6.67.

If instead the investor had entered a quanto forward with notional amount of \$0.01, the payoff on the contract would be $V_T = N^{USD}(S_T - F^Q) = 0.01(11,000 - 10,161) = \$8.39$. Thus, the quanto has succeeded in extracting the stock price appreciation only.

An example of such contracts is the dollar-denominated Nikkei futures contract traded on the CME, where the payoff is given by \$5 times the change in the Nikkei futures. The CME also has traditional yen-denominated Nikkei futures contracts, where the value is 500 yen times the Nikkei futures. Unlike the quanto, profits on such contracts are affected by currency swings. Effectively, the quanto has an embedded currency forward contract with a variable notional amount.

EXAMPLE 9.5: FRM EXAM 2003—QUESTION 2

The current spot CHF/USD rate is 1.3680CHF. The three-month USD interest rate is 1.05%, the three-month Swiss interest rate is 0.35%, both continuously compounded and per annum. A currency trader notices that the three-month forward price is USD 0.7350. In order to arbitrage, the trader should

a. Borrow CHF, buy USD spot, go long Swiss franc forward
b. Borrow CHF, sell Swiss franc spot, go short Swiss franc forward
c. Borrow USD, buy Swiss franc spot, go short Swiss franc forward
d. Borrow USD, sell USD spot, go long Swiss franc forward

9.5 CURRENCY SWAPS

Currency swaps are agreements by two parties to exchange a stream of cash flows in different currencies according to a prearranged formula.

9.5.1 Instruments

Consider two counterparties, company A and company B that can raise funds either in dollars or in yen, $100 million or Y10 billion at the current rate of 100Y/$, over 10 years. Company A wants to raise dollars, and company B wants to raise yen. Table 9.4a displays borrowing costs. This example is similar to that of interest rate swaps, except that rates are now in different currencies.

Company A has an absolute advantage in the two markets as it can raise funds at rates systematically lower than company B. Company B, however, has a comparative advantage in raising dollars as the cost is only 0.50% higher than for company A, compared to the cost difference of 1.50% in yen. Conversely, company A must have a comparative advantage in raising yen.

This provides the basis for a swap which will be to the mutual advantage of both parties. If both institutions directly issue funds in their final desired market, the total cost will be 9.50% (for A) and 6.50% (for B), for a total of 16.00%. In contrast, the total cost of raising capital where each has a comparative advantage is 5.00% (for A) and 10.00% (for B), for a total of 15.00%. The gain to both parties from entering a swap is 16.00 -- 15.00 = 1.00%. For instance, the swap described in Tables 9.4b and 9.4c splits the benefit equally between the two parties.

Company A issues yen debt at 5.00%, then enters a swap whereby it promises to pay 9.00% in dollars in exchange for receiving 5.00% yen payments. Its effective funding cost is therefore 9.00%, which is less than the direct cost by 50bp.

TABLE 9.4a   Cost of Capital Comparison

Company	Yen	Dollar
A	5.00%	9.50%
B	6.50%	10.00%

TABLE 9.4b   Swap to Company A

Operation	Yen	Dollar
Issue debt	Pay yen 5.00%	
Enter swap	Receive yen 5.00%	Pay dollar 9.00%
Net		Pay dollar 9.00%
Direct cost		Pay dollar 9.50%
Savings		0.50%

TABLE 9.4c Swap to Company B

Operation	Dollar	Yen
Issue debt	Pay dollar 10.00%	
Enter swap	Receive dollar 9.00%	Pay yen 5.00%
Net		Pay yen 6.00%
Direct cost		Pay yen 6.50%
Savings		0.50%

Similarly, company B issues dollar debt at 10.00%, then enters a swap whereby it receives 9.00% in dollars in exchange for paying 5.00% yen. If we add up the difference in dollar funding cost of 1.00% to the 5.00% yen funding costs, the effective funding cost is therefore 6.00%, which is less than the direct cost by 50bp.[3] Both parties benefit from the swap.

While payments are typically netted for an interest rate swap, because they are in the same currency, this is not the case for currency swaps. Full interest payments are made in different currencies. In addition, at initiation and termination, there is exchange of principal in different currencies. For instance, assuming annual payments, company A will receive 5.0% on a notional of Y10b, which is Y500 million in exchange for paying 9.0% on a notional of $100 million, or $9 million every year.

9.5.2 Pricing

Consider now the pricing of the swap to company A. This involves receiving 5.00% yen in exchange for paying 9.00% dollars. As with interest rate swaps, we can price the swap using either of two approaches, taking the difference between two bond prices or valuing a sequence of forward contracts.

This swap is equivalent to a long position in a fixed-rate, a 5% 10-year yen denominated bond and a short position in a 10-year 9% dollar denominated bond. The value of the swap is that of a long yen bond minus a dollar bond. Defining S as the dollar price of the yen and P and P^* as the dollar and yen bond, we have:

$$V = S(\$/Y)P^*(Y) - P(\$) \tag{9.13}$$

Here, we indicate the value of the yen bond by an asterisk, P^*.

---

[3] Note that B is somewhat exposed to currency risk, as funding costs cannot be simply added when they are denominated in different currencies. The error, however, is of a second-order magnitude.

In general, the bond value can be written as $P(c, r, F)$ where the coupon is c, the yield is r and the face value is F. Our swap is initially worth (in millions)

$$V = \frac{1}{100} P(5\%, 5\%, Y10{,}000) - P(9\%, 9\%, \$100)$$

$$= \frac{\$1}{Y100} Y10{,}000 - \$100 = \$0$$

Thus, the initial value of the swap is zero, assuming a flat term structure for both countries and no credit risk.

We can identify three conditions under which the swap will be in-the-money. This will happen if the value of the yen S appreciates, or if the yen interest rate r^* falls, or if the dollar interest rate r goes up.

Thus the swap is exposed to three risk factors, the spot rate, and two interest rates. The latter exposures are given by the duration of the equivalent bond.

KEY CONCEPT

A position in a receive-foreign currency swap is equivalent to a long position in a foreign currency bond offset by a short position in a dollar bond.

The swap can be alternatively valued as a sequence of forward contracts. Recall that the valuation of a forward contract on one yen is given by

$$V_i = (F_i - K)\exp(-r_i \tau_i) \tag{9.14}$$

using continuous compounding. Here, r_i is the dollar interest rate, F_i is the prevailing forward rate (in \$/yen), and K is the locked-in rate of exchange defined as the ratio of the dollar to yen payment on this maturity. Extending this to multiple maturities, the swap is valued as

$$V = \sum_i n_i (F_i - K)\exp(-r_i \tau_i) \tag{9.15}$$

where $n_i F_i$ is the dollar value of the yen payments translated at the forward rate and the other term $n_i K$ is the dollar payment in exchange.

Table 9.5 compares the two approaches for a three-year swap with annual payments. Market rates have now changed and $r = 8\%$ for U.S. yields and $r^* = 4\%$ for yen yields. We assume annual compounding. The spot exchange rate has moved from 100Y/\$ to 95Y/\$, reflecting a depreciation of the dollar (or appreciation of the yen).

TABLE 9.5 Pricing a Currency Swap

	Specifications		Market Data
	Notional Amount (millions)	Contract Rates	Market Rates
Dollar	$100	9%	8%
Yen	Y10,000	5%	4%
Exchange rate		100Y/$	95Y/$

Valuation Using Bond Approach (millions)

	Dollar Bond			Yen Bond		
Time (year)	Dollar Payment	PV($1)	PV(CF)	Yen Payment	PV(Y1)	PV(CF)
1	9	0.9259	8.333	500	0.9615	480.769
2	9	0.8573	7.716	500	0.9246	462.278
3	109	0.7938	86.528	10,500	0.8890	9,334.462
Total			$102.58			Y10,277.51
Swap ($)			−$102.58			$108.18
Value						$5.61

Valuation Using Forward Contract Approach (millions)

Time (year)	Forward Rate (Y/$)	Yen Receipt (Y)	Yen Receipt (Y)	Dollar Payment ($)	Difference CF ($)	PV(CF) ($)
1	91.48	500	5.47	−9.00	−3.534	−3.273
2	88.09	500	5.68	−9.00	−3.324	−2.850
3	84.83	10,500	123.78	−109.00	14.776	11.730
Value						$5.61

The middle panel shows the valuation using the difference between the two bonds. First, we discount the cash flows in each currency at the newly prevailing yield. This gives $P = \$102.58$ for the dollar bond and Y10,277.51 for the yen bond. Translating the latter at the new spot rate of Y95, we get $108.18. The swap is now valued at $108.18 − $102.58, which is a positive value of $V = \$5.61$ million. The appreciation of the swap is principally driven by the appreciation of the yen.

The bottom panel shows how the swap can be valued by a sequence of forward contracts. First, we compute the forward rates for the three maturities. For example, the one-year rate is $95 \times (1 + 4\%)/(1 + 8\%) = 91.48$ Y/$, by interest rate parity. Next, we convert each yen receipt into dollars at the forward rate, for example Y500 million in one year, which is $5.47 million. This is offset against a payment of $9 million, for a net planned cash outflow of −$3.53 million. Discounting and adding up the planned cash flows, we get $V = \$5.61$ million, which must be exactly equal to the value found using the alternative approach.

EXAMPLE 9.6: FRM EXAM 2004—QUESTION 54

Which of the following statements is correct when comparing the differences between an interest rate swap and a currency swap?

a. At maturity, there is no exchange of principal between the counterparties in interest rate swaps and there is an exchange of principal in currency swaps.

b. At maturity, there is no exchange of principal between the counterparties in currency swaps and there is an exchange of principle in interest rate swaps.

c. The counterparties in a interest rate swap need to consider fluctuations in exchange rates, while currency swap counterparties are only exposed to fluctuations in interest rates.

d. Currency swap counterparties are exposed to less counterparty credit risk due to the offsetting effect of currency and interest rate risk in the transaction.

EXAMPLE 9.7: FRM EXAM 2006—QUESTION 88

You have entered into a currency swap in which you receive 4%pa in yen and pay 6%pa in dollars once a year. The principals are 1,000 million yen and 10 million dollars. The swap will last for another two years, and the current exchange rate is 115 yen/$. The annualized spot rates (with continuous compounding) are 2.00% and 2.50% in yen for one- and two-year maturities, and 4.50% and 4.75% in dollars. What is the value of the swap to you in million dollars?

a. −1.270

b. −0.447

c. 0.447

d. 1.270

EXAMPLE 9.8: FRM EXAM 2007—QUESTION 87

Your company is expecting a major export order from a London-based client. The receivables under the contract are to be billed in GBP, while your reporting currency is USD. Since the order is a large sum, your company does not want to bear the exchange risk and wishes to hedge it using derivatives. To minimize the cost of hedging, which of the following is the most suitable contract?

a. A chooser option for GBP/USD pair

b. A currency swap where you pay fixed in USD and receive floating in GBP

c. A barrier put option to sell GBP against USD

d. An Asian call option on GBP against USD

9.6 COMMODITIES

9.6.1 Products

Commodities are typically traded on exchanges. Contracts include spot, futures, and options on futures. There is also an OTC market for long-term commodity swaps, where payments are tied to the price of a commodity against a fixed or floating rate.

Commodity contracts can be classified into:

- Agricultural products, including grains and oilseeds (corn, wheat, soybean) food and fiber (cocoa, coffee, sugar, orange juice)
- Livestock and meat (cattle, hogs)
- Base metals (aluminum, copper, nickel, and zinc)
- Precious metals (gold, silver, platinum)
- Energy products (natural gas, heating oil, unleaded gasoline, crude oil)

The Goldman Sachs Commodity Index (GSCI) is a broad production-weighted index of commodity price performance, which is composed of 24 liquid exchange-traded futures contracts as of 2008. The index contains 72% energy products, 7% industrial metals, 2% precious metals, 14% agricultural products, and 5% livestock products. The CME trades futures and options contracts on the GSCI.

In the last few years, active markets have developed for electricity products, electricity futures for delivery at specific locations, for instance California/Oregon border (COB), Palo Verde, and so on. These markets have mushroomed following the deregulation of electricity prices, which has led to more variability in electricity prices.

More recently, OTC markets and exchanges have introduced weather derivatives, where the payout is indexed to temperature or precipitation. On the CME, for instance, contract payouts are based on the Degree Day Index over a calendar month. This index measures the extent to which the daily temperature deviates from the average. These contracts allow users to hedge situations where their income is negatively affected by extreme weather. Markets are also evolving in newer products, such as indices of consumer bankruptcy and catastrophe insurance contracts.

Such commodity markets allow participants to exchange risks. Farmers, for instance, can sell their crops at a fixed price on a future date, insuring themselves against variations in crop prices. Likewise, consumers can buy these crops at a fixed price.

9.6.2 Pricing of Futures

Commodities differ from financial assets in two notable dimensions: they may be expensive, even impossible, to store and they may generate a flow of benefits that are not directly measurable.

The first dimension involves the cost of carrying a physical inventory of commodities. For most financial instruments, this cost is negligible. For bulky commodities, this cost may be high. Other commodities, like electricity cannot be stored easily.

The second dimension involves the benefit from holding the physical commodity. For instance, a company that manufactures copper pipes benefits from an inventory of copper which is used up in its production process. This flow is also called convenience yield for the holder. For a financial asset, this flow would be a monetary income payment for the investor. When an asset such as gold can be lent out for a profit, the yield represents the lease rate, which is the return to lending gold short-term.

Consider the first factor, storage cost only. The cash-and-carry relationship should be modified as follows. We compare two positions. In the first, we buy the commodity spot plus pay up-front the present value of storage costs PV(C). In the second, we enter a forward contract and invest the present value of the forward price. Since the two positions are identical at expiration, they must have the same initial value:

$$F_t e^{-r\tau} = S_t + \text{PV}(C) \qquad (9.16)$$

where $e^{-r\tau}$ is the present value factor. Alternatively, if storage costs are incurred per unit time and defined as c, we can restate this relationship as

$$F_t e^{-r\tau} = S_t e^{c\tau} \qquad (9.17)$$

Due to these costs, the forward rate should be much greater than the spot rate, as the holder of a forward contract benefits not only from the time value of money but also from avoiding storage costs.

Example: Computing the Forward Price of Gold

Let us use data from December 1999. The spot price of gold is $S = \$288$, the one-year interest rate is $r = 5.73\%$ (continuously compounded), and storage costs are $2 per ounce per year, paid up front. The fair price for a one-year forward contract should be $F = [S + \text{PV}(C)]e^{r\tau} = [\$288 + \$2]e^{5.73\%} = \$307.1$.

Let us now turn to the convenience yield, which can be expressed as y per unit time. In fact, y represents the net benefit from holding the commodity, after storage costs. Following the same reasoning as before, the forward price on a commodity should be given by

$$F_t e^{-r\tau} = S_t e^{-y\tau} \qquad (9.18)$$

where $e^{-y\tau}$ is an actualization factor. This factor may have an economically identifiable meaning, reflecting demand and supply conditions in the cash and futures

markets. Alternatively, it can be viewed as a *plug-in* that, given F, S, and e^{-rt}, will make Equation (9.18) balance.

Figure 9.3, for example, displays the shape of the term structure of spot and futures prices for the New York Mercantile Exchange (NYMEX) crude oil contract. On December 1997, the term structure is relatively flat. On December 1998, the term structure becomes strongly upward-sloping. Part of this slope can be explained by the time value of money (the term e^{-rt} in the equation). In contrast, the term structure is downward-sloping on December 1999. This can be interpreted in terms of a large convenience yield from holding the physical asset (in other words, the term e^{-yt} in the equation dominates).

Let us focus for example on the one-year contract. Using $S = \$25.60$, $F = \$20.47$, $r = 5.73\%$ and solving for y,

$$y = r - \frac{1}{\tau}\ln(F/S) \tag{9.19}$$

we find $y = 28.10\%$, which is quite large. In fact, variations in y can be substantial. Just one year before, a similar calculation would have given $y = -9\%$, which implies a negative convenience yield, or a storage cost.

Table 9.6 displays futures prices for selected contracts. Futures prices are generally increasing with maturity, reflecting the time value of money, storage cost, and low convenience yields. There are some irregularities, however, reflecting anticipated imbalances between demand and supply. For instance, gasoline futures prices increase in the summer due to increased automobile driving. Natural gas displays the opposite pattern, where prices increase during the winter due to the demand for heating. Agricultural products can also be highly seasonal. In contrast, futures prices for gold are going up monotonically with time, since this is a perfectly storable good.
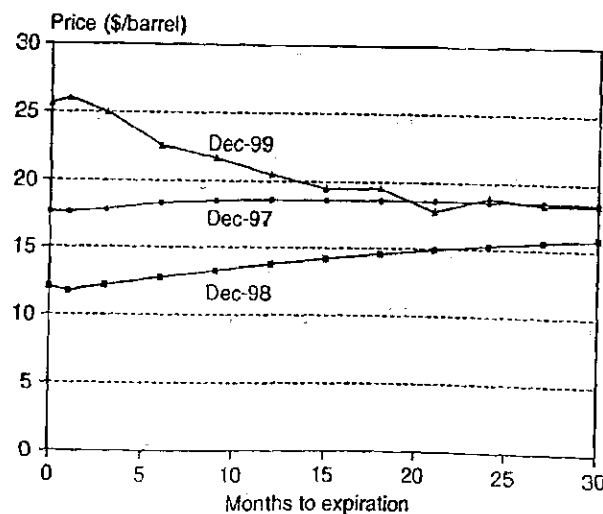
FIGURE 9.3   Spot and Futures Prices for Crude Oil

TABLE 9.6   Futures Prices as of December 30, 1999

Maturity	Corn	Sugar	Copper	Gold	Nat.Gas	Gasoline
Jan			85.25	288.5		0.6910
Mar	204.5	18.24	86.30	290.6	2.328	0.6750
July	218.0	19.00	87.10	294.9	2.377	0.6675
Sept	224.0	19.85	87.90	297.0	2.418	0.6245
Dec	233.8	18.91	88.45	300.1	2.689	
Mar 01	241.5	18.90	88.75	303.2	2.494	
...						
Dec 01	253.5			312.9	2.688	

9.6.3 Futures and Expected Spot Prices

An interesting issue is whether today's futures price gives the best forecast of the future spot price. If so, it satisfies the expectations hypothesis, which can be written as:

$$F_t = E_t[S_T] \tag{9.20}$$

The reason this relationship may hold is as follows. Say that the one-year oil futures price is $F = \$20.47$. If the market forecasts that oil prices in one year will be at $25, one could make a profit by going long a futures contract at the cheap futures price of $F = \$20.47$, waiting a year, then buying oil at $20.47, and reselling it at the higher price of $25. In other words, deviations from this relationship imply speculative profits.

To be sure, these profits are not risk-free. Hence, they may represent some compensation for risk. For instance, if the market is dominated by producers who want to hedge by selling oil futures, F will be abnormally low compared with expectations. Thus the relationship between futures prices and expected spot prices can be complex.

For financial assets for which the arbitrage between cash and futures is easy, the futures or forward rate is solely determined by the cash-and-carry relationship; i.e., the interest rate and income on the asset. For commodities, however, the arbitrage may not be so easy. As a result, the futures price may deviate from the cash-and-carry relationship through this convenience yield factor. Such prices may reflect expectations of futures spot prices, as well as speculative and hedging pressures.

A market is said to be in contango when the futures price trades at a premium relative to the spot price, as shown in Figure 9.4. Using Equation (9.19), this implies that the convenience yield is smaller than the interest rate $y < r$.

Normally, the size of the premium should be limited by arbitrage opportunities. If this became too large, traders could buy the commodity spot, put it in storage, and simultaneously sell it for future delivery at the higher forward price. In December 2008, however, the premium for one-year oil contracts reached an

Futures price

Backwardation

Contango

Maturity ————▶
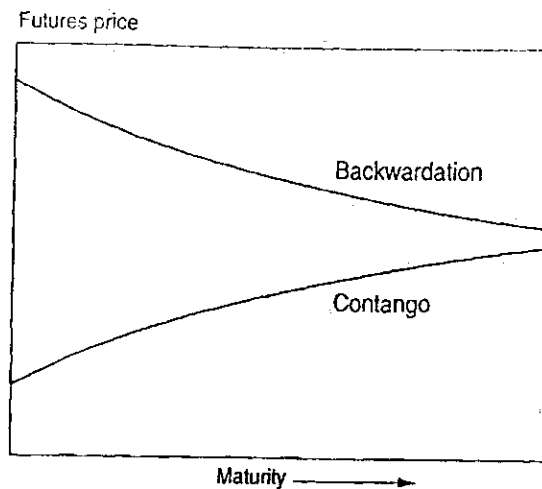
FIGURE 8.4 Patterns of Contango and Backwardation

all-time high of $13 per barrel. This was explained by the credit crunch, which prevented oil traders to secure loans to finance oil storage.

A market is said to be in backwardation (or inverted) when forward prices trade at a discount relative to spot prices. This implies that the convenience yield is greater than the interest rate $y > r$. In other words, a high convenience yields puts a higher price on the cash market, as there is great demand for immediate consumption of the commodity.

With backwardation, the futures price tends to increase as the contract nears maturity. In such a situation, a roll-over strategy should be profitable, provided that prices do not move too much. This involves buying a long maturity contract, waiting, and then selling it at a higher price in exchange for buying a cheaper, longer-term contract.

This strategy is comparable to riding the yield curve when upward-sloping. This involves buying long maturities and waiting to have yields fall due to the passage of time. If the shape of the yield curve does not change too much, this will generate a capital gain from bond price appreciation. Because of the negative price-yield relationship, a positively sloped yield curve is equivalent to backwardation in bond prices.

This was basically the strategy followed by Metallgesellschaft Refining & Marketing (MGRM), the U.S. subsidiary of Metallgesellschaft, which had made large sales of long-term oil to clients on the OTC market. These were hedged by rolling over long positions in West Texas Intermediate (WTI) crude oil futures. This made money as long as the market was in backwardation. When the market turned to contango, however, the long positions started to lose money as they got closer to maturity. In addition, the positions were so large that they moved markets against MG. These losses caused cash-flow, or liquidity problems. MGRM ended up liquidating the positions, which led to a realized loss of $1.3 billion.

A similar problem afflicted Amaranth, a hedge fund that lost $6.6 billion as a result of bad bets against natural gas futures. In September 2006, the price

of natural gas fell sharply. In addition, the spread between prices in winter and summer months collapsed. As the size of the positions were huge, this led to large losses that worsened when the fund attempted to liquidate the contracts.

KEY CONCEPT

Markets are in contango if spot prices are lower than forward prices. Markets are in backwardation if spot prices are higher than forward prices. Backwardation occurs when there is high current demand for the commodity, which implies high convenience yields.

EXAMPLE 9.9: FRM EXAM 2007—QUESTION 29

On January 1, a risk manager observes that the one-year continuously compounded interest rate is 5% and storage costs of a commodity product A is USD 0.05 per quarter (payable at each quarter end). He further observes the following forward prices for product A: March, 5.35; June, 5.90; September, 5.30; December, 5.22. Given the following explanation of supply and demand for this product, how would you best describe its forward price curve from June to December?

a. Backwardation as the supply of product A is expected to decline after summer
b. Contango as the supply of product A is expected to decline after summer
c. Contango as there is excess demand for product A in early summer
d. Backwardation as there is excess demand for product A in early summer

EXAMPLE 9.10: FRM EXAM 2007—QUESTION 30

Continuing with the previous question, what is the annualized rate of return earned on a cash-and-carry trade entered into in March and closed out in June?

a. 9.8%
b. 8.9%
c. 39.1%
d. 35.7%

EXAMPLE 9.11: FRM EXAM 2004—QUESTION 5

Which of the following causes led MGRM into severe financial distress?

I. There was a mismatch of cash flows from hedge and physical transactions.
II. MGRM failed to consider hedging market risk from fixed price physical sales contracts.
III. MGRM held a great percentage of the total open interest on the NYMEX.
IV. The futures market went from backwardation to contango.

    a. I and III
    b. I and IV
    c. I, III, and IV
    d. II, III and IV

9.7 IMPORTANT FORMULAS

Gordon-growth model for valuation of stocks: $P = \frac{D}{r-g}$

Warrant valuation: $W_0 = n \times c(S_0 + \frac{M}{N}W_0, K, \tau, \sigma, r, d)$

Stock index futures: $F_t e^{-r\tau} = S_t e^{-y\tau}$

Payoff on a variance swap: $V_T = (\sigma^2 - K_V)N$

Valuation of an outstanding variance swap:
$$V_t = Ne^{-r\tau}[w(\sigma_{t_0,t}^2 - K_V) + (1 - w)(K_t - K_V)]$$

Pricing a currency swap as two bond positions: $V = S(\$/Y)P^*(Y) - P(\$)$

Pricing a currency swap as a sequence of forwards:
$$V = \sum_i n_i(F_i - K)\exp(-r_i\tau_i)$$

Pricing of commodity futures with storage costs:
$$F_t e^{-r\tau} = S_t + PV(C), \text{ or } F_t e^{-r\tau} = S_t e^{c\tau}$$

Expectations hypothesis: $F_t = E_t[S_T]$

9.8 ANSWERS TO CHAPTER EXAMPLES

Example 9.1: Leverage and Return on Equity

c. The fund borrows $200 million and invests $300 million, which creates a yield of $300 \times 14\% = \$42$ million. Borrowing costs are $200 \times 8\% = \$16$ million, for a difference of $26 million on equity of $100 million, or 26%. Note that this is a yield, not expected rate of return if we expect some losses from default. This higher yield also implies higher risk.

Example 9.2: FRM Exam 2001—Question 119

a. The conversion rate is expressed here in terms of the conversion price. The conversion rate for this bond is $100 into $40, or one bond into 2.5 shares. Immediate conversion will yield $2.5 \times \$45 = \$112.5$. The call price is $106. Since the market price is higher than the call price and the conversion value, and the bond is being called, the best value is achieved by selling the bond.

Example 9.3: FRM Exam 2001—Question 117

d. Companies issue convertible bonds because the coupon is lower than for regular bonds. In addition, these bonds are callable in order to force conversion into the stock at a favorable ratio. In the previous question, for instance, conversion would provide equity capital to the firm at the price of $40, while the market price is at $45.

Example 9.4: FRM Exam 2000—Question 12

a. This is the cash-and-carry relationship, solved for S. We have $Se^{-y\tau} = Fe^{-r\tau}$, or $S = 552.3 \times \exp(-7.5/200)/\exp(-4.2/200) = 543.26$. We verify that the forward price is greater than the spot price since the dividend yield is less than the risk-free rate.

Example 9.5: FRM Exam 2003—Question 2

c. For consistency, translate the spot rate in dollars, $S = 0.7310$. The CHF interest rate is lower than the USD rate, so the CHF must be selling at a forward premium. The fair forward price is $F = S \exp((r - r^*)\tau) = 0.7310 \exp((0.0105 - 0.0035) \, 0.25) = 0.7323$. Because this is less than the observed price of 0.7350, we sell at the expensive forward price and borrow USD, buy CHF spot, invest in CHF. At maturity, we liquidate the CHF investment to satisfy the forward sale into dollars, repay the loan, and make a tidy profit.

Example 9.6: FRM Exam 2004—Question 54

a. Because payments on currency swaps are in different currencies, they cannot be netted.

Example 9.7: FRM Exam 2006—Question 88

a. The net present values of the payoffs in two currencies are described in the next table. As a result, the value of the currency swap is given by the dollar value of a long position in the yen bond minus a position in the dollar bond, or $(1/115)$ $1,000(102.85/100) - 10(102.13/100) = \$8.943 - \$10.213 = -\$1.270$.

T	Yen Rate	CF	NPV	USD Rate	CF	NPV
1	2.00%	4	3.92	4.50%	6	5.74
2	2.50%	104	98.93	4.75%	106	96.39
Sum			102.85			102.13

Example 9.8: FRM Exam 2007—Question 87

c. A cross-currency swap is inappropriate because there is no stream of payment but just one. Also, one would want to pay GBP, not receive it. An Asian option is generally cheap, but this should be a put option, not a call. Among the two remaining choices, the chooser option is more expensive because it involves a call and put.

Example 9.9: FRM Exam 2007—Question 29

d. From June to December, prices go down, which is backwardation. June prices are abnormally high because of excess demand, which pushes prices up.

Example 9.10: FRM Exam 2007—Question 30

d. The trade involves now going long a March contract and short a June contract. In practice, this means taking delivery of the commodity and holding it for three months until resale in June. The final payout is $5.90 - 0.05$ on a base of 5.35. This gives an annualized rate of return of $r = 4 \ln(5.85/5.35) = 35.7\%$.

Example 9.11: FRM Exam 2003—Question 5

c. MGRM did consider hedging its OTC contracts with futures but was hit with liquidity risk as the long futures positions lost money due to the move into contango. In addition, the positions were very large, which led to losses on the unwinding of the hedges.

Market Risk Management
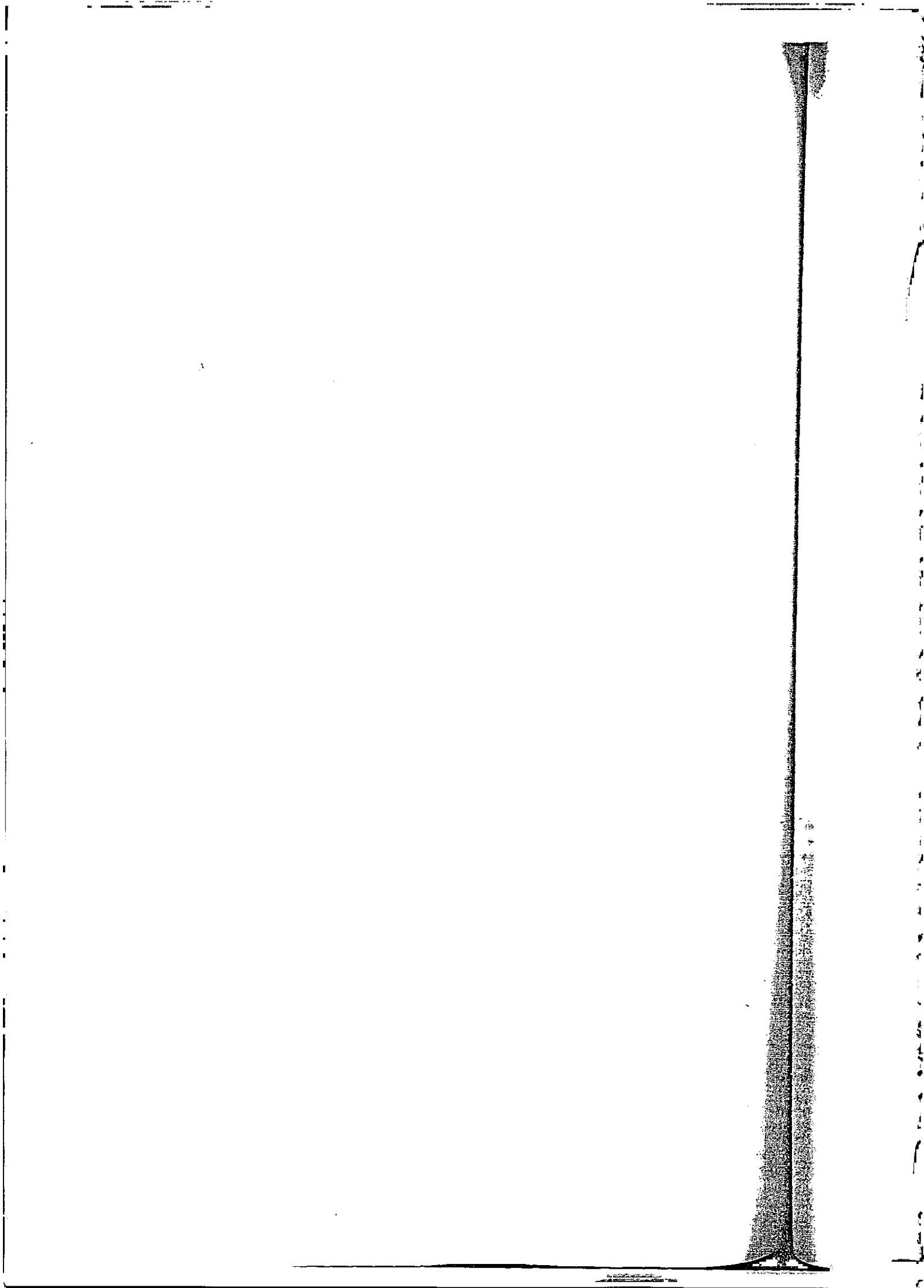
Introduction to Market Risk

This chapter provides an introduction to market risk. Market risk is primarily measured with position-based risk measures such as value at risk (VAR). VAR is a statistical measure of *total* portfolio risk, based on the most current positions, which takes into account portfolio diversification and leverage.

In theory, risk managers should consider the entire distribution of profits and losses over the specified horizon. In practice, this distribution is summarized by one number, the worst loss at a specified confidence level, such as 99%. VAR, however, is only one of the measures that risk managers focus on. It should be complemented by stress-testing, which identifies potential losses under extreme market conditions.

Section 10.1 gives a brief overview of financial market risks and the history of risk measurement systems. Section 10.2 then introduces measures of downside risk. It shows how to compute VAR for a very simple portfolio. It also discusses caveats, or pitfalls to be aware of when interpreting VAR numbers. Section 10.3 extends VAR methods to cash flow at risk. Section 10.4 turns to the choice of VAR parameters, that is, the confidence level and horizon. Next, Section 10.5 describes the broad components of a VAR system. Finally, Section 10.6 shows how to complement VAR by stress tests.

10.1 INTRODUCTION TO FINANCIAL MARKET RISKS

10.1.1 Types of Financial Risks

Financial risks include market risk, credit risk, and operational risk. Market risk is the risk of losses due to movements in financial market prices or volatilities. This usually includes liquidity risk, which is the risk of losses due to the need to liquidate positions to meet funding requirements. Liquidity risk, unfortunately, is not amenable to formal quantification. Because of its importance, it will be covered in Chapter 25. Credit risk is the risk of losses due to the fact that counterparties may be unwilling or unable to fulfill their contractual obligations. Operational risk is the risk of loss resulting from failed or inadequate internal processes, systems, and people, or from external events. Oftentimes, however, these three categories interact with each other, so that any classification is, to some extent, arbitrary.

For example, credit risk can interact with other types of risks. At the most basic level, it involves the risk of default on the asset, such as a loan or bond. When the asset is traded, however, market risk also reflects credit risk. Take a corporate bond, for example. Some of the price movement may be due to movements in risk-free interest rates, which is pure market risk. The remainder will reflect the market's changing perception of the likelihood of default. Thus, for traded assets, there is no clear-cut delineation of market and credit risk. Some arbitrary classification must take place. Furthermore, operational risk is often involved as well.

Consider a simple transaction whereby a trader purchases 1 million worth of a British pound (BP) spot from Bank A. The current rate is $1.5/BP, for settlement in two business days. So, our bank will have to deliver $1.5 million in two days in exchange for receiving BP 1 million. This simple transaction involves a series of risks.

- *Market risk:* During the day, the spot rate could change. Say that after a few hours the rate moves to $1.4/BP. The trader cuts the position and enters a spot sale with another bank, Bank B. The million pounds is now worth only $1.4 million, for a loss of $100,000 to be realized in two days. The loss is the change in the market value of the investment.
- *Credit risk:* The next day, Bank B goes bankrupt. The trader must now enter a new, replacement trade with Bank C. If the spot rate has dropped further from $1.4/BP to $1.35/BP, the gain of $50,000 on the spot sale with Bank B is now at risk. The loss is the change in the market value of the investment, if positive. Thus there is interaction between market and credit risk.
- *Settlement risk:* The next day, our bank wires the $1.5 million to Bank A in the morning, which defaults at noon and does not deliver the promised BP 1 million. This is also known as Herstatt risk because this German bank defaulted on such obligations in 1974, potentially destabilizing the whole financial system. The loss is now potentially the whole principal in dollars.
- *Operational risk:* Suppose that our bank wired the $1.5 million to a wrong bank, Bank D. After two days, our back office gets the money back, which is then wired to Bank A plus compensatory interest. The loss is the interest on the amount due.

10.1.2 Risk Management Tools

In the past, risks were measured using a variety of ad hoc tools, none of which was satisfactory. These included notional amounts, sensitivity measures, and scenarios. While these measures provide some intuition of risk, they do not measure what matters, that is, the potential for downside loss for the total portfolio. They fail to take into account differences in volatilities across markets, correlations across risk factors, as well as the probability of adverse moves in the risk factors.

Consider for instance a five-year inverse floater, which pays a coupon equal to 16 percent minus twice current LIBOR, if positive, on a notional principal of $100 million. The initial market value of the note is $100 million. This investment is

extremely sensitive to movements in interest rates. If rates go up, the present value of the cash flows will drop sharply. In addition, the discount rate also increases. The combination of a decrease in the numerator terms and an increase in the denominator terms will push the price down sharply.

The question is, how much could an investor lose on this investment over a specified horizon? The *notional amount* only provides an indication of the potential loss. The worst case scenario is one where interest rates rise above 8%. In this situation, the coupon will drop to $16 - 2 \times 8 =$ zero. The bond becomes a zero-coupon bond, whose value is $68 million, discounted at 8%. This gives a loss of $100 - $68 = $32 million. While sizable, this is still less than the notional.

A *sensitivity measure* such as duration is more helpful. As we have seen in Chapter 7, the bond has three times the duration of a similar five-year note. Assume the latter is 4.5 years. This gives a modified duration of $D = 3 \times 4.5 = 13.5$ years. This duration measure reveals the extreme sensitivity of the bond to interest rates but does not answer the question of whether such a disastrous movement in interest rates is likely. It also ignores the nonlinearity between the note price and yields.

Scenario analysis provides some improvement, as it allows the investor to investigate nonlinear, extreme effects in price. But again, the method does not associate the loss with a probability.

Another general problem is that these sensitivity or scenario measures do not allow the investor to aggregate risk across different markets. Let us say that this investor also holds a position in a bond denominated in another currency, the euro. Do the risks add up, or diversify each other?

The great beauty of VAR is that it provides a neat answer to all these questions. One number aggregates the risks across the whole portfolio, taking into account leverage and diversification, and providing a risk measure with an associated probability.

If the worst increase in yield at the 95% level is 1.65%, we can compute VAR as

$$VAR = \text{Market value} \times \text{Modified Duration} \times \text{Worst yield increase} \qquad (10.1)$$

In this case, VAR $= \$100 \times 13.5 \times 0.0165 = \$22$ million. The investor can now make a statement such as: The worst loss at the 95% confidence level is approximately $22 million. With appropriate caveats, this is a huge improvement over traditional risk measurement methods, as it expresses risk in an intuitive fashion. Once the VAR apparatus is in place, it is easy to implement stress-tests, which provide information about extreme losses that is complementary to VAR.

Measures such as notional amounts and exposures have been, and are still used to set limits, in an attempt to control risk before it occurs, or "ex ante." These measures should be supplemented by VAR, which is an ex ante measure of the potential dollar loss. Other risk management tools include stop losses, which are rules enforcing position cuts after losses occur, that is "ex post." While stop losses are useful, especially in trending markets, they only provide partial protection because they are applied *after* a loss.

The VAR revolution started in 1993 when it was endorsed by the Group of Thirty (G-30) as part of "best practices" for dealing with derivatives. The methodology behind VAR, however, is not new. It results from a merging of finance theory, which focuses on the pricing and sensitivity of financial instruments, and statistics, which studies the behavior of the risk factors. The idea behind VAR, or total portfolio risk, can be traced to the pioneering work of Markowitz in 1952.

EXAMPLE 10.1: FRM EXAM 2005—QUESTION 82

Which of the following statements about trader limits are *correct?*

I. Stop-loss limits are useful if markets are trending.
II. Exposure limits do not allow for diversification.
III. VAR limits are not susceptible to arbitrage.
IV. Stop-loss limits are effective in preventing losses.
  a. I and II
  b. III and IV
  c. I and III
  d. II and IV

10.2 DOWNSIDE RISK MEASURES

10.2.1 VAR: Definition

Value at Risk VAR is a summary measure of downside risk expressed in dollars, or in the reference currency. A general definition is

VAR is the maximum loss over a target horizon such that there is a low, prespecified probability that the actual loss will be larger.

Consider for instance a position of $4 billion short the yen, long the dollar. This position corresponds to a well-known hedge fund that took a bet that the yen would fall in value against the dollar. How much could this position lose over a day?

To answer this question, we could use 10 years of historical daily data on the yen/dollar rate and simulate a daily return. The simulated daily return in dollars is then

$$R_t(\$) = Q_0(\$)[S_t - S_{t-1}]/S_{t-1} \qquad (10.2)$$

where Q_0 is the current dollar value of the position and S is the spot rate in yen per dollar measured over two consecutive days.

For instance, for two hypothetical days $S_1 = 112.0$ and $S_2 = 111.8$. The simulated return is

$$R_2(\$) = \$4,000 \text{ million} \times [111.8 - 112.0]/112.0 = -\$7.2 \text{ million}$$

Repeating this operation over the whole sample, or 2,527 trading days, creates a time-series of fictitious returns, which is plotted in Figure 10.1.

We can now construct a frequency distribution of daily returns. For instance, there are four losses below \$160 million, three losses between \$160 million and \$120 million, and so on. The histogram, or frequency distribution, is graphed in Figure 10.2. We can also order the losses from worst to best return.

Define x as the dollar profit or loss. VAR is typically reported as a positive number, even if it is a loss. It is defined implicitly by

$$c = \int_{-\text{VAR}}^{\infty} f(x)dx \qquad (10.3)$$

When the outcomes are discrete, VAR is the smallest loss such that the right-tail probability is at least c.

Sometimes, VAR is reported as the deviation between the mean and the quantile. This second definition is more consistent than the usual one. Because it considers the deviation between two values on the target date, it takes into account the time value of money. In most applications, however, the time horizon is very
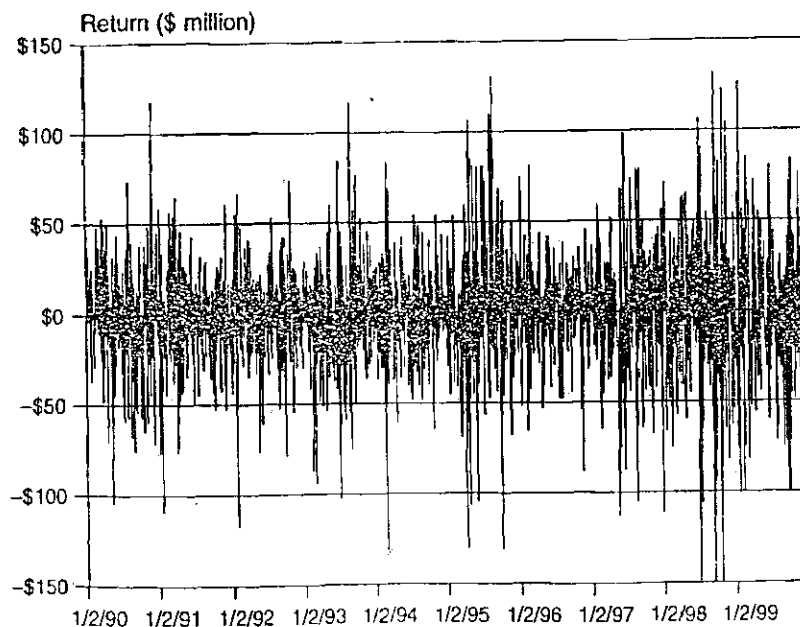
FIGURE 10.1 Simulated Daily Returns
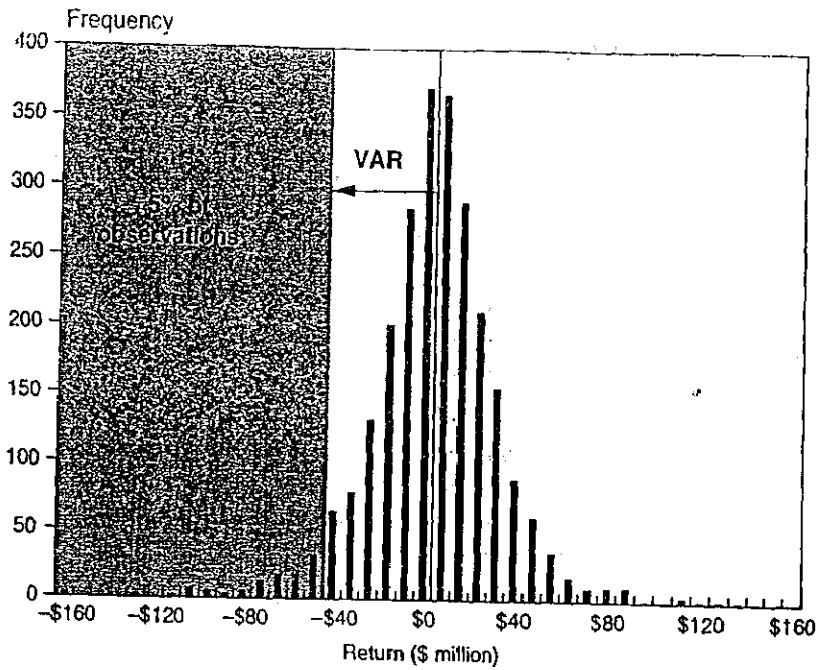
FIGURE 10.2 Distribution of Daily Return

short in which case the average return on financial series is close to zero. As a result, the two definitions usually give similar values.

In this hedge fund example, we want to find the cutoff value R^* such that the probability of a loss worse than $-R^*$ is $p = 1 - c = 5\%$. With a total of $T = 2,527$ observations, this corresponds to a total of $pT = 0.05 \times 2,527 = 126$ observations in the left tail. We pick from the ordered distribution the cutoff value, which is $R^* = \$47.1$ million. We can now make a statement such as: The maximum loss over one day is about $47 million at the 95% confidence level. This describes risk in a way that notional amounts or exposures cannot convey.

We now wish to summarize the distribution by one number. We could describe the quantile, that is, the level of loss that will not be exceeded at some high confidence level. Select for instance this confidence level as $c = 95\%$. This corresponds to a right-tail probability. We could as well define VAR in terms of a left-tail probability, which we write as $p = 1 - c$.

It is essential to verify the quality of VAR forecasts by checking whether the number of losses worse than VAR, also called exceptions, is in line with expectations. From the confidence level, we can determine the number of expected exceedences n over a period of N days:

$$n = p \times N \tag{10.4}$$

Say that we want to backtest VAR measured at the 99% confidence level over the past 250 days. The expected number of exceptions is $n = 0.01 \times 250 = 2.5$.

Under the null hypothesis that the model is correctly calibrated, or that the true probability of exceptions is p, the distribution of number of exceptions x follows a binomial distribution

$$f(x) = \binom{N}{x} p^x (1-p)^{N-x}, \quad x = 0, 1, \dots, n \qquad (10.5)$$

For example, the probability of having exactly zero exceptions is $f(x = 0) = (0.99)^{250} = 0.081$. The probability of having five exceptions is $f(x = 5) = \binom{250}{5} 0.01^5 (0.99)^{245} = 0.0666$. As shown in Chapter 2, this distribution can be used to compute a cutoff value for the number of exceptions, beyond which we would have to conclude that the model is flawed. Overall, there is only a probability of 10.8% of observing five or more exceptions if the VAR model were correctly specified.

Finally, we should note that exception tests only focus on the frequency of occurrences of exceptions. They do not take into account the size of losses.

EXAMPLE 10.2: FRM EXAM 2005—QUESTION 43

The 10-Q report of ABC Bank states that the monthly VAR of ABC Bank is USD 10 million at 95% confidence level. What is the proper interpretation of this statement?

a. If we collect 100 monthly gain/loss data of ABC Bank, we will always see five months with losses larger than $10m.
b. There is a 95% probability that the bank will lose less than $10m over a month.
c. There is a 5% probability that the bank will gain less than $10m each month.
d. There is a 5% probability that the bank will lose less than $10m over a month.

EXAMPLE 10.3: FRM EXAM 2003—QUESTION 11

Based on a 90% confidence level, how many exceptions in backtesting a VAR would be expected over a 250-day trading year?

a. 10
b. 15
c. 25
d. 50

EXAMPLE 10.4: FRM EXAM 2007—QUESTION 101

A large, international bank has a trading book whose size depends on the opportunities perceived by its traders. The market risk manager estimates the one-day VAR, at the 95% confidence level, to be USD 50 million. You are asked to evaluate how good of a job the manager is doing in estimating the one-day VAR. Which of the following would be the most convincing evidence that the manager is doing a poor job, assuming that losses are identically independently distributed?

a. Over the last 250 days, there are eight exceedences.
b. Over the last 250 days, the largest loss is USD 500 million.
c. Over the last 250 days, the mean loss is USD 60 million.
d. Over the last 250 days, there is no exceedence.

10.2.2 VAR: Caveats

VAR is a useful summary measure of risk but is subject to caveats:

- *VAR does not describe the worst loss.* This is not what VAR is designed to measure. Indeed we would expect the VAR number to be exceeded with a frequency of p, that is, five days out of a hundred for a 95% confidence level. This is perfectly normal. In fact, backtesting procedures are designed to check whether the frequency of exceedences is in line with p.

- *VAR does not describe the losses in the left tail.* VAR does not say anything about the distribution of losses in its left tail. It just indicates the probability of such a value occurring. For the same VAR number, however, we can have very different distribution shapes. In the case of Figure 10.2, the average value of the losses worse than $47 million is around $74 million, which is 60% worse than the VAR. So, it would be unusual to sustain many losses beyond $200 million.

  Other distributions are possible, however, while maintaining the same VAR. Figure 10.3 illustrates a distribution with 125 occurrences of large losses of $160 million. Because there is still one observation left at $47 million, the VAR is unchanged at $47 million. Yet this distribution implies a high probability of sustaining very large losses, unlike the original one.

  This can create other strange results. For instance, one can construct examples, albeit stretched, where the VAR of a portfolio is greater than the sum of the VARs for its components. As a result, VAR fails to qualify as a *subadditive* risk measure, which is one of the desirable properties listed in the appendix. Subadditivity implies that the risk of a portfolio must be less than the sum of risks for portfolio components.

- *VAR is measured with some error.* The VAR number itself is subject to normal sampling variation. In our example, we used 10 years of daily data. Another
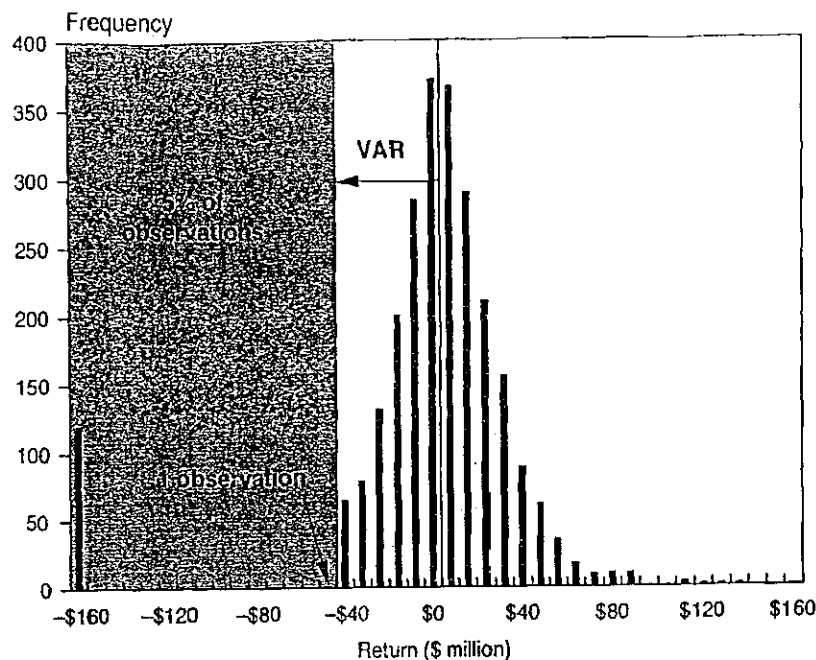
FIGURE 10.3   Altered Distribution with Same VAR

sample period, or a period of different length, will lead to a different VAR number. Different statistical methodologies or simplifications can also lead to different VAR numbers. One can experiment with sample periods and methodologies to get a sense of the precision in VAR. Hence, it is useful to remember that there is limited precision in VAR numbers. What matters is the first-order magnitude.

In addition, VAR measures are subject to the same problems that affect all risk measures based on a "window" of recent historical data. Ideally, the past window should reflect the range of future outcomes. If not, all risk measures based on recent historical data may be misleading.

10.2.3   Alternative Measures of Risk

The conventional VAR measure is the *quantile* of the distribution measured in dollars. This single number is a convenient summary, but its very simplicity may be dangerous. We have seen in Figure 10.3 that the same VAR can hide very different distribution patterns. The appendix reviews desirable properties for risk measures and shows that VAR may display some undesirable properties under some conditions. In particular, the VAR of a portfolio can be greater than the sum of subportfolios VARs. If so, merging portfolios can increase risk, which is an unexpected result. Alternative measures of risk are described below.

The Entire Distribution   In our example, VAR is simply one quantile in the distribution. The risk manager, however, has access to the whole distribution and could report a range of VAR numbers for increasing confidence levels.

The Conditional VAR A related concept is the expected value of the loss when it exceeds VAR. This measures the average of the loss conditional on the fact that it is greater than VAR. Define the VAR number as $-q$. Formally, the conditional VAR (CVAR) is the negative of

$$E[X \mid X < q] = \int_{-\infty}^{q} x f(x) dx \bigg/ \int_{-\infty}^{q} f(x) dx \qquad (10.6)$$

Note that the denominator represents the probability of a loss exceeding VAR, which is also $p = 1 - c$. This ratio is also called expected shortfall, tail conditional expectation, conditional loss, or expected tail loss. CVAR indicates the potential loss if the portfolio is "hit" beyond VAR. Because CVAR is an average of the tail loss, one can show that it qualifies as a *subadditive* risk measure. For our yen position, the average loss beyond the $47 million VAR is CVAR = $74 million.

The Standard Deviation A simple summary measure of the distribution is the usual standard deviation (SD)

$$SD(X) = \sqrt{\frac{1}{(N-1)} \sum_{i=1}^{N} [x_i - E(X)]^2} \qquad (10.7)$$

The advantage of this measure is that it takes into account all observations, not just the few around the quantile. Any large negative value, for example, will affect the computation of the variance, increasing $SD(X)$. If we are willing to take a stand on the shape of the distribution, say normal or Student's t, we do know that the standard deviation is the most efficient measure of dispersion. For example, for our yen position, this value is SD = $29.7 million.

Furthermore, we can translate this standard deviation into a VAR measure, using a multiplier $\alpha(c)$ that depends on the distribution and the selected confidence level c

$$VAR = \alpha \sigma W \qquad (10.8)$$

where σ is the volatility of the rate of return, which is unitless, and W is the amount invested, measured in the reference currency. Here SD $= \sigma W$.

With a normal distribution and $c = 95\%$, we have $\alpha = 1.645$. This gives a VAR estimate of $1.645 \times 29.7 = $49 million, which is not far from the empirical quantile of $47 million.

Note that Equation (10.8) measures VAR relative to the mean, because the standard deviation is a measure of dispersion around the mean. If it is important to measure the loss relative to the initial value, VAR is then

$$VAR = (\alpha \sigma - \mu) W \qquad (10.9)$$

where μ is the expected rate of return over the horizon. In this case, the mean is very small, at $-\$0.4$ million, which leads to very close VAR measures.

Under these conditions, VAR inherits all properties of the standard deviation. In particular, the SD of a portfolio must be smaller than the sum of the SDs of subportfolios, so it is subadditive.

The disadvantage of the standard deviation is that it is symmetrical and cannot distinguish between large losses or gains. Also, computing VAR from SD requires a distributional assumption, which may not be valid.

The Semistandard Deviation   This is a simple extension of the usual standard deviation that considers only data points that represent a loss. Define N_L as the number of such points. The measure is

$$\mathrm{SD}_L(X) = \sqrt{\frac{1}{(N_L)} \sum_{i=1}^{N} [\mathrm{Min}(x_i, 0)]^2} \qquad (10.10)$$

The advantage of this measure is that it accounts for asymmetries in the distribution, e.g., negative skewness which is especially dangerous. The semistandard deviation is sometimes used to report downside risk, but is much less intuitive and less popular than VAR.

The Drawdown   Drawdown is the decline from peak over a fixed time interval. Define x^{MAX} as the local maximum over this period $[0, T]$, which occurs at time $t_{MAX} \in [0, T]$. Relative to this value, the drawdown at time t is

$$\mathrm{DD}(X) = \frac{(x^{MAX} - x_t)}{x^{MAX}} \qquad (10.11)$$

The maximum drawdown is the largest such value over the period, or decline from peak to trough (local maximum to local minimum).

This measure is useful if returns are not independent from period to period. When a market trends, for example, the cumulative loss over a longer period is greater than the loss extrapolated from a shorter period. Alternatively, drawdowns are useful measures of risk if the portfolio is actively managed. A portfolio insurance program, for example, should have lower drawdowns relative to a fixed position in the risky asset because it cuts the position as losses accumulate.

The disadvantage of this measure is that it is backward-looking. It cannot be constructed from the current position, as in the case of VAR. In addition, the maximum drawdown corresponds to different time intervals, i.e., $t_{MAX} - t_{MIN}$. As a result, maximum drawdown measures are not directly comparable across portfolios, in contrast with VAR or the standard deviation, which are defined over a fixed horizon or in annual terms.

EXAMPLE 10.5: FRM EXAM 2003—QUESTION 5

Given the following 30 ordered percentage returns of an asset, calculate the VAR and expected shortfall at a 90% confidence level: $-16, -14, -10, -7, -7, -5, -4, -4, -4, -3, -1, -1, 0, 0, 0, 1, 2, 2, 4, 6, 7, 8, 9, 11, 12, 12, 14, 18, 21, 23$.

  a. VAR (90%) = 10, expected shortfall = 14
  b. VAR (90%) = 10, expected shortfall = 15
  c. VAR (90%) = 14, expected shortfall = 15
  d. VAR (90%) = 18, expected shortfall = 22

10.3 CASH FLOW AT RISK

VAR methods have been developed to measure the mark-to-market risk of commercial bank portfolios. By now, these methods have spread to other financial institutions (e.g., investment banks, savings and loans), and the investment management industry.

In each case, the objective function is the market value of the portfolio, assuming fixed positions. VAR methods, however, are now also spreading to other sectors (e.g., nonfinancial corporations), where the emphasis is on periodic earnings. Cash flow at risk (CFAR) measures the worst shortfall in cash flows due to unfavorable movements in market risk factors. This involves quantities, Q, unit revenues, P, and unit costs, C. Simplifying, we can write

$$CF = Q \times (P - C) \tag{10.12}$$

Suppose we focus on the exchange rate, S, as the market risk factor. Each of these variables can be affected by S. Revenues and costs can be denominated in the foreign currency, partially or wholly. Quantities can also be affected by the exchange rate through foreign competition effects. Because quantities are random, this creates quantity uncertainty. The risk manager needs to model the relationship between quantities and risk factors. Once this is done, simulations can be used to project the cash-flow distribution and identify the worst loss at some confidence level. If this is unacceptably high, the company should hedge its exposure. The role of the risk manager is then to evaluate the effect of different instruments (e.g., forwards, options) and help the company decide on a risk management program.

Over the long-term, the firm may have strategic options. For example, the firm may decide to withdraw from a foreign market if it becomes unprofitable due to a depreciation of the country's currency. Alternatively, the firm may decide to open up an assembly operation in the foreign country to take advantage of lower local costs. Such strategic options can be incorporated in the risk analysis.

A classic example is the value of a farmer's harvest, say corn. At the beginning of the year, costs are fixed. The price of corn and the size of the harvest in the fall, however, are unknown. Suppose price movements are primarily driven by supply shocks, such as the weather. If there is a drought during the summer, quantities will fall and prices will increase, conversely if there is an exceptionally abundant harvest. Because of the negative correlation between Q and P, total revenues will fluctuate less than if quantities were fixed. Such relationships need to be factored into the risk measurement system because they will affect the hedging program.

EXAMPLE 10.6: FRM EXAM 2007—QUESTION 112

The Chief Risk Officer (CRO) of an exporting firm is attempting to estimate the firm's one-year cash flow at risk. Which of the following issues describes an approach that is *irrelevant* to the task of the CRO?

a. Because cash flow at risk is generally estimated over a quarter or over a year, it is necessary to forecast the future values of risk factors.

b. To the extent that the firm's income from exports is best approximated by a real option because the firm does not have to export when the price of the foreign currency is unexpectedly low, the CRO can use option analysis and does not have to worry about forecasting exchange rates.

c. A parametric approach can be used if exposures to foreign exchange risk factors are linear, if there are no other risk factors, and if exchange rate changes are normally distributed.

d. Using a Monte Carlo approach will help the CRO if the firm's foreign currency income is a nonlinear function of exchange rates.

10.4 VAR PARAMETERS

To measure VAR, we first need to define two quantitative parameters, the confidence level and the horizon.

10.4.1 Confidence Level

The higher the confidence level c, the greater the VAR measure. Varying the confidence level provides useful information about the return distribution and potential extreme losses. It is not clear, however, whether one should stop at 99%, 99.9%, 99.99%, and so on. Each of these values will create an increasingly larger loss, but less likely.

Another problem is that as c increases, the number of occurrences below VAR shrinks, leading to poor measures of high quantiles. With 1,000 observations, for example, VAR can be taken as the tenth lowest observation for a 99% confidence level. If the confidence level increases to 99.9%, VAR is taken from the lowest

observation only. Finally, there is no simple way to estimate a 99.99% VAR from this sample because it has too few observations.

The choice of the confidence level depends on the use of VAR. For most applications, VAR is simply a benchmark measure of downside risk. If so, what really matters is *consistency* of the VAR confidence level across trading desks or time.

In contrast, if the VAR number is being used to decide how much capital to set aside to avoid bankruptcy, then a high confidence level is advisable. Obviously, institutions would prefer to go bankrupt very infrequently. This capital adequacy use, however, applies to the overall institution and not to trading desks.

Another important point is that VAR models are only useful insofar as they can be verified. This is the purpose of backtesting, which systematically checks whether the frequency of losses exceeding VAR is in line with $p = 1 - c$. For this purpose, the risk manager should choose a value of c that is not too high. Picking, for instance, $c = 99.99\%$ should lead, on average, to one exceedence out of 10,000 trading days, or 40 years. In other words, it is going to be impossible to verify if the true probability associated with VAR is indeed 99.99%. For all these reasons, the usual recommendation is to pick a confidence level that is not too high, such as 95% to 99%.

10.4.2 Horizon

The longer the horizon T, the greater the VAR measure. This extrapolation is driven by two factors, the behavior of the risk factors, and the portfolio positions.

To extrapolate from a one-day horizon to a longer horizon, we need to assume that returns are independently and identically distributed. If so, the daily volatility can be transformed into to a multiple-day volatility by multiplication by the square root of time. We also need to assume that the distribution of daily returns is unchanged for longer horizons, which restricts the class of distribution to the so-called "stable" family, of which the normal is a member. If so, we have

$$\text{VAR}(T \text{ days}) = \text{VAR}(1 \text{ day}) \times \sqrt{T} \qquad (10.13)$$

This requires (1) the distribution to be invariant to the horizon (i.e., the same α, as for the normal), (2) the distribution to be the same for various horizons (i.e., no time decay in variances), and (3) innovations to be independent across days.

KEY CONCEPT

VAR can be extended from a one-day horizon to T days by multiplication by the square root of time. This adjustment is valid with independent and identically distributed (i.i.d.) returns that have a normal distribution.

The choice of the horizon also depends on the characteristics of the portfolio. If the positions change quickly, or if exposures (e.g., option deltas) change as underlying prices change, increasing the horizon will create "slippage" in the VAR measure.

Again, the choice of the horizon depends on the use of VAR. If the purpose is to provide an accurate benchmark measure of downside risk, the horizon should be relatively short, ideally less than the average period for major portfolio rebalancing.

In contrast, if the VAR number is being used to decide how much capital to set aside to avoid bankruptcy, then a long horizon is advisable. Institutions will want to have enough time for corrective action as problems start to develop.

In practice, the horizon cannot be less than the frequency of reporting of profits and losses. Typically, banks measure P&L on a daily basis, and corporates on a longer interval (ranging from daily to monthly). This interval is the minimum horizon for VAR.

Another criterion relates to the backtesting issue. Shorter time intervals create more data points matching the forecast VAR with the actual, subsequent P&L. As the power of the statistical tests increases with the number of observations, it is advisable to have a horizon as short as possible.

For all these reasons, the usual recommendation is to pick a horizon that is as short as feasible, for instance one day for trading desks. The horizon needs to be appropriate to the asset classes and the purpose of risk management. For institutions such as pension funds, for instance, a one-month horizon may be more appropriate.

For capital adequacy purposes, institutions should select a high confidence level and a long horizon. There is a trade-off, however, between these two parameters. Increasing one *or* the other will increase VAR.

EXAMPLE 10.7: FRM EXAM 2003—QUESTION 19

The VAR on a portfolio using a 1-day horizon is USD 100 million. The VAR using a 10-day horizon is

a. USD 316 million if returns are not independently and identically distributed

b. USD 316 million if returns are independently and identically distributed

c. USD 100 million since VAR does not depend on any day horizon

d. USD 31.6 million irrespective of any other factors

10.4.3 Application: The Basel Rules

The Basel Committee on Banking Supervision has laid out minimum capital requirements for commercial banks to cover market risk. This **Market Risk Charge** (MRC) is based on VAR models computed with the following parameters:

- A horizon of 10 trading days, or two calendar weeks
- A 99% confidence interval
- An observation period based on at least a year of historical data and updated at least once a quarter

Under the Internal Models Approach (IMA), the MRC is computed as the sum of a general market risk charge (GMRC) plus a specific market risk charge (SRC):

$$\text{MRC}_t^{IMA} = \text{Max}\left(k\frac{1}{60}\sum_{i=1}^{60}\text{VAR}_{t-i},\ \text{VAR}_{t-1}\right) + \text{SRC}_t \qquad (10.14)$$

The GMRC involves the average of the trading VAR over the last 60 days, times a supervisor-determined multiplier k (with a minimum value of 3), as well as yesterday's VAR, The Basel Committee allows the 10-day VAR to be obtained from an extrapolation of one-day VAR figures. Thus, VAR is really

$$\text{VAR}_t(10, 99\%) = \sqrt{10} \times \text{VAR}_t(1, 99\%)$$

Presumably, the 10-day period corresponds to the time required for corrective action by bank regulators, should an institution start to run into trouble. Presumably as well, the 99% confidence level corresponds to a low probability of bank failure due to market risk. Even so, one occurrence every 100 periods implies a high frequency of failure. There are $52/2 = 26$ two-week periods in one year. Thus, one failure should be expected to happen every $100/26 = 3.8$ years, which is still much too frequent. This explains why the Basel Committee has applied a multiplier factor, $k \geq 3$ to guarantee further safety. In addition, this factor is supposed to protect against fat tails, unstable parameters, changing positions, and more generally, model risk.

The specific risk charge is designed to provide a buffer against losses due to idiosyncractic factors related to the individual issuer of the security. It includes the risk that an individual debt or equity moves by more or less than the general market factors, reflecting basis risk or event risk such as a downgrade or default. Chapter 30 gives more details on the market risk framework, which is being updated following the credit crisis that started in 2007.

EXAMPLE 10.8: MARKET RISK CHARGE

The 95%, one-day RiskMetrics VAR for a bank trading portfolio is $1,000,000. What is the approximate general market risk charge?

a. $3,000,000
b. $9,500,000
c. $4,200,000
d. $13,400,000

10.5 COMPONENTS OF RISK MEASUREMENT SYSTEMS

As described in Figure 10.4, a risk measurement system combines the following steps:

1. From market data, construct the distribution of risk factors (e.g., normal, empirical, or other).
2. Collect the portfolio positions and map them onto the risk factors.
3. Construct the portfolio VAR using one of the three methods (delta-normal, historical, Monte Carlo), which will be explained in Chapter 15.

10.5.1 Portfolio Positions

We start with portfolio positions. The assumption is that all positions are constant over the horizon. This, of course, cannot be true in an environment where traders turn over their portfolio actively. Rather, it is a simplification.

The true risk can be greater or lower than the VAR measure. It can be greater if VAR is based on close-to-close positions that reflect lower trader limits. If traders take more risks during the day, the true risk will be greater than indicated by VAR. Conversely, the true risk can be lower if management enforces loss limits, in other words, cuts down the risk that traders can take if losses develop.

10.5.2 Risk Factors

The risk factors represent a subset of all market variables that adequately span the risks of the current, or allowed, portfolio. There are literally tens of thousands of securities available, but a much more restricted set of useful risk factors.
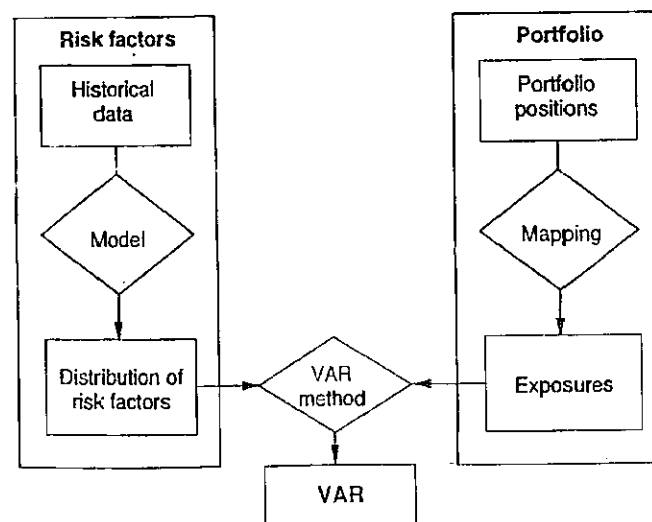
FIGURE 10.4 Elements of a VAR System

The key is to choose market factors that are adequate for the portfolio. For a simple fixed-income portfolio, one bond market risk factor may be enough. In contrast, for a highly leveraged portfolio, multiple risk factors are needed. For an option portfolio, volatilities should be added as risk factors. In general, the more complex the strategies, the greater the number of risk factors that should be used.

10.5.3 VAR Methods

Similarly, the choice of the VAR method depends on the nature of the portfolio. A simple method may be sufficient for simple portfolios. For a fixed-income portfolio, a linear method may be adequate. In contrast, if the portfolio contains options, we need to include nonlinear effects. For simple, plain-vanilla options, we may be able to approximate their price behavior with a first and second derivative (delta and gamma). For more complex options, such as digital or barrier options, this may not be sufficient.

This is why risk management is as much an art as it is a science. Risk managers need to make reasonable approximations to come up with a cost-efficient measure of risk. They also need to be aware of the fact that traders could be induced to find "holes" in the risk management system.

Once this risk measurement system is in place, it can also be used to perform stress-tests. The risk manager can easily submit the current portfolio to various scenarios, which are simply predefined movements in the risk factors. Therefore such systems are essential to evaluate the risks of the current portfolio.

EXAMPLE 10.9: POSITION-BASED RISK MEASURES

The standard VAR calculation for extension to multiple periods also assumes that positions are fixed. If risk management enforces loss limits, the true VAR will be

a. The same
b. Greater than calculated
c. Less than calculated
d. Unable to be determined

10.6 STRESS-TESTING

The Counterparty Risk Management Policy Group (2008) said that "Risk monitoring and risk management cannot be left to quantitative risk metrics, which by

nature are backward looking."[1] This is why VAR should be complemented by stress-testing, which aims at identifying situations that could create extraordinary losses for the institution. As shown in the yen example in Figure 10.2, VAR does not purport to account for extreme losses.

Stress-testing is a key risk management process, which includes (1) scenario analysis, (2) stressing models, volatilities and correlations, and (3) developing policy responses. Scenario analysis submits the portfolio to large movements in financial market variables. These scenarios can be created using a number of methods.

- *Moving key variables one at a time*, which is a simple and intuitive method. Unfortunately, it is difficult to assess realistic comovements in financial variables. It is unlikely that all variables will move in the worst possible direction at the same time.
- *Using historical scenarios*, for instance the 1987 stock market crash, the devaluation of the British pound in 1992, the bond market debacle of 1984, and so on.
- *Creating prospective scenarios*, for instance working through the effects, direct and indirect, of a U.S. stock market crash. Ideally, the scenario should be tailored to the portfolio at hand, assessing the worst thing that could happen to current positions.
- *Reverse stress-tests* start from assuming a large loss and then explore the conditions that would lead to this loss. This type of analysis forces institutions to think of other scenarios and to address issues not normally covered in regular stress-tests, such as financial contagion.

Stress-testing is useful to guard against event risk, which is the risk of loss due to an observable political or economic event. The problem (from the viewpoint of stress-testing) is that such events are relatively rare and may be difficult to anticipate. These include

- *Changes in governments* leading to changes in economic policies
- *Changes in economic policies*, such as default, capital controls, inconvertibility, changes in tax laws, expropriations, and so on
- *Coups, civil wars, invasions*, or other signs of political instability
- *Currency devaluations*, which are usually accompanied by other drastic changes in market variables

These risks often arise in emerging markets,[2] perhaps due to their lack of relative political stability. To guard against event risk, risk managers should

---

[1] Counterparty Risk Management Policy Group (2008), *Containing Systemic Risk: The Road to Reform*, New York: CRMPG.

[2] The term "emerging stock market" was coined by the International Finance Corporation (IFC), in 1981. IFC defines an emerging stock market as one located in a developing country. Using the World Bank's definition, this includes all countries with a GNP per capita less than $8,625 in 1993.

construct prospective events and analyze their impact on portfolio values. Even so, this is not an easy matter. Recent years have demonstrated that markets seem to be systematically taken by surprise. Few people seem to have anticipated the Russian default, for instance. The Argentinian default was also unique in many respects.

Example: Turmoil in Argentina

Argentina is a good example of political risk in emerging markets. Up to 2001, the Argentine peso was fixed to the U.S. dollar at a one-to-one exchange rate. The government had promised it would defend the currency at all costs. Argentina, however, suffered from the worst economic crisis in decades, compounded by the cost of excessive borrowing.

In December 2001, Argentina announced it would stop paying interest on its $135 billion foreign debt. This was the largest sovereign default recorded so far. Economy Minister Cavallo also announced sweeping restrictions on withdrawals from bank deposits to avoid capital flight. On December 20, President Fernando de la Rua resigned after 25 people died in street protest and rioting. President Duhalde took office on January 2 and devalued the currency on January 6. The exchange rate promptly moved from 1 peso/dollar to more than 3 pesos.

Such moves could have been factored into risk management systems by scenario analysis. What was totally unexpected, however, was the government's announcement that it would treat differentially bank loans and deposits. Dollar-denominated bank deposits were converted into devalued pesos, but dollar-denominated bank loans were converted into pesos at a one-to-one rate. This mismatch rendered much of the banking system technically insolvent, because loans (bank assets) overnight became less valuable than deposits (bank liabilities). Whereas risk managers had contemplated the market risk effect of a devaluation, few had considered this possibility of such political actions.

By 2005, the Argentinian government proposed to pay back about 30% of the face value of its debt. This recovery rate was very low by historical standards.

The goal of stress-testing is to identify areas of potential vulnerability. This is not to say that the institution should be totally protected against every possible contingency, as this would make it impossible to take any risk. Rather, the objective of stress-testing and management response should be to ensure that the institution can withstand likely scenarios without going bankrupt. Stress-testing can be easily implemented once the VAR structure is in place. In Figure 10.4, all that is needed is to enter the scenario values into the risk factor inputs.

EXAMPLE 10.10: FRM EXAM 2000—QUESTION 105

Value-at-risk analysis should be complemented by stress-testing because stress-testing

a. Provides a maximum loss, expressed in dollars
b. Summarizes the expected loss over a target horizon within a minimum confidence interval
c. Assesses the behavior of portfolio at a 99% confidence level
d. Identifies losses that go beyond the normal losses measured by VAR

EXAMPLE 10.11: FRM EXAM 2006—QUESTION 87

Which of the following is true about stress-testing?

a. It is used to evaluate the potential impact on portfolio values of unlikely, although plausible, events or movements in a set of financial variables.
b. It is a risk management tool that directly compares predicted results to observed actual results. Predicted values are also compared with historical data.
c. Both a. and b. above are true.
d. None of the above are true.

10.7 IMPORTANT FORMULAS

VAR: $c = \int_{-VAR}^{\infty} f(x)dx$

CVAR: $E[X \mid X < q] = \int_{-\infty}^{q} xf(x)dx / \int_{-\infty}^{q} f(x)dx$

Drawdown: $DD(X) = \frac{(x^{MAX} - x_t)}{x^{MAX}}$

Square root of time adjustment: $VAR(T \text{ days}) = VAR(1 \text{ day}) \times \sqrt{T}$

Market Risk Charge: $MRC_t^{IMA} = \text{Max}\left(k\frac{1}{60}\sum_{i=1}^{60} VAR_{t-i},\ VAR_{t-1}\right) + SRC_t$

10.8 ANSWERS TO CHAPTER EXAMPLES

Example 10.1: FRM Exam 2005—Question 32

a. Stop-loss limits cut down the positions after a loss is incurred, which is useful if market are trending. Exposure limits do not allow for diversification because correlations are not considered. VAR limits can be arbitraged, especially with weak

VAR models. Finally, stop-loss limits are put in place after losses are incurred, so cannot prevent losses. As a result, statement I. and II. are correct.

Example 10.2: FRM Exam 2005—Question 43

b. VAR is the worst loss, such that there is a 95% probability that the losses will be less severe. Alternatively, there is a 5% probability that the loss will be worse. So b. is correct. Answer d. says "lose less" and therefore is incorrect.

Example 10.3: FRM Exam 2003—Question 11

b. Based on Equation (10.4), this is $10\% \times 250 = 25$.

Example 10.4: FRM Exam 2007—Question 101

d. We should expect $(1 - 95\%)250 = 12.5$ exceptions on average. Having eight exceptions is too few, but the difference could be due to luck. Having zero exceptions, however, would be very unusual, with a probability of $1 - (1 - 5\%)^{250}$, which is very low. This means that the risk manager is providing VAR estimates that are much too high. Otherwise, the largest or mean losses are not directly useful without more information on the distribution of profits.

Example 10.5: FRM Exam 2003—Question 5

b. The 10% lower cutoff point is the third lowest observation, which is VAR = 10. The expected shortfall is then the average of the observations in the tails, which is 15.

Example 10.6: FRM Exam 2007—Question 27

b. Conditional VAR is coherent, so statement a. is correct. A low conditional VAR means that when a loss exceeds VAR, it should be small on average, so statement c. is correct. Conditional VAR must be greater than VAR, so statement d. is correct. Economic capital is a VAR measure, not a conditional VAR measure.

Example 10.6: FRM Exam 2007—Question 112

b. A risk analysis requires constructing the distribution of risk factors, so statement a. is correct. Statement c. suggests that the delta-normal VAR method may be appropriate in some contexts, which is correct. Statement d. explains that a Monte Carlo can be used to measure risk, which is correct. Statement b. is false because the distribution of cash flows should be evaluated even in the presence of strategic options.

Example 10.7: FRM Exam 2003—Question 19

b. The square root of time $\sqrt{10} = 3.16$ adjustment applies if the distribution is i.i.d. (and normal).

Example 10.8: Market Risk Charge

d. First, we have to convert the 95% VAR to a 99% measure, assuming a normal distribution in the absence of other information. The GMRC is then $3 \times VAR \times \sqrt{10} = 3 \times \$1,000,000(2.33)/1.65) \times \sqrt{10} = \$13,396,000$.

Example 10.9: Position-Based Risk Measures

c. Less than calculated. Loss limits cut down the positions as losses accumulate. This is similar to a long position in an option, where the delta increases as the price increases, and vice versa. Long positions in options have shortened left tails, and hence involve less risk than an unprotected position.

Example 10.10: FRM Exam 2000—Question 105

d. Stress-testing identifies low-probability losses beyond the usual VAR measures. It does not, however, provide a maximum loss.

Example 10.11: FRM Exam 2006—Question 87

a. Stress-testing is indeed used to evaluate the effect of extreme events. Answer b. is about backtesting, not stress-testing.

APPENDIX: DESIRABLE PROPERTIES FOR RISK MEASURES

The purpose of a risk measure is to summarize the entire distribution of dollar returns X by one number, $\rho(X)$. Artzner et al. (1999) list four desirable properties of risk measures for capital adequacy purposes.[3]

1. Monotonicity: if $X_1 \leq X_2$, $\rho(X_1) \geq \rho(X_2)$.
   In other words, if a portfolio has systematically lower values than another (in each state of the world), it must have greater risk.
2. Translation Invariance: $\rho(X + k) = \rho(X) - k$.
   In other words, adding cash k to a portfolio should reduce its risk by k. This reduces the lowest portfolio value. As with X, k is measured in dollars.

---

[3] See Artzner, P., Delbaen F., Eber J.-M., and Heath D. (1999), Coherent Measures of Risk. *Mathematical Finance*, 9 (July), 203–228.

3. Homogeneity: $\rho(bX) = b\rho(X)$.

In other words, increasing the size of a portfolio by a factor b should scale its risk measure by the same factor b. This property applies to the standard deviation.[4]

4. Subadditivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.

In other words, the risk of a portfolio must be less than the sum of separate risks. Merging portfolios cannot increase risk.

The usefulness of these criteria is that they force us to think about ideal properties and, more importantly, potential problems with simplified risk measures. Indeed, Artzner et al. show that the quantile-based VAR measure fails to satisfy the last property. They give some pathological examples of positions that combine to create portfolios with larger VAR. They also show that the conditional VAR, $E[-X \mid X \leq -\text{VAR}]$, satisfies all these desirable coherence properties.

Assuming a normal distribution, however, the standard deviation-based VAR satisfies the subadditivity property. This is because the volatility of a portfolio is less than the sum of volatilities: $\sigma(X_1 + X_2) \leq \sigma(X_1) + \sigma(X_2)$. We only have a strict equality when the correlation is perfect (positive for long positions). More generally, this property holds for elliptical distributions, for which contours of equal density are ellipsoids.

Example: Why VAR Is Not Necessarily Subadditive

Consider a trader with an investment in a corporate bond with face value of $100,000 and default probability of 0.5%. Over the next period, we can either have no default, with a return of zero, or default with a loss of $100,000. The payoffs are thus −$100,000 with probability of 0.5% and +$0 with probability of 99.5%. Since the probability of getting $0 is greater than 99%, the VAR at the 99% confidence level is $0, without taking the mean into account. This is consistent with the definition that VAR is the smallest loss, such that the right-tail probability is at least 99%.

Now, consider a portfolio invested in three bonds (A,B,C) with the same characteristics and independent payoffs. The VAR numbers add up to $\sum_i \text{VAR}_i = \$0$. To compute the portfolio VAR, we tabulate the payoffs and probabilities:

State	Bonds	Probability	Payoff
No default		$0.995 \times 0.995 \times 0.995 = 0.9850749$	$0
1 default	A,B,C	$3 \times 0.005 \times 0.995 \times 0.995 = 0.0148504$	−$100,000
2 defaults	AB,AC,BC	$3 \times 0.005 \times 0.005 \times 0.995 = 0.0000746$	−$200,000
3 defaults	ABC	$0.005 \times 0.005 \times 0.005 = 0.0000001$	−$300,000

[4] This assumption, however, may be questionable in the case of huge portfolios that could not be liquidated without substantial market impact. Thus, it ignores liquidity risk.

Here, the probability of zero or one default is $0.9851 + 0.0148 = 99.99\%$. The portfolio VAR is therefore \$100,000, which is the lowest number, such that the probability exceeds 99%. Thus, the portfolio VAR is greater than the sum of individual VARs. In this example, VAR is not subadditive. This is an undesirable property because it creates disincentives to aggregate the portfolio, since it appears to have higher risk.

Admittedly, this example is a bit contrived. Nevertheless, it illustrates the danger of focusing on VAR as a sole measure of risk. The portfolio may be structured to display a low VAR. When a loss occurs, however, this may be a huge loss. This is an issue with asymmetrical positions, such as short positions in options or undiversified portfolios exposed to credit risk.

CHAPTER 11
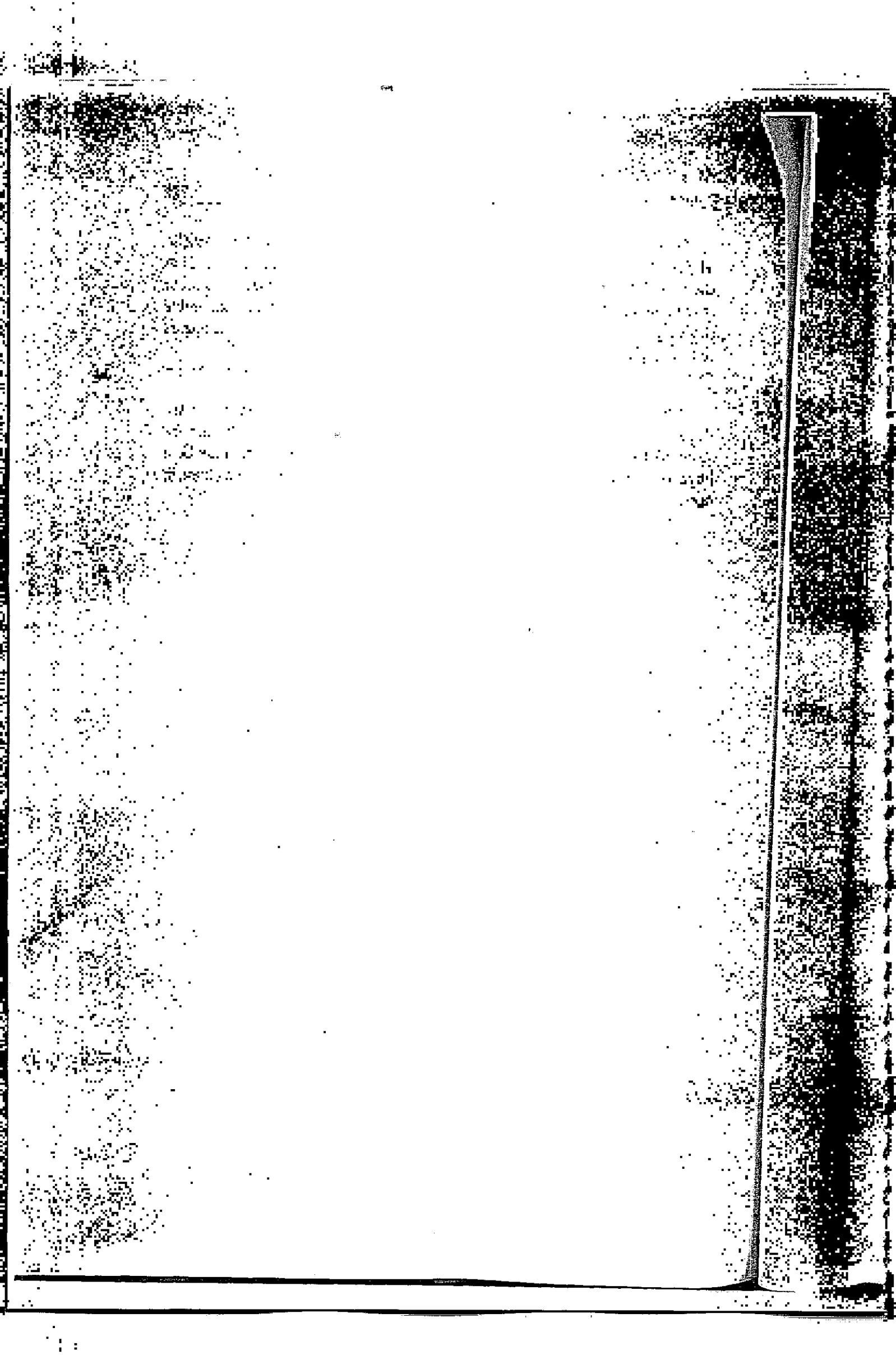
Sources of Market Risk

We now turn to a systematic analysis of the major financial market risk factors. Section 11.1 presents a general overview of financial market risks. Downside risk can be decomposed into two types of drivers: exposures and risk factors. This decomposition is useful because it separates risk into a component over which the risk manager has control (exposure) and another component that is exogenous (the risk factors). We illustrate this decomposition in the context of a simple asset, a fixed-coupon bond.

The next four sections then turn to the major categories of market risk. Currency, fixed-income, equity, and commodities risk are analyzed in Sections 11.2, 11.3, 11.4, and 11.5, respectively. Currency risk refers to the volatility of floating exchange rates and devaluation risk, for fixed currencies. Fixed-income risk relates to term-structure risk, global interest rate risk, real yield risk, credit spread risk, and prepayment risk. Equity risk can be described in terms of country risk, industry risk, and stock-specific risk. Commodity risk includes volatility risk, convenience yield risk, delivery and liquidity risk. This chapter primarily focuses on volatility and correlation measures as drivers of risk. Also important is the shape of the distribution, however.

Finally, Section 11.6 discusses simplifications in risk models. We explain how the multitude of risk factors can be summarized into a few essential drivers. Such factor models include the diagonal model, which decomposes returns into a market-wide factor and residual risk.

11.1 SOURCES OF LOSS: A DECOMPOSITION

To examine sources of losses, consider a plain fixed-coupon bond. The potential for loss can be decomposed into the effect of *dollar duration* $D^* P$ and the changes in the yield dy. The bond's value change is given by

$$dP = -(D^* P) \times dy \tag{11.1}$$

This illustrates the general principle that losses can occur because of a combination of two factors:

1. The exposure to the factor, or dollar duration (a choice variable)
2. The movement in the factor itself (which is external to the portfolio)

273

This linear characterization also applies to *systematic risk* and option *delta*. We can, for instance, decompose the return on stock i, R_i into a component due to the market R_M and some residual risk, which we ignore for now because its effect washes out in a large portfolio:

$$R_i = \alpha_i + \beta_i \times R_M + \epsilon_i \approx \beta_i \times R_M \qquad (11.2)$$

We ignore the constant α_i because it does not contribute to risk, as well as the residual ϵ_i, which is diversified. Specific risk can be defined as risk that is due to issuer-specific price movements, after accounting for general market factors.

Note that R_i is expressed here in terms of rate of return and, hence, has no dimension. To get a change in a dollar price, we write

$$dP_i = R_i P_i \approx (\beta P_i) \times R_M \qquad (11.3)$$

Similarly, the change in the value of a derivative f can be expressed in terms of the change in the price of the underlying asset S,

$$df = \Delta \times dS \qquad (11.4)$$

To avoid confusion, we use the conventional notations of Δ for the first partial derivative of the option. Changes are expressed in infinitesimal amounts df and dS.

Equations (11.1), (11.2), and (11.4) all reveal that the change in value is linked to an exposure coefficient and a change in market variable:

Market Loss = Exposure × Adverse Movement in Financial Variable

To have a loss, we need to have some exposure *and* an unfavorable move in the risk factor. This decomposition is also useful to understand the drivers of discontinuities in the portfolio value. These can come from either discontinuous payoffs, or from jumps in the risk factors. Discontinuous payoffs arise with some instruments, such as binary options, which pay a fixed amount if the options ends up in the money and none otherwise. Discontinuities can also arise if there are jumps in the risk factors, such as the 1987 stock market crash, or due to event risk.

11.2 CURRENCY RISK

Currency risk arises from potential movements in the value of foreign currencies. This includes currency-specific volatility, correlations across currencies, and devaluation risk. Currency risk arises in the following environments.

* In a *pure currency float*, the external value of a currency is free to move, to depreciate or appreciate, as pushed by market forces. An example is the dollar/euro exchange rate.

■ In a *fixed currency system*, a currency's external value is fixed (or pegged) to another currency. An example is the Hong Kong dollar, which is fixed against the U.S. dollar. This does not mean there is no risk, however, due to possible readjustments in the parity value, called devaluations or revaluations.

■ In a *change in currency regime*, a currency that was previously fixed becomes flexible, or vice versa. For instance, the Argentinian peso was fixed against the dollar until 2001, and floated thereafter. Changes in regime can also lower currency risk, as in the recent case of the euro.[1]

11.2.1 Currency Volatility

Table 11.1 compares the RiskMetrics volatility forecasts for a group of 21 currencies.[2] Ten of these correspond to "industrialized" countries, the others to "emerging" markets.

These numbers are standard deviations, adapted from value-at-risk (VAR) forecasts at the 95% confidence level divided by 1.645. The table reports daily, monthly, and annualized standard deviations at the end of 2006 and 1996. Annual volatilities are obtained from monthly volatilities multiplied by the square root of 12.

TABLE 11.1   Currency Volatility Against U.S. Dollar (Percent)

Country	Currency Code	End 2006			End 1996
		Daily	Monthly	Annual	Annual
Argentina	ARS	0.269	1.073	3.72	0.42
Australia	AUD	0.374	1.995	6.91	8.50
Canada	CAD	0.338	1.840	6.37	3.60
Switzerland	CHF	0.403	2.223	7.70	10.16
Denmark	DKK	0.361	1.933	6.70	7.78
United Kingdom	GBP	0.383	2.072	7.18	9.14
Hong Kong	HKD	0.035	0.161	0.56	0.26
Indonesia	IDR	0.286	1.443	5.00	1.61
Japan	JPY	0.363	2.040	7.07	6.63
South Korea	KRW	0.325	1.675	5.80	4.49
Mexico	MXN	0.324	1.856	6.43	6.94
Malaysia	MYR	0.311	1.430	4.95	1.60
Norway	NOK	0.520	2.760	9.56	7.60
New Zealand	NZD	0.455	2.642	9.15	7.89
Philippines	PHP	0.197	1.087	3.76	0.57
Sweden	SEK	0.498	2.535	8.78	6.38
Singapore	SGD	0.183	0.991	3.43	1.79
Thailand	THB	0.645	2.647	9.17	1.23
Taiwan	TWD	0.217	1.093	3.79	0.94
South Africa	ZAR	0.666	4.064	14.08	8.37
Euro	EUR	0.360	1.934	6.70	8.26

---

[1] As of 2007, the Eurozone includes a block of 13 countries, Austria, Belgium/Luxembourg, Finland, France, Germany, Ireland, Italy, Netherlands, Portugal, and Spain. Greece joined on January 1, 2001. Slovenia joined on January 1, 2007. Cyprus joined on January 1, 2008. Currency risk is not totally eliminated, however, as there is always a possibility that the currency union could dissolve.

[2] For updates, see www.riskmetrics.com.
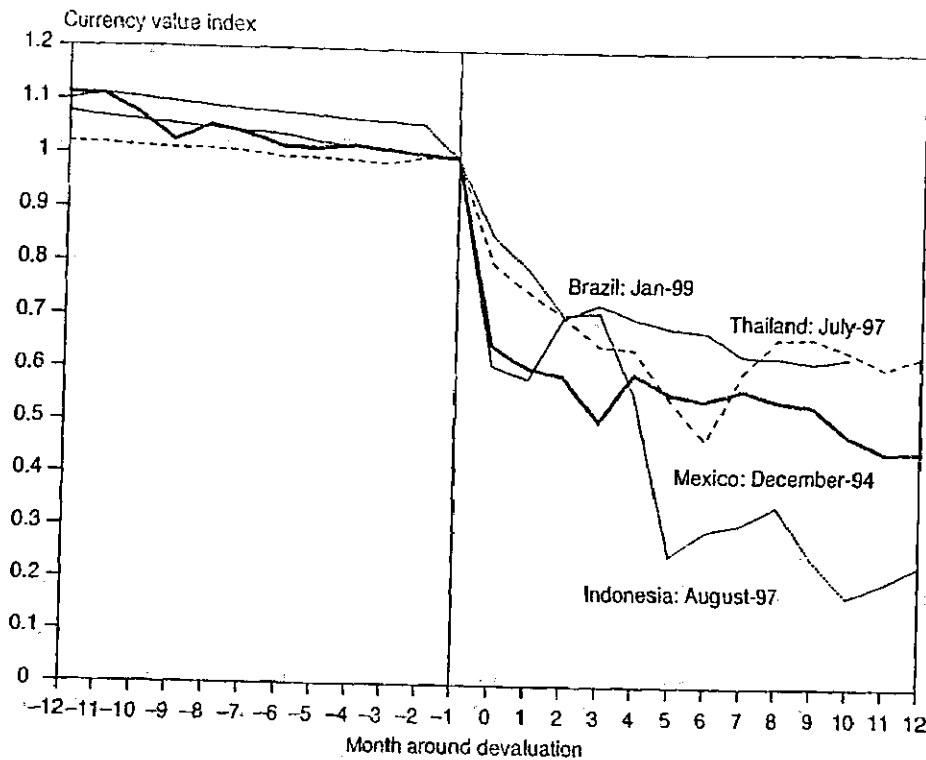
FIGURE 11.1   Effect of Currency Devaluation

Across developed markets, volatility typically ranges from 6% to 10% per annum. The Canadian dollar is notably lower at 4% to 6% volatility. Some currencies, such as the Hong Kong dollar have very low volatility, reflecting their pegging to the dollar. This does not mean that they have low risk, however. They are subject to devaluation risk, which is the risk that the currency peg could fail. This has happened to Thailand and Indonesia, which in 1996 had low volatility but converted to a floating exchange rate regime, with much higher volatility in the latter period.

The typical impact of a currency devaluation is illustrated in Figure 11.1. Each currency has been scaled to a unit value at the end of the month just before the devaluation. In previous months, we observe only small variations in exchange rates. In contrast, the devaluation itself leads to a dramatic drop in value ranging from 20% to an extreme 80% in the case of the rupiah.

Currency risk is also related to other financial risks, in particular interest rate risk. Often, interest rates are raised in an effort to stem the depreciation of a currency, resulting in a positive correlation between the currency and the bond market. These interactions should be taken into account when designing scenarios for stress-tests.

11.2.2   Correlations

Next, we briefly describe the correlations between these currencies against the U.S. dollar. Generally, correlations are low, mostly in the range of −0.10 to 0.20. This indicates substantial benefits from holding a well-diversified currency portfolio.

There are, however, blocks of currencies with very high correlations. European currencies, such as the DKK, SEK, NOK, CHF, have high correlation with each other and the Euro, on the order of 0.90. The GBP also has high correlations with European currencies, around 0.60–0.70. As a result, investing across European currencies does little to diversify risk, from the viewpoint of a U.S. dollar-based investor.

11.2.3 Cross-Rate Volatility

Exchange rates are expressed relative to a base currency, usually the dollar. The cross rate is the exchange rate between two currencies other than the reference currency. For instance, say that S_1 represents the dollar/pound rate and that S_z represents the dollar/euro (EUR) rate. Then the euro/pound rate is given by the ratio

$$S_3(EUR/BP) = \frac{S_1(\$/BP)}{S_2(\$/EUR)} \tag{11.5}$$

Using logs, we can write

$$\ln[S_3] = \ln[S_1] - \ln[S_2] \tag{11.6}$$

The volatility of the cross rate is

$$\sigma_3^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2 \tag{11.7}$$

Alternatively, this shows that we could infer the correlation coefficient ρ_{12} from the triplet of variances. Note that this assumes both the numerator and denominator are in the same currency. Otherwise, the log of the cross rate is the sum of the logs, and the negative sign in Equation (11.7) must be changed to a positive sign.

EXAMPLE 11.1: IMPLIED CORRELATION

What is the implied correlation between JPY/EUR and EUR/USD when given the following volatilities for foreign exchange rates? JPY/USD at 8%; JPY/EUR at 10%; EUR/USD at 6%.

a. 60%

b. 30%

c. −30%

d. −60%

11.3 FIXED-INCOME RISK

Fixed-income risk arises from potential movements in the level and volatility of bond yields. Yield curves move up and down and in various other ways. For the risk manager, this will creates yield curve risk for fixed-income portfolios.

11.3.1 Factors Affecting Yields

Movements in yields reflects economic fundamentals. The primary determinant of movements in interest rates is inflationary expectations. Any perceived increase in the forecast rate of inflation will make bonds with fixed nominal coupons less attractive, thereby increasing their yield.

Figure 11.2 compares the level of short-term U.S. interest rates with the concurrent level of inflation. The graphs show that most of the long-term movements in nominal rates can be explained by inflation. In more recent years, however, inflation has been subdued. Rates have fallen accordingly.

The real interest rate is defined as the nominal rate minus the rate of inflation over the same period. This is generally positive but in recent years has been negative as the Federal Reserve has kept nominal rates to very low levels in order to stimulate economic activity.

Movements in the term structure of interest rates can be summarized by two maturities. In practice, market observers focus on a long-term rate (e.g., from the 10-year Treasury note) and a short-term rate (e.g., from the three-month Treasury bill), as shown in Figure 11.3.

Shaded areas indicate periods of U.S. economic recessions. Recession periods are officially defined by the National Bureau of Economic Research (NBER),
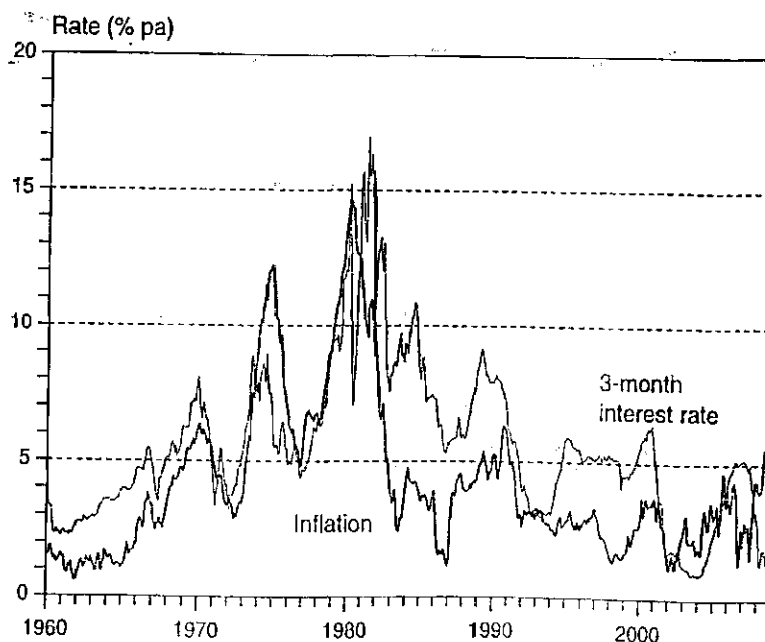
FIGURE 11.2 Inflation and Interest Rates

Yield (% pa)
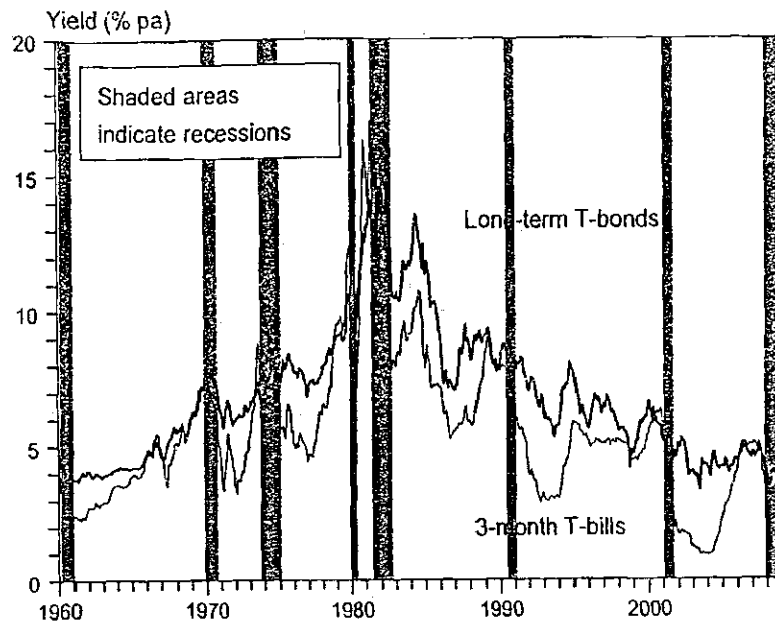
FIGURE 11.3 Movements in the Term Structure

always with a substantial delay. This explains why the 2008 recession is not yet recorded in Figure 11.3.

Generally, the two rates move in tandem, although the short-term rate displays more variability. The term spread is defined as the difference between the long rate and the short rate. Figure 11.4 relates the term spread to economic activity. As the graph shows, periods of recessions usually witness an increase in the term spread. Slow economic activity decreases the demand for capital, which in turn
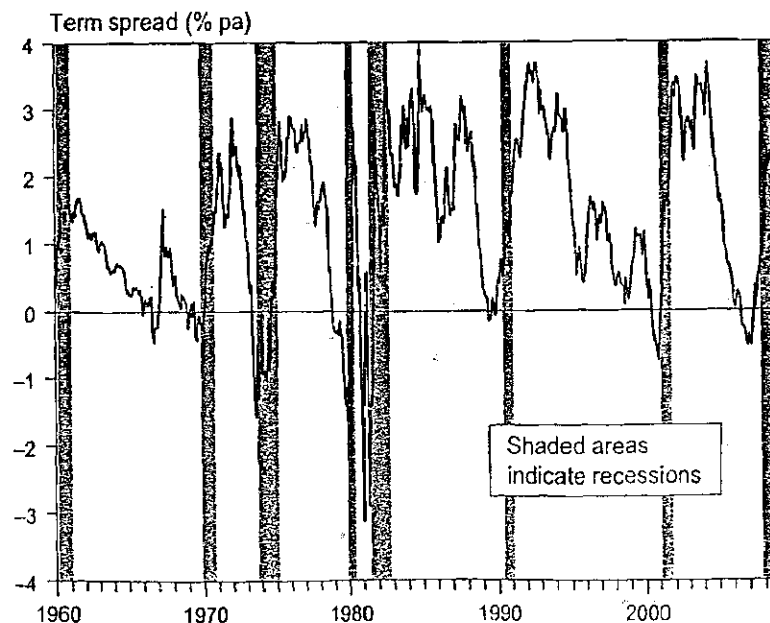
FIGURE 11.4 Term Structure Spread

decreases short-term rates and increases the term spread. Equivalently, the central banks lowers short-term rates to stimulate the economy.

11.3.2 Bond Price and Yield Volatility

Table 11.2 compares the RiskMetrics volatility forecasts for U.S. bond prices as of 2006 and 1996. The table includes Eurodeposits, fixed swap rates, and zero-coupon Treasury rates, for maturities ranging from 30 day to 30 years.

Risk can be measured as either return volatility or yield volatility. Using the duration approximation, the volatility of the rate of return in the bond price is

$$\sigma\left(\frac{dP}{P}\right) = |\ D^*\ | \times \sigma(dy) \tag{11.8}$$

The table shows that short-term investments have very little price risk, as expected, due to their short maturity and duration. The price risk of 10-year bonds is around 6% to 10% annually, which is similar to that of floating currencies. The risk of 30-year bonds is higher, around 20%, which is similar to that of equities.

Instead of measuring price volatility, it is more intuitive to consider yield volatility, $\sigma(dy)$. This is because yield volatility remains more constant over time

TABLE 11.2   U.S. Fixed-Income Return Volatility (Percent)

Type/ Maturity	Code	Yield Level	End 2006			End 1996
			Daily	Monthly	Annual	Annual
Euro-30d	R030	5.325	0.001	0.004	0.01	0.05
Euro-90d	R090	5.365	0.002	0.010	0.04	0.08
Euro-180d	R180	5.375	0.005	0.030	0.11	0.19
Euro-360d	R360	5.338	0.028	0.148	0.51	0.58
Swap-2Y	S02	5.158	0.081	0.420	1.45	1.57
Swap-3Y	S03	5.100	0.127	0.657	2.27	2.59
Swap-4Y	S04	5.062	0.172	0.890	3.08	3.59
Swap-5Y	S05	5.075	0.219	1.120	3.88	4.70
Swap-7Y	S07	5.116	0.283	1.460	5.06	6.69
Swap-10Y	S10	5.177	0.383	1.965	6.81	9.82
Zero-2Y	Z02	4.811	0.088	0.444	1.54	1.64
Zero-3Y	Z03	4.716	0.130	0.663	2.30	2.64
Zero-4Y	Z04	4.698	0.173	0.871	3.02	3.69
Zero-5Y	Z05	4.688	0.216	1.084	3.76	4.67
Zero-7Y	Z07	4.692	0.279	1.395	4.83	6.81
Zero-9Y	Z09	4.695	0.343	1.714	5.94	8.64
Zero-10Y	Z10	4.698	0.375	1.874	6.49	9.31
Zero-15Y	Z15	4.772	0.531	2.647	9.17	13.82
Zero-20Y	Z20	4.810	0.690	3.441	11.92	17.48
Zero-30Y	Z30	4.847	1.014	5.049	17.49	23.53

TABLE 11.3   U.S. Fixed-Income Yield Volatility, 2006 (Percent)

Type/ Maturity	Code	Yield Level	End 2006		
			Daily	Monthly	Annual
Euro-30d	R030	5.325	0.008	0.048	0.17
Euro-90d	R090	5.365	0.006	0.041	0.14
Euro-180d	R180	5.375	0.010	0.062	0.22
Euro-360d	R360	5.338	0.030	0.157	0.54
Swap-2Y	S02	5.158	0.033	0.181	0.63
Swap-3Y	S03	5.100	0.020	0.115	0.40
Swap-4Y	S04	5.062	0.023	0.131	0.46
Swap-5Y	S05	5.075	0.022	0.132	0.46
Swap-7Y	S07	5.116	0.020	0.096	0.33
Swap-10Y	S10	5.177	0.008	0.048	0.17
Zero-2Y	Z02	4.811	0.012	0.064	0.22
Zero-3Y	Z03	4.716	0.020	0.113	0.39
Zero-4Y	Z04	4.698	0.020	0.110	0.38
Zero-5Y	Z05	4.688	0.024	0.126	0.44
Zero-7Y	Z07	4.692	0.023	0.126	0.44
Zero-9Y	Z09	4.695	0.024	0.130	0.45
Zero-10Y	Z10	4.698	0.026	0.139	0.48
Zero-15Y	Z15	4.772	0.027	0.143	0.49
Zero-20Y	Z20	4.810	0.036	0.183	0.63
Zero-30Y	Z30	4.847	0.036	0.184	0.64

than price volatility, which must approach zero as the bond approaches maturity. Yield volatilities are displayed in Table 11.3. Yield volatility averages around 0.50 percent per annum for swaps and zeros. As shown from Table 11.2, however, volatility was fairly low at the end of 2006 compared to 1996, and the historical average. Typical yield volatility is around 1%.

EXAMPLE 11.2: FRM EXAM 2007—QUESTION 50

A portfolio consists of two zero coupon bonds, each with a current value of $10. The first bond has a modified duration of one year and the second has a modified duration of nine years. The yield curve is flat, and all yields are 5%. Assume all moves of the yield curve are parallel shifts. Given that the daily volatility of the yield is 1%, which of the following is the best estimate of the portfolio daily VAR at the 95% confidence level?

a. USD 1.65

b. USD 2.33

c. USD 1.16

d. USD 0.82

11.3.3 Correlations

Table 11.4 displays correlation coefficients for all maturity pairs at a one-day horizon. Correlations are generally very high, suggesting that yields are affected by a common factor.

If the yield curve were to move in strict parallel fashion, all correlations should be equal to 1.000. In practice, the yield curve displays more complex patterns but still satisfies some smoothness conditions. This implies that movements in adjoining maturities are highly correlated. For instance, the correlation between the nine-year zero and 10-year zero is 0.9998, which is very high. Correlations are the lowest for maturities further apart, for instance 0.848 between the two-year and 30-year zero.

These high correlations give risk managers an opportunity to simplify the number of risk factors they have to deal with. Suppose, for instance, that the portfolio consists of global bonds in 17 different currencies. Initially, the risk manager decides to keep 14 risk factors in each market. This leads to a very large number of correlations within, but also across all markets. With 17 currencies, and 14 maturities, the total number of risk factors is $N = 17 \times 14 = 238$. The correlation matrix has $N \times (N - 1) = 238 \times 237 = 56,406$ elements off the diagonal. Surely some of this information is superfluous.

The matrix in Table 11.4 can be simplified using principal components. Principal components is a statistical technique that extracts linear combinations of the original variables that explain the highest proportion of diagonal components of the matrix. For this matrix, the first principal component explains 94% of the total variance and has similar weights on all maturities. Hence, it could be called a level risk factor. The second principal component explains 4% of the total variance. As it is associated with opposite positions on short and long maturities, it could be called a slope risk factor (or *twist*). Sometimes a third factor is found that represents curvature risk factor, or a bend risk factor (also called a *butterfly*).

Previous research has indeed found that, in the United States and other fixed-income markets, movements in yields could be usefully summarized by two to three factors that typically explain over 95% of the total variance.

TABLE 11.4   U.S. Fixed-Income Return Correlations, 2006 (Daily)

	Z02	Z03	Z04	Z05	Z07	Z09	Z10	Z15	Z20	Z30
Z02	1.000									
Z03	0.991	1.000								
Z04	0.980	0.994	1.000							
Z05	0.968	0.985	0.998	1.000						
Z07	0.949	0.972	0.991	0.996	1.000					
Z09	0.934	0.960	0.982	0.990	0.998	1.000				
Z10	0.927	0.954	0.978	0.987	0.997	0.9998	1.000			
Z15	0.896	0.931	0.960	0.971	0.986	0.992	0.994	1.000		
Z20	0.874	0.913	0.944	0.958	0.976	0.983	0.985	0.998	1.000	
Z30	0.848	0.891	0.925	0.940	0.960	0.969	0.972	0.992	0.998	1.000

EXAMPLE 11.3: FRM EXAM 2005—QUESTION 109

In using principal components analysis (PCA) to analyze movements in yields for fixed-income investments, which of the following is *correct*?

a. Changes in long-term yields tend to be larger than changes in short-term yields.

b. The "twist" factor explains more of the variance than the "butterfly" factor.

c. Movements in yields can be usefully summarized by two to three factors that typically explain approximately 50% of the variance in developed markets.

d. PCA extracts nonlinear combinations of the original variables that explain the highest proportion of diagonal components of the matrix.

11.3.4 Global Interest Rate Risk

Different fixed-income markets are exposed to their own sources of risk. The Japanese government bond market, for example is exposed to yen interest rates. Yet in all markets, we observe similar patterns. To illustrate, Table 11.5 shows price and yield volatilities for 17 different fixed-income markets, focusing only on 10-year zeros.

TABLE 11.5   Global Fixed-Income Volatility, 2006 (Percent)

Country	Code	Yield Level	Return Vol. Daily	Return Vol. Monthly	Return Vol. Annual	Yield Vol. $\sigma(dy)$ Daily	Yield Vol. $\sigma(dy)$ Monthly	Yield Vol. $\sigma(dy)$ Annual
Austrl.	AUD	5.847	0.355	1.859	6.44	0.038	0.198	0.691
Belgium	BEF	4.018	0.230	1.257	4.35	0.025	0.137	0.471
Canada	CAD	4.096	0.288	1.485	5.14	0.030	0.153	0.532
Germany	DEM	3.963	0.225	1.246	4.32	0.024	0.135	0.471
Denmark	DKK	3.952	0.244	1.318	4.56	0.026	0.141	0.492
Spain	ESP	4.029	0.239	1.288	4.46	0.026	0.140	0.482
France	FRF	4.006	0.229	1.250	4.33	0.025	0.136	0.471
U.K.	GBP	4.695	0.287	1.458	5.05	0.030	0.153	0.531
Ireland	IEP	3.980	0.244	1.300	4.50	0.026	0.141	0.492
Italy	ITL	4.251	0.253	1.358	4.70	0.027	0.144	0.501
Japan	JPY	1.715	0.287	1.410	4.89	0.030	0.143	0.498
Nether.	NLG	4.004	0.226	1.240	4.29	0.024	0.134	0.471
New Zl.	NZD	5.894	0.267	1.474	5.11	0.028	0.155	0.549
Sweden	SEK	3.793	0.253	1.345	4.66	0.027	0.143	0.503
U.S.	USD	4.698	0.375	1.874	6.49	0.039	0.195	0.684
S.Afr.	ZAR	7.727	0.459	2.881	9.98	0.046	0.285	0.992
Euro	EUR	3.962	0.222	1.230	4.26	0.024	0.133	0.461

The level of yields falls within a remarkably narrow range, 4% to 6%. This reflects the fact that yields are primarily driven by inflationary expectations, which have become similar across all these markets. Indeed central banks across all these countries have proved their common determination to keep inflation in check. Two notable exceptions are South Africa, where yields are higher and Japan, where yields are lower. These two countries are experiencing much higher and lower inflation, respectively, than the rest of the group.

The table also shows that most countries have an annual volatility of yield changes around 0.60%. In fact, we would expect this volatility to decrease as yields drop toward zero and to be higher when yields are higher. This can be modeled by relating the volatility of yield changes to a function of the yield level, as explained in Chapter 4. One such function is the Cox, Ingersoll, and Ross (1985) model, which posits that movements in yields should be proportional to the square root of the yield level. Thus neither the normal nor the lognormal model is totally appropriate.

Finally, correlations are very high across continental European bond markets that are part of the euro. For example, the correlation between French and German bonds is above 0.986. These markets are now moving in synchronization, as monetary policy is dictated by the European Central Bank (ECB). Eurozone bonds only differ in terms of credit risk.

11.3.5 Real Yield Risk

So far, the analysis has only considered nominal interest rate risk, as most bonds represent obligations in nominal terms, i.e., in dollars for the coupon and principal payment. Recently, however, many countries have issued inflation-protected bonds, which make payments that are fixed in real terms but indexed to the rate of inflation.

In this case, the source of risk is real interest rate risk. This real yield can be viewed as the internal rate of return that will make the discounted value of promised real bond payments equal to the current real price. This is a new source of risk, as movements in real interest rates may not correlate perfectly with movements in nominal yields.

These two markets can be used to infer inflationary expectations. The implied rate of inflation can be measured as the nominal yield minus the real yield.

Example: Real and Nominal Yields

Consider for example the 10-year Treasury Inflation Protected (TIP) note paying a 3% coupon in real terms. The actual coupon and principal payments are indexed to the increase in Consumer Price Index (CPI).

The TIP is now trading at a clean real price of 108-23+. Discounting the coupon payments and the principal gives a real yield of $r = 1.98\%$. Note that since the bond is trading at a premium, the real yield must be lower than the coupon.

Projecting the rate of inflation at $\pi = 2\%$, semiannually compounded, we infer the projected nominal yield as $(1 + y/200) = (1 + r/200)(1 + \pi/200)$, which gives 4.00%. This is the same order of magnitude as the current nominal yield on the 10-year Treasury note, which is 3.95%. The two bonds have a very different risk profile, however. If the rate of inflation moves to 5%, payments on the TIP will grow at 5% plus 2%, while the coupon on the regular note will stay fixed.

11.3.6 Credit Spread Risk

Credit spread risk is the risk that yields on duration-matched credit-sensitive bond and Treasury bonds could move differently. The topic of credit risk will be analyzed in more detail in another part of this book. Suffice to say that the credit spread represent a compensation for the loss due to default, plus perhaps a risk premium that reflects investor risk aversion.

A position in a credit spread can be established by investing in credit-sensitive bonds, such as corporates, agencies, mortgage-backed securities (MBSs), and shorting Treasuries with the appropriate duration. This type of position benefits from a stable or shrinking credit spread, but loses from a widening of spreads. Because credit spreads cannot turn negative, their distribution is asymmetric, however. When spreads are tight, large moves imply increases in spreads rather than decreases. Thus positions in credit spreads can be exposed to large losses.

Figure 11.5 displays the time-series of credit spreads since 1960, taken from the Baa-Treasury spread. The graph shows that credit spreads display cyclical patterns, increasing during a recession and decreasing during economic expansions. Greater spreads during recessions reflect the greater number of defaults during difficult times. In 2008, the spread reached 5%, a record over this period.
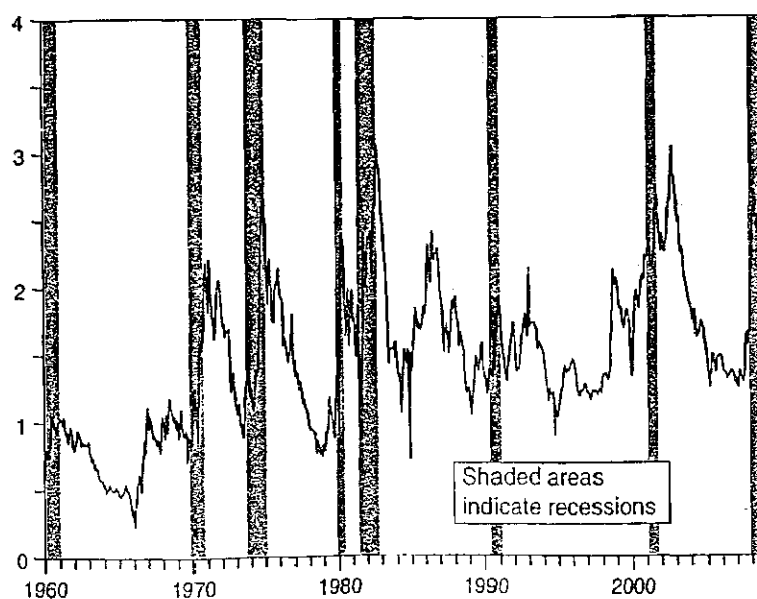
FIGURE 11.5   Credit Spreads

11.3.7 Prepayment Risk

Prepayment risk arises in the context of home mortgages when there is uncertainty about whether the homeowner will refinance his loan early. It is a prominent feature of mortgage-backed securities, where the investor has granted the borrower an option to repay the debt early.

This option, however, is much more complex than an ordinary option, due to the multiplicity of factors involved. We have seen in Chapter 7 that it depends on the age of the loan (seasoning), the current level of interest rates, the previous path of interest rates (burnout), economic activity, and seasonal patterns. Assuming that the prepayment model adequately captures all these features, investors can evaluate the attractiveness of MBSs by calculating their option-adjusted spread (OAS). This represents the spread over the equivalent Treasury minus the cost of the option component.

EXAMPLE 11.4: FRM EXAM 2002—QUESTION 128

During 2002, an Argentinean pension fund with 80% of its assets in dollar-denominated debt lost more than 40% of its value. Which of the following reasons could explain all of the 40% loss:

a. The assets were invested in a diversified portfolio of AAA firms in the U.S.

b. The assets invested in local currency in Argentina lost all their value and the value of the dollar-denominated assets stayed constant.

c. The dollar-denominated assets were invested in U.S. Treasury debt, but the fund had bought credit protection on sovereign debt from Argentina.

d. The fund had invested 80% of its funds in dollar-denominated sovereign debt from Argentina.

11.4 EQUITY RISK

Equity risk arises from potential movements in the value of stock prices. We will show that we can usefully decompose the total risk into a marketwide risk and stock-specific risk.

11.4.1 Stock Market Volatility

Table 11.6 compares the RiskMetrics volatility forecasts for a group of 31 stock markets. The selected indices are those most recognized in each market, for example the S&P 500 in the United States, Nikkei 225 in Japan, and FTSE-100 in Britain. Most of these have an associated futures contract, so positions can be

TABLE 11.8   Equity Volatility (Percent)

| Stock Market Country | Code | End 2006 | | | End 1996 |
		Daily	Monthly	Annual	Annual
Australia	AUD	0.598	3.302	11.4	13.4
Canada	CAD	0.586	3.380	11.7	13.8
Switzerland	CHF	0.606	3.212	11.1	11.1
Germany	DEM	0.701	3.616	12.5	18.6
France	FRF	0.719	3.690	12.8	16.1
U.K.	GBP	0.474	2.691	9.3	11.1
Hong Kong	HKD	0.983	4.604	15.9	17.3
Japan	JPY	0.686	4.046	14.0	19.9
South Korea	KRW	0.731	3.873	13.4	25.5
U.S.	USD	0.444	2.380	8.2	12.9
South Africa	ZAR	0.769	4.268	14.8	11.9

taken or hedged in futures. Nearly all of these indices are weighted by market capitalization, although there is now a trend toward weighting by market float, which is the market value of freely traded shares. For some companies, a large fraction of outstanding shares may be kept by strategic investors (company management, or a government, for example).

We immediately note that risk is much greater than for currencies, typically ranging from 12% to 20%. Markets that are less diversified or are exposed to greater fluctuations in economic fundamentals are more volatile. Concentration refers to the proportion of the index due to the biggest stocks. In Finland, for instance, half of the index represents one firm only, Nokia, which makes the index more volatile than otherwise.

11.5 COMMODITY RISK

Commodity risk arises from potential movements in the value of commodity contracts, which include agricultural products, metals, and energy products.

11.5.1 Commodity Volatility

Table 11.7 displays the volatility of the commodity contracts currently covered by the RiskMetrics system. These can be grouped into *precious metals* (gold, platinum, silver), *base metals* (aluminum, copper, nickel, zinc), and *energy products* (natural gas, heating oil, unleaded gasoline, crude oil–West Texas Intermediate).

Precious metals have an annual volatility ranging from 20% to 30% in 2006, on the same order of magnitude as equity markets. Among base metals, spot volatility are similarly volatile. Energy products, in contrast, are much more volatile with numbers ranging from 20% to 70% in 2006. This is because energy products are less storable than metals and as a result, are much more affected by variations in demand and supply.

TABLE 11.7 Commodity Volatility (Percent)

| Commodity Term | Code | End 2006 | | | End 1996 |
		Daily	Monthly	Annual	Annual
Gold, spot	GLD.C00	0.841	5.03	17.4	5.5
Platinum, spot	PLA.C00	1.508	8.45	29.3	6.5
Silver, spot	SLV.C00	1.722	9.40	32.6	18.1
Aluminium, spot	ALU.C00	1.409	7.99	27.7	16.8
15-month	ALU.C15	0.966	5.65	19.6	13.9
Copper, spot	COP.C00	1.479	8.75	30.3	35.4
15-month	COP.C15	1.300	7.82	27.1	21.5
Nickel, spot	NIC.C00	2.060	11.73	40.7	22.7
15-month	NIC.C15	2.327	12.93	44.8	22.7
Zinc, spot	ZNC.C00	1.751	10.24	35.5	12.4
15-month	ZNC.C15	1.282	8.12	28.1	11.6
Natural gas, 1m	GAS.C01	3.910	20.74	71.9	95.8
15-month	GAS.C06	2.396	11.86	41.1	34.4
Heating oil, 1m	HTO.C01	1.824	9.52	33.0	34.4
12-month	HTO.C12	1.075	5.86	20.3	22.7
Unleaded gas, 1m	UNL.C01	1.910	10.25	35.5	31.0
6-month	UNL.C06	1.220	7.08	24.5	23.5
Crude oil, 1m	WTI.C01	1.467	8.21	28.4	32.8
12-month	WTI.C12	1.045	5.60	19.4	28.9

11.5.2 Futures Risk

The forward or futures price on a commodity can be expressed as

$$F_t e^{-r\tau} = S_t e^{-y\tau} \qquad (11.9)$$

where $e^{-r\tau}$ is the present value factor in the base currency and $e^{-y\tau}$ includes a convenience yield y. Any storage cost should be deducted from y. This represents an implicit flow benefit from holding the commodity, as was explained in Chapter 9. For precious metals, this convenience yield is close to zero.

While this convenience yield is conceptually similar to that of a dividend yield on a stock index, it cannot be measured as regular income. Rather, it should be viewed as a "plug-in" that, given F, S, and $e^{-r\tau}$, will make Equation (11.9) balance. Further, it can be quite volatile.

As Table 11.7 shows, futures prices are less volatile for longer maturities. This decreasing term structure of volatility is more marked for energy products and less so for base metals. In addition, movements in futures prices are much less tightly related to spot prices than for financial contracts.

This is illustrated in Table 11.8, which displays correlations for copper contracts as well as for natural gas contracts. The correlations for natural gas are much lower than for copper. Thus variations in the basis are much more important for energy products than for financial products, or even metals.

TABLE 11.8 Correlations across Maturities, 2006 (Daily)

Copper	COP.C00	COP.C03	COP.C15	COP.C27
COP.C00	1			
COP.C03	.998	1		
COP.C15	.979	.985	1	
COP.C27	.923	.935	.974	1

Nat.Gas	GAS.C01	GAS.C03	GAS.C06	GAS.C12
GAS.C01	1			
GAS.C03	.916	1		
GAS.C06	.897	.968	1	
GAS.C12	.884	.824	.878	1

This is confirmed by Figure 11.6, which compares the spot and futures prices for crude oil. There is much more variation in the basis between the spot and futures prices for crude oil. The market switches from backwardation ($S > F$) to contango ($S < F$). As a result, the futures contract represents a separate risk factor. Energy risk measurement systems require separate risk factors for each maturity.

In addition to traditional market sources of risk, positions in commodity futures are also exposed to delivery and liquidity risks. Asset liquidity risk is due to the relative low volume in some of these markets, relative to other financial products.

Also, taking delivery or having to deliver on a futures contract that is carried to expiration is costly. Transportation, storage and insurance costs can be quite high. Futures delivery also requires complying with the type and location of the commodity that is to be delivered.
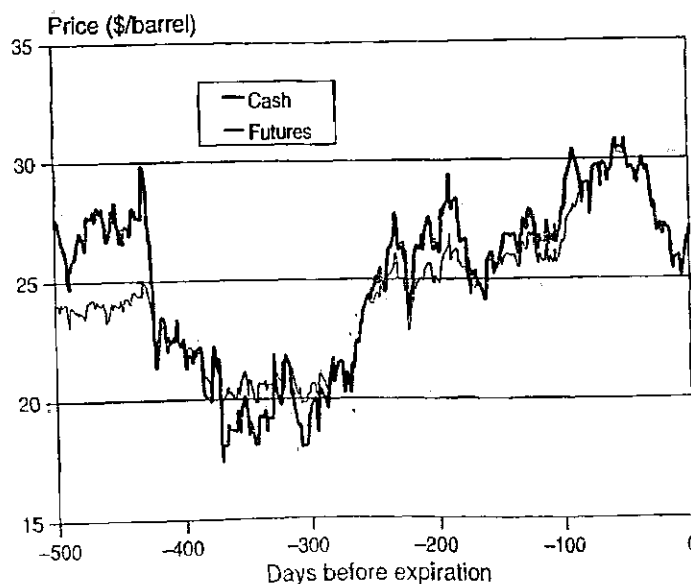
FIGURE 11.6 Futures and Spot for Crude Oil

EXAMPLE 11.5: FRM EXAM 2006—QUESTION 115

Assume the risk-free rate is 5% per annum, the cost of storing oil for a year is 1% per annum, the convenience yield for owning oil for a year is 2% per annum, and the current price of oil is USD 50 per barrel. All rates are continuously compounded. What is the forward price of oil in a year?

a. USD 49.01
b. USD 52.04
c. USD 47.56
d. USD 49.50

EXAMPLE 11.6: FRM EXAM 2006—QUESTION 138

Imagine a stack-and-roll hedge of monthly commodity deliveries that you continue for the next five years. Assume the hedge ratio is adjusted to take into effect the mistiming of cash flows but is not adjusted for the basis risk of the hedge. In which of the following situations is your calendar basis risk likely to be greatest?

a. Stack-and-roll in the front month in oil futures
b. Stack-and-roll in the 12-month contract in natural gas futures
c. Stack-and-roll in the three-year contract in gold futures
d. All four situations will have the same basis risk

11.6 RISK SIMPLIFICATION

The fundamental idea behind modern risk measurement methods is to aggregate the portfolio risk at the highest level. In practice, it would be too complex to model each risk factor individually. Instead, some simplification is required. We have seen, for example, that movements in the terms structure of interest rates could be simplified to a few major risk factors. This approach expands on the diagonal model proposed by Professor William Sharpe. This was initially applied to stocks, but the methodology can be used in any market.

11.6.1 Diagonal Model

The diagonal model starts with a statistical decomposition of the return on stock *i* into a marketwide return and an idiosyncratic risk. The diagonal model adds the

assumption that all specific risks are uncorrelated. Hence, any correlation across two stocks must come from the joint effect of the market.

We decompose the return on stock i, R_i, into a constant, a component due to the market, R_M, and some residual risk:

$$R_i = \alpha_i + \beta_i \times R_M + \epsilon_i \tag{11.10}$$

where β_i is called systematic risk of stock i. It is also the regression slope ratio:

$$\beta_i = \frac{Cov[R_i, R_M]}{V[R_M]} = \rho_{iM}\frac{\sigma(R_i)}{\sigma(R_M)} \tag{11.11}$$

Note that the residual is uncorrelated with R_M by assumption. The contribution of William Sharpe was to show that equilibrium in capital markets imposes restrictions on the α_i. For risk managers, who primarily focus on risk, however, the diagonal model allows considerable simplifications in the risk model, so we ignore the intercept in what follows.

Consider a portfolio that consists of positions w_i on the various assets. We have

$$R_p = \sum_{i=1}^{N} w_i R_i \tag{11.12}$$

Using Equation (11.10), the portfolio return is also

$$R_p = \sum_{i=1}^{N} (w_i \beta_i R_M + w_i \epsilon_i) = \beta_p R_M + \sum_{i=1}^{N} (w_i \epsilon_i) \tag{11.13}$$

The portfolio variance is

$$V[R_p] = \beta_p^2 V[R_M] + \sum_{i=1}^{N} \left(w_i^2 V[\epsilon_i] \right) \tag{11.14}$$

since all the residual terms are uncorrelated. Suppose that, for simplicity, the portfolio is equally weighted and that the residual variances are all the same $V[\epsilon_i] = V$. This implies $w_i = w = 1/N$. As the number of assets, N, increases, the second term will tend to

$$\sum_{i=1}^{N} \left(w_i^2 V[\epsilon_i] \right) \to N \times [(1/N)^2 V] = (V/N)$$

which should vanish as N increases. In this situation, the only remaining risk is the general market risk, consisting of the beta squared times the variance of the market:

$$V[R_p] \to \beta_p^2 V[R_M]$$

So, this justifies ignoring specific risk in large, well-diversified portfolios. For portfolio with a small number of stocks concentrated in one sector, this approach will underestimate risk.

More generally, this approach can be expanded to multiple factors. The appendix shows how this approach can be used to build a covariance matrix from general market factors. Each security is "mapped" on the selected risk factors. Exposures are then added up across the entire portfolio, for which risk is aggregated at the top level. This mapping approach is particularly useful when there is no history of returns for some positions. Instead, these positions can be mapped on the risk factors.

11.6.2 Fixed-Income Portfolio Risk

As an example of portfolio simplification, we turn to the analysis of a corporate bond portfolio with N individual bonds. Each "name" is potentially a source of risk. Instead of modeling all securities, the risk manager should attempt to simplify the risk profile of the portfolio. Potential major risk factors are movements in a set of J Treasury zero-coupon rates, z_j, and in K credit spreads, s_k, sorted by credit rating. The goal is to provide a good approximation to the risk of the portfolio.

In addition, it is not practical to model the risk of all bonds. The bonds may not have a sufficient history. Even if they do, the history may not be relevant if it does not account for the probability of default.

We model the movement in each corporate bond yield y_i by a movement in the Treasury factor z_j at the closest maturity and in the credit rating s_k class to which it belongs. The remaining component is ϵ_i, which is assumed to be independent across i. We have $y_i = z_j + s_k + \epsilon_i$. This decomposition is illustrated in Figure 11.7 for a corporate bond rated BBB with a 20-year maturity.

The movement in the bond price is

$$\Delta P_i = -\text{DVBP}_i \Delta y_i = -\text{DVBP}_i \, \Delta z_j - \text{DVBP}_i \, \Delta s_k - \text{DVBP}_i \, \Delta \epsilon_i \qquad (11.15)$$
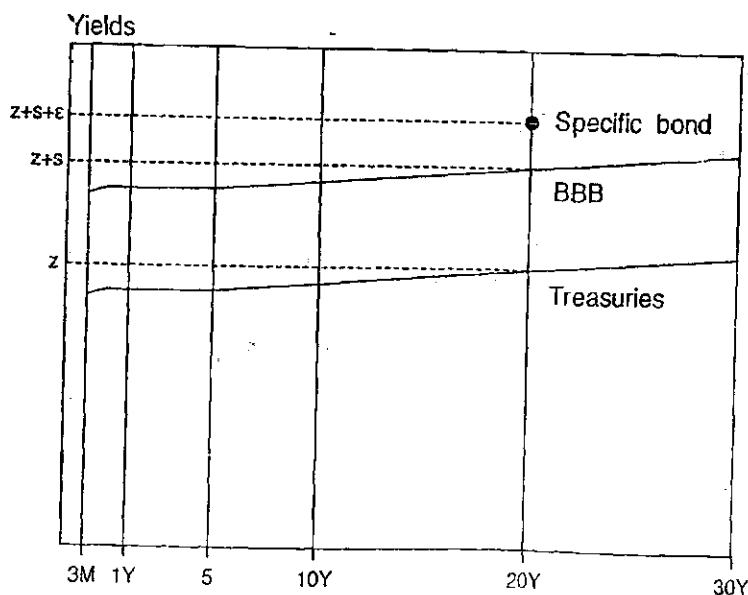
FIGURE 11.7 Yield Decomposition

where DVBP is the total dollar value of a basis point for the associated risk factor. We hold n_i units of this bond, so that its value is

$$P = \sum_{i=1}^{N} n_i P_i \qquad (11.16)$$

Expanding the portfolio into its components, we have

$$\Delta P = -\sum_{i=1}^{N} n_i \Delta P_i = -\sum_{i=1}^{N} n_i \text{DVBP}_i \Delta y_i \qquad (11.17)$$

Using the risk factor decomposition, the portfolio price movement is

$$\Delta P = -\sum_{j=1}^{J} \text{DVBP}_j^z \Delta z_j - \sum_{k=1}^{K} \text{DVBP}_k^s \Delta s_k - \sum_{i=1}^{N} n_i \text{DVBP}_i \Delta \epsilon_i \qquad (11.18)$$

where DVBP_j^z results from the summation of $n_i \text{DVBP}_i$ for all bonds that are exposed to the jth maturity. As in Equation (11.14), The total variance can be decomposed into

$$V[\Delta P] = \text{General Risk} + \sum_{i=1}^{N} n_i^2 \text{DVBP}_i^2 \; V[\Delta \epsilon_i] \qquad (11.19)$$

If the portfolio is well diversified, the general risk term should dominate. So, we could simply ignore the second term. Ignoring specific risk, a portfolio composed of thousands of securities can be characterized by its exposure to just a few government maturities and credit spreads. This is a considerable simplification.

EXAMPLE 11.7: FRM EXAM 2002—QUESTION 44

The historical simulation (HS) approach is based on the empirical distributions and a large number of risk factors. The RiskMetrics approach assumes normal distributions and uses mapping on equity indices. The HS approach is more likely to provide an accurate estimate of VAR than the RiskMetrics approach for a portfolio that consists of

a. A small number of emerging market securities

b. A small number of broad market indexes

c. A large number of emerging market securities

d. A large number of broad market indexes

EXAMPLE 11.8: FRM EXAM 2007—QUESTION 11

A hedge fund manager has to choose a risk model for a large "equity market neutral" portfolio, which has zero beta. Many of the stocks held are recent IPOs. Among the following alternatives, the best is

a. A single index model with no specific risk, estimated over the last year

b. A diagonal index model with idiosyncratic risk, estimated over the last year

c. A model that maps positions on industry and style factors

d. A full covariance matrix model using a very short window

11.7 IMPORTANT FORMULAS

Cross-exchange rate: $S_3(\text{EUR/BP}) = \frac{S_1(\$/\text{BP})}{S_2(\$/\text{EUR})}$, $\sigma_3^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2$

Volatility of the rate of return in the bond price: $\sigma\left(\frac{dP}{P}\right) = |D^*| \times \sigma(dy)$

Futures price on a commodity: $F_t e^{-r\tau} = S_t e^{-y\tau}$

Diagonal Model: $R_i = \alpha_i + \beta_i \times R_M + \epsilon_i$

11.8 ANSWERS TO CHAPTER EXAMPLES

Example 11.1: Implied Correlation

d. The logs of JPY/EUR and EUR/USD add up to that of JPY/USD: ln[JPY/USD] = ln[JPY/EUR] + ln[EUR/USD]. So, $\sigma^2(\text{JPY/USD}) = \sigma^2(\text{JPY/EUR}) + \sigma^2(\text{EUR/USD}) + 2\rho\sigma(\text{JPY/EUR})\sigma(\text{EUR/USD})$, or $8^2 = 10^2 + 6^2 + 2\rho10 \times 6$, or $2\rho10 \times 6 = -72$, or $\rho = -0.60$.

Example 11.2: FRM Exam 2007—Question 50

a. The dollar duration of the portfolio is $1 \times \$10 + 9 \times \$10 = \$100$. Multiplied by 0.01 and 1.65, this gives $\$1.65$.

Example 11.3: FRM Exam 2005—Question 109

b. Answer a. is incorrect, as short-term rates tend to move more than long-term rates. Answer c. understates the explanatory power, which is 95%, not 50%. Answer d. is incorrect, because PCA is a linear combination of the factors.

Example 11.4: FRM Exam 2002—Question 128

d. In 2001, Argentina defaulted on its debt, both in the local currency and in dollars. Answer a. is wrong because a diversified portfolio could not have lost

so much. The funds were invested at 80% in dollar-denominated assets, so b. is wrong. Even a total wipeout of the local-currency portion could not explain a loss of 40% on the portfolio. If the fund had bought credit protection, it would have not lost as much, so c. is wrong. The fund must have had credit exposure to Argentina, so answer d. is correct.

Example 11.5: FRM Exam 2006—Question 115

b. Using $F_t e^{-r\tau} = S_t e^{-y\tau}$, we have $F = S \exp(-(y-c)\tau + r\tau) = 50 \exp(-(0.02 - 0.01) + 0.05) = 52.04$.

Example 11.6: FRM Exam 2006—Question 138

a. For gold, forward rates closely follow spot rates, so there is little basis risk. For oil and natural gas, there is most movement at the short end of the term structure of futures prices. So using short maturities, or the front month, has the greatest basis risk.

Example 11.7: FRM Exam 2002—Question 44

a. The question deals with the distribution of the assets and the effect of diversification. Emerging market securities are more volatile and less likely to be normally distributed than broad market indices. In addition, a small portfolio is less likely to be well represented by a mapping approach, and is less likely to be normal. The RiskMetrics approach assumes that the conditional distribution is normal and simplifies risk by mapping. This will be acceptable with a large number of securities with distributions close to the normal, which is answer d. Answer a. describes the least diversified portfolio, for which the HS method is best.

Example 11.8: FRM Exam 2007—Question 11

c. Answer a. is incorrect because it only considers the portfolio beta, which is zero by construction. So, it would erroneously conclude that there is no risk. Answer b. is better but would miss the risk of the IPO positions because they have no history. Answer c. will produce unreliable numbers because of the short window. The best solution is to replace the IPO positions by exposures on industry and style factors.

APPENDIX: SIMPLIFICATION OF THE COVARIANCE MATRIX

This appendix shows how the diagonal model can be used to construct a simplified covariance matrix, which is useful for some VAR approaches. Say that we have $N = 100$ assets, which implies a covariance matrix with $N(N+1)/2 = 5,050$ entries.

First, we derive the covariance between any two stocks under the one-factor model

$$\text{Cov}[R_i, R_j] = \text{Cov}[\beta_i R_M + \epsilon_i, \beta_j R_M + \epsilon_j] = \beta_i \beta_j \sigma_M^2 \qquad (11.20)$$

using the assumption that the residual components are uncorrelated with each other and with the market. Also, the variance of a stock is

$$\text{Cov}[R_i, R_i] = \beta_i^2 \sigma_M^2 + \sigma_{\epsilon,i}^2 \qquad (11.21)$$

The covariance matrix is then

$$\Sigma = \begin{bmatrix} \beta_1^2 \sigma_M^2 + \sigma_{\epsilon,1}^2 & \beta_1 \beta_2 \sigma_M^2 & \cdots & \beta_1 \beta_N \sigma_M^2 \\ \vdots & & & \\ \beta_N \beta_1 \sigma_M^2 & \beta_N \beta_2 \sigma_M^2 & \cdots & \beta_N^2 \sigma_M^2 + \sigma_{\epsilon,N}^2 \end{bmatrix}$$

which can also be written as

$$\Sigma = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix} [\beta_1 \ldots \beta_N] \sigma_M^2 + \begin{bmatrix} \sigma_{\epsilon,1}^2 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \sigma_{\epsilon,N}^2 \end{bmatrix}$$

Using matrix notation, we have

$$\Sigma = \beta \beta' \sigma_M^2 + D_\epsilon \qquad (11.22)$$

This consists of N elements in the vector β, of N elements on the diagonal of the matrix D_ϵ, plus the variance of the market itself. The diagonal model reduces the number of parameters from $N \times (N+1)/2$ to $2N+1$, a considerable improvement. For example, with 100 assets the number is reduced from 5,050 to 201.

In summary, this diagonal model substantially simplifies the risk structure of an equity portfolio. Risk managers can proceed in two steps: first, managing the overall market risk of the portfolios, and second, managing the concentration risk of individual securities.

Still, this one-factor model could miss common effects among groups of stocks, such as industry effects. To account for these, Equation (11.10) can be generalized to K factors

$$R_i = \alpha_i + \beta_{i1} y_1 + \cdots + \beta_{iK} y_K + \epsilon_i \qquad (11.23)$$

where $y_1, \ldots, y_K$ are the factors, which are assumed independent of each other for simplification. The covariance matrix generalizes Equation (11.22) to

$$\Sigma = \beta_1 \beta_1' \sigma_1^2 + \cdots + \beta_K \beta_K' \sigma_K^2 + D_\epsilon \qquad (11.24)$$

The number of parameters is now $(N \times K + K + N)$. For example, with 100 assets and five factors, this number is 605, which is still much lower than 5,050 for the unrestricted model.

Hedging Linear Risk

Risk that has been measured can be managed. This chapter turns to the active management of market risks.

The traditional approach to market risk management includes hedging. Hedging consists of taking positions that lower the risk profile of the portfolio. The techniques for hedging have been developed in the futures markets, where farmers, for instance, use financial instruments to hedge the price risk of their products.

This implementation of hedging is quite narrow, however. Its objective is to find the optimal position in a futures contract that minimizes the variance, or more generally the VAR, of the total position. The portfolio consists of two position, a fixed inventory to be hedged and a "hedging" instrument. In this chapter, the value of the hedging instrument is linearly related to the underlying risk factor.

More generally, we can distinguish between

- Static hedging, which consists of putting on, and leaving, a position until the hedging horizon.
- Dynamic hedging, which consists of continuously rebalancing the portfolio to the horizon. This can create a risk profile similar to positions in options.

Dynamic hedging is associated with options, which will be examined in the next chapter. Since options have nonlinear payoffs in the underlying asset, the hedge ratio, which can be viewed as the slope of the tangent to the payoff function, . must be readjusted as the price moves.

Even with static hedging, hedging will create hedge slippage, or basis risk. Basis risk arises when changes in payoffs on the hedging instrument do not perfectly offset changes in the value of the inventory position.

A final note on hedging is in order. Obviously, if the objective of hedging is to lower volatility, hedging will eliminate downside risk but also any upside potential. The objective of hedging is to lower risk, not to make profits, so this is a double-edged sword. Whether hedging is beneficial should be examined in the context of the trade-off between risk and return.

This chapter discusses linear hedging. A particularly important application is hedging with futures. Section 12.1 presents an introduction to futures hedging with a unit hedge ratio. Section 12.2 then turns to a general method for finding the optimal hedge ratio. This method is applied in Section 12.3 for hedging bonds and equities.

12.1 INTRODUCTION TO FUTURES HEDGING

12.1.1 Unitary Hedging

Consider the situation of a U.S. exporter who has been promised a payment of 125 million Japanese yen in seven months. This is cash position, or anticipated inventory. The perfect hedge would be to enter a seven-month forward contract over-the-counter (OTC). Assume for this illustration that this OTC contract is not convenient. Instead, the exporter decides to turn to an exchange-traded futures contract, which can be bought or sold easily.

The Chicago Mercantile Exchange (CME) lists yen contracts with face amount of Y12,500,000 that expire in nine months. The exporter places an order to sell 10 contracts, with the intention of reversing the position in seven months, when the contract will still have two months to maturity.[1] Because the amount sold is the same as the underlying, this is called a unitary hedge.

Table 12.1 describes the initial and final conditions for the contract. At each date, the futures price is determined by interest parity. Suppose that the yen depreciates sharply, or that the dollar goes up from Y125 to Y150. This lead to a loss on the anticipated cash position of $Y125,000,000 \times (0.006667 - 0.00800) = -\$166,667$. This loss, however, is offset by a gain on the futures, which is $(-10) \times Y12,500,000 \times (0.006711 - 0.00806) = \$168,621$. The net is a small gain of $1,954. Overall, the exporter has been hedged.

This example shows that futures hedging can be quite effective, removing the effect of fluctuations in the risk factor. Define Q as the amount of yen transacted and S and F as the spot and futures rates, indexed by 1 at the initial time and by 2

TABLE 12.1 A Futures Hedge

Item	Initial Time	Exit Time	Gain or Loss
Market Data:			
Maturity (months)	9	2	
US interest rate	6%	6%	
Yen interest rate	5%	2%	
Spot (Y/$)	Y125.00	Y150.00	
Futures (Y/$)	Y124.07	Y149.00	
Contract Data:			
Spot ($/Y)	0.008000	0.006667	−$166,667
Futures ($/Y)	0.008060	0.006711	$168,621
Basis ($/Y)	−0.000060	−0.000045	$1,954

[1] In practice, if the liquidity of long-dated contracts is not adequate, the exporter could use nearby contracts and roll them over prior to expiration into the next contracts. When there are multiple exposures, this practice is known as a stack hedge. Another type of hedge is the strip hedge, which involves hedging the exposures with a number of different contracts. While a stack hedge has superior liquidity, it also entails greater basis risk than a strip hedge. Hedgers must decide whether the greater liquidity of a stack hedge is worth the additional basis risk.

at the exit time. The P&L on the unhedged transaction is

$$Q[S_2 - S_1] \qquad (12.1)$$

Instead, the hedged profit is

$$Q[(S_2 - S_1) - (F_2 - F_1)] = Q[(S_2 - F_2) - (S_1 - F_1)] = Q[b_2 - b_1] \qquad (12.2)$$

where $b = S - F$ is the basis. The profit only depends on the movement in the basis. Hence the effect of hedging is to transform price risk into basis risk. A short hedge position is said to be *long the basis*, since it benefits from an increase in the basis.

In this case, the basis risk is minimal for a number of reasons. First, the cash and futures correspond to the same asset. Second, the cash-and-carry relationship holds very well for currencies. Third, the remaining maturity at exit is rather short. This is not always the case, however.

12.1.2 Basis Risk

Basis risk arises when the characteristics of the futures contract differ from those of the underlying position. Futures contracts are standardized to a particular grade, say West Texas Intermediate (WTI) for oil futures traded on the NYMEX. This defines the grade of crude oil deliverable against the contract. A hedger, however, may have a position in a different grade, which may not be perfectly correlated with WTI. Thus basis risk is the uncertainty whether the cash-futures spread will widen or narrow during the hedging period. Hedging can be effective, however, if movements in the basis are dominated by movements in cash markets.

For most commodities, basis risk is inevitable. Organized exchanges strive to create enough trading and liquidity in their listed contracts, which requires standardization. Speculators also help to increase trading volumes and provide market liquidity. Thus there is a trade-off between liquidity and basis risk.

Basis risk is higher with cross-hedging, which involves using a futures on a totally different asset or commodity than the cash position. For instance, a U.S. exporter who is due to receive a payment in Norwegian Kroner (NK) could hedge using a futures contract on the $/euro exchange rate. Relative to the dollar, the euro and the NK should behave similarly, but there is still some basis risk.

Basis risk is lowest when the underlying position and the futures correspond to the same asset. Even so, some basis risk remains because of differing maturities. As we have seen in the yen hedging example, the maturity of the futures contract is 9 instead of 7 months. As a result, the liquidation price of the futures is uncertain.

Figure 12.1 describes the various time components for a hedge using T-bond futures. The first component is the *maturity of the underlying bond*, say 20 years. The second component is the *time to futures expiration*, say nine months. The third component is the *hedge horizon*, say seven months. Basis risk occurs when the hedge horizon does not match the time to futures expiration.
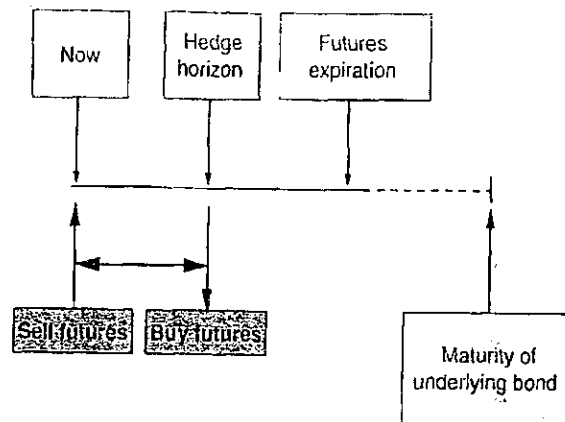
FIGURE 12.1   Hedging Horizon and Contract Maturity

EXAMPLE 12.1: FRM EXAM 2000—QUESTION 79

Under which scenario is basis risk likely to exist?

a. A hedge (which was initially matched to the maturity of the underlying) is lifted before expiration.
b. The correlation of the underlying and the hedge vehicle is less than one and their volatilities are unequal.
c. The underlying instrument and the hedge vehicle are dissimilar.
d. All of the above are correct.

EXAMPLE 12.2: FRM EXAM 2007—QUESTION 99

Which of the following trade(s) contain mainly basis risk?

I. Long 1,000 lots Nov 07 ICE Brent Oil contracts and short 1,000 lots Nov 07 NYMEX WTI Crude Oil contracts
II. Long 1,000 lots Nov 07 ICE Brent Oil contracts and long 2,000 lots Nov 07 ICE Brent Oil at-the-money put
III. Long 1,000 lots Nov 07 ICE Brent Oil contracts and short 1,000 lots Dec 07 ICE Brent Oil contracts
IV. Long 1,000 lots Nov 07 ICE Brent Oil contracts and short 1,000 lots Dec 07 NYMEX WTI Crude Oil contracts
   a. II and IV only
   b. I and III only
   c. I, III, and IV only
   d. III and IV only

12.2 OPTIMAL HEDGING

The previous section gave an example of a unit hedge, where the amounts transacted are identical in the two markets. In general, this is not appropriate. We have to decide how much of the hedging instrument to transact.

Consider a situation where a portfolio manager has an inventory of carefully selected corporate bonds that should do better than their benchmark. The manager wants to guard against interest rate increases, however, over the next three months. In this situation, it would be too costly to sell the entire portfolio only to buy it back later. Instead, the manager can implement a temporary hedge using derivative contracts, for instance T-bond futures.

Here, we note that the only risk is price risk, as the quantity of the inventory is known. This may not always be the case, however. Farmers, for instance, have uncertainty over both prices and the size of their crop. If so, the hedging problem is substantially more complex as it involves hedging *revenues*, which involves analyzing demand and supply conditions.

12.2.1 The Optimal Hedge Ratio

Define ΔS as the change in the dollar value of the inventory and ΔF as the change in the dollar value of one futures contract. The inventory, or position to be hedged, can be existing or anticipatory, that is, to be received in the future with a great degree of certainty. The manager is worried about potential movements in the value of the inventory ΔS.

If the manager goes long N futures contracts, the total change in the value of the portfolio is

$$\Delta V = \Delta S + N \Delta F \qquad (12.3)$$

One should try to find the hedge that reduces risk to the minimum level. The variance of total profits is equal to

$$\sigma_{\Delta V}^2 = \sigma_{\Delta S}^2 + N^2 \sigma_{\Delta F}^2 + 2 N \sigma_{\Delta S, \Delta F} \qquad (12.4)$$

Note that volatilities are initially expressed in dollars, not in rates of return, as we attempt to stabilize dollar values.

Taking the derivative with respect to N

$$\frac{\partial \sigma_{\Delta V}^2}{\partial N} = 2 N \sigma_{\Delta F}^2 + 2 \sigma_{\Delta S, \Delta F} \qquad (12.5)$$

For simplicity, drop the Δ in the subscripts. Setting Equation (12.5) equal to zero and solving for N, we get

$$N^* = -\frac{\sigma_{\Delta S, \Delta F}}{\sigma_{\Delta F}^2} = -\frac{\sigma_{SF}}{\sigma_F^2} = -\rho_{SF} \frac{\sigma_S}{\sigma_F} \qquad (12.6)$$

where σ_{SF} is the covariance between futures and spot price changes. Here, N^* is the minimum variance hedge ratio.

In practice, there is often confusion about the definition of the portfolio value and unit prices. Here S consists of the number of units (shares, bonds, bushels, gallons) times the unit price (stock price, bond price, wheat price, fuel price).

It is sometimes easier to deal with unit prices and to express volatilities in terms of *rates of changes in unit prices*, which are unitless. Defining quantities Q and unit prices s, we have $S = Qs$. Similarly, the notional amount of one futures contract is $F = Q_f f$. We can then write

$$\sigma_{\Delta S} = Q\sigma(\Delta s) = Qs\sigma(\Delta s/s)$$
$$\sigma_{\Delta F} = Q_f\sigma(\Delta f) = Q_f f\sigma(\Delta f/f)$$
$$\sigma_{\Delta S, \Delta F} = \rho_{sf}[Qs\sigma(\Delta s/s)][Q_f f\sigma(\Delta f/f)]$$

Using Equation (12.6), the optimal hedge ratio N^* can also be expressed as

$$N^* = -\rho_{SF}\frac{Qs\sigma(\Delta s/s)}{Q_f f\sigma(\Delta f/f)} = -\rho_{SF}\frac{\sigma(\Delta s/s)}{\sigma(\Delta f/f)}\frac{Qs}{Q_f f} = -\beta_{sf}\frac{Q \times s}{Q_f \times f} \qquad (12.7)$$

where β_{sf} is the coefficient in the regression of $\Delta s/s$ over $\Delta f/f$. The second term represents an adjustment factor for the size of the cash position and of the futures contract.

The optimal amount N^* can be derived from the slope coefficient of a regression of $\Delta s/s$ on $\Delta f/f$:

$$\frac{\Delta s}{s} = \alpha + \beta_{sf}\frac{\Delta f}{f} + \epsilon \qquad (12.8)$$

As seen in Chapter 3, standard regression theory shows that

$$\beta_{sf} = \frac{\sigma_{sf}}{\sigma_f^2} = \rho_{sf}\frac{\sigma_s}{\sigma_f} \qquad (12.9)$$

Thus, the best hedge is obtained from a regression of the (change in the) value of the inventory on the value of the hedge instrument.

KEY CONCEPT

The optimal hedge is given by the negative of the beta coefficient of a regression of changes in the cash value on changes in the payoff on the hedging instrument.

We can do more than this, though. At the optimum, we can find the variance of profits by replacing N in Equation (12.4) by N^*, which gives

$$\sigma_V^{*2} = \sigma_S^2 + \left(\frac{\sigma_{SF}}{\sigma_F^2}\right)^2 \sigma_F^2 + 2\left(\frac{-\sigma_{SF}}{\sigma_F^2}\right)\sigma_{SF}$$

$$= \sigma_S^2 + \frac{\sigma_{SF}^2}{\sigma_F^2} + 2\frac{-\sigma_{SF}^2}{\sigma_F^2} = \sigma_S^2 - \frac{\sigma_{SF}^2}{\sigma_F^2} \tag{12.10}$$

We can measure the quality of the optimal hedge ratio in terms of the amount by which we decreased the variance of the original portfolio:

$$R^2 = \frac{(\sigma_S^2 - \sigma_V^{*2})}{\sigma_S^2} \tag{12.11}$$

After substitution of Equation (12.10), we find that $R^2 = (\sigma_S^2 - \sigma_S^2 + \sigma_{SF}^2/\sigma_F^2)/\sigma_S^2 = \sigma_{SF}^2/(\sigma_F^2\sigma_S^2) = \rho_{SF}^2$. This unitless number is also the coefficient of determination, or the percentage of variance in $\Delta s/s$ explained by the independent variable $\Delta f/f$. Thus this regression also gives us the effectiveness of the hedge, which is measured by the proportion of variance eliminated.

We can also express the volatility of the hedged position from Equation (12.10) using the R^2 as

$$\sigma_V^* = \sigma_S\sqrt{(1 - R^2)} \tag{12.12}$$

This shows that if $R^2 = 1$, the regression fit is perfect, and the resulting portfolio has zero risk. In this situation, the portfolio has no basis risk. However, if the R^2 is very low, the hedge is not effective.

12.2.2 Example

An airline knows that it will need to purchase 10,000 metric tons of jet fuel in three months. It wants some protection against an upturn in prices using futures contracts.

The company can hedge using heating oil futures contracts traded on NYMEX. The notional for one contract is 42,000 gallons. As there is no futures contract on jet fuel, the risk manager wants to check if heating oil could provide an efficient hedge instead. The current price of jet fuel is $277/metric ton. The futures price of heating oil is $0.6903/gallon. The standard deviation of the rate of change in jet fuel prices over three months is 21.17%, that of futures is 18.59%, and the correlation is 0.8243.

Compute

a. The notional and standard deviation of the unhedged fuel cost in dollars
b. The optimal number of futures contract to buy/sell, rounded to the closest integer
c. The standard deviation of the hedged fuel cost in dollars

Answer

a. The position notional is $Qs = \$2,770,000$. The standard deviation in dollars is

$$\sigma(\Delta s/s)s\,Q = 0.2117 \times \$277 \times 10,000 = \$586,409$$

For reference, that of one futures contract is

$$\sigma(\Delta f/f)f\,Q_f = 0.1859 \times \$0.6903 \times 42,000 = \$5,389.72$$

with a futures notional of $fQ_f = \$0.6903 \times 42,000 = \$28,992.60$.

b. The cash position corresponds to a payment, or liability. Hence, the company will have to *buy* futures as protection. First, we compute beta, which is $\beta_{sf} = 0.8243(0.2117/0.1859) = 0.9387$. The corresponding covariance term is $\sigma_{sf} = 0.8243 \times 0.2117 \times 0.1859 = 0.03244$. Adjusting for the notionals, this is $\sigma_{SF} = 0.03244 \times \$2,770,000 \times \$28,993 = 2,605,268,452$. The optimal hedge ratio is, using Equation (12 7),

$$N^* = \beta_{sf}\frac{Q \times s}{Q_f \times f} = 0.9387\frac{10,000 \times \$277}{42,000 \times \$0.69} = 89.7$$

or 90 contracts after rounding (which we ignore in what follows).

c. To find the risk of the hedged position, we use Equation (12.10). The volatility of the unhedged position is $\sigma_S = \$586,409$. The variance of the hedged position is

$\sigma_S^2 = (\$586,409)^2$	$= +343,875,515,281$
$-\sigma_{SF}^2/\sigma_F^2 = -(2,605,268,452/5,390)^2$	$= -233,653,264,867 \cdot$
V(hedged)	$= +110,222,250,414$

Taking the square root, the volatility of the hedged position is $\sigma_V^* = \$331,997$. Thus the hedge has reduced the risk from $586,409 to $331,997. Computing the R^2, we find that one minus the ratio of the hedged and unhedged variances is $(1 - 110,222,250,414/343,875,515,281) = 67.95\%$. This is exactly the square of the correlation coefficient, $0.8243^2 = 0.6795$, or effectiveness of the hedge.
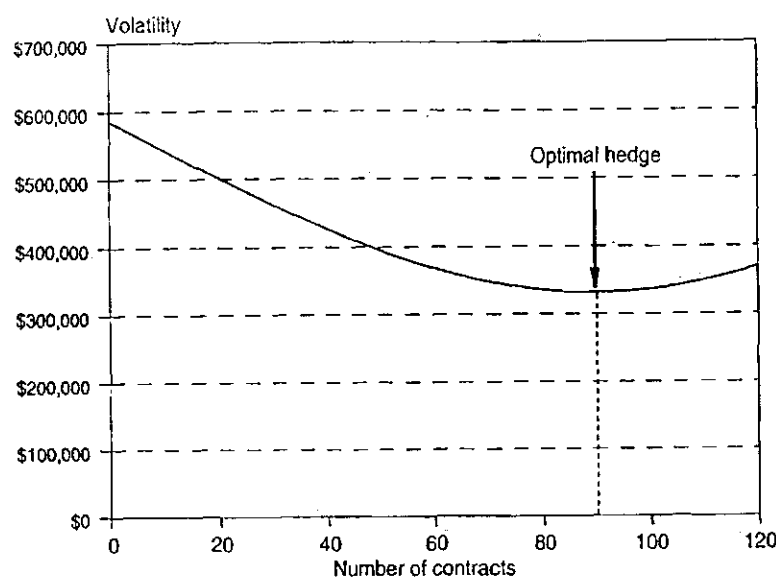
FIGURE 12.2   Risk of Hedged Position and Number of Contracts

Figure 12.2 displays the relationship between the risk of the hedged position and the number of contracts. With no hedging, the volatility is $586,409. As N increases, the risk decreases, reaching a minimum for $N^* = 90$ contracts. The graph also shows that the quadratic relationship is relatively flat for a range of values around the minimum. Choosing anywhere between 80 and 100 contracts will have little effect on the total risk. Given the substantial reduction in risk, the risk manager could choose to implement the hedge.

12.2.3 Liquidity Issues

Although futures hedging can be successful at mitigating market risk, it can create other risks. Futures contracts are marked to market daily. Hence they can involve large cash inflows or outflows. Cash outflows, in particular, can create liquidity problems, especially when they are not offset by cash inflows from the underlying position.

EXAMPLE 12.3: FRM EXAM 2001—QUESTION 86

If two securities have the same volatility and a correlation equal to −0.5, their minimum variance hedge ratio is

a. 1:1
b. 2:1
c. 4:1
d. 16:1

EXAMPLE 12.4: FRM EXAM 2007—QUESTION 125

A firm is going to buy 10,000 barrels of West Texas Intermediate Crude Oil. It plans to hedge the purchase using the Brent Crude Oil futures contract. The correlation between the spot and futures prices is 0.72. The volatility of the spot price is 0.35 per year. The volatility of the Brent Crude Oil futures price is 0.27 per year. What is the hedge ratio for the firm?

a. 0.9333
b. 0.5554
c. 0.8198
d. 1.2099

EXAMPLE 12.5: FRM EXAM 2003—QUESTION 14

A bronze producer will sell 1,000 mt (metric tons) of bronze in three months at the prevailing market price at that time. The standard deviation of the price of bronze over a three-month period is 2.6%. The company decides to use three-month futures on copper to hedge. The copper futures contract is for 25 mt of copper. The standard deviation of the futures price is 3.2%. The correlation between three-month changes in the futures price and the price of bronze is 0.77. To hedge its price exposure, how many futures contracts should the company buy/sell?

a. Sell 38 futures
b. Buy 25 futures
c. Buy 63 futures
d. Sell 25 futures

12.3 APPLICATIONS OF OPTIMAL HEDGING

The linear framework presented here is completely general. We now specialize it to two important cases, duration and beta hedging. The first applies to the bond market, the second to the stock market.

12.3.1 Duration Hedging

Modified duration can be viewed as a measure of the exposure of relative changes in prices to movements in yields. Using the definitions in Chapter 1, we can write

$$\Delta P = (-D^* P)\Delta y \tag{12.13}$$

where D^* is the modified duration. The dollar duration is defined as $(D^* P)$.

Assuming the duration model holds, which implies that the change in yield Δy does not depend on maturity, we can rewrite this expression for the cash and futures positions

$$\Delta S = (-D_S^* S)\Delta y \qquad \Delta F = (-D_F^* F)\Delta y$$

where D_S^* and D_F^* are the modified durations of S and F, respectively. Note that these relationships are supposed to be perfect, without an error term. The variances and covariance are then

$$\sigma_S^2 = (D_S^* S)^2 \sigma^2(\Delta y) \qquad \sigma_F^2 = (D_F^* F)^2 \sigma^2(\Delta y) \qquad \sigma_{SF} = (D_F^* F)(D_S^* S)\sigma^2(\Delta y)$$

We can replace these in Equation (12.6):

$$N^* = -\frac{\sigma_{SF}}{\sigma_F^2} = -\frac{(D_F^* F)(D_S^* S)}{(D_F^* F)^2} = -\frac{(D_S^* S)}{(D_F^* F)} \qquad (12.14)$$

Alternatively, this can be derived as follows. Write the total portfolio payoff as

$$\begin{aligned}
\Delta V &= \Delta S + N\Delta F \\
&= (-D_S^* S)\Delta y + N(-D_F^* F)\Delta y \\
&= -[(D_S^* S) + N(D_F^* F)] \times \Delta y
\end{aligned}$$

which is zero when the net exposure, represented by the term between brackets, is zero. In other words, the optimal hedge ratio is simply minus the ratio of the dollar duration of cash relative to the dollar duration of the hedge. This ratio can also be expressed in dollar value of a basis point (DVBP).

More generally, we can use N as a tool to modify the total duration of the portfolio. If we have a target duration of D_V, this can be achieved by setting $[(D_S^* S) + N(D_F^* F)] = D_V^* V$, or

$$N = \frac{(D_V^* V - D_S^* S)}{(D_F^* F)} \qquad (12.15)$$

of which Equation (12.14) is a special case.

KEY CONCEPT

The optimal duration hedge is given by the ratio of the dollar duration of the position to that of the hedging instrument.

Example 1

A portfolio manager holds a bond portfolio worth $10 million with a modified duration of 6.8 years, to be hedged for three months. The current futures price is 93-02, with a notional of $100,000. We assume that its duration can be measured by that of the cheapest-to-deliver, which is 9.2 years.

Compute

a. The notional of the futures contract
b. The number of contracts to buy/sell for optimal protection

Answer

a. The notional is $[93 + (2/32)]/100 \times \$100,000 = \$93,062.5$.
b. The optimal number to *sell* is from Equation (12.14)

$$N^* = -\frac{(D_S^* S)}{(D_F^* F)} = -\frac{6.8 \times \$10,000,000}{9.2 \times \$93,062.5} = -79.4$$

or 79 contracts after rounding. Note that the DVBP of the futures is about $9.2 \times \$93,000 \times 0.01\% = \$85$.

Example 2

On February 2, a corporate Treasurer wants to hedge a July 17 issue of $5 million of commercial paper with a maturity of 180 days, leading to anticipated proceeds of $4.52 million. The September Eurodollar futures trades at 92, and has a notional amount of $1 million.

Compute

a. The current dollar value of the futures contract
b. The number of contracts to buy/sell for optimal protection

Answer

a. The current dollar price is given by $\$10,000[100 - 0.25(100 - 92)] = \$980,000$. Note that the duration of the futures is always three months (90 days), since the contract refers to three-month LIBOR.
b. If rates increase, the cost of borrowing will be higher. We need to offset this by a gain, or a short position in the futures. The optimal number is from Equation (12.14)

$$N^* = -\frac{(D_S^* S)}{(D_F^* F)} = -\frac{180 \times \$4,520,000}{90 \times \$980,000} = -9.2$$

or nine contracts after rounding. Note that the DVBP of the futures is about $0.25 \times \$1,000,000 \times 0.01\% = \$25$.

EXAMPLE 12.6: FRM EXAM 2007—QUESTION 17

On June 2, a fund manager with USD 10 million invested in government bonds is concerned that interest rates will be highly volatile over the next three months. The manager decides to use the September Treasury bond futures contract to hedge the portfolio. The current futures price is USD 95.0625. Each contract is for the delivery of USD 100,000 face value of bonds. The duration of the manager's bond portfolio in three months will be 7.8 years. The cheapest-to-deliver bond in the Treasury bond futures contract is expected to have a duration of 8.4 years at maturity of the contract. At the maturity of the Treasury bond futures contract, the duration of the underlying benchmark Treasury bond is nine years. What position should the fund manager undertake to mitigate his interest rate risk exposure?

a. Short 94 contracts
b. Short 98 contracts
c. Short 105 contracts
d. Short 113 contracts

EXAMPLE 12.7: DURATION HEDGING

What assumptions does a duration-based hedging scheme make about the way in which interest rates move?

a. All interest rates change by the same amount.
b. A small parallel shift occurs in the yield curve.
c. Any parallel shift occurs in the term structure.
d. Interest rates movements are highly correlated.

EXAMPLE 12.8: HEDGING WITH EURODOLLAR FUTURES

If all spot interest rates are increased by one basis point, a value of a portfolio of swaps will increase by $1,100. How many Eurodollar futures contracts are needed to hedge the portfolio?

a. 44
b. 22
c. 11
d. 1,100

EXAMPLE 12.9: FRM EXAM 2004—QUESTION 4

Albert Henri is the fixed-income manager of a large Canadian pension fund. The present value of the pension fund's portfolio of assets is CAD 4 billion while the expected present value of the fund's liabilities is CAD 5 billion. The respective modified durations are 8.254 and 6.825 years. The fund currently has an actuarial deficit (assets < liabilities) and Albert must avoid widening this gap. There are currently two scenarios for the yield curve: the first scenario is an upward shift of 25 bps, with the second scenario a downward shift of 25 bps. The most liquid interest rate futures contract has a present value of CAD 68,336 and a duration of 2.1468 years. Analyzing both scenarios separately, what should Albert Henry do to avoid widening the pension fund gap? Choose the best option.

First Scenario	Second Scenario
a. Do nothing.	Buy 7,559 contracts.
b. Do nothing.	Sell 7,559 contracts.
c. Buy 7,559 contracts.	Do nothing.
d. Do nothing.	Do nothing.

12.3.2 Beta Hedging

We now turn to equity hedging using stock index futures. Beta, or systematic risk can be viewed as a measure of the exposure of the rate of return on a portfolio i to movements in the "market" m:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it} \tag{12.16}$$

where β represents the systematic risk, α the intercept (which is not a source of risk and therefore ignored for risk management purposes), and ϵ the residual component, which is uncorrelated with the market. We can also write, in line with the previous sections and ignoring the residual and intercept,

$$(\Delta S/S) \approx \beta(\Delta M/M) \tag{12.17}$$

Now, assume that we have at our disposal a stock-index futures contract, which has a beta of unity $(\Delta F/F) = 1(\Delta M/M)$. For options, the beta is replaced by the net delta, $(\Delta C) = \delta(\Delta M)$.

As in the case of bond duration, we can write the total portfolio payoff as

$$
\begin{aligned}
\Delta V &= \Delta S + N\Delta F \\
&= (\beta S)(\Delta M/M) + NF(\Delta M/M), \\
&= [(\beta S) + NF] \times (\Delta M/M)
\end{aligned}
$$

which is set to zero when the net exposure, represented by the term between brackets is zero. The optimal number of contracts to short is

$$
N^* = -\frac{\beta S}{F} \tag{12.18}
$$

KEY CONCEPT

The optimal hedge with stock index futures is given by the beta of the cash position times its value divided by the notional of the futures contract.

Example

A portfolio manager holds a stock portfolio worth $10 million with a beta of 1.5 relative to the S&P 500. The current futures price is 1,400, with a multiplier of $250.

Compute

a. The notional of the futures contract
b. The number of contracts to sell short for optimal protection

Answer

a. The notional amount of the futures contract is $250 × 1400 = $350,000.
b. The optimal number of contract to short is, from Equation (12.18),

$$
N^* = -\frac{\beta S}{F} = -\frac{1.5 \times \$10,000,000}{1 \times \$350,000} = -42.9
$$

or 43 contracts after rounding.

The quality of the hedge will depend on the size of the residual risk in the market model of Equation (12.16). For large portfolios, the approximation may be good. In contrast, hedging an individual stock with stock index futures may give poor results.

For instance, the correlation of a typical U.S. stock with the S&P 500 is 0.50. For an industry index, it is typically 0.75. Using the regression effectiveness in Equation (12.12), we find that the volatility of the hedged portfolio is

still about $\sqrt{1 - 0.5^2} = 87\%$ of the unhedged volatility for a typical stock and about 66% of the unhedged volatility for a typical industry. The lower number shows that hedging with general stock index futures is more effective for large portfolios. To obtain finer coverage of equity risks, hedgers could use futures contracts on industrial sectors, or exchange-traded funds (ETFs), or even single stock futures.

EXAMPLE 12.10: FRM EXAM 2005—QUESTION 97

Suppose that the benchmark for an equity portfolio of USD 12 million is the S&P 500. Also suppose the current value of the S&P 500 is 1,040 and the portfolio beta relative to the S&P 500 is 1.4. If the portfolio manager wants to completely hedge the portfolio over the next three months using the S&P 500 index futures (that has a multiplier of 250), which of the following is the correct hedging strategy?

a. Long 46 contracts
b. Short 46 contracts
c. Long 65 contracts
d. Short 65 contracts

EXAMPLE 12.11: FRM EXAM 2007—QUESTION 107

The current value of the S&P 500 index is 1,457, and each S&P futures contract is for delivery of 250 times the index. A long-only equity portfolio with market value of USD 300,100,000 has beta of 1.1. To reduce the portfolio beta to 0.75, how many S&P futures contract should you sell?

a. 288 contracts
b. 618 contracts
c. 906 contracts
d. 574 contracts

12.4 IMPORTANT FORMULAS

Profit on position with unit hedge: $Q[(S_2 - S_1) - (F_2 - F_1)] = Q[b_2 - b_1]$

Optimal hedge ratio: $N^* = -\beta_{sf} \frac{Q \times s}{Q_f \times f}$

Optimal hedge ratio (unitless): $\beta_{sf} = \frac{\sigma_{sf}}{\sigma_f^2} = \rho_{sf} \frac{\sigma_s}{\sigma_f}$

Volatility of the hedged position: $\sigma_V^* = \sigma_S \sqrt{(1 - R^2)}$

Duration hedge: $N^* = -\dfrac{(D_S^* S)}{(D_F^* F)}$

$N^* = -\beta \dfrac{S}{F}$

12.5 ANSWERS TO CHAPTER EXAMPLES

Example 12.1: FRM Exam 2000—Question 79

d. Basis risk occurs if movements in the value of the cash and hedged positions do not offset each other perfectly. This can happen if the instruments are dissimilar or if the correlation is not unity. Even with similar instruments, if the hedge is lifted before the maturity of the underlying, there is some basis risk.

Example 12.2: FRM Exam 2007—Question 99

c. There is mainly basis risk for positions that are both long and short either different months or contracts. Position II) is long twice the same contract and thus has no basis risk (but a lot of directional risk).

Example 12.3: FRM Exam 2001—Question 86

b. Set x as the amount to invest in the second security, relative to that in the first (or the hedge ratio). The variance is then proportional to $1 + x^2 + 2x\rho$. Taking the derivative and setting to zero, we have $x = -rho = 0.5$. Thus, one security must have twice the amount in the other. Alternatively, the hedge ratio is given by $N^* = -\rho \frac{\sigma_S}{\sigma_F}$, which gives 0.5. Answer b. is the only one which is consistent with this number or its inverse.

Example 12.4: FRM Exam 2007—Question 125

a. The optimal hedge ratio is $\beta_{sf} = \rho_{sf} \frac{\sigma_s}{\sigma_f} = 0.72 \; 0.35/0.27 - 0.933$.

Example 12.5: FRM Exam 2003—Question 14

b. The optimal hedge ratio is $\rho \sigma_s / \sigma_f = 0.77 \times 2.6/3.2 = 0.626$. Taking into account the size of the position, the number of contracts to buy is $0.626 \times 1,000/25 = 15.03$.

Example 12.6: FRM Exam 2007—Question 17

b. The number of contracts to short is $N^* = -\frac{(D_S^* S)}{(D_F^* F)} = -(7.8 \times 10,000,000)/(8.4 \times (95.0625) \times 1,000) = -97.7$, or 98 contracts. Note that the relevant duration for the futures is that of the CTD; other numbers are irrelevant.

Example 12.7: Duration Hedging

b. The assumption is that of (1) parallel and (2) small moves in the yield curve. Answers a. and c. are the same, and omit the size of the move. Answer d. would require perfect, not high, correlation plus small moves.

Example 12.8: Hedging with Eurodollar Futures

a. The DVBP of the portfolio is $1100. That of the futures is $25. Hence the ratio is 1100/25 = 44.

Example 12.9: FRM Exam 2004—Question 4

a. We first have to compute the dollar duration of assets and liabilities, which gives, in millions, $4{,}000 \times 8.254 = 33{,}016$ and $5{,}000 \times 6.825 = 34{,}125$, respectively. Because the DD of liabilities exceeds that of assets, a decrease in rates will increase the liabilities more than the assets, leading to a worsening deficit. Mr. Henri needs to buy interest rate futures as an offset. The number of contracts is $(34{,}125 - 33{,}016)/(68{,}336 \times 2.1468/1{,}000{,}000) = 7{,}559$.

Example 12.10: FRM Exam 2005—Question 97

d. To hedge, the portfolio manager should sell index futures, to create a profit if the portfolio loses value. The number of contracts is $N^* = -\beta S/F = -(1.4 \times 12{,}000{,}000)/(1{,}400 \times 250) = -64.6$, or 65 contracts.

Example 12.11: FRM Exam 2007—Question 107

a. This is as in the previous question, but the hedge is partial, i.e. for a change of 1.10 to 0.75. So, $N^* = -\beta S/F = -(1.10 - 0.75)300{,}100{,}000/(1457 \times 250) = -288.3$ contracts.

Nonlinear Risk: Options

The previous chapter focused on "linear" hedging, using contracts such as forwards and futures whose values are linearly related to the underlying risk factors. Positions in these contracts are fixed over the hedge horizon. Because linear combinations of normal random variables are also normally distributed, linear hedging maintains normal distributions, albeit with lower variances.

Hedging nonlinear risks, however, is much more complex. Because options have nonlinear payoffs, the distribution of option values can be sharply asymmetrical. Due to the ubiquitous nature of options, risk managers need to be able to evaluate the risk of positions with options. Since options can be replicated by dynamic trading of the underlying instruments, this also provides insights into the risks of active trading strategies.

In a previous chapter, we have seen that market losses can be ascribed to the combination of two factors: exposure and adverse movements in the risk factor. Thus a large loss could occur because of the risk factor, which is bad luck. Too often, however, losses occur because the exposure profile is similar to a short option position. This is less forgivable, because exposure is under the control of the portfolio manager.

The challenge is to develop measures that provide an intuitive understanding of the exposure profile. Section 13.1 introduces option pricing and the Taylor approximation.[1] It starts from the Black–Scholes formula that was presented in Chapter 6. Partial derivatives, also known as "Greeks," are analyzed in Section 13.2. Section 13.3 then turns to the interpretation of dynamic hedging and discusses the distribution profile of option positions.

13.1 EVALUATING OPTIONS

13.1.1 Definitions

We consider a derivative instrument whose value depends on an underlying asset, which can be a price, an index, or a rate. As an example, consider a call option

---

[1] The reader should be forewarned that this chapter is more technical than others. It presupposes some exposure to option pricing and hedging.

where the underlying asset is a foreign currency. We use these definitions:

S_t = current spot price of the asset in dollars

F_t = current forward price of the asset

K = exercise price of option contract

f_t = current value of derivative instrument

r_t = domestic risk-free rate

r_t^* = foreign risk-free rate (also written as y)

σ_t = annual volatility of the rate of change in S

τ = time to maturity.

More generally, r^* represents the income payment on the asset, which represents the *annual rate* of dividend or coupon payments on a stock index or bond. For most options, we can write the value of the derivative as the function

$$f_t = f(S_t, r_t, r_t^*, \sigma_t, K, \tau) \tag{13.1}$$

The contract specifications are represented by K and the time to maturity τ. The other factors are affected by market movements, creating volatility in the value of the derivative. For simplicity, we drop the time subscripts in what follows.

Derivatives pricing is all about finding the value of f, given the characteristics of the option at expiration and some assumptions about the behavior of markets. For a forward contract, for instance, the expression is very simple. It reduces to

$$f = Se^{-r^*\tau} - Ke^{-r\tau} \tag{13.2}$$

More generally, we may not be able to derive an analytical expression for the function f, requiring numerical methods.

13.1.2 Taylor Expansion

We are interested in describing the movements in f. The exposure profile of the derivative can be described *locally* by taking a Taylor expansion,

$$df = \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}dS^2 + \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial r^*}dr^* + \frac{\partial f}{\partial \sigma}d\sigma + \frac{\partial f}{\partial \tau}d\tau + \cdots \tag{13.3}$$

Because the value depends on S in a nonlinear fashion, we added a quadratic term for S. The terms in Equation (13.3) approximate a nonlinear function by linear and quadratic polynomials.

Option pricing is about finding f. Option hedging uses the partial derivatives. Risk management is about combining those with the movements in the risk factors.

Figure 13.1 describes the relationship between the value of a European call and the underlying asset. The actual price is the solid line. The straight thin line is the linear (delta) estimate, which is the tangent at the initial point. The other
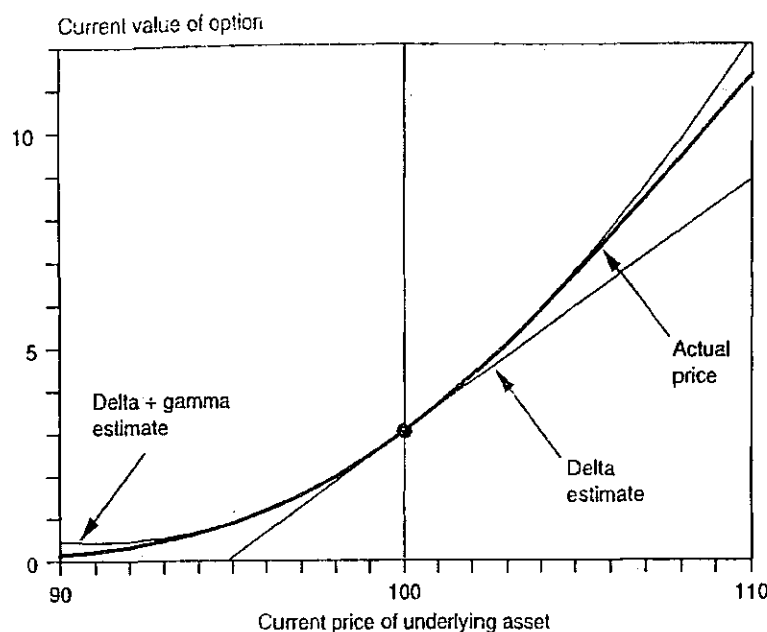
FIGURE 13.1 Delta–Gamma Approximation for a Long Call

line is the quadratic (delta plus gamma) estimates, which gives a much better fit because it has more parameters.

Note that, because we are dealing with sums of local price movements, we can aggregate the sensitivities at the portfolio level. This is similar to computing the portfolio duration from the sum of durations of individual securities, appropriately weighted.

Defining $\Delta = \frac{\partial f}{\partial S}$, for example, we can summarize the portfolio, or "book" Δ_P in terms of the total sensitivity,

$$\Delta_P = \sum_{i=1}^{N} x_i \Delta_i \qquad (13.4)$$

where x_i is the number of options of type i in the portfolio. To hedge against first-order price risk, it is sufficient to hedge the *net* portfolio delta. This is more efficient than trying to hedge every single instrument individually.

The Taylor expansion will provide a bad approximation in a number of cases:

- *Large movements in the underlying risk factor*
- *Highly nonlinear exposures*, such as options near expiry or exotic options
- *Cross-partial effects*, such as σ changing in relation with S

If this is the case, we need to turn to a full revaluation of the instrument. Using the subscripts 0 and 1 as the initial and final values, the change in the option value is

$$f_1 - f_0 = f(S_1, r_1, r_1^*, \sigma_1, K, \tau_1) - f(S_0, r_0, r_0^*, \sigma_0, K, \tau_0) \qquad (13.5)$$

13.1.3 Option Pricing

We now present the various partial derivatives for conventional European call and put options. As we have seen in Chapter 6, the Black–Scholes (BS) model provides a closed-form solution, from which these derivatives can be computed analytically.

The key point of the BS derivation is that a position in the option can be replicated by a "delta" position in the underlying asset. Hence, a portfolio combining the asset and the option in appropriate proportions is risk-free "locally," that is, for small movements in prices. To avoid arbitrage, this portfolio must return the risk-free rate. The option value is the discounted expected payoff:

$$f_t = E_{RN}[e^{-r\tau}F(S_T)] \tag{13.6}$$

where E_{RN} represents the expectation of the future payoff in a "risk-neutral" world, that is, assuming the underlying asset grows at the risk-free rate and the discounting also employs the risk-free rate.

In the case of a European call, the final payoff is $F(S_T) = Max(S_T - K, 0)$, and the current value of the call is given by

$$c = Se^{-r^*\tau}N(d_1) - Ke^{-r\tau}N(d_2) \tag{13.7}$$

where $N(d)$ is the cumulative distribution function for the standard normal distribution:

$$N(d) = \int_{-\infty}^{d} \Phi(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-\frac{1}{2}x^2}dx$$

with Φ defined as the standard normal density function. $N(d)$ is also the area to the left of a standard normal variable with value equal to d. The values of d_1 and d_2 are

$$d_1 = \frac{ln(Se^{-r^*\tau}/Ke^{-r\tau})}{\sigma\sqrt{\tau}} + \frac{\sigma\sqrt{\tau}}{2}, \quad d_2 = d_1 - \sigma\sqrt{\tau}$$

By put–call parity, the European put option value is

$$p = Se^{-r^*\tau}[N(d_1) - 1] - Ke^{-r\tau}[N(d_2) - 1] \tag{13.8}$$

13.2 OPTION "GREEKS"

13.2.1 Option Sensitivities: Delta and Gamma

Given these closed-form solutions for European options, we can derive all partial derivatives. The most important sensitivity is the delta, which is the first partial
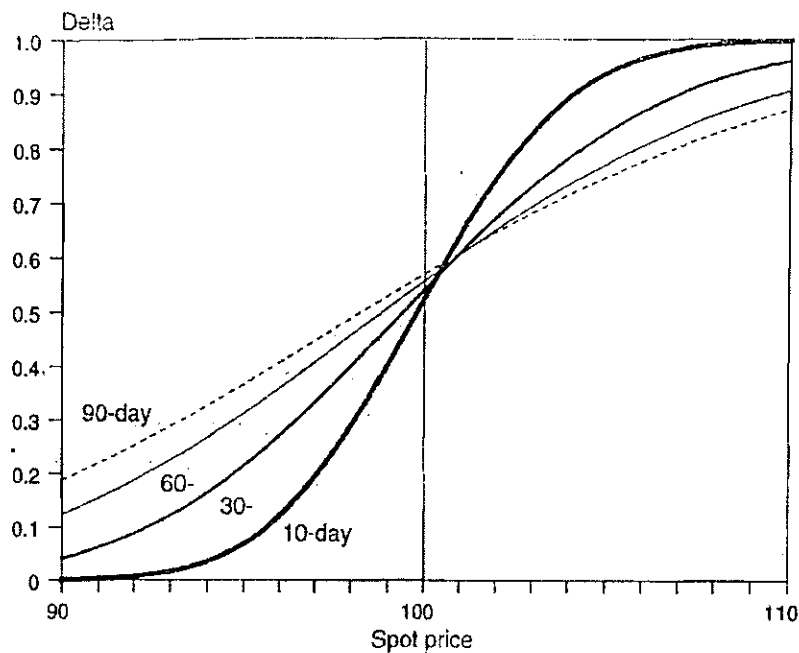
FIGURE 13.2 Option Delta

derivative with respect to the price. For a call option, this can be written explicitly as:

$$\Delta_c = \frac{\partial c}{\partial S} = e^{-r^* \tau} N(d_1) \tag{13.9}$$

which is always positive and below unity.

Figure 13.2 relates delta to the current value of S, for various maturities. The essential feature of this figure is that Δ varies substantially with the spot price and with time. As the spot price increases, d_1 and d_2 become very large, and Δ tends toward $e^{-r^* \tau}$, close to one for short maturities. In this situation, the option behaves like an outright position in the asset. Indeed the limit of Equation (13.7) is $c = Se^{-r^* \tau} - Ke^{-r \tau}$, which is exactly the value of our forward contract, Equation (13.2).

At the other extreme, if S is very low, Δ is close to zero and the option is not very sensitive to S. When S is close to the strike price K, Δ is close to 0.5, and the option behaves like a position of 0.5 in the underlying asset.

KEY CONCEPT

The delta of an at-the-money call option is close to 0.5. Delta moves to 1 as the call goes deep in-the-money. It moves to zero as the call goes deep out of the money.

The delta of a put option is

$$\Delta_p = \frac{\partial p}{\partial S} = e^{-r\tau}[N(d_1) - 1] \qquad (13.10)$$

which is always negative. It behaves similarly to the call Δ, except for the sign. The delta of an at-the-money put is about -0.5.

KEY CONCEPT

The delta of an at-the-money put option is close to -0.5. Delta moves to -1 as the put goes deep in the money. It moves to zero as the put goes deep out-of-the-money.

The figure also shows that, as the option nears maturity, the Δ function becomes more curved. The function converges to a step function, i.e., 0 when $S < K$, and 1 otherwise. Close-to-maturity options have unstable deltas.

For a European call or put, gamma (Γ) is the second-order term,

$$\Gamma = \frac{\partial^2 c}{\partial S^2} = \frac{e^{-r\tau}\Phi(d_1)}{S\sigma\sqrt{\tau}} \qquad (13.11)$$

which has the "bell shape" of the normal density function Φ. This is also the derivative of Δ with respect to S. Thus Γ measures the "instability" in Δ. Note that gamma is identical for a call and put with identical characteristics.

Figure 13.3 plots the call option gamma. At-the-money options have the highest gamma, which indicates that Δ changes very fast as S changes. In contrast, both in-the-money options and out-of-the-money options have low gammas because their delta is constant, close to one or zero, respectively. The figure also shows that as the maturity nears, the option gamma increases. This leads to the useful rule in the box.

KEY CONCEPT

For vanilla options, gamma is the highest, or nonlinearities are most pronounced, for short-term at-the-money options.

Thus, gamma is similar to the concept of convexity developed for bonds. Fixed-coupon bonds, however, always have positive convexity, whereas options can create positive or negative convexity. Positive convexity or gamma is beneficial, as it implies that the value of the asset drops more slowly and increases more
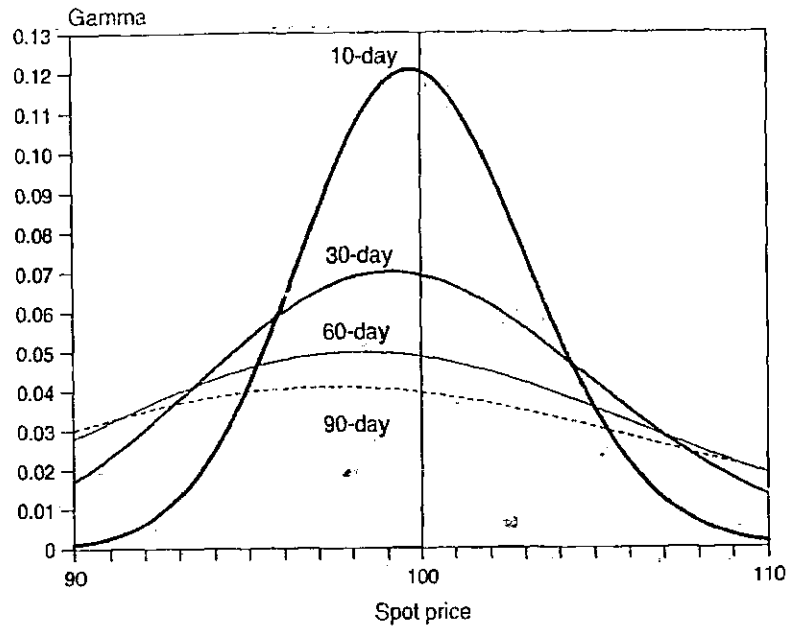
FIGURE 13.3 Option Gamma

quickly than otherwise. In contrast, negative convexity can be dangerous because it implies faster price falls and slower price increases.

Figure 13.4 summarizes the delta and gamma exposures of positions in options. Long positions in options, whether calls or puts, create positive convexity. Short positions create negative convexity. In exchange for assuming the harmful effect of this negative convexity, option sellers receive the premium.
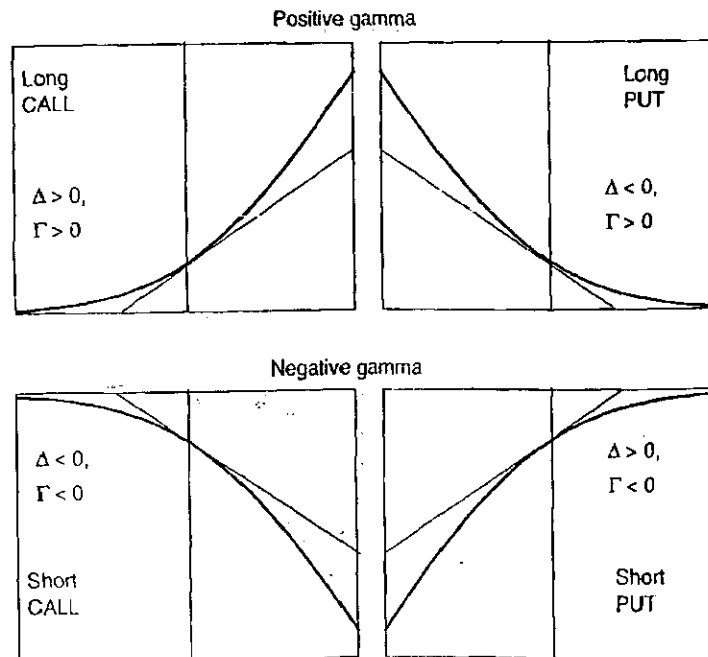
FIGURE 13.4 Delta and Gamma of Option Positions

EXAMPLE 13.1: FRM EXAM 2006—QUESTION 91

The dividend yield of an asset is 10% per annum. What is the delta of a long forward contract on the asset with 6 months to maturity?

a. 0.95

b. 1.00

c. 1.05

d. Cannot determine without additional information

EXAMPLE 13.2: FRM EXAM 2004—QUESTION 21

A 90-day European put option on Microsoft has an exercise price of $30. The current market price for Microsoft is $30. The delta for this option is close to

a. −1

b. −0.5

c. 0.5

d. 1

EXAMPLE 13.3: FRM EXAM 2006—QUESTION 80

You are given the following information about a European call option: Time to maturity = two years; continuous risk-free rate = 4%; continuous dividend yield = 1%; $N(d1) = 0.64$. Calculate the delta of this option.

a. −0.64

b. 0.36

c. 0.63

d. 0.64

EXAMPLE 13.4: FRM EXAM 2003—QUESTION 94

Which of the following IBM options has the highest gamma with the current market price of IBM common stock at USD 68?

a. Call option expiring in 10 days with strike USD 70

b. Call option expiring in 10 days with strike USD 50

c. Put option expiring in 10 days with strike USD 50

d. Put option expiring in two months with strike USD 70

EXAMPLE 13.5: FRM EXAM 2001—QUESTION 79

A bank has sold USD 300,000 of call options on 100,000 equities. The equities trade at 50, the option strike price is 49, the maturity is in three months, volatility is 20%, and the interest rate is 5%. How does the bank delta hedge?

a. Buy 65,000 shares
b. Buy 100,000 shares
c. Buy 21,000 shares
d. Sell 100,000 shares

EXAMPLE 13.6: FRM EXAM 2006—QUESTION 106

Suppose an existing short option position is delta-neutral, but has a gamma of −600. Also assume that there exists a traded option with a delta of 0.75 and a gamma of 1.50. In order to maintain the position gamma-neutral and delta-neutral, which of the following is the appropriate strategy to implement?

a. Buy 400 options and sell 300 shares of the underlying asset.
b. Buy 300 options and sell 400 shares of the underlying asset.
c. Sell 400 options and buy 300 shares of the underlying asset.
d. Sell 300 options and buy 400 shares of the underlying asset.

13.2.2 Option Sensitivities: Vega

Unlike linear contracts, options are exposed not only to movements in the direction of the spot price, but also in its volatility. Options therefore can be viewed as "volatility bets."

The sensitivity of an option to volatility is called the option vega (sometimes also called lambda, or kappa). For European calls and puts, this is

$$\Lambda = \frac{\partial c}{\partial \sigma} = Se^{-r^*\tau}\sqrt{\tau}\ \Phi(d_1) \tag{13.12}$$

which also has the "bell shape" of the normal density function Φ. As with gamma, vega is identical for similar call and put positions. Vega must be positive for long option positions.

Figure 13.5 plots the call option vega. The graph shows that at-the-money options are the most sensitive to volatility. The time effect, however, is different from that for gamma, because the term $\sqrt{\tau}$ appears in the numerator instead of
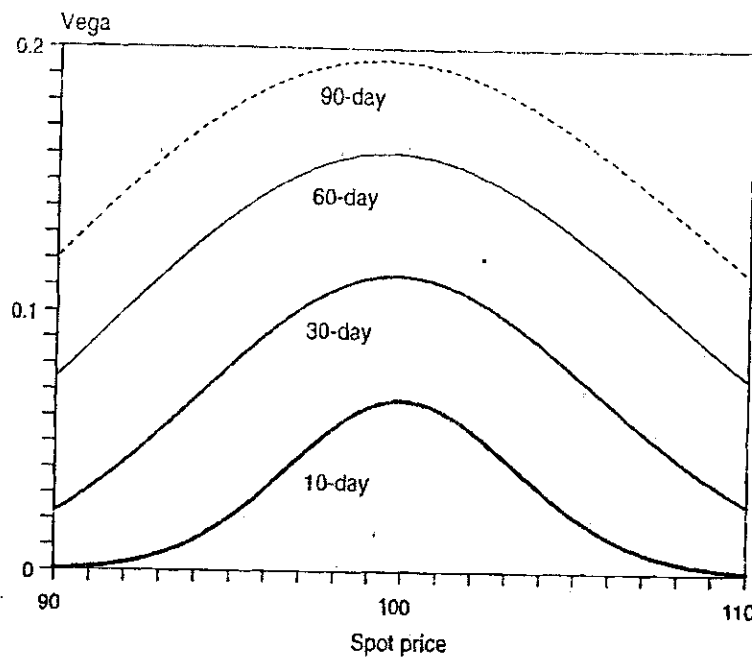
FIGURE 13.5 Option Vega

denominator. Thus, vega decreases with maturity, unlike gamma, which increases with maturity.

KEY CONCEPT

Vega is highest for long-term at-the-money options.

Changes in the volatility parameter can be a substantial source of risk. Figure 13.6 illustrates the time-variation in the implied volatility for options on the S&P stock index, also known as Volatility Index (VIX).[2] Over this period, the average value of the VIX index was 21%. The volatility in the daily change in VIX was about 2.4%.[3]

The VIX index experiences sharp spikes on a regular basis, however, reflecting increased uncertainty. In particular, VIX came close to or exceeded 50% during the crash of October 1987, during the LTCM crisis of September 1998, after the World Trade Center attack of September 2001, at the bottom of the 2000–2002 bear market in July 2002, and during the credit crisis that suddenly worsened in September 2008.

---

[2] The implied volatility is derived from the market prices of at-the-money near-term options on the S&P100 index and is calculated by the Chicago Board Options Exchange. In 2003, the methodology was changed; the new VIX index is derived from the prices of S&P500 index options across a wide range of strike prices.

[3] There is strong mean reversion in these data, so that daily volatilities cannot be extrapolated to annual data.

VIX Index (%)
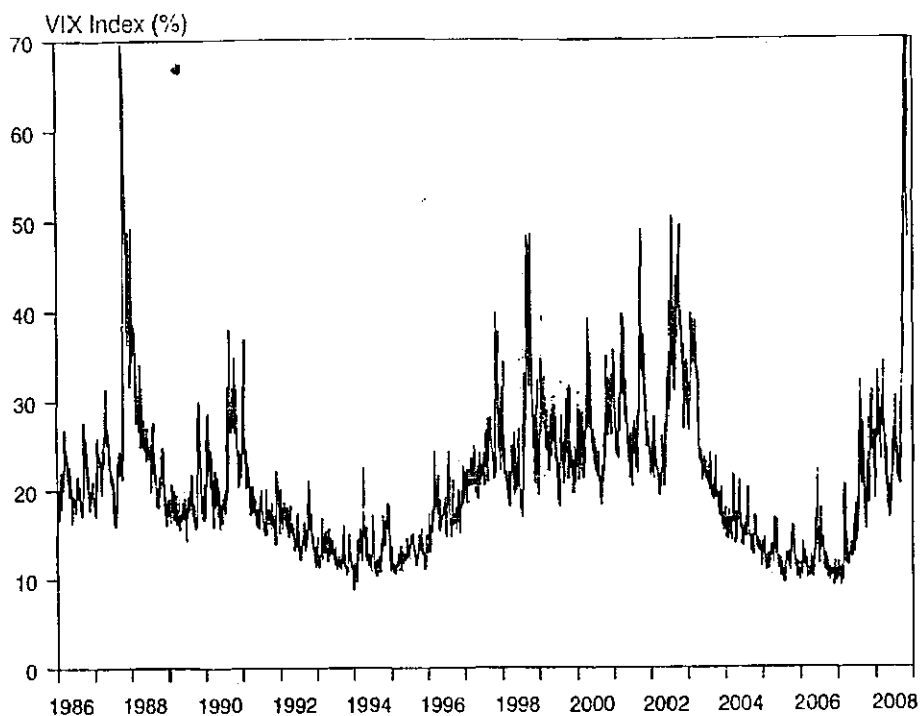
FIGURE 13.6 Movements in Implied Volatility

13.2.3 Option Sensitivities: Rho

The sensitivity to the domestic interest rate, also called rho, is

$$\rho_c = \frac{\partial c}{\partial r} = K e^{-r\tau} \tau N(d_2) \tag{13.13}$$

For a put,

$$\rho_p = \frac{\partial p}{\partial r} = -K e^{-r\tau} \tau N(-d_2) \tag{13.14}$$

An increase in the rate of interest increases the value of the call, as the underlying asset grows at a higher rate, which increases the probability of exercising the call, with a fixed strike price K. In the limit, for an infinite interest rate, the probability of exercise is one and the call option is equivalent to the stock itself. The reasoning is opposite for a put option.

The exposure to the yield on the asset is, for calls and puts, respectively,

$$\rho_C^* = \frac{\partial c}{\partial r^*} = -S e^{-r^*\tau} \tau N(d_1) \tag{13.15}$$

$$\rho_P^* = \frac{\partial p}{\partial r^*} = S e^{-r^*\tau} \tau N(-d_1) \tag{13.16}$$

An increase in the dividend yield decreases the growth rate of the underlying asset, which is harmful to the value of the call. Again, the reasoning is opposite for a put option.

13.2.4 Option Sensitivities: Theta

Finally, the variation in option value due to the passage of time is also called theta. This is also the time decay. Unlike other factors, however, the movement in remaining maturity is perfectly predictable. Time is not a risk factor.

For a European call, this is

$$\Theta_c = \frac{\partial c}{\partial t} = -\frac{\partial c}{\partial \tau} = -\frac{Se^{-r^*\tau}\sigma\Phi(d_1)}{2\sqrt{\tau}} + r^*Se^{-r^*\tau}N(d_1) - rKe^{-r\tau}N(d_2) \qquad (13.17)$$

For a European put, this is

$$\Theta_p = \frac{\partial p}{\partial t} = -\frac{\partial p}{\partial \tau} = -\frac{Se^{-r^*\tau}\sigma\Phi(d_1)}{2\sqrt{\tau}} - r^*Se^{-r^*\tau}N(-d_1) + rKe^{-r\tau}N(-d_2)$$

$$(13.18)$$

Theta is generally negative for long positions in both calls and puts. This means that the option loses value as time goes by.

For American options, however, Θ is *always* negative. Because they give their holder the choice to exercise early, shorter-term American options are unambiguously less valuable than longer-term options.

Figure 13.7 displays the behavior of a call option theta for various prices of the underlying asset and maturities. For long positions in options, theta is negative,
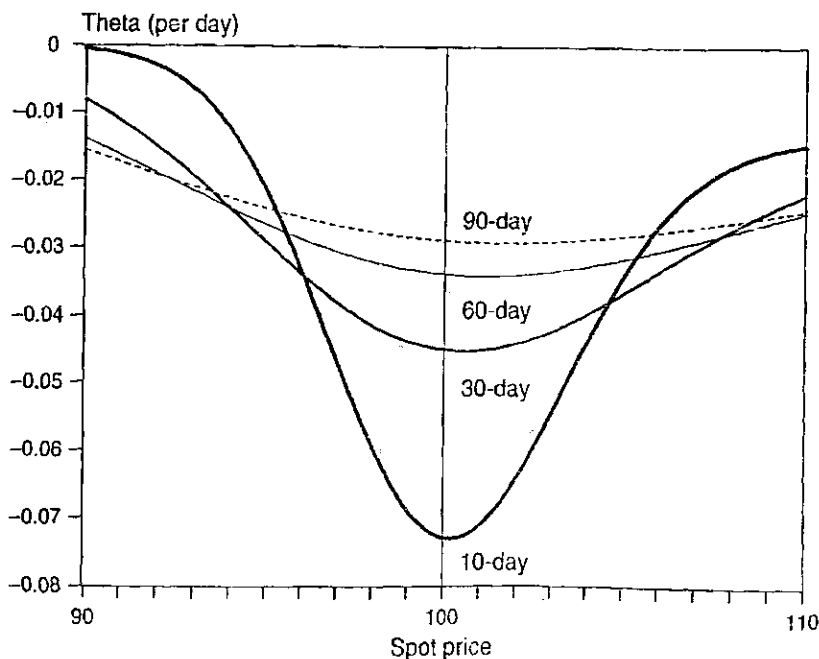
FIGURE 13.7   Option Theta

which reflects the fact that the option is a wasting asset. Like gamma, theta is greatest for short-term at-the-money options, when measured in absolute value. At-the-money options lose a great proportion of their value when the maturity is near.

13.2.5 Option Pricing and the "Greeks"

Having defined the option sensitivities, we can illustrate an alternative approach to the derivation of the Black–Scholes formula. Recall that the underlying process for the asset follows a stochastic process known as a geometric Brownian motion (GBM),

$$dS = \mu S dt + \sigma S dz \tag{13.19}$$

where dz has a normal distribution with mean zero and variance dt.

Considering only this *single* source of risk, we can return to the Taylor expansion in Equation (13.3). The value of the derivative is a function of S and time, which we can write as $f(S, t)$. The question is, how does f evolve over time?

We can relate the stochastic process of f to that of S using Ito's lemma, named after its creator. This can be viewed as an extension of the Taylor approximation to a stochastic environment. Applied to the GBM, this gives

$$df = \left(\tfrac{\partial f}{\partial S}\mu S + \tfrac{1}{2}\tfrac{\partial^2 f}{\partial S^2}\sigma^2 S^2 + \tfrac{\partial f}{\partial \tau}\right)dt + \left(\tfrac{\partial f}{\partial S}\sigma S\right)dz \tag{13.20}$$

This is also

$$df = (\Delta \mu S + \frac{1}{2}\Gamma\sigma^2 S^2 + \Theta)dt + (\Delta\sigma S)dz \tag{13.21}$$

The first term, including dt, is the trend. The second, including dz, is the stochastic component.

Next, we construct a portfolio delicately balanced between S and f that has no exposure to dz. Define this portfolio as

$$\Pi = f - \Delta S \tag{13.22}$$

Using (13.19) and (13.21), its stochastic process is

$$
\begin{aligned}
d\Pi &= [(\Delta\mu S + \tfrac{1}{2}\Gamma\sigma^2 S^2 + \Theta)dt + (\Delta\sigma S)dz] - \Delta[\mu S dt + \sigma S dz] \\
&= (\Delta\mu S)dt + (\tfrac{1}{2}\Gamma\sigma^2 S^2)dt + \Theta dt + (\Delta\sigma S)dz - (\Delta\mu S)dt - (\Delta\sigma S)dz \\
&= (\tfrac{1}{2}\Gamma\sigma^2 S^2 + \Theta)dt
\end{aligned}
$$

$$\tag{13.23}$$

This simplification is extremely important. Note how the terms involving dz cancel out each other. The portfolio has been immunized against this source of risk. At the same time, the terms in μS also happened to cancel out each other.

The fact that μ disappears from the trend in the portfolio is important, as it explains why the trend of the underlying asset does not appear in the Black–Scholes formula.

Continuing, we note that the portfolio Π has no risk. To avoid arbitrage, it must return the risk-free rate:

$$d\Pi = [r\Pi]dt = r(f - \Delta S)dt \tag{13.24}$$

If the underlying asset has a dividend yield of y, this must be adjusted to

$$d\Pi = (r\Pi)dt + y\Delta Sdt = r(f - \Delta S)dt + y\Delta Sdt \tag{13.25}$$

Setting the trends in Equations (13.23) and (15.24) equal to each other, we must have

$$(r - y)\Delta S + \frac{1}{2}\Gamma\sigma^2 S^2 + \Theta = rf \tag{13.26}$$

This is the Black–Scholes partial differential equation (PDE), which applies to any contract, or portfolio, that derives its value from S. The solution of this equation, with appropriate boundary conditions, leads to the BS formula for a European call, Equation (13.7).

We can use this relationship to understand how the sensitivities relate to each other. Consider a portfolio of derivatives, all on the same underlying asset, that is delta-hedged. Setting $\Delta = 0$ in Equation (13.26), we have

$$\frac{1}{2}\Gamma\sigma^2 S^2 + \Theta = rf \tag{13.27}$$

This shows that, for such portfolio, when Γ is large and positive, Θ must be negative if rf is small. In other words, a delta-hedged position with positive gamma, which is beneficial in terms of price risk, must have negative theta, or time decay. An example is the long straddle examined in Chapter 6. Such position is delta-neutral and has large gamma or convexity. It would benefit from a large move in S, whether up or down. This portfolio, however, involves buying options whose value decay very quickly with time. Thus, there is an intrinsic trade-off between Γ and Θ.

KEY CONCEPT

For delta-hedged portfolios, Γ and Θ must have opposite signs. Portfolios with positive convexity, for example, must experience time decay.

13.2.6 Option Sensitivities: Summary

We now summarize the sensitivities of option positions with some illustrative data in Table 13.1. Three strike prices are considered, $K = 90, 100$, and 110. We verify

TABLE 13.1   Derivatives for a European Call Parameters: $S = \$100, \sigma = 20\%, r = 5\%, y = 3\%, \tau = 3$ month

	Variable	Unit	Strike			Worst Loss	
			K = 90	K = 100	K = 110	Variable	Loss
c		Dollars	$11.02	$4.22	$1.05		
		Change per:					
Δ	spot price	dollar	0.868	0.536	0.197	−$2.08	−$1.114
Γ	spot price	dollar	0.020	0.039	0.028	4.33	$0.084
Λ	volatility	(% pa)	0.103	0.198	0.139	−2.5	−$0.495
ρ	interest rate	(% pa)	0.191	0.124	0.047	−0.10	−$0.013
ρ*	asset yield	(% pa)	−0.220	−0.135	−0.049	0.10	−$0.014
Θ	time	day	−0.014	−0.024	−0.016		

that the Γ, Λ, and Θ measures are all highest when the option is at-the-money ($K = 100$). Such options have the most nonlinear patterns.

The table also shows the loss for the worst daily movement in each risk factor at the 95% confidence level. For S, this is $dS = -1.645 \times 20\% \times \$100/\sqrt{252} = -\$2.08$. We combine this with delta, which gives a potential loss of $\Delta \times dS = -\$1.114$, or about a fourth of the option value.

Next, we examine the second order term, S^2. The worst squared daily movement is $dS^2 = 2.08^2 = 4.33$ in the risk factor at the 95% confidence level. We combine this with gamma, which gives a potential gain of $\frac{1}{2}\Gamma \times dS^2 = 0.5 \times 0.039 \times 4.33 = \$0.084$. Note that this is a gain because gamma is positive, but much smaller than the first-order effect. Thus the worst loss due to S would be $-\$1.114 + \$0.084 = -\$1.030$ using the linear and quadratic effects.

For σ, we observe a volatility of daily changes in σ on the order of 1.5%. The worst daily move is therefore $-1.645 \times 1.5 = -2.5$, expressed in percent, which gives a worst loss of $-\$0.495$. Finally, for r, we assume an annual volatility of changes in rates of 1%. The worst daily move is then $-1.645 \times 1/\sqrt{252} = -0.10$, in percent, which gives a worst loss of $-\$0.013$. So, most of the risk originates from S. In this case, a linear approximation using Δ only would capture most of the downside risk. For near-term at-the-money options, however, the quadratic effect is more important.

EXAMPLE 13.7: FRM EXAM 2004—QUESTION 65

Which of the following statements is *true* regarding options' Greeks?

a. Theta tends to be large and positive when buying at-the-money options.

b. Gamma is greatest for in-the-money options with long maturities.

c. Vega is greatest for at-the-money options with long maturities.

d. Delta of deep in-the-money put options tends towards +1.

EXAMPLE 13.8: FRM EXAM 2006—QUESTION 33

Steve, a market risk manager at Marcat Securities, is analyzing the risk of its S&P 500 index options trading desk. His risk report shows the desk is net long gamma and short vega. Which of the following portfolios of options shows exposures consistent with this report?

a. The desk has substantial long-expiry long call positions and substantial short-expiry short put positions.
b. The desk has substantial long-expiry long put positions and substantial long-expiry short call positions.
c. The desk has substantial long-expiry long call positions and substantial short-expiry short call positions.
d. The desk has substantial short-expiry long call positions and substantial long-expiry short call positions.

EXAMPLE 13.9: FRM EXAM 2006—QUESTION 54

Which of the following statements is *incorrect*?

a. The vega of a European call option is highest when the option is at-the-money.
b. The delta of a European-styled put option on an underlying stock moves toward zero as the price of the underlying stock rises.
c. The gamma of an at-the-money European-styled option tends to increase as the remaining maturity of the option decreases.
d. Compared to an at-the-money European-styled call option, an out-of-the-money European-styled option with the same strike price and remaining maturity has a greater negative value for theta.

EXAMPLE 13.10: FRM EXAM 2000—QUESTION 76

How can a trader produce a short vega, long gamma position?

a. Buy short-maturity options, sell long-maturity options.
b. Buy long-maturity options, sell short-maturity options.
c. Buy and sell options of long maturity.
d. Buy and sell options of short maturity.

EXAMPLE 13.11: FRM EXAM 2001—QUESTION 113

An option portfolio exhibits high unfavorable sensitivity to increases in implied volatility and while experiencing significant daily losses with the passage of time. Which strategy would the trader most likely employ to hedge his portfolio?

    a. Sell short dated options and buy long dated options
    b. Buy short dated options and sell long dated options
    c. Sell short dated options and sell long dated options
    d. Buy short dated options and buy long dated options

13.3 DYNAMIC HEDGING

The BS derivation taught us how to price and hedge options. Perhaps even more importantly, it showed that holding a call option is equivalent to holding a fraction of the underlying asset, where the fraction dynamically changes over time.

13.3.1 Delta and Dynamic Hedging

This equivalence is illustrated in Figure 13.8, which displays the current value of a call as a function of the current spot price. The long position in a call is replicated
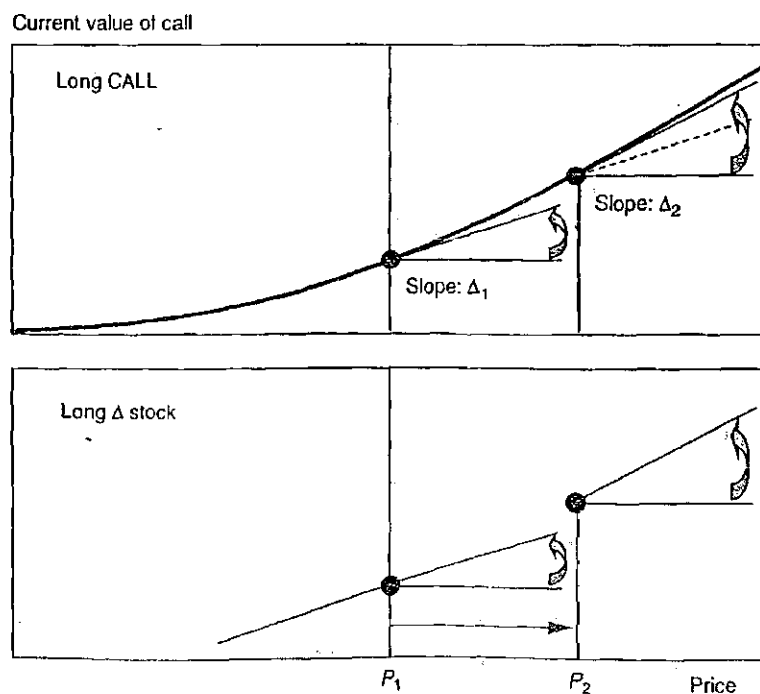
FIGURE 13.8 Dynamic Replication of a Call Option

Current value of put

Long PUT

Slope: Δ

0

Delta

-0.5

-1.0

Spot price
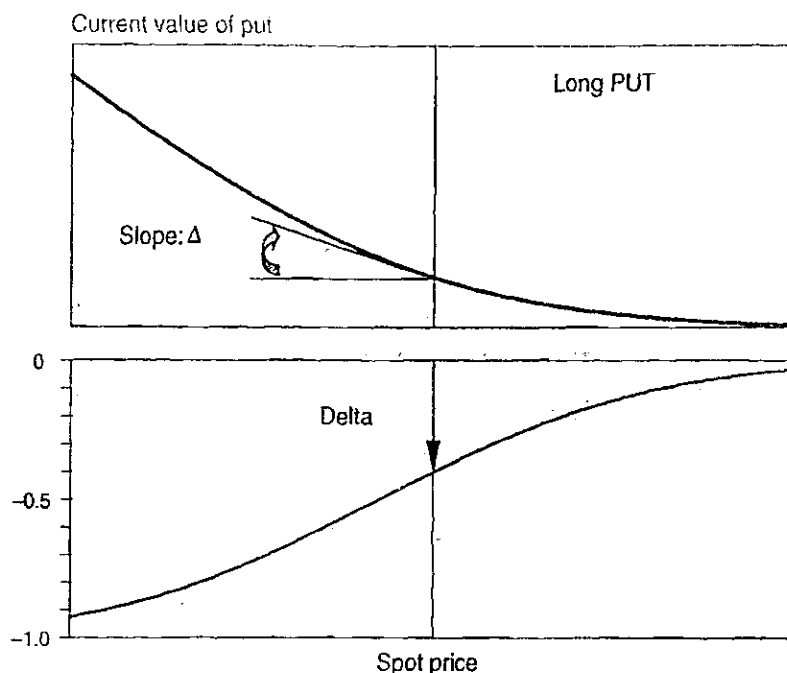
FIGURE 13.9 Dynamic Replication of a Put Option

by a partial position in the underlying asset. For an at-the-money position, the initial delta is about 0.5.

As the stock price increases from P_1 to P_2, the slope of the option curve, or delta, increases from Δ_1 to Δ_2. As a result, the option can be replicated by a larger position in the underlying asset. Conversely, when the stock price decreases, the size of the position is cut, as in a graduated stop-loss order. Thus the dynamic adjustment buys more of the asset as its price goes up, and conversely, sells it after a fall.

Figure 13.9 shows the dynamic replication of a put. We start at-the-money with Δ close to -0.5. As the price S goes up, Δ increases toward 0. Note that this is an increase since the initial delta was negative. As with the long call position, we *buy* more of the asset *after* its price has gone up. In contrast, short positions in calls and puts imply opposite patterns. Dynamic replication of a short option position implies buying more of the asset after its price has gone down.

13.3.2 Implications

For risk managers, these patterns are extremely important for a number of reasons. First, dynamic replication of a long option position is bound to lose money. This is because it buys the asset *after* the price has gone up—in other words, too late. Each transaction loses a small amount of money, which will accumulate precisely to the option premium.

A second point is that these automatic trading systems, if applied on large scale, have the potential to be destabilizing. Selling on a downturn in price can

exacerbate the downside move. Some have argued that the crash of 1987 was due to the large-scale selling of portfolio insurers in a falling market. These portfolio insurers were, in effect, replicating a long position in puts, blindly selling when the market was falling.[4]

A third point is that this pattern of selling an asset after its price went down is similar to prudent risk-management practices. Typically, traders must cut down their positions after they incur large losses. This is similar to decreasing Δ when S drops. Thus, loss-limit policies bear some resemblance to a long position in an option.

Finally, the success of this replication strategy critically hinges on the assumption of a continuous GBM price process. With this process, it is theoretically possible to rebalance the portfolio as often as needed. In practice, the replication may fail if prices experience drastic jumps.

13.3.3 Distribution of Option Payoffs

Unlike linear derivatives such as forwards and futures, payoffs on options are intrinsically asymmetric. This is not necessarily because of the distribution of the underlying factor, which is often symmetric, but rather is due to the exposure profile. Long positions in options, whether calls or puts, have positive gamma, positive skewness, or long right tails. In contrast, short positions in options are short gamma and hence have negative skewness or long left tails. This is illustrated in Figure 13.10.

Long option:
long gamma,
long right tail

Short option:
short gamma,
long left tail
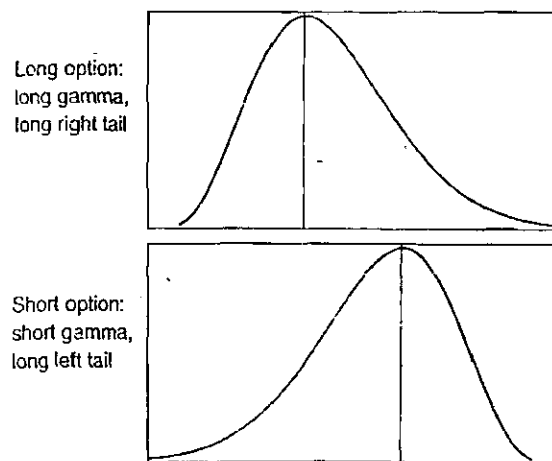
FIGURE 13.10   Distributions of Payoffs on Long and Short Options

---

[4] The exact role of portfolio insurance, however, is still hotly debated. Others have argued that the crash was aggravated by a breakdown in market structures, i.e., the additional uncertainty due to the inability of the stock exchanges to handle abnormal trading volumes.

We now summarize VAR formulas for simple option positions. Assuming a normal distribution, the VAR of the underlying asset is

$$VAR(dS) = \alpha S \sigma(dS/S) \tag{13.28}$$

where α corresponds to the desired confidence level, e.g. $\alpha = 1.645$ for a 95% confidence level. The linear VAR for an option is

$$VAR_1(dc) = \Delta \times VAR(dS) \tag{13.29}$$

The quadratic VAR for an option is

$$VAR_2(dc) = \Delta \times VAR(dS) - \frac{1}{2}\Gamma \times VAR(dS)^2 \tag{13.30}$$

Long option positions have positive gammas and hence lower risk than using a linear model. Conversely, negative gammas translate into higher VARs.

Lest we think that such options require sophisticated risk management methods, what matters is the *extent* of nonlinearity. Figure 13.11 illustrates the risk of a call option with a maturity of three months. It shows that the degree of nonlinearity also depends on the horizon. With a VAR horizon of two weeks, the range

Option value

Distribution of
option values

Spot price

2 weeks

2 months

Distribution of
spot prices
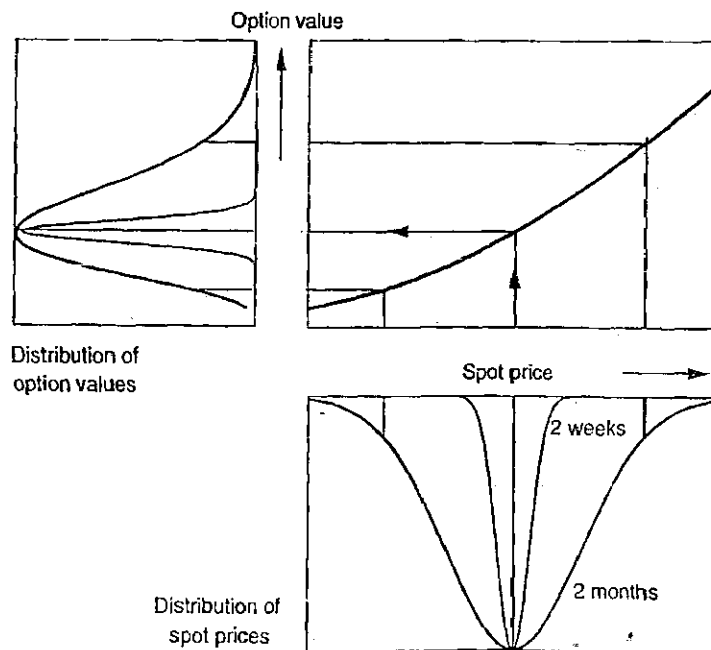
FIGURE 13.11   Skewness and VAR Horizon

of possible values for S is quite narrow. If S follows a normal distribution, the option value will be approximately normal. However, if the VAR horizon is set at two months, the nonlinearities in the exposure combine with the greater range of price movements to create a heavily skewed distribution.

So, for plain-vanilla options, the linear approximation may be adequate as long as the VAR horizon is kept short. For more exotic options, or longer VAR horizons, risk managers must account for nonlinearities.

EXAMPLE 13.12: FRM EXAM 2006—QUESTION 31

You are implementing a portfolio insurance strategy using index futures designed to protect the value of a portfolio of stocks not paying any dividends. Assuming the value of your stock portfolio decreases, which trade would you make to protect your portfolio?

a. Buy an amount of index futures equivalent to the change in the call delta times the original portfolio value.
b. Sell an amount of index futures equivalent to the change in the call delta times the original portfolio value.
c. Buy an amount of index futures equivalent to the change in the put delta times the original portfolio value.
d. Sell an amount of index futures equivalent to the change in the put delta times the original portfolio value.

EXAMPLE 13.13: FRM EXAM 2000—QUESTION 97

A trader buys an at-the-money call option with the intention of delta-hedging it to maturity. Which one of the following is likely to be the most profitable over the life of the option?

a. An increase in implied volatility
b. The underlying price steadily rising over the life of the option
c. The underlying price steadily decreasing over the life of the option
d. The underlying price drifting back and forth around the strike over the life of the option

EXAMPLE 13.14: FRM EXAM 2004—QUESTION 26

A non-dividend-paying stock has a current price of $100 per share. You have just sold a six-month European call option contract on 100 shares of this stock at a strike price of $101 per share. You want to implement a dynamic delta hedging scheme to hedge the risk of having sold the option. The option has a delta of 0.50. You believe that delta would fall to 0.44 if the stock price falls to $99 per share. Identify what action you should take now (i.e., when you have just written the option contract) to make your position delta neutral. After the option is written, if the stock price falls to $99 per share, identify what action should be taken at that time; i.e., later, to rebalance your delta-hedged position.

a. Now: buy 50 shares of stock; later: buy 6 shares of stock.
b. Now: buy 50 shares of stock; later: sell 6 shares of stock.
c. Now: sell 50 shares of stock; later: buy 6 shares of stock.
d. Now: sell 50 shares of stock; later: sell 6 shares of stock.

EXAMPLE 13.15: FRM EXAM 2005—QUESTION 130

An option on the Bovespa stock index is struck on 3,000 Brazilian Real (BRL). The delta of the option is 0.6, and the annual volatility of the index is 24%. Using delta-normal assumptions, what is the 10-day VAR at the 95% confidence level? Assume 260 days per year.

a. 44 BRL
b. 139 BRL
c. 2240 BRL
d. 278 BRL

13.4 IMPORTANT FORMULAS

Black–Scholes option pricing model: $c = Se^{-r^*\tau}N(d_1) - Ke^{-r\tau}N(d_2)$
Taylor series expansion:

$$df = \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}dS^2 + \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial r^*}dr^* + \frac{\partial f}{\partial \sigma}d\sigma + \frac{\partial f}{\partial \tau}d\tau + \cdots$$

$$df = \Delta dS + \frac{1}{2}\Gamma dS^2 + \rho dr + \rho^* dr^* + \Lambda d\sigma + \Theta d\tau + \cdots$$

Delta: $\Delta_c = \frac{\partial c}{\partial S} = e^{-r^*\tau}N(d_1)$, $\Delta_p = \frac{\partial p}{\partial S} = e^{-r^*\tau}[N(d_1) - 1]$

Gamma (for calls and puts): $\Gamma = \frac{\partial^2 c}{\partial S^2} = \frac{e^{-r^*\tau}}{S\sigma\sqrt{\tau}}\Phi(d_1)$

Vega (for calls and puts): $\Lambda = \frac{\partial c}{\partial \sigma} = Se^{-r^*\tau}\sqrt{\tau}\,\Phi(d_1)$

	Long Call			Long Put		
	OTM	ATM	ITM	ITM	ATM	OTM
Δ	$\to 0$	0.5	$\to 1$	$\to -1$	-0.5	$\to 0$
Γ	Low	High, > 0 esp. short-term	Low	Low	High, > 0 esp. short-term	Low
Λ	Low	High, > 0 esp. long-term	Low	Low	High, > 0 esp. long-term	Low
Θ	Low	High, < 0 esp. short-term	Low	Low	High, < 0 esp. short-term	Low

Black–Scholes PDE: $(r - y)\Delta S + \frac{1}{2}\Gamma\sigma^2 S^2 + \Theta = rf$

Linear VAR for an option: $\text{VAR}_1(dc) = \Delta \times \text{VAR}(dS)$

Quadratic VAR for an option: $\text{VAR}_2(dc) = \Delta \times \text{VAR}(dS) - \frac{1}{2}\Gamma \times \text{VAR}(dS)^2$

13.5 ANSWERS TO CHAPTER EXAMPLES

Example 13.1: FRM Exam 2006—Question 91

a. The delta of a long forward contract is $e^{-r^*\tau} = \exp(-0.10 \times 0.5) = 0.95$.

Example 13.2: FRM Exam 2004—Question 21

b. The option is ATM because the strike price is close to the spot price. This is a put, so the delta must be close to -0.5.

Example 13.3: FRM Exam 2006—Question 80

c. This is a call option, so delta must be positive. This is given by $\Delta = \exp(-r^*\tau)$ $N(d_1) = \exp(-0.01 \times 2) \times 0.64 = 0.63$.

Example 13.4: FRM Exam 2003—Question 94

a. Gamma is highest for short-term ATM options. The first answer has a strike price close to $S = 78$ and short maturity.

Example 13.5: FRM Exam 2001—Question 79

a. This is an at-the-money option with a delta of about 0.5. Since the bank sold calls, it needs to delta-hedge by buying the shares. With a delta of 0.54, it would need to buy approximately 50,000 shares. Answer a. is the closest. Note that most other information is superfluous.

Example 13.6: FRM Exam 2005—Question 106

a. Because gamma is negative, we need to buy a call to increase the portfolio gamma back to zero. The number is $600/1.5 = 400$ calls. This, however, will increase the delta from zero to $400 \times 0.75 = 300$. Hence, we must sell 300 shares to bring back the delta to zero. Note that positions in shares have zero gamma.

Example 13.7: FRM Exam 2004—Question 65

c. Theta is negative for long positions in ATM options, so a. is incorrect. Gamma is small for ITM options, so b. is incorrect. Delta of ITM puts tends to -1, so d. is incorrect.

Example 13.8: FRM Exam 2006—Question 33

d. Long gamma means that the portfolio is long options with high gamma, typically short-term (short-expiry) ATM options. Short vega means that the portfolio is short options with high vega, typically long-term (long-expiry) ATM options.

Example 13.9: FRM Exam 2006—Question 54

d. Vega is highest for ATM European options, so answer a. is correct. Delta is negative and moves to zero as S increases, so answer b. is correct. Gamma increases as the maturity of an ATM option decreases, so answer c. is correct. Theta is greater (in absolute value) for short-term ATM options, so statement d. is incorrect.

Example 13.10: FRM Exam 2000—Question 76

a. Long positions in options have positive gamma and vega. Gamma (or instability in delta) increases near maturity; vega decreases near maturity. So, to obtain positive gamma and negative vega, we need to buy short-maturity options and sell long-maturity options.

Example 13.11: FRM Exam 2001—Question 113

a. Such a portfolio is short vega (volatility) and short theta (time). We need to implement a hedge that is delta-neutral and involves buying and selling options with different maturities. Long positions in short-dated options have high negative theta and low positive vega. Hedging can be achieved by selling short-term options and buying long-term options.

Example 13.12: FRM Exam 2006—Question 31

d. Portfolio insurance is a form of dynamic hedging that replicates a long position in a put option. If the value of the portfolio decreases, one should sell the index futures in the amount that represents the change in the put delta.

Example 13.13: FRM Exam 2000—Question 97

d. An important aspect of the question is the fact that the option is held to maturity. Answer a. is incorrect because changes in the implied volatility would change the value of the option, but this has no effect when holding to maturity. The profit from the dynamic portfolio will depend on whether the actual volatility differs from the initial implied volatility. It does not depend on whether the option ends up in-the-money or not, so answers b. and c. are incorrect. The portfolio will be profitable if the actual volatility is small, which implies small moves around the strike price.

Example 13.14: FRM Exam 2004—Question 26

b. The dynamic hedge should replicate a long position in the call. Due to the positive delta, this implies a long position of $\Delta \times 100 = 50$ shares. If the delta falls, the position needs to be adjusted by selling $(0.5 - 0.44) \times 100 = 6$ shares.

Example 13.15: FRM Exam 2005—Question 130

b. The linear VAR is derived from the worst move in the index value, which is $\alpha S \sigma \sqrt{T} = 1.645 \times 3,000(24\%/\sqrt{260})\sqrt{10} = 232.3$. Multiplying by the delta of 0.6 gives 139.
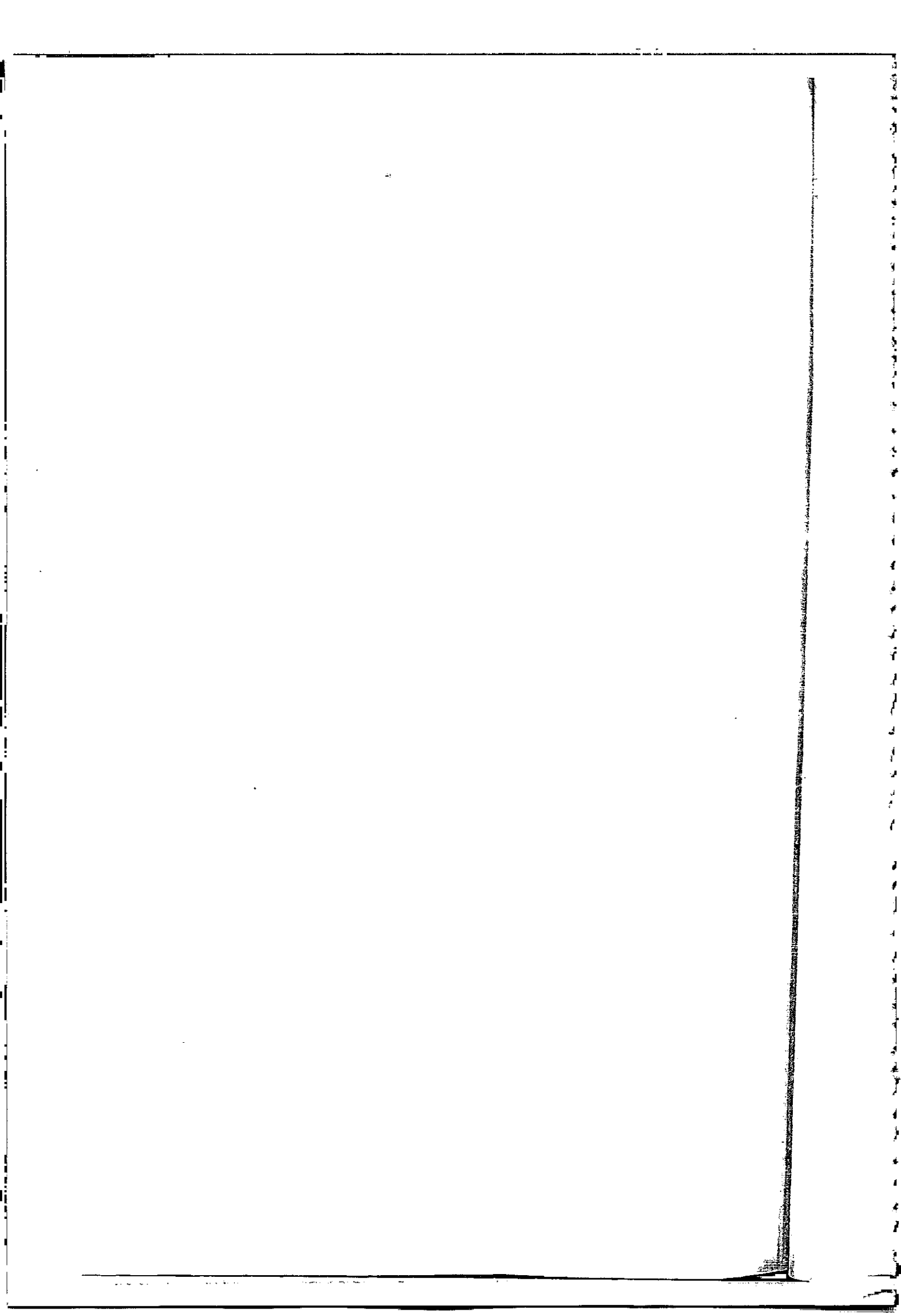
CHAPTER 14

Modeling Risk Factors

We now turn to an analysis of the distribution of risk factors used in risk mea-surement. A previous chapter has described the major risk factors, including fixed-income, equity, currency, and commodity price risk. The emphasis was on the volatility as a measure of dispersion. More generally, risk managers need to consider the whole shape of the distribution, which is not necessarily normal, as well as potential time variation in this distribution.

In fact, most financial time series are characterized by fatter tails than the normal distribution. In addition, there is ample empirical evidence that, over short horizons, risk changes in a predictable fashion. This time variation could potentially explain the observed high frequency of extreme observations, which could be drawn from distributions with temporarily higher volatility.

Section 14.1 starts by describing the normal distribution. We compare the normal and lognormal distributions and explain why this choice is so popular. A major failing of this distribution, however, is its inability to represent the fre-quency of large observations found in financial data. Section 14.2 discusses other distributions that have fatter tails than the normal.

Section 14.3 then turns to time-variation in risk. We describe the generalized autoregressive conditional heteroskedastic (GARCH) model and a special case, which is RiskMetrics' exponentially weighted moving average (EWMA).

14.1 NORMAL AND LOGNORMAL DISTRIBUTIONS

14.1.1 Why the Normal?

The normal, or Gaussian, distribution is usually the first choice when modeling asset returns. This distribution plays a special role in statistics, as it is easy to handle and is stable under addition, meaning that a sum of normal variables is itself normal. It also provides the limiting distribution of the average of *independent* random variables (through the central limit theorem).

Empirically, the normal distribution provides a rough, first-order approxima-tion to the distribution of many random variables: rates of changes in currency prices, rates of changes in stock prices, rates of changes in bond prices, changes in yields, and rates of changes in commodity prices. All of these are characterized

by many cases of small moves and fewer cases of large moves, which provides a rationale for the bell-shaped, normal distribution.

14.1.2 Computing Returns

In what follows, the random variable is the new price P_1, given the current price P_0. Defining $r = (P_1 - P_0)/P_0$ as the rate of return in the price, we can start with the assumption is that this random variable is drawn from a normal distribution,

$$r \sim \Phi(\mu, \sigma) \tag{14.1}$$

with some mean μ and standard deviation σ. Turning to prices, we have $P_1 = P_0(1 + r)$ and

$$P_1 \sim P_0 + \Phi(P_0\mu, P_0\sigma) \tag{14.2}$$

For instance, starting from a stock price of $100, if $\mu = 0\%$ and $\sigma = 15\%$, we have $P_1 \sim \$100 + \Phi(\$0, \$15)$.

In this case, however, the normal distribution cannot even be theoretically correct. Because of limited liability, stock prices cannot go below zero. Similarly, commodity prices and yields cannot turn negative. This is why another popular distribution is the lognormal distribution, which is such that

$$R = \ln(P_1/P_0) \sim \Phi(\mu, \sigma) \tag{14.3}$$

By taking the logarithm, the price is given by $P_1 = P_0 \exp(R)$, which precludes prices from turning negative as the exponential function is always positive. Figure 14.1 compares the normal and lognormal distributions over a one-year
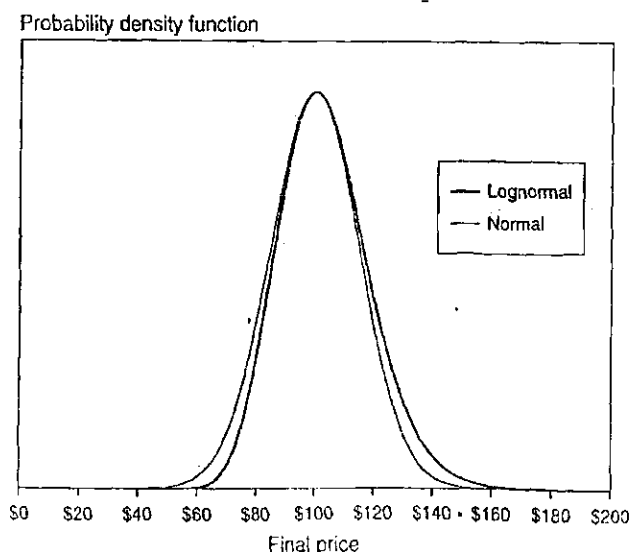
FIGURE 14.1 Normal and Lognormal Distributions—Annual Horizon

horizon with $\sigma = 15\%$ annually. The distributions are very similar, except for the tails. The lognormal is skewed to the right.

The difference between the two distributions is driven by the size of the volatility parameter over the horizon. Small values of this parameter imply that the distributions are virtually identical. This can happen either when the asset is not very risky, that is, when the annual volatility is small, or when the horizon is very short. In this situation, there is very little chance of prices turning negative. The limited liability constraint is not important.

KEY CONCEPT

The normal and lognormal distributions are very similar for short horizons or low volatilities.

As an example, Table 14.1 compares the computation of returns over a one-day and one-year horizon. The one-day returns are 1.000% and 0.995% for discrete and log-returns, respectively, which translates into a relative difference of 0.5%, which is minor. In contrast, the difference is more significant over longer horizons.

14.1.3 Time Aggregation

Longer horizons can be accommodated assuming a constant lognormal distribution across horizons. Over two periods, for instance, the price movement can be described as the sum of the price movements over each day:

$$R_{t,2} = \ln(P_t/P_{t-2}) = \ln(P_t/P_{t-1}) + \ln(P_{t-1}/P_{t-2}) = R_{t-1} + R_t \qquad (14.4)$$

More generally, define T as the number of steps. If returns are identically and independently distributed (i.i.d.), the variance of multiple-period returns is

$$V[R(0,T)] = V[R(0,1)] + V[R(1,2)] + \cdots + V[R(T-1,T)] = V[R(0,1)]T \qquad (14.5)$$

TABLE 14.1 Comparison between Discrete and Log Returns

	Daily	Annual
Initial Price	100	100
Ending Price	101	115
Discrete Return	1.0000	15.0000
Log Return	0.9950	13.9762
Relative Difference	0.50%	7.33%

since the variances are all the same and all the covariance terms are zero because of the independence assumption. Similarly, the mean of multiple-period returns is

$$E[R(0, T)] = E[R(0, 1)] + E[R(1, 2)] + \cdots + E[R(T - 1, T)] = E[R(0, 1)]T$$

$$(14.6)$$

assuming expected returns are the same for each day.

Thus the multiple-period volatility is

$$\sigma_T = \sigma\sqrt{T} \qquad (14.7)$$

If the distribution is stable under addition, i.e., we can use the same multiplier for a one-period and T-period return, we have a multiple-period VAR of

$$\text{VAR} = \alpha(\sigma\sqrt{T})W \qquad (14.8)$$

In other words, extension to a multiple period follows a square root of time rule. Figure 14.2 shows how VAR grows with the length of the horizon and for various confidence levels. This is scaled to an annual standard deviation of 1, which is a 84.1% VAR. The figure shows that VAR increases more slowly than time. The one-month 99% VAR is 0.67, but increases only to 2.33 at a one-year horizon.

In summary, the square root of time rule applies under the following conditions:

- The distribution is the same at each period (i.e., there is no predictable time variation in expected return nor in risk).
- Returns are uncorrelated/independent across each period, so that all covariances terms disappear.
- The distribution is the same for one- or T-period, or is stable under addition, such as the normal.
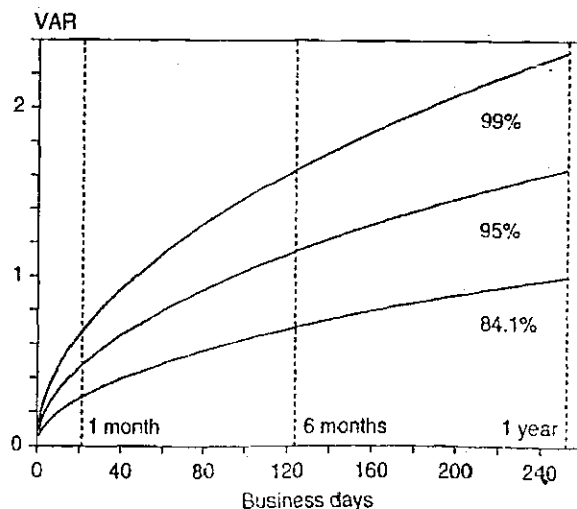
FIGURE 14.2   VAR at Increasing Horizons

If returns are not independent, we may be able to characterize the risk in some cases. For instance, when returns follow a first-order autoregressive process,

$$R_t = \rho R_{t-1} + u_t \tag{14.9}$$

we can write the variance of two-day returns as

$$V[R_t + R_{t-1}] = V[R_t] + V[R_{t-1}] + 2\text{Cov}[R_t, R_{t-1}] = \sigma^2 + \sigma^2 + 2\rho\sigma^2 \tag{14.10}$$

or

$$V[R_t + R_{t-1}] = \sigma^2 \times 2[1 + \rho] \tag{14.11}$$

A positive value for ρ describes a situation where a movement in one direction is likely to be followed by another in the same direction. This implies that markets are trending. In this case, the longer-term volatility increases faster than with the usual square root of time rule.

On the other hand, a negative value for ρ describes a situation where a movement in one direction is likely to be reversed later. In this mean-reversion case, the longer-term volatility increases more slowly than with the usual square root of time rule.

EXAMPLE 14.1: TIME SCALING

Consider a portfolio with a one-day VAR of $1 million. Assume that the market is trending with an autocorrelation of 0.1. Under this scenario, what would you expect the two-day VAR to be?

a. $2 million
b. $1.414 million
c. $1.483 million
d. $1.449 million

14.2 FAT TAILS

Perhaps the most serious problem with the normal distribution is the fact that its tails "disappear" too fast, at least faster than what is empirically observed in financial data. We typically observe that every market experiences one or more daily moves of 4 standard deviations or more per year. Such frequency is incompatible with a normal distribution. With a normal distribution, the probability of this happening is 0.0032% for one day, which implies a frequency of once every 125 years.

KEY RULE OF THUMB

Every financial market experiences one or more daily price moves of 4 standard deviations or more each year. And in any year, there is usually at least one market that has a daily move greater than 10 standard deviations.

This empirical observation can be explained in a number of ways: (1) The true distribution has fatter tails (e.g., the Student's t), or (2) the observations are drawn from a mix of distributions (e.g., a mix of two normals, one with low risk, the other with high risk), or (3) the distribution is nonstationary.

The first explanation is certainly a possibility. Figure 14.3 displays the density function of the normal and Student's t distribution, with four and six degrees of freedom (df). The student density has fatter tails, which better reflect the occurrences of extreme observations in empirical financial data.

The distributions are further compared in Table 14.2. The left-side panel reports the tail probability of an observation lower than the deviate. For instance, the probability of observing a draw less than -3 is 0.001, or 0.1% for the normal, 0.012 for the Student's t with six degrees of freedom, and 0.020 for the Student's t with four degrees of freedom. There is a greater probability of observing an extreme move when the data is drawn from a Student's t rather than a normal distribution.

We can transform these into an expected number of occurrences in one year, or 250 business days. The right-side panel shows that the corresponding numbers are 0.34, 3.00, and 4.99 for the respective distributions. In other words, with a normal distribution, we should expect that this extreme movement below $z = -3$ will occur on average one day or less. With a Student's t with $df = 4$, the expected number is five in a year, which is closer to reality.

Probability density function
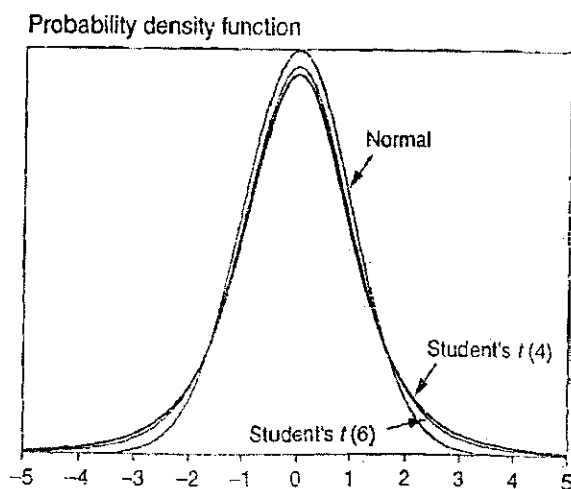
FIGURE 14.3   Normal and Student Distributions

TABLE 14.2 Comparison of the Normal and Student's t Distributions

Deviate	Tail probability			Expected Number in 250 days		
	Normal	t df = 6	t df = 4	Normal	t df = 6	t df = 4
−5	0.00000	0.00123	0.00375	0.00	0.31	0.94
−4	0.00003	0.00356	0.00807	0.01	0.89	2.02
−3	0.00135	0.01200	0.01997	0.34	3.00	4.99
−2	0.02275	0.04621	0.05806	5.69	11.55	14.51
−1	0.15866	0.17796	0.18695	39.66	44.49	46.74
					Deviate (alpha)	
Probability = 1%				2.33	3.14	3.75
Ratio to normal				1.00	1.35	1.61

The bottom panel reports the deviate that corresponds to a 99% right-tail confidence level, or 1% left tail. For the normal distribution, this is the usual 2.33. For the Student's t with df = 4, α is 3.75, much higher. The ratio of the two is 1.61. Thus a rule of thumb would be to correct the VAR measure from a normal distribution by a ratio of 1.61 to achieve the desired coverage in the presence of fat tails. More generally, this explains why "safety factors" are used to multiply VAR measures, such as the Basel multiplicative factor of three.

14.3 TIME-VARIATION IN RISK

An alternative class of explanation is that financial data can be viewed as drawn from a normal distribution with time-varying parameters. This is useful only if this time variation has some predictability.

14.3.1 GARCH

A specification that has proved quite successful in practice is the generalized autoregressive conditional heteroskedastic (GARCH) model developed by Engle (1982) and Bollerslev (1986).

This class of models assumes that the return at time t has a normal distribution, for example, conditional on parameters μ_t and σ_t:

$$r_t \sim \Phi(\mu_t, \sigma_t) \tag{14.12}$$

The important point is that σ is indexed by time. In this context, we define the conditional variance as that conditional on current information. This may differ from the unconditional variance, which is the same for the whole sample. Thus the average variance is unconditional, whereas a time-varying variance is conditional.

There is substantial empirical evidence that conditional volatility models successfully forecast risk. The general assumption is that the conditional returns have

a normal distribution, although this could be extended to other distributions such as the Student's t.

The GARCH model assumes that the conditional variance depends on the latest innovation, and on the previous conditional variance. Define $h_t = \sigma_t^2$ as the conditional variance, using information up to time $t - 1$, and r_{t-1} as the previous day's return, also called innovation. The simplest such model is the GARCH(1,1) process,

$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta h_{t-1} \qquad (14.13)$$

which involves one lag of the innovation and one lag of the previous forecast. The β term is important because it allows persistence in the variance, which is a realistic feature of the data.

The average, unconditional variance is found by setting $E[r_{t-1}^2] = h_t = h_{t-1} = h$. Solving for h, we find

$$h = \frac{\alpha_0}{1 - \alpha_1 - \beta} \qquad (14.14)$$

This model is stationary when the sum of parameters $\gamma = \alpha_1 + \beta$ are less than unity. This sum is also called the Persistence, as it defines the speed at which shocks to the variance revert to their long-run values.

To understand how the process works, consider Table 14.3. The parameters are $\alpha_0 = 0.01$, $\alpha_1 = 0.03$, $\beta = 0.95$. The unconditional variance is $0.01/(1 - 0.03 - 0.95) = 0.7$ daily, which is typical of a currency series, as it translates into an annualized volatility of 11%. The process is stationary because $\alpha_1 + \beta = 0.98 < 1$.

At time 0, we start with the variance at $h_0 = 1.1$ (expressed in percent squared). The conditional volatility is $\sqrt{h_0} = 1.05\%$. The next day, there is a large return of 3%. The new variance forecast is then $h_1 = 0.01 + 0.03 \times 3^2 + 0.95 \times 1.1 = 1.32$. The conditional volatility just went up to 1.15%. If nothing happens the following days, the next variance forecast is $h_2 = 0.01 + 0.03 \times 0^2 + 0.95 \times 1.32 = 1.27$. And so on.

TABLE 14.3 Building a GARCH Forecast

Time	Return	Conditional Variance	Conditional Risk	Conditional 95% Limit
$t-1$	r_{t-1}	h_t	$\sqrt{h_t}$	$2\sqrt{h_t}$
0	0.0	1.10	1.05	± 2.10
1	3.0	1.32	1.15	± 2.30
2	0.0	1.27	1.13	± 2.25
3	0.0	1.22	1.10	± 2.20

The GARCH process can be extrapolated to later days. For the next-day forecast,

$$E_{t-1}(r_{t+1}^2) = \alpha_0 + \alpha_1 E_{t-1}(r_t^2) + \beta h_t = \alpha_0 + \alpha_1 h_t + \beta h_t = \alpha_0 + \gamma h_t$$

For the following day,

$$E_{t-1}(r_{t+2}^2) = \alpha_0 + \alpha_1 E_{t-1}(r_{t+1}^2) + \beta E_{t-1}(h_{t+1}) = \alpha_0 + (\alpha_1 + \beta)E_{t-1}(r_{t+1}^2)$$

$$E_{t-1}(r_{t+2}^2) = \alpha_0 + \gamma(\alpha_0 + \gamma h_t)$$

Generally,

$$E_{t-1}(r_{t+n}^2) = \alpha_0(1 + \gamma + \gamma^2 + \cdots + \gamma^{n-1}) + \gamma^n h_t$$

Figure 14.4 illustrates the dynamics of shocks to a GARCH process for various values of the persistence parameter. As the conditional variance deviates from the starting value, it slowly reverts to the long-run value at a speed determined by $\alpha_1 + \beta$.

Note that these are forecasts of one-day variances at forward points in time. The total variance over the horizon is the sum of one-day variances. The *average* variance is marked with a black rectangle on the graph.

The graph also shows why the square root of time rule for extrapolating returns does not apply when risk is time-varying. If the initial value of the variance is greater than the long-run average, simply extrapolating the one-day variance to a longer horizon will overstate the average variance. Conversely, starting from a lower value and applying the square root of time rule will understate risk.
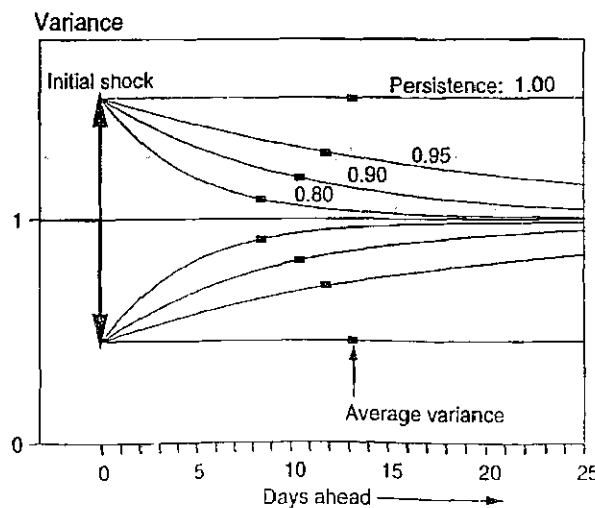
FIGURE 14.4 Shocks to a GARCH Process

KEY CONCEPT

The square root of time rule used to scale one-day returns into longer horizons is generally inappropriate when risk is time-varying.

EXAMPLE 14.2: FRM EXAM 2006—QUESTION 36

Which of the following GARCH models will take the shortest time to revert to its long-run value?

a. $h_t = 0.05 + 0.03r_{t-1}^2 + 0.96h_{t-1}$
b. $h_t = 0.03 + 0.02r_{t-1}^2 + 0.95h_{t-1}$
c. $h_t = 0.02 + 0.01r_{t-1}^2 + 0.97h_{t-1}$
d. $h_t = 0.01 + 0.01r_{t-1}^2 + 0.98h_{t-1}$

EXAMPLE 14.3: FRM EXAM 2006—QUESTION 132

Assume you are using a GARCH model to forecast volatility that you use to calculate the one-day VAR. If volatility is mean reverting, what can you say about the T-day VAR?

a. It is less than the $\sqrt{T} \times$ one-day VAR.
b. It is equal to $\sqrt{T} \times$ one-day VAR.
c. It is greater than the $\sqrt{T} \times$ one-day VAR.
d. It could be greater or less than the $\sqrt{T} \times$ one-day VAR.

EXAMPLE 14.4: FRM EXAM 2007—QUESTION 34

A risk manager estimates daily variance h_t using a GARCH model on daily returns r_t: $h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta h_{t-1}$, with $\alpha_0 = 0.005$, $\alpha_1 = 0.04$, $\beta = 0.94$. The long-run *annualized* volatility is approximately

a. 13.54%
b. 7.94%
c. 72.72%
d. 25.00%

14.3.2 EWMA

The RiskMetrics approach is a specific case of the GARCH process and is particularly simple and convenient to use. Variances are modeled using an exponentially weighted moving average (EWMA) forecast. The forecast is a weighted average of the previous forecast, with weight λ, and of the latest squared innovation, with weight $(1 - \lambda)$:

$$h_t = \lambda h_{t-1} + (1 - \lambda)r_{t-1}^2 \qquad (14.15)$$

The λ parameter, also called the decay factor, determines the relative weights placed on previous observations. The EWMA model places geometrically declining weights on past observations, assigning greater importance to recent observations. By recursively replacing h_{t-1} in Equation (14.15), we have

$$h_t = (1 - \lambda)[r_{t-1}^2 + \lambda r_{t-2}^2 + \lambda^2 r_{t-3}^2 + \cdots] \qquad (14.16)$$

The weights therefore decrease at a geometric rate. The lower λ, the more quickly older observations are forgotten. RiskMetrics has chosen $\lambda = 0.94$ for daily data and $\lambda = 0.97$ for monthly data.

Table 14.4 shows how to build the EWMA forecast using a parameter of $\lambda = 0.95$, which is consistent with the previous GARCH example. At time 0, we start with the variance at $h_0 = 1.1$, as before. The next day, we have a return of 3%. The new variance forecast is then $h_1 = 0.05 \times 3^2 + 0.95 \times 1.1 = 1.50$. The next day, this moves to $h_2 = 0.05 \times 0^2 + 0.95 \times 1.50 = 1.42$. And so on.

This model is a special case of the GARCH process, where α_0 is set to 0, and α_1 and β sum to unity. The model therefore has permanent persistence. Shocks to the volatility do not decay, as shown in Figure 14.4 when the persistence is 1.00. Thus longer-term extrapolation from the GARCH and EWMA models may give quite different forecasts. Over a one-day horizon, however, the two models are quite similar and often indistinguishable from each other.

Figure 14.5 displays the pattern of weights for previous observations. With $\lambda = 0.94$, the weights decay quickly. The weight on the last day is $(1 - \lambda) = (1 - 0.94) = 0.06$. The weight on the previous day is $(1 - \lambda)\lambda = 0.0564$, and so on. The weight drops below 0.00012 for data more than 100 days old. With $\lambda = 0.97$, the weights start at a lower level but decay more slowly. In comparison, moving average (MA) models have a fixed window, with equal weights within

TABLE 14.4   Building an EWMA Forecast

Time	Return	Conditional Variance	Conditional Risk	Conditional 95% Limit
$t-1$	r_{t-1}	h_t	$\sqrt{h_t}$	$2\sqrt{h_t}$
0	0.0	1.10	1.05	±2.1
1	3.0	1.50	1.22	±2.4
2	0.0	1.42	1.19	±2.4
3	0.0	1.35	1.16	±2.3
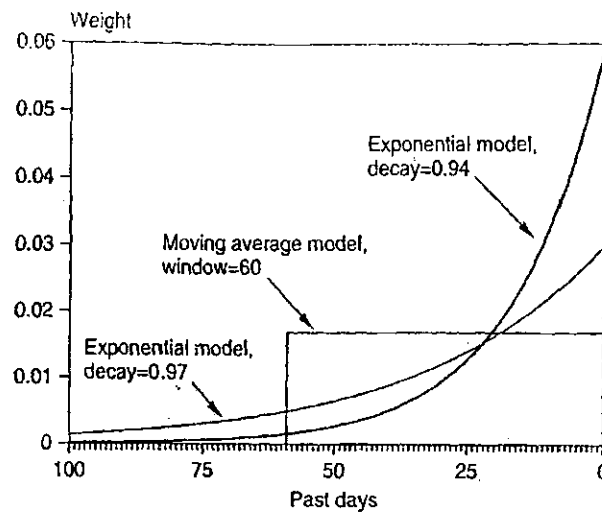
FIGURE 14.5   Weights on Past Observations

the window but otherwise zero. MA models with shorter windows give a greater weight to recent observations. As a result, they are more responsive to current events and more volatile.

EXAMPLE 14.5: FRM EXAM 2002—QUESTION 13

The GARCH model is useful for simulating asset returns. Which of the following statements about this model is *false*?

a. The Exponentially Weighted Moving Average (EWMA) approach of RiskMetrics is a particular case of a GARCH process.
b. The GARCH allows for time-varying volatility.
c. The GARCH can produce fat tails in the return distribution.
d. The GARCH imposes a positive conditional mean return.

EXAMPLE 14.6: FRM EXAM 2007—QUESTION 46

A bank uses the Exponentially Weighted Moving Average (EWMA) technique with λ of 0.9 to model the daily volatility of a security. The current estimate of the daily volatility is 1.5%. The closing price of the security is USD 20 yesterday and USD 18 today. Using continuously compounded returns, what is the updated estimate of the volatility?

a. 3.62%
b. 1.31%
c. 2.96%
d. 5.44%

EXAMPLE 14.7: FRM EXAM 2006—QUESTION 40

Using a daily RiskMetrics EWMA model with a decay factor $\lambda = 0.95$ to develop a forecast of the conditional variance, which weight will be applied to the return that is four days old?

a. 0.000
b. 0.043
c. 0.048
d. 0.950

EXAMPLE 14.8: EFFECT OF WEIGHTS ON OBSERVATIONS

Until January 1999 the historical volatility for the Brazilian real versus the U.S. dollar had been very small for several years. On January 13, Brazil abandoned the defense of the currency peg. Using the data from the close of business on January 13, which of the following methods for calculating volatility would have shown the greatest jump in measured historical volatility?

a. 250 day equal weight
b. Exponentially weighted with a daily decay factor of 0.94
c. 60 day equal weight
d. All of the above

14.3.3 Option Data

All the previous models were based on historical data. Although conditional volatility models are a substantial improvement over models that assume constant risk, they are always, by definition, one step too late.

These models start to react *after* a big shock has occurred. In many situations, this may be too late—hence the quest for forward-looking risk measures.

Such forward-looking measures are contained in option implied standard deviations (ISD). ISD are obtained by, first, assuming an option pricing model and, next, inverting the model, that is, solving for the parameter that will make the model price equal to the observed market price.

Define $f(\cdot)$ as an option pricing function, such as the Black–Scholes model for European options. Normally, we input σ into f along with other parameters and then solve for the option price. However, if the market trades these options and if all the other inputs are observable, we can recover σ_{ISD} by setting the model price equal to the market price:

$$c_{MARKET} = f(\sigma_{ISD}) \tag{14.17}$$

MARKET RISK MANAGEMENT

This assumes that the model firs the data perfectly, which may not be the case for out-of-the-money options. Hence, this method works best for short-term (two weeks to three months) at-the-money options.

This approach can even be generalized to implied correlations. For this, we need triplets of options, such as $/yen, $/euro, yen/euro. The first one can be used to recover σ_1, the second σ_2, and the third the covariance σ_{12}, from which the implied correlation ρ_{12} can be recovered.

There is much empirical evidence that ISD provide superior forecasts of future risk. This is expected, as the essence of option trading is to place volatility bets. The main drawback of these methods is that risk measures recovered from market prices are defined in a risk-neutral space. For forecasting risk, we need actual, real-world, physical distributions. Implied volatility may be systematically higher than forecast volatility due to a risk premium. If this risk premium is stable, however, changes in ISDs should prove informative for predicting changes in risk.

In practice, while historical time-series models can be applied systematically to all series for which we have data, we do not have actively traded options for all risk factors. In addition, we have even fewer combinations of options that permit us to compute implied correlations. Thus, it is difficult to integrate ISD with time-series models.

14.3.4 Implied Distributions

Options can be used to derive more information about future distributions than the volatility alone. Recently, option watchers have observed some inconsistencies in the pricing of options, especially for stock index options. In particular, options that differ only by their strike prices are characterized by different ISDs. Options that are out-of-the-money have higher ISDs than at-the-money options. This has become known as the smile effect in ISDs, which is shown in Figure 14.6, which
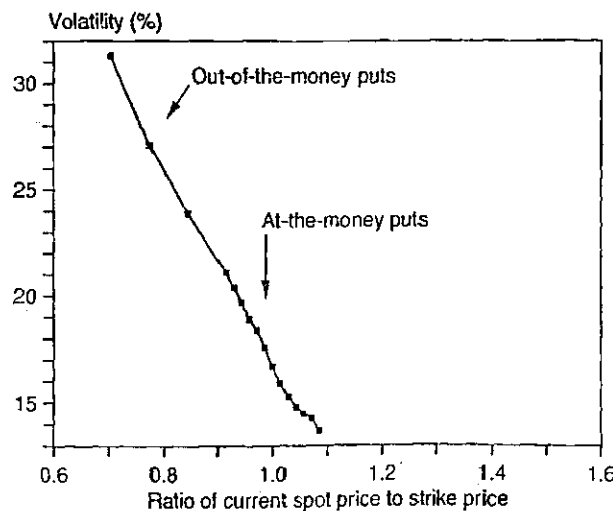
FIGURE 14.6 Smile Effect
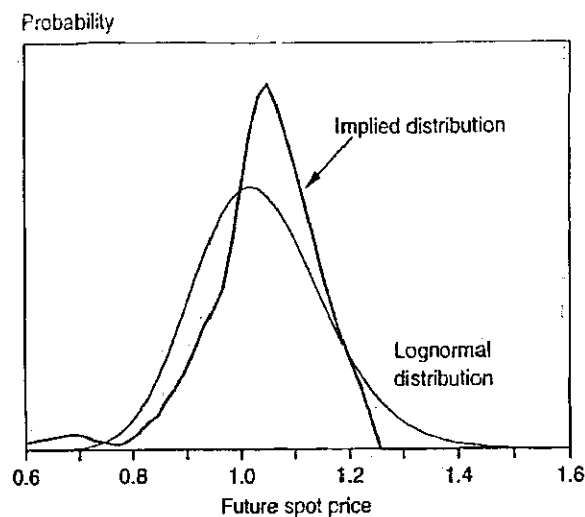
FIGURE 14.7 Implied Distribution

plots equity ISDs against the ratio of the strike price over the current spot price. In this case, the smile is totally asymmetric (more like a smirk).

Low values of the ratio, describing out-of-the-money puts, are associated with high ISDs. In other words, out-of-the-money puts appear overpriced relative to others. Different ISDs are clearly inconsistent with the joint assumption of a lognormal distribution for prices and efficient markets. Perhaps the data are trying to tell a story. This effect became most pronounced after the stock market crash of 1987, raising the possibility that the market expected another crash, although with low probability.

Recently, Rubinstein (1994) has extended the concept of ISD to the whole implied distribution of future prices. By judiciously choosing options with sufficiently spaced strike prices, one can recover the entire implied distribution that is consistent with option prices. This distribution, shown in Figure 14.7, displays a hump for values of the future price 30% below the current price. This hump is nowhere apparent from the usual log-normal distribution.

We can give two interpretations to this result. The first is that the market indeed predicts a small probability of a future crash. The second has to do with the fact that this distribution derived from option prices assumes risk-neutrality, since the Black–Scholes approach assumes that investors are risk neutral. Thus this distribution may differ from the true, objective distribution due to a risk premium. Intuitively, investors may be very averse to a situation where they have to suffer a large fall in the value of their stock portfolios. As a result, they will bid up the price of put options, which is reflected in a higher than otherwise implied volatility.

This is currently an area of active research. The consensus, however, is that options should contain valuable information about future distributions since, after all, option traders bet good money on their forecasts.

EXAMPLE 14.9: FRM EXAM 2006—QUESTION 29

Risk-neutral default probability and real-world (or physical) default probability are used in the analysis of credit risk. Which one of the following statements on their uses is correct?

a. Real-world default probability should be used in scenario analyses of potential future losses from defaults, and real-world default probability should also be used in valuing credit derivatives.
b. Real-world default probability should be used in scenario analyses of potential future losses from defaults, but risk-neutral default probability should be used in valuing credit derivatives.
c. Risk-neutral default probability should be used in scenario analyses of potential future losses from defaults, and risk-neutral default probability should be used in valuing credit derivatives.
d. Risk-neutral default probability should be used in scenario analyses of potential future losses from defaults, but real-world default probability should be used in valuing credit derivatives.

14.4 IMPORTANT FORMULAS

VAR assuming i.i.d. returns: $VAR = \alpha(\sigma\sqrt{T})W$

Time aggregation of variance with nonzero autocorrelation:
$$V[R_t + R_{t-1}] = \sigma^2 \times 2[1 + \rho]$$

GARCH process: $h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta h_{t-1}$

GARCH long-run mean: $h = \frac{\alpha_0}{(1-\alpha_1-\beta)}$

EWMA process: $h_t = \lambda h_{t-1} + (1-\lambda)r_{t-1}^2$

14.5 ANSWERS TO CHAPTER EXAMPLES

Example 14.1: Time Scaling

c. Knowing that the variance is $V(2\text{-day}) = V(1\text{-day})\ [2 + 2\rho]$, we find $VAR(2\text{-day}) = VAR(1\text{-day})\sqrt{2 + 2\rho} = \$1\sqrt{2 + 0.2} = \$1.483$, assuming the same distribution for the different horizons.

Example 14.2: FRM Exam 2006—Question 36

b. The persistence parameter $\alpha_1 + \beta$ is, respectively, $0.99, 0.97, 0.98, 0.99$. Model b. has the lowest parameter and hence will revert the fastest to the mean.

Example 14.3: FRM Exam 2006—Question 132

d. If the initial volatility were equal to the long-run volatility, then the T-day VAR could be computed using the square root of time rule, assuming normal distributions. If the starting volatility were higher, then the T-day VAR should be less than the $\sqrt{T}$ × one-day VAR. Conversely if the starting volatility were lower than the long-run value. However, the question does not indicate the starting point. Hence, answer d. is correct.

Example 14.4: FRM Exam 2007—Question 34

b. The long-run mean variance is $h = \alpha_0/(1 - \alpha_1 - \beta) = 0.006/(1 - 0.04 - 0.94) = 0.25$. Taking the square root, this gives 0.5 for daily volatility. Multiplying by $\sqrt{252}$, we have an annualized volatility of 7.937%.

Example 14.5: FRM Exam 2002—Question 13

d. The GARCH model allows for time variation in volatility and includes the EWMA model as a special case. It can also induce fat tails in the return distribution, but says nothing about the mean, so answer d. is false.

Example 14.6: FRM Exam 2007—Question 46

a. The log return is $\ln(18/20) = -10.54\%$. The new variance forecasts is $h = 0.90 \times (1.5^2) + (1 - 0.90) \times 10.54^2 = 0.001313$, or taking the square root, 3.62%.

Example 14.7: FRM Exam 2006—Question 40

b. The weight of the last day is $(1 - 0.95) = 0.050$. For the day before, this is 0.05×0.95, and for four days ago, $0.05 \times 0.95^3 = 0.04287$.

Example 14.8: Effect of Weights on Observations

b. The EWMA model puts a weight of 0.06 on the latest observation, which is higher than the weight of $(1/60) = 0.0167$ for the 60-day MA and $(1/250) = 0.004$ for the 250-day MA.

Example 14.9: FRM Exam 2007—Question 29

b. Real-world probabilities should be used for risk management, or to devise scenarios. In contrast, risk-neutral probabilities should be used to price assets, such as credit derivatives.
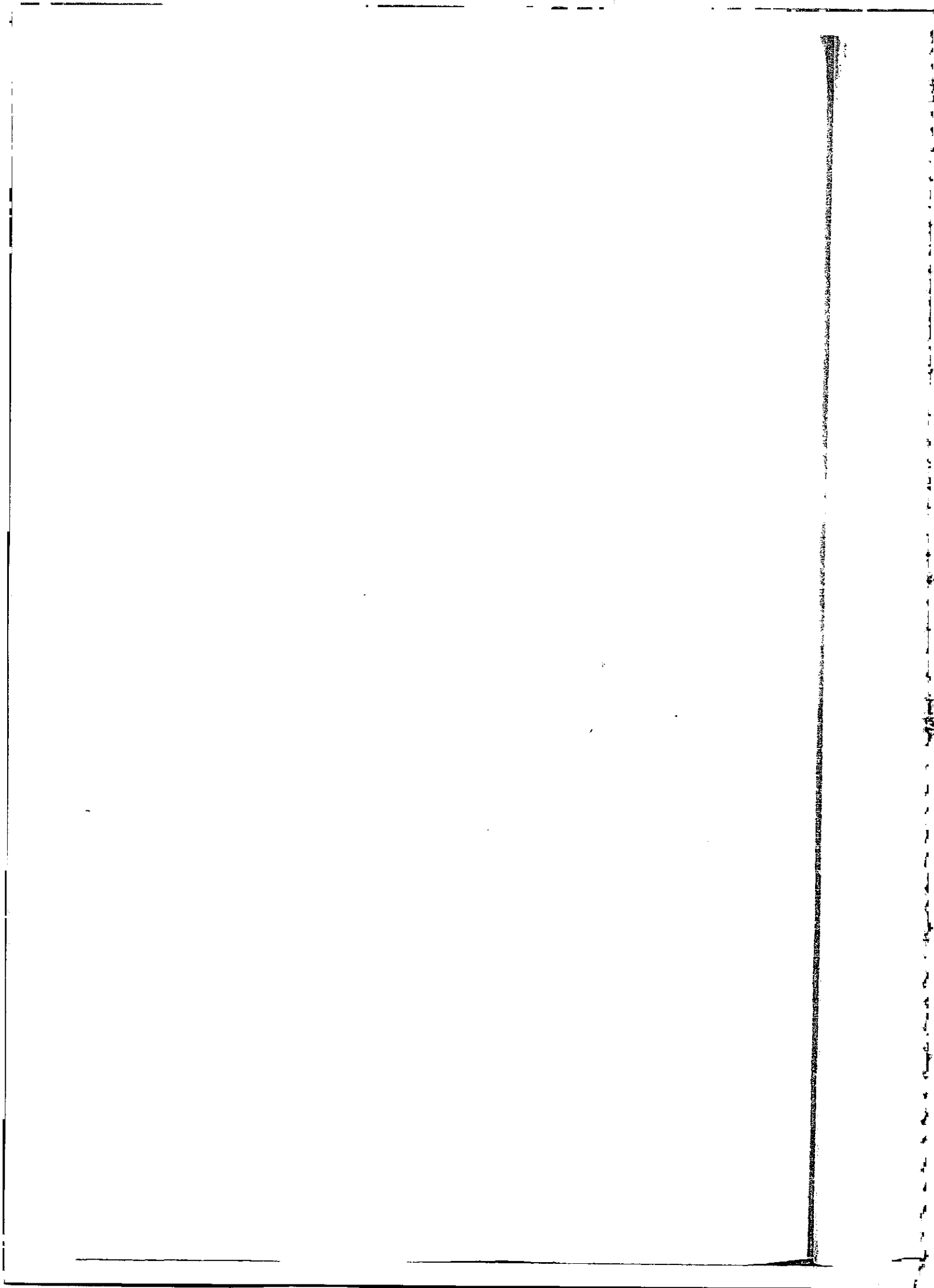
VAR Methods

So far, we have considered sources of risk in isolation. This approach reflects the state of the art up to the beginning of the 1990s. Until then, risk was measured and managed at the level of a desk or business unit. The finance profession was basically compartmentalized. This approach, however, fails to take advantage of portfolio theory, which has taught us that risk should be measured at the level of the portfolio.

This chapter turns to firm-wide VAR methods. These can be separated into local valuation and full valuation methods. Local valuation methods make use of the valuation of the instruments at the current point, along with the first and perhaps the second partial derivatives. Full valuation methods, in contrast, reprice the instruments over a broad range of values for the risk factors.

These methods are discussed in Section 15.1. Section 15.2 describes the three main VAR methods. The first step in all methods is mapping, which consists of replacing each instrument by its exposures on selected risk factors.

This considerably simplifies the risk measurement process. It would be infeasible to model all instruments individually, because there are too many. The art of risk management consists of choosing a set of limited risk factors that will adequately cover the spectrum of risks for the portfolio at hand. Thus, risk management is truly the art of the approximation. Sometimes, however, these approximations fail, as shown in Section 15.3, which discusses the performance of risk systems during the recent credit crisis.

Finally, Section 15.4 works through a detailed example, which is a forward currency contract.

15.1 VAR: LOCAL VERSUS FULL VALUATION

The various approaches to VAR described in Figure 15.1. The left branch describes local valuation methods, also known as analytical methods. These include linear models and nonlinear models. Linear models are based on the covariance matrix approach. This matrix can be simplified using factor models, or even a diagonal model, which is a one-factor model.

Nonlinear models take into account the first and second partial derivatives. The latter are called gamma or convexity. Next, the right branch describes full valuation methods and include historical or Monte Carlo simulations.
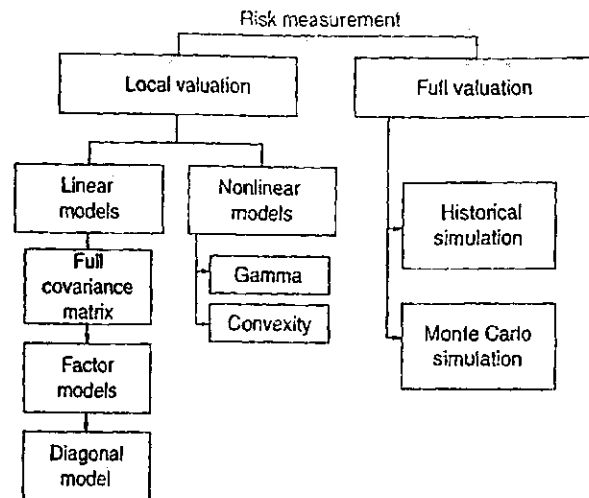
FIGURE 15.1   VAR Methods

15.1.1 Local Valuation

VAR was born from the recognition that we need an estimate that accounts for various sources of risk and expresses loss in terms of probability. Extending the duration equation to the worst change in yield at some confidence level dy, we have

$$(\text{Worst } dP) = (-D^*P) \times (\text{Worst } dy) \qquad (15.1)$$

where D^* is modified duration. For a long position in the bond, the worst movement in yield is an increase at say, the 95% confidence level. This will lead to a fall in the bond value at the same confidence level. We call this approach local valuation, because it uses information about the initial price and the exposure at the initial point. As a result, the VAR for the bond is given by

$$\text{VAR}(dP) = |-D^*P| \times \text{VAR}(dy) \qquad (15.2)$$

The main advantage of this approach is its simplicity: The distribution of the price is the same as that of the change in yield. This is particularly convenient for portfolios with numerous sources of risks, because linear combinations of normal distributions are normally distributed. Figure 15.2, for example, shows how the linear exposure combined with the normal density (in the right panel) combines to create a normal density.

15.1.2 Full Valuation

More generally, to take into account nonlinear relationships, one would have to reprice the bond under different scenarios for the yield. Defining y_0 as the initial yield,

$$(\text{Worst } dP) = P[y_0 + (\text{Worst } dy)] - P[y_0] \qquad (15.3)$$

We call this approach full valuation, because it requires repricing the asset.
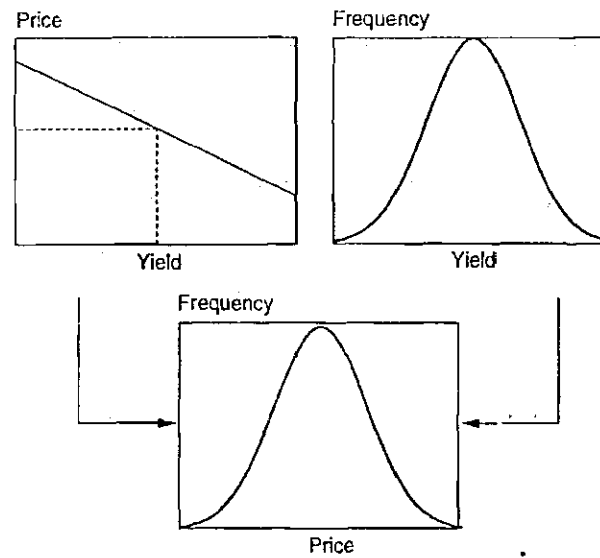
FIGURE 15.2 Distribution with Linear Exposures

This approach is illustrated in Figure 15.3, where the nonlinear exposure combined with the normal density creates a distribution that is not symmetrical any more, but skewed to the right. This is more precise but, unfortunately, is more complex than a simple, linear valuation method.

15.1.3 Delta–Gamma Method

Ideally, we would like to keep the simplicity of the local valuation while accounting for nonlinearities in the payoffs patterns. Using the Taylor expansion,

$$dP \approx \frac{\partial P}{\partial y}dy + (1/2)\frac{\partial^2 P}{\partial y^2}(dy)^2 = (-D^*P)dy + (1/2)CP(dy)^2 \qquad (15.4)$$
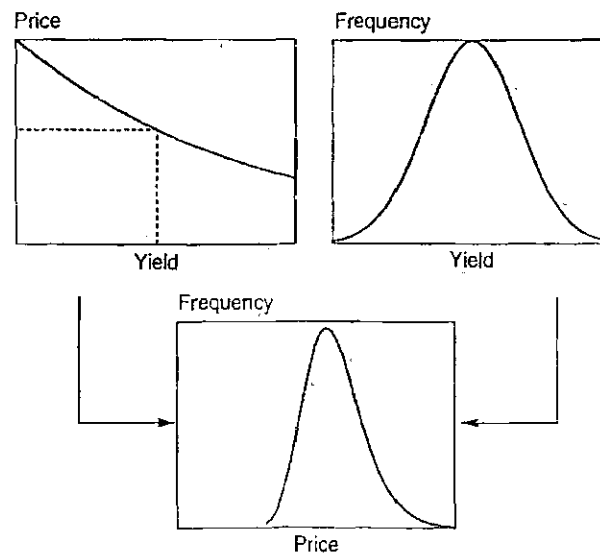
FIGURE 15.3 Distribution with Nonlinear Exposures

where the second-order term involves convexity C. Note that the valuation is still local because we only value the bond once, at the original point. The first and second derivatives are also evaluated at the local point.

Because the price is a monotonic function of the underlying yield, we can use the Taylor expansion to find the worst downmove in the bond price from the worst move in the yield. Calling this $dy^* = \text{VAR}(dy)$, we have

$$(\text{Worst } dP) = P(y_0 + dy^*) - P(y_0) \approx (-D^* P)(dy^*) + (1/2)(C\,P)(dy^*)^2 \quad (15.5)$$

This leads to a simple adjustment for VAR

$$\text{VAR}(dP) = | -D^* P | \times \text{VAR}(dy) - (1/2)(C\,P) \times \text{VAR}(dy)^2 \quad (15.6)$$

More generally, this method can be applied to derivatives, for which we write the Taylor approximation as

$$df \approx \frac{\partial f}{\partial S} dS + (1/2)\frac{\partial^2 f}{\partial S^2} dS^2 = \Delta dS + (1/2)\Gamma dS^2 \quad (15.7)$$

where Γ is now the second derivative, or gamma, like convexity.

For a long call option, the worst value is achieved as the underlying price moves down by $\text{VAR}(dS)$. With $\Delta > 0$ and $\Gamma > 0$, the VAR for the derivative is now

$$\text{VAR}(df) = | \Delta | \times \text{VAR}(dS) - (1/2)\Gamma \times \text{VAR}(dS)^2 \quad (15.8)$$

This method is called delta–gamma because it provides an analytical, second-order correction to the delta–normal VAR. This explains why long positions in options, with positive gamma, have less risk than with a linear model. Conversely, short positions in options have greater risk than implied by a linear model.

This simple adjustment, unfortunately, only works when the payoff function is monotonic, that is, involves a one-to-one relationship between the option value f and S. More generally, the delta–gamma–delta VAR method involves, first, computing the moments of df using Equation (15.7) and, second, choosing the normal distribution that provides the best fit to these moments.

The improvement brought about by this method depends on the size of the second-order coefficient, as well as the size of the worst move in the risk factor. For forward contracts, for instance, $\Gamma = 0$, and there is no point in adding second-order terms. Similarly, for most fixed-income instruments over a short horizon, the convexity effect is relatively small and can be ignored.

EXAMPLE 15.1: FRM EXAM 2004—QUESTION 60

Which of the following methodologies would be most appropriate for stress testing your portfolio?

a. Delta–gamma valuation
b. Full revaluation
c. Marked to market
d. Delta–normal VAR

EXAMPLE 15.2: FRM EXAM 2002—QUESTION 38

If you use delta–VAR for a portfolio of options, which of the following statements is *always* correct?

a. It necessarily understates VAR because it uses a linear approximation.
b. It can sometimes overstate VAR.
c. It performs most poorly for a portfolio of deep in-the-money options.
d. It performs most poorly for a portfolio of deep out-of-the-money options.

15.2 VAR METHODS: OVERVIEW

15.2.1 Mapping

This section provides an introduction to the three VAR methods. The portfolio could consist of a large number of instruments, say M. Because it would be too complex to model each instrument separately, the first step is mapping, which consists of replacing the instruments by positions on a limited number of risk factors. Say we have N risk factors. The positions are then aggregated across instruments, which yields dollar exposures x_i.

For instance, we could reduce the large spectrum of maturities in the U.S. fixed-income market by 14 maturities. We then replace the positions in every bond by exposures on these 14 risk factors. Perhaps this can be reduced further. For some portfolios, one interest rate risk factor may be sufficient.

Figure 15.4 displays the mapping process. We have six instruments, say different forward contracts on the same currency but with different maturities. These can be replaced by positions on three risk factors only. In the next section, we provide a fully worked-out example.

The distribution of the portfolio return $R_{p,i+1}$ is then derived from the exposures and movements in risk factors, Δf. Some care has to be taken defining the risk factors (in gross return, change in yield, rate of return, and so on); the

Instruments

| #1 | #2 | #3 | #4 | #5 | #6 |

Risk
factors   #1   #2   #3

Risk aggregation
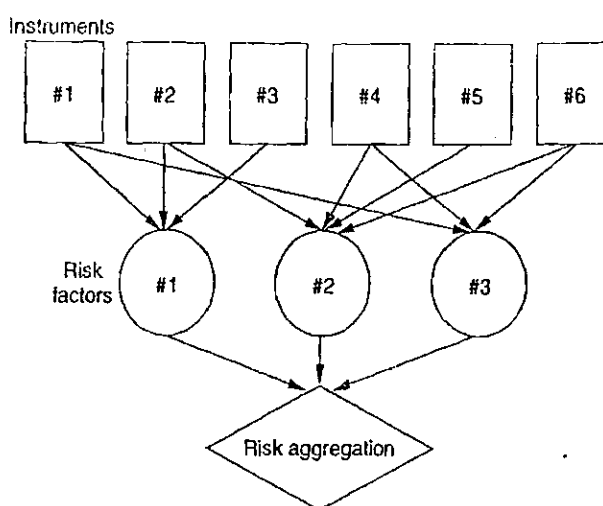
FIGURE 15.4   Mapping Approach

exposures x have to be consistently defined. Here, R_p must be measured as the change in *dollar* value of the portfolio (or whichever base currency is used).

15.2.2 Delta–Normal Method

The delta–normal method is the simplest VAR approach. It assumes that the portfolio exposures are linear and that the risk factors are jointly normally distributed. As such, it is a local valuation method.

Because the portfolio return is a linear combination of normal variables, it is normally distributed. Using matrix notations, the portfolio variance is given by

$$\sigma^2(R_{p,t+1}) = x_t' \Sigma_{t+1} x_t \qquad (15.9)$$

where Σ_{t+1} is the forecast of the covariance matrix over the horizon.

If the portfolio volatility is measured in dollars, VAR is directly obtained from the standard normal deviate α that corresponds to the confidence level c:

$$\text{VAR} = \alpha\sigma(R_{p,t+1}) \qquad (15.10)$$

This is called the diversified VAR, because it accounts for diversification effects. In contrast, the undiversified VAR is simply the sum of the individual VARs for each risk factor. It assumes that all prices will move in the worst direction simultaneously, which is unrealistic.

The RiskMetrics approach is similar to the delta–normal approach. The only difference is that the risk factor returns are measured as logarithms of the price ratios, instead of rates of returns.

The main benefit of this approach is its appealing simplicity. This is also its drawback. The delta–normal method cannot account for nonlinear effects such as encountered with options. It may also underestimate the occurrence of large observations because of its reliance on a normal distribution.

15.2.3 Historical Simulation Method

The historical-simulation (HS) method is a full valuation method. It consists of going back in time (e.g., over the last 250 days), and applying current weights to a time series of historical asset returns. It replays a "tape" of history with current weights.

Define the current time as t; we observe data from 1 to t. The current portfolio value is P_t, which is a function of the current risk factors

$$P_t = P[f_{1,t}, f_{2,t}, \ldots, f_{N,t}] \tag{15.11}$$

We sample the factor movements from the historical distribution, without replacement

$$\Delta f_i^k = \{\Delta f_{i,1}, \Delta f_{i,2}, \ldots, \Delta f_{i,t}\} \tag{15.12}$$

From this we can construct hypothetical factor values, starting from the current one

$$f_i^k = f_{i,t} + \Delta f_i^k \tag{15.13}$$

which are used to construct a hypothetical value of the current portfolio under the new scenario, using Equation (15.11)

$$P^k = P[f_1^k, f_2^k, \ldots, f_N^k] \tag{15.14}$$

We can now compute changes in portfolio values from the current position $R^k = (P^k - P_t)/P_t$.

We sort the t returns and pick the one that corresponds to the cth quantile, $R_p(c)$. VAR is obtained from the difference between the average and the quantile,

$$\text{VAR} = \text{AVE}[R_p] - R_p(c) \tag{15.15}$$

The advantage of this method is that it makes no specific distributional assumption about return distribution, other than relying on historical data. This is an improvement over the normal distribution because historical data typically contain fat tails. The main drawback of the method is its reliance on a short historical moving window to infer movements in market prices. If this window does not contain some market moves that are likely, it may miss some risks.

15.2.4 Monte Carlo Simulation Method

The Monte Carlo simulation method is basically similar to the historical simulation, except that the movements in risk factors are generated by drawings from

some prespecified distribution. Instead of Equation (15.12), we have

$$\Delta f^k \sim g(\theta), \quad k = 1, \dots K \tag{15.16}$$

where g is the joint distribution (e.g. a normal or Student's t) and θ the required parameters. The risk manager samples pseudo-random numbers from this distribution and then generates pseudo-dollar returns as before. Finally, the returns are sorted to produce the desired VAR.

This method is the most flexible, but also carries an enormous computational burden. It requires users to make assumptions about the stochastic process and to understand the sensitivity of the results to these assumptions. Thus, it is subject to model risk.

Monte Carlo methods also create inherent sampling variability because of the randomization. Different random numbers will lead to different results. It may take a large number of iterations to converge to a stable VAR measure. It should be noted that when all risk factors have a normal distribution and exposures are linear, the method should converge to the VAR produced by the delta–normal VAR.

15.2.5 Comparison of Methods

Table 15.1 provides a summary comparison of the three mainstream VAR methods. Among these methods, the delta–normal is by far the easiest to implement and communicate. For simple portfolios with little optionality, this may be perfectly appropriate. In contrast, the presence of options may require a full valuation method.

TABLE 15.1 Comparison of Approaches to VAR

Features	Delta–normal	Historical simulation	Monte Carlo simulation
Valuation	Linear	Full	Full
Distribution			
Shape	Normal	Actual	General
Extreme events	Low probability	In recent data	Possible
Implementation			
Ease of computation	Yes	Intermediate	No
Communicability	Easy	Easy	Difficult
VAR precision	Excellent	Poor with short window	Good with many iterations
Major pitfalls	Nonlinearities, fat tails	Time variation in risk, unusual events	Model risk

EXAMPLE 15.3: FRM EXAM 2001—QUESTION 92

Under usually accepted rules of market behavior, the relationship between parametric delta–normal VAR and historical VAR will tend to be

a. Parametric VAR will be higher.
b. Parametric VAR will be lower.
c. It depends on the correlations.
d. None of the above are correct.

EXAMPLE 15.4: FRM EXAM 2004—QUESTION 51

In early 2000, a risk manager calculates the VAR for a technology stock fund based on the last three years of data. The strategy of the fund is to buy stocks and write out-of-the-money puts. The manager needs to compute VAR. Which of the following methods would yield results that are *least* representative of the risks inherent in the portfolio?

a. Historical simulation with full repricing
b. Delta–normal VAR assuming zero drift
c. Monte Carlo style VAR assuming zero drift with full repricing
d. Historical simulation using delta–equivalents for all positions

EXAMPLE 15.5: FRM EXAM 2006—QUESTION 114

Which of the following is most accurate with respect to delta–normal VAR?

a. The delta–normal method provides accurate estimates of VAR for assets that can be expressed as a linear or nonlinear combination of normally distributed risk factors.
b. The delta-normal method provides accurate estimates of VAR for options that are at or near-the-money and close to expiration.
c. The delta–normal method provides estimates of VAR by generating a covariance matrix and measuring VAR using relatively simple matrix multiplication.
d. The delta–normal method provides accurate estimates of VAR for options and other derivatives over ranges even if deltas are unstable.

EXAMPLE 15.6: FRM EXAM 2005—QUESTION 94

Which of the following statements about VAR estimation methods is *wrong*?

a. The delta–normal VAR method is more reliable for portfolios that implement portfolio insurance through dynamic hedging than for portfolios that implement portfolio insurance through the purchase of put options.

b. The full valuation VAR method based on historical data is more reliable for large portfolios that contain significant option-like investments than the delta–normal VAR method.

c. The delta–normal VAR method can understate the true VAR for stock portfolios when the distribution of the return of the stocks has high kurtosis.

d. Full valuation VAR methods based on historical data take into account nonlinear relationships between risk factors and security prices.

EXAMPLE 15.7: FRM EXAM 2005—QUESTION 128

Natural gas prices exhibit seasonal volatility. Specifically the entire forward curve is more volatile during the wintertime. Which of the following statements concerning VAR is correct if the VAR is estimated using unweighted historical simulation and a three-year sample period?

a. We will overstate VAR in the summer and understate VAR in the winter.

b. We will overstate VAR in the summer and overstate VAR in the winter.

c. We will understate VAR in the summer and understate VAR in the winter.

d. We will understate VAR in the summer and overstate VAR in the winter.

EXAMPLE 15.8: FRM EXAM 2004—QUESTION 30

You are given the following information about the returns of stock P and stock Q: Variance of return of stock P = 100.0. Variance of return of stock Q = 225.0. Covariance between the return of stock P and the return of stock Q = 53.2. At the end of 1999, you are holding USD 4 million in stock P. You are considering a strategy of shifting USD 1 million into stock Q and keeping USD 3 million in stock P. What percentage of risk, as measured by standard deviation of return, can be reduced by this strategy?

a. 0.5%

b. 5.0%

c. 7.4%

d. 9.7%

15.3 LIMITATIONS OF VAR SYSTEMS

The goal of risk measurement systems is to describe the distribution of potential losses on the portfolio. VAR is a single summary measure of dispersion in portfolio returns and consequently has limitations that should be obvious.

As explained in Chapter 10, VAR cannot be viewed as a worst-loss measure. Instead, it should be viewed as a measure of dispersion that should be exceeded with some regularity, e.g., in 1% of the cases with the usual 99% confidence level. In addition, VAR does not describe the extent of losses in the left tail. Instruments such as short position in options could generate infrequent but extreme losses when they occur. To detect such vulnerabilities, the distribution of losses beyond VAR should be examined as well.

The traditional application of VAR models, such as historical simulation, involves moving windows, typically using one to three years of historical data. Such windows may not represent the range of potential movements in the risk factors, however, which is why stress tests are needed.

In addition, as we have seen in this chapter, the implementation of VAR systems often requires simplifications, obtained by mapping the positions on the selected risk factors. Thus, risk managers should be cognizant of weaknesses in their risk systems.

During the credit crisis that started in 2007, risk management systems failed at many banks. Some banks suffered losses that were much more frequent and much worse than they had anticipated. In 2007 alone, for example, UBS suffered losses of $19 billion from positions in mortgage-backed securities. Instead of experiencing the expected number of 2 or 3 exceptions (i.e., 1% of 250 days), this bank suffered 29 exceptions, or losses worse than VAR.[1]

This was due to a number of factors. First, markets were extremely volatile after an extended period of relative stability. Figure 15.5, for example, plots the daily volatility forecast for the S&P stock index using an Exponentially Weighed Moving Average (EWMA) with decay of 0.94. This model shows that during 2004 to 2006, the volatility was very low, averaging 0.7% daily. As a result, many financial institutions entered 2007 with high levels of leverage. When volatility started to spike during 2007, risk models experienced many exceptions. The graph also shows a volatility forecast derived from the usual moving average (MA) model with a window of one year, which is typical of most VAR models based on historical simulation. The figure shows that the MA model systematically underestimated the EWMA volatility starting in mid-2007, which is when banks' risk models started to slip.

In addition to the effects of heightened volatility in all financial markets, many banks experienced large losses on super senior, triple-A rated, tranches of securities backed by subprime mortgages. As will be seen in Chapter 22, these

---

[1] For a lucid explanation of risk management weaknesses, see the *Shareholder Report on UBS's Write-Downs* (April 2008).
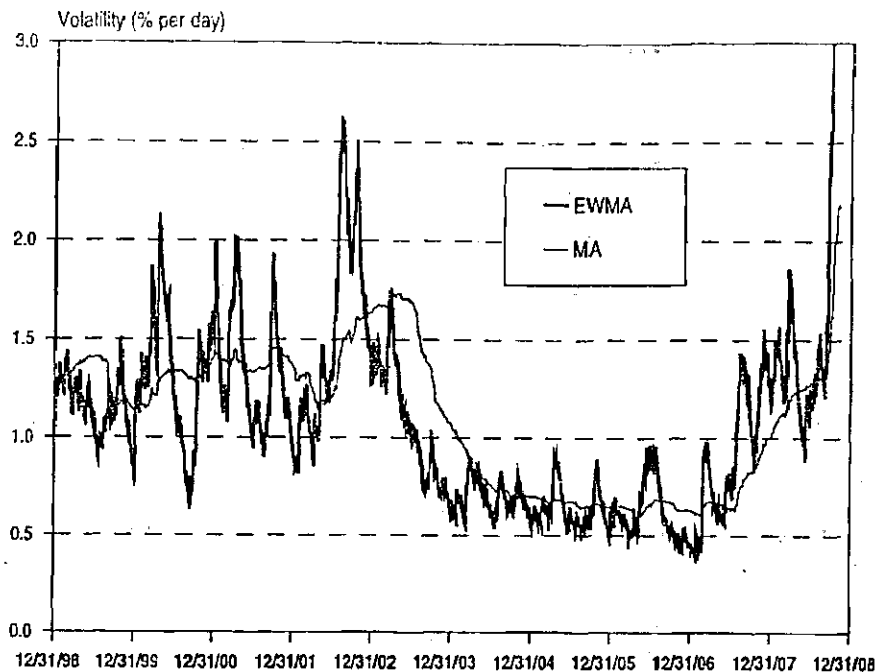
FIGURE 15.5 Volatility of S&P Stock Index (EWMA)

structures are fairly complex to model due to the need to estimate joint default probabilities. Investing in super senior tranches can be viewed as selling out-of-the-money put options, which, as we have seen, involve nonlinear payoffs. As long as the real estate market continued to go up, the default rate on subprime debt was relatively low and the super senior debt was safe, experiencing no price volatility. However, as the real estate market corrected sharply, the put options moved in-the-money, which led to large losses on the super senior debt. Of course, none of these movements showed up in the recent historical data because this only reflected a sustained appreciation in the housing market but also because of the inherent nonlinearity in these securities.

Instead of modeling these complexities, some banks simply mapped the super senior debt on AAA-rated corporate bond curves. This ignored the nonlinearities in the securities and was an act of blind faith in the credit rating. In this case, the mapping process was flawed and gave no warning sign of the impending risks.

By now, a number of reports have been written on the risk management practices at major financial institutions.[2] A striking observation is the range of quality of risk management practices. The characteristics of winners and losers are compared in Table 15.2.

---

[2] See Senior Supervisor Group (March 6, 2008), *Observations on Risk Management Practices during the Recent Market Turbulence.*

TABLE 15.2 Differences in Risk Management Practices

Practice	Winners	Losers
Business model	■ Avoided CDOs, SIVs	■ Exposed to CDOs, SIVs
Organizational structure	■ Cooperative	■ Hierarchical
Firm-wide risk analysis	■ Shared information across the firm	■ No prompt discussion of risks across the firm
Valuations	■ Developed in-house expertise	■ Relied on credit ratings
Management of liquidity	■ Charged business lines for liquidity risk	■ Did not consider contingent exposures
Risk measurement	■ Qualitative and quantitative analysis	■ Strict model application
	■ Varied assumptions	■ Mapped to corporate AAA
	■ Tested correlations	■ No test of correlations

In general, institutions that lost the most had a hierarchical business structure where top management wanted to expand the structured credit business, which involved collateralized debt obligations (CDOs) and structured investment vehicles (SIVs), due to their perceived profitability. Top management did not encourage feedback and often did not pay attention to warning signals given by risk managers. Many of these institutions failed to develop their own valuations models for these complex structures and instead relied on credit ratings. In addition, they did not consider contingent exposures and did not charge business lines for potential claims on the bank's balance sheet, which encouraged expansion into structured credit. These institutions blindly applied models without consideration of their weaknesses and typically did not perform stress tests of correlations.

15.4 EXAMPLE

15.4.1 Mark-to-Market

We now illustrate the computation of VAR for a simple example. The problem at hand is to evaluate the one-day downside risk of a currency forward contract. We will show that to compute VAR we need first to value the portfolio, mapping the value of the portfolio on fundamental risk factors, then to generate movements in these risk factors, and finally, to combine the risk factors with the valuation model to simulate movements in the contract value.

Assume that on December 31, 1998, we have a forward contract to buy £10 million in exchange for delivering $16.5 million in three months.

As before, we use these definitions:

S_t = current spot price of the pound in dollars
F_t = current forward price
K = purchase price set in contract

$$f_t = \text{current value of contract}$$
$$r_t = \text{domestic risk-free rate}$$
$$r_t^* = \text{foreign risk-free rate}$$
$$\tau = \text{time to maturity}$$

To be consistent with conventions in the foreign exchange market, we define the present value factors using discrete compounding

$$P_t = PV(\$1) = \frac{1}{1+r_t\tau} \qquad P_t^* = PV(\$1) = \frac{1}{1+r_t^*\tau} \qquad (15.17)$$

The current market value of a forward contract to buy one pound is given by

$$f_t = S_t\frac{1}{1+r_t^*\tau} - K\frac{1}{1+r_t\tau} = S_t P_t^* - K P_t \qquad (15.18)$$

which is exposed to three risk factors: the spot rate and the two interest rates. In addition, we can use this equation to derive the exposures on the risk factors. After differentiation, we have

$$df = \frac{\partial f}{\partial S}dS + \frac{\partial f}{\partial P^*}dP^* + \frac{\partial f}{\partial P}dP = P^*dS + SdP^* - KdP \qquad (15.19)$$

Alternatively,

$$df = (SP^*)\frac{dS}{S} + (SP^*)\frac{dP^*}{P^*} - (KP)\frac{dP}{P} \qquad (15.20)$$

Intuitively, the forward contract is equivalent to

- A long position of (SP^*) on the spot rate
- A long position of (SP^*) in the foreign bill
- A short position of (KP) in the domestic bill (borrowing)

We can now mark to market our contract. If Q represents our quantity, £10 million, the current market value of our contract is

$$V_t = Qf_t = \$10,000,000 \, S_t\frac{1}{1+r_t^*\tau} - \$16,500,000\frac{1}{1+r_t\tau} \qquad (15.21)$$

On the valuation date, we have $S_t = 1.6637, r_t = 4.9375\%,$ and $r_t^* = 5.9688\%$. Hence

$$P_t = \frac{1}{1 + r_t \tau} = \frac{1}{(1 + 4.9375\% \times 90/360)} = 0.9879$$

and similarly, $P_t^* = 0.9854$. The current market value of our contract is

$$V_t = \$10,000,000 \times 1.6637 \times 0.9854 - \$16,500,000 \times 0.9879 = \$93,581$$

which is slightly in-the-money. We are going to use this formula to derive the distribution of contract values under different scenarios for the risk factors.

15.4.2 Risk Factors

Assume now that we only consider the last 100 days to be representative of movements in market prices. Table 15.3 displays quotations on the spot and 3-month rates for the last 100 business days, starting on August 10.

We first need to convert these quotes into true random variables, that is, with zero mean and constant dispersion. Table 15.4 displays the one-day changes in

TABLE 15.3 Historical Market Factors

Date	$ Eurorate (3mo-%pa)	£ Eurorate (3mo-%pa)	Spot Rate S($/£)	Number
		Market Factors		
8/10/98	5.5938	7.4375	1.6341	
8/11/98	5.5625	7.5938	1.6315	1
8/12/98	6.0000	7.5625	1.6287	2
8/13/98	5.5625	7.4688	1.6267	3
8/14/98	5.5625	7.6562	1.6191	4
8/17/98	5.5625	7.6562	1.6177	5
8/18/98	5.5625	7.6562	1.6165	6
8/19/98	5.5625	7.5625	1.6239	7
8/20/98	5.5625	7.6562	1.6277	8
8/21/98	5.5625	7.6562	1.6387	9
8/24/98	5.5625	7.6562	1.6407	10
...				
12/15/98	5.1875	6.3125	1.6849	90
12/16/98	5.1250	6.2188	1.6759	91
12/17/98	5.0938	6.3438	1.6755	92
12/18/98	5.1250	6.1250	1.6801	93
12/21/98	5.1250	6.2812	1.6807	94
12/22/98	5.2500	6.1875	1.6789	95
12/23/98	5.2500	6.1875	1.6769	96
12/24/98	5.1562	6.1875	1.6737	97
12/29/98	5.1875	6.1250	1.6835	98
12/30/98	4.9688	6.0000	1.6667	99
12/31/98	4.9375	5.9688	1.6637	100

TABLE 15.4 Movements in Market Factors

	Movements in Market Factors				
Number	$dr(\$1)$	$dr(£1)$	$dP/P(\$1)$	$dP/P(£1)$	$dS(\$/£)/S$
1	−0.0313	0.1563	0.00000	−0.00046	−0.0016
2	0.4375	−0.0313	−0.00116	0.00000	−0.0017
3	−0.4375	−0.0937	0.00100	0.00015	−0.0012
4	0.0000	0.1874	−0.00008	−0.00054	−0.0047
5	0.0000	0.0000	−0.00008	−0.00008	−0.0009
6	0.0000	0.0000	−0.00008	−0.00008	−0.0007
7	0.0000	−0.0937	−0.00008	0.00015	0.0046
8	0.0000	0.0937	−0.00008	−0.00031	0.0023
9	0.0000	0.0000	−0.00008	−0.00008	0.0068
10	0.0000	0.0000	−0.00008	−0.00008	0.0012
...					
90	0.0937	0.0625	−0.00031	−0.00023	−0.0044
91	−0.0625	−0.0937	0.00008	0.00015	−0.0053
92	−0.0312	0.1250	0.00000	−0.00038	−0.0002
93	0.0312	0.2188	−0.00015	0.00046	0.0027
94	0.0000	0.1562	−0.00008	−0.00046	0.0004
95	0.1250	−0.0937	−0.00039	0.00015	−0.0011
96	0.0000	0.0000	−0.00008	−0.00008	−0.0012
97	−0.0938	0.0000	0.00015	−0.00008	−0.0019
98	0.0313	−0.0625	−0.00015	0.00008	0.0059
99	−0.2187	−0.1250	0.00046	0.00023	−0.0100
100	−0.0313	−0.0312	0.00000	0.00000	−0.0018

interest rates dr, as well as the relative changes in the associated present value factors dP/P and in spot rates dS/S. For instance, for the first day,

This information is now used to construct the distribution of risk factors.

$$dr_1 = 5.5625 - 5.5938 = -0.0313 \text{ and}$$

$$dS/S_1 = (1.6315 - 1.6341)/1.6341 = -0.0016$$

15.4.3 VAR: Historical Simulation

The historical-simulation method takes historical movements in the risk factors to simulate potential future movements. For instance, one possible scenario for the U.S. interest rate is that, starting from the current value $r_0 = 4.9375$, the movement the next day could be similar to that observed on August 11, which is a decrease of $dr_1 = -0.0313$. The new value is $r(1) = 4.9062$.

We compute the simulated values of other variables as

$$r^*(1) = 5.9688 + 0.1563 = 6.1251 \text{ and}$$

$$S(1) = 1.6637 \times (1 - 0.0016) = 1.6611.$$

Armed with these new values, we can reprice the forward contract, now worth

$$V_t = \$10,000,000 \times 1.6611 \times 0.9849 - \$16,500,000 \times 0.9879 = \$59,941$$

TABLE 15.5 Simulated Market Factors

Number	Simulated Market Factors					Hypothetical MTM Contract
	$r(\$1)$	$r(\pounds 1)$	$S(\$/\pounds)$	PV($1)	PV(£1)	
1	4.9062	6.1251	1.6611	0.9879	0.9849	$59,941
2	5.3750	5.9375	1.6608	0.9867	0.9854	$84,301
3	4.5000	5.8751	1.6617	0.9889	0.9855	$59,603
4	4.9375	5.1562	1.6559	0.9878	0.9848	$9,467
5	4.9375	5.9688	1.6623	0.9878	0.9853	$79,407
6	4.9375	5.9688	1.6625	0.9878	0.9853	$81,421
7	4.9375	5.8751	1.6713	0.9878	0.9855	$172,424
8	4.9375	6.0625	1.6676	0.9878	0.9851	$128,149
9	4.9375	5.9688	1.6749	0.9878	0.9853	$204,361
10	4.9375	5.9688	1.6657	0.9878	0.9853	$113,588
...						
90	5.0312	6.0313	1.6564	0.9876	0.9851	$23,160
91	4.8750	5.8751	1.6548	0.9880	0.9855	$7,268
92	4.9063	6.0938	1.6633	0.9879	0.9850	$83,368
93	4.9687	5.7500	1.6683	0.9877	0.9858	$148,705
94	4.9375	6.1250	1.6643	0.9878	0.9849	$93,128
95	5.0625	5.8751	1.6619	0.9875	0.9855	$84,835
96	4.9375	5.9688	1.6617	0.9878	0.9853	$74,054
97	4.8437	5.9688	1.6605	0.9880	0.9853	$58,524
98	4.9688	5.9063	1.6734	0.9877	0.9854	$193,362
99	4.7188	5.8438	1.6471	0.9883	0.9856	−$73,811
100	4.9062	5.9376	1.6607	0.9879	0.9854	$64,073
	4.9375	5.9688	1.6637	0.9879	0.9854.	$93,581

Note that, because the contract is long the pound that fell in value, the current value of the contract has decreased relative to the initial value of $93,581.

We record the new contract value and repeat this process for all the movements from day 1 to day 100. This creates a distribution of contract values, which is reported in the last column of Table 15.5.

The final step consists of sorting the contract values, as shown in Table 15.6. Suppose we want to report VAR relative to the initial value (instead of relative to the average on the target date.) The last column in the table reports the *change* in the portfolio value, i.e., $V(k) - V_0$. These range from a loss of $200,752 to a gain of $280,074.

We can now characterize the risk of the forward contract by its entire distribution, which is shown in Figure 15.6. The purpose of VAR is to report a single number as a downside risk measure. Let us take, for instance, the 95% lower quantile. From Table 15.6, we identify the fifth-lowest value out of 100, which is $127,232. Ignoring the mean, the 95% VAR is $VAR_{HS} = \$127,232$.

15.4.4 VAR: Delta–Normal Method

The delta–normal approach takes a different approach to constructing the distribution of the portfolio value. We assume that the three risk factors (dS/S), (dP/P), (dP^*/P^*) are jointly normally distributed.

TABLE 15.6   Distribution of Portfolio Values

	Sorted Values	
Number	Hypothetical MTM	Change in MTM
1	−$107,171	−$200,752
2	−$73,811	−$167,392
3	−$46,294	−$139,875
4	−$37,357	−$130,938
5	−$33,651	−$127,232
6	−$22,304	−$115,885
7	−$11,694	−$105,275
8	$7,268	−$86,313
9	$9,467	−$84,114
10	$13,744	−$79,837
...		
90	$193,362	$99,781
91	$194,405	$100,824
92	$204,361	$110,780
93	$221,097	$127,515
94	$225,101	$131,520
95	$228,272	$134,691
96	$233,479	$139,897
97	$241,007	$147,426
98	$279,672	$186,091
99	$297,028	$203,447
100	$373,655	$280,074
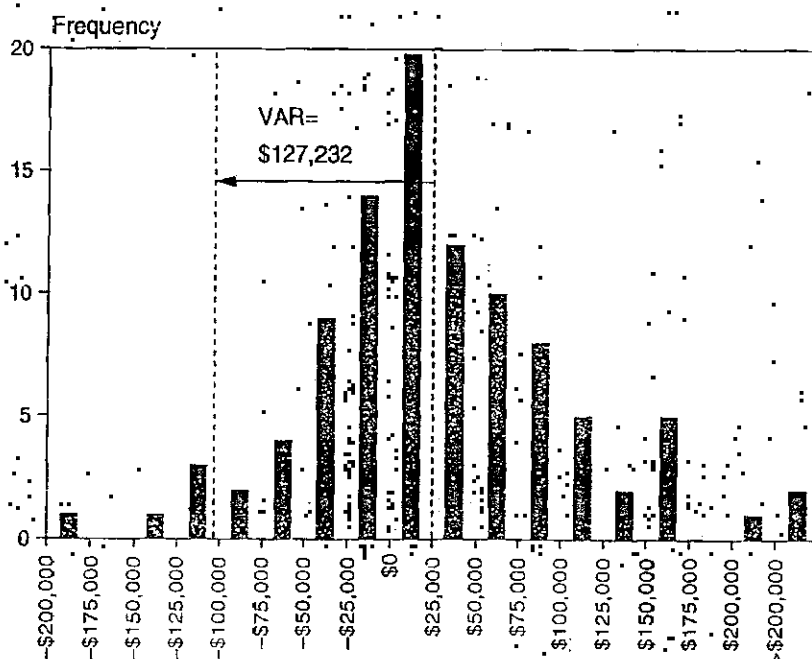
FIGURE 15.6   Empirical Distribution of Value Changes

TABLE 15.7 Covariance Matrix Approach

	dP/P($1)	dP/P(£1)	dS($/£)/S
Standard Deviation:	0.022%	0.026%	0.473%
Correlation Matrix:	**dP/P($1)**	**dP/P(£1)**	**dS($/£)/S**
dP/P($1)	1.000	0.137	0.040
dP/P(£1)	0.137	1.000	−0.063
dS($/£)/S	0.040	−0.063	1.000
Covariance Matrix:	**dP/P($1)**	**dP/P(£1)**	**dS($/£)/S**
Σ dP/P($1)	4.839E-08	7.809E-09	4.155E-08
dP/P (£1)	7.809E-09	6.720E-08	−7.688E-08
dS($/£)/S	4.155E-08	−7.688E-08	2.237E-08

Exposure:

x′   −$16,300,071 $16,393,653 $16,393,653

Σx

4.839E-08	7.809E-09	4.155E-08		−$16,300,071		$0.020
7.809E-09	6.720E-08	−7.688E-08	×	$16,393,653	=	−$0.286
4.155E-08	−7.688E-08	2.237E-05		$16,393,653		$364.852

$\sigma^2 = x'(\Sigma x)$ Variance:

				$0.020		
−$16,300,071	$16,393,653	$16,393,653	×	−$0.286	=	$5,976,242,188
				$364.852		

σ   Standard deviation $77,306

We can write Equation (15.20) as

$$df = (SP^*)\frac{dS}{S} + (SP^*)\frac{dP^*}{P^*} - (KP)\frac{dP}{P} = x_1 dz_1 + x_2 dz_2 + x_3 dz_3 \qquad (15.22)$$

where the dz are normal variables and x are exposures.

Define Σ as the (3 by 3) covariance matrix of the dz, and x as the vector of exposures. We compute VAR from $\sigma^2(df) = x'\Sigma x$. Table 15.7 details the steps. First, we compute the covariance matrix of the three risk factors. The top of the table shows the standard deviation of daily returns as well as correlations. From these, we construct the covariance matrix.

Next, the table shows the vector of exposures, x'. The matrix multiplication Σx is shown on the following lines. After that, we compute $x'(\Sigma x)$, which yields the variance. Taking the square root, we have $\sigma(df) = \$77,306$. Finally, we transform into a 95% quantile by multiplying by 1.645, which gives $\text{VAR}_{DN} = \$127,169$.

Note how close this number is to the VAR_{HS} of $127,232 we found previously. This suggests that the distribution of these variables is close to a normal distribution. Indeed, the empirical distribution in Figure 15.6 roughly looks like a normal. The fitted distribution is shown in Figure 15.7.
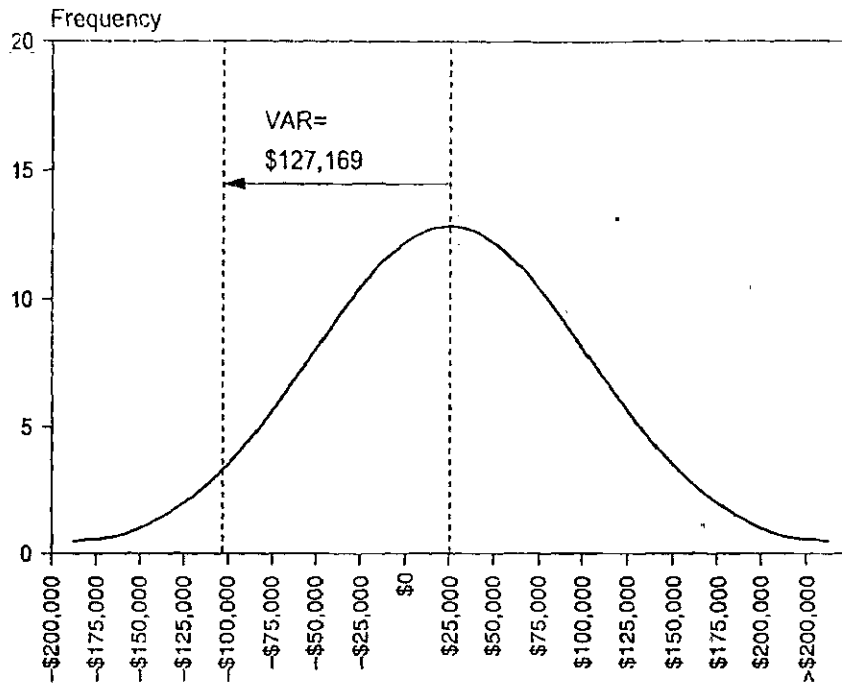
FIGURE 15.7   Normal Distribution of Value Changes

15.5 IMPORTANT FORMULAS

Linear VAR, fixed-income: $\text{VAR}(dP) = |-D^*P| \times \text{VAR}(dy)$
Quadratic VAR, fixed-income:

$$\text{VAR}(dP) = |-D^*P| \times \text{VAR}(dy) - (1/2)(C\,P) \times \text{VAR}(dy)^2$$

Full-valuation VAR, fixed-income: $(\text{Worst } dP) = P[y_0 + (\text{Worst } dy)] - P[y_0]$
Delta–VAR: $\text{VAR}(df) = |\Delta| \, \text{VAR}(dS)$
Delta–gamma VAR: $\text{VAR}(df) = |\Delta| \, \text{VAR}(dS) - (1/2)\Gamma \times \text{VAR}(dS)^2$
Delta–normal VAR: $\text{VAR} = \alpha\sigma(R_{p,t+1})$, $\sigma^2(R_{p,t+1}) = x_t'\Sigma_{t+1}x_t$
Historical simulation VAR: $\Delta f_i^k = \{\Delta f_{i,1}, \Delta f_{i,2}, \ldots, \Delta f_{i,t}\}$
Monte Carlo simulation VAR: $\Delta f^k \sim g(\theta)$

15.6 ANSWERS TO CHAPTER EXAMPLES

Example 15.1: FRM Exam 2004—Question 60

b. By definition, stress-testing involves large movements in the risk factors. This requires a full revaluation of the portfolio.

Example 15.2: FRM Exam 2002—Question 38

b. This question has to be read very carefully in view of the "always" characterization. The Delta–VAR could understate or overstate the true VAR, depending on whether the position is net long or short options, so a. is incorrect. The Delta–VAR is generally better for in-the-money options, because these have low gamma, so c. is false. For out-of-the-money options, delta is close to zero, so the Delta–VAR method would predict zero risk. The risk could indeed be very small, so d. is incorrect. So, b. is the most general statement.

Example 15.3: FRM Exam 2001—Question 92

b. Parametric VAR usually assumes a normal distribution. Given that actual distributions of financial variables have fatter tails than the normal distribution, parametric VAR at high confidence levels will generally underestimate VAR.

Example 15.4: FRM Exam 2004—Question 51

d. Because the portfolio has options, methods a. or c. based on full repricing would be appropriate. Next, recall that technology stocks have had a big increase in price until March 2000. From 1996 to 1999, the NASDAQ index went from 1300 to 4000. This creates a positive drift in the series of returns. So, historical simulation without an adjustment for this drift would bias the simulated returns upward, thereby underestimating VAR.

Example 15.5: FRM Exam 2006—Question 114

c. The delta–normal approach will perform poorly with nonlinear payoffs, so answer a. is false. Similarly, the approach will fail to measure risk properly for options if the delta changes, which is the case for at-the-money options, so answers b. and d. are false.

Example 15.6: FRM Exam 2005—Question 94

a. Full valuation methods are more precise for portfolios with options, so answers b. and d. are correct. The delta–normal VAR understates the risk when distributions have fat tails, so answer c. is correct. Answer a. is indeed wrong. The delta–normal method will be poor for outright positions in options, or their dynamic replication.

Example 15.7: FRM Exam 2005—Question 128

a. This method essentially estimates the average volatility over a three-year window, ignoring seasonality. As a result, if the conditional volatility is higher

during the winter, the method will understate the true risk, and conversely for the summer.

Example 15.8: FRM Exam 2004—Question 30

b. The variance of the original portfolio is 1,600, implying a volatility of 40. The new portfolio has variance of $3^2 \times 100 + 1^2 \times 225 + 2 \times 53.2 \times 3 \times 1 = 1,444$. This gives a volatility of 38, which is a reduction of 5%.
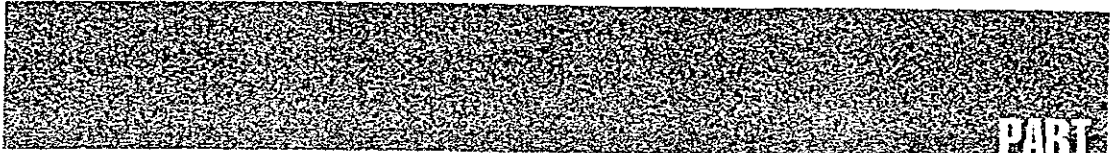
Investment Risk Management

Portfolio Management

Value at Risk techniques were developed in the early 1990s as position-based risk measures to control the risk of proprietary trading desks of commercial banks. The advent of these methods was spurred by commercial bank regulation but quickly spread to investment banks, which also have large trading operations. These techniques have been incorporated in the panoply of risk measurement tools used in the investment management industry. Institutional investors pay particular attention to the control of risk in their investment portfolio.

Risk that can be measured can be managed better. Even so, risk management accounts for one facet of the investment process only, which is risk. Investors only assume risk because they expect to be compensated for it in the form of higher returns. The real issue is how to balance risk against expected return.

This trade-off is the subject of portfolio management. So, this is much broader than risk management. Once a broad portfolio allocation into asset classes is selected, reflecting the best trade-off between risk and return, the total fund risk can be allocated to various managers using a process called risk budgeting.

At the end of the investment process, it is important to assess whether realized returns were in line with the risks assumed. The purpose of performance attribution methods is to decompose the investment performance into various components, where the goal is to identify whether the active manager really adds value. Part of the returns represents general market factors, also called "beta bets"; the remainder represents true value added, or "alpha bets."

The purpose of this chapter is to present risk and performance measurement tools in the investment management industry. Section 16.1 gives a brief introduction to institutional investors. Risk and performance measurement techniques are developed in Section 16.2. Finally, Section 16.3 discusses risk budgeting. Hedge funds, because of their importance, will be covered in the next chapter.

16.1 INSTITUTIONAL INVESTORS

Institutional investors are entities that have large amounts of funds to invest for an organization, or on behalf of others. This is in contrast with *private* investors.[1] As shown in Table 16.1, institutional investors can be classified into investment

---

[1] The SEC has formal definitions of, e.g., "qualified institutional buyers" under Rule 144a.

TABLE 16.1   Classification of Institutional Investors

Investment companies	Open-end funds
	Closed-end funds
Pension funds	Defined-benefit
	Defined-contribution
Insurance funds	Life
	Nonlife
Others	Foundations and endowment funds
	Non-pension funds managed by banks
	Private partnerships

companies, pension funds, insurance funds, and others. The latter category includes endowment funds, bank-managed funds, and private partnerships, also known as hedge funds. Hedge funds are private partnership funds that can take long and short positions in various markets and are accessible only to large investors.

Even though institutional investors and bank proprietary desks are generally exposed to similar risk factors, their philosophy is quite different. Bank trading desks employ high leverage and are aggressive investors. They typically have short horizons and engage in active trading in generally liquid markets. Financial institutions, such as commercial banks, investment banks, and broker-dealers, are sometimes called the sell side because they are primarily geared toward selling financial services.

On the other hand, institutional investors are part of the buy side because they are buying financial services from the sell side, in other words Wall Street for the United States. In contrast to the sell side, institutional investors have little or no leverage and are more conservative. Most have longer time horizons and can invest in less liquid markets. Many hedge funds, however, have greater leverage and trade actively.

16.2 PERFORMANCE MEASUREMENT

Performance measurement should properly adjust for the risks taken. This can be done with a number of metrics, typically based on the standard deviation and regression coefficients. At an even more basic level, however, the first question is how to define the risks that matter to the investor or the manager. In particular, should risk be measured in absolute terms or relative to some benchmark?

16.2.1 Risk Measurement

- Absolute risk is measured in terms of shortfall relative to the initial value of the investment, or perhaps an investment in cash. It can be expressed in dollar terms (or in the relevant base currency). Let us use the standard deviation as the risk measure and define P as the initial portfolio value and R_P as the rate of return. Absolute risk in dollar terms is

$$\sigma(\Delta P) = \sigma(\Delta P/P) \times P = \sigma(R_P) \times P \tag{16.1}$$

16.2.2 Surplus Risk

As is sometimes said, "risk is in the eye of the beholder." For investors with fixed future liabilities, the risk is not being able to perform on these liabilities. For pension funds with defined benefits, these liabilities consist of promised payments to current and future pensioners, and are called defined benefit obligations. In this case, the investment risk falls on the entity promising the benefits. In contrast, employees covered by a defined contribution plan are subject to investment risk.

For life insurance companies, these liabilities represent the likely pattern of future claim payments. These liabilities can be represented by their net present value. In general, the present value of long-term fixed payments behaves very much like a *short position in a fixed-rate bond*. If the payments are indexed to inflation, the analogous instrument is an inflation-protected bond.

The difference between the current values of assets and liabilities is called the surplus, S, defined as the difference between the value of assets A and liabilities L. The change is then $\Delta S = \Delta A - \Delta L$. Normalizing by the initial value of assets, we have

$$R_S = \frac{\Delta S}{A} = \frac{\Delta A}{A} - \frac{\Delta L}{L}\frac{L}{A} = R_{asset} - R_{liabilities}\frac{L}{A} \tag{16.4}$$

The duration of liabilities is long, typically 12 years. Using the duration approximation, the return on liabilities can be measured from changes in yields y, as $R_{liabilities} = -D^*\Delta y$. The worst combination of movements in market values is when asset fall due to a fall in equities, in a year when yields decrease. Immunization occurs when the asset portfolio, or part of it, provides a perfect hedge against changes in the value of the liabilities. Thus, investments in long-term bonds help to hedge movements in liabilities.

In this case, risk should be measured as the potential shortfall in surplus over the horizon. This is sometimes called surplus at risk. This VAR-type measure is an application of relative risk, where the benchmark is the present value of liabilities.

EXAMPLE 16.2: PENSION LIABILITIES

The AT&T pension plan reports a projected benefit obligation of $17.4 billion. If the discount rate decreases by 0.5%, this liability will increase by $0.8 billion. Based on this information, the liabilities behave like a

a. Short position in the stock market
b. Short position in cash
c. Short position in a bond with maturity of about nine years
d. Short position in a bond with duration of about nine years

---

EXAMPLE 16.3: FRM EXAM 2006—QUESTION 25

The DataSoft Corporation has an employee pension scheme with fixed liabilities and a long time horizon reflecting its young workforce. The fund's assets are $9 billion and the present value of its liabilities is $8.8 billion. Which of the following statements are incorrect?

   I. The present value of long-term fixed payments behaves very much like a long position in a fixed-rate bond.
  II. Surplus at Risk is a measure of relative risk.
 III. The DataSoft Corporation will be able to immunize its liabilities by investing $8 billion in long-term fixed-rate bonds.
    a. I and II
    b. II and III
    c. I and III
    d. I, II, and III

16.2.3 Risk-Adjusted Performance Measurement

This dichotomy, absolute versus relative returns, carries through performance measurement, which evaluates the risk-adjusted performance of the fund. The Sharpe ratio (SR) measures the ratio of the average rate of return, $\mu(R_P)$, in excess of the risk-free rate R_F, to the absolute risk

$$SR = \frac{[\mu(R_P) - R_F]}{\sigma(R_P)} \tag{16.5}$$

This approach can be extended to include VAR, or the quantile of returns, in the denominator instead of the volatility of returns. The Sharpe ratio focuses on total risk measured in absolute terms. Because total risk includes both systematic and idiosyncratic risk, this measures is appropriate for portfolios that are not very diversified, i.e., which have large idiosyncratic risk.

A related measure is the Sortino ratio (SOR). This replaces the standard deviation in the denominator by the semi-standard deviation, $\sigma_L(R_P)$, which considers only data points that represent a loss. The ratio is

$$SOR = \frac{[\mu(R_P) - R_F]}{\sigma_L(R_P)} \tag{16.6}$$

where $\sigma_L(R_P) = \sqrt{\frac{1}{(N_L)} \sum_{i=1}^{N} [Min(R_{P,i}, 0)]^2}$, and N_L is the number of observed losses. The Sortino ratio is more relevant than the Sharpe ratio when the return distribution is skewed to the left. It is much less widely used, however.

In contrast, the information ratio (IR) measures the ratio of the average rate of return in excess of the benchmark to the TEV

$$\text{IR} = \frac{[\mu(R_P) - \mu(R_B)]}{\omega} \tag{16.7}$$

Table 16.2 presents an illustration. The risk-free interest rate is $R_F = 3\%$ and the portfolio average return is -6%, with volatility of 25%. Hence, the Sharpe Ratio of the portfolio is SR $= [(-6\%) - (3\%)]/25\% = -0.36$. Because this is negative, the absolute performance is poor.

Assume now that the benchmark returned -10% over the same period and that the tracking error volatility was 8%. Hence, the Information Ratio is IR $= [(-6\%) - (-10\%)]/8\% = 0.50$, which is positive. The relative performance is good even though the absolute performance is poor. Note that this information ratio of 0.50 is typical of the performance of the top 25th percentile of money managers and is considered "good."[3]

Dealing with ratios, however, is rather abstract. It is more intuitive to express performance in terms of a rate of return, adjusted for risk. Suppose we use a reference benchmark, R_B, for which we measure first its average return and risk. We can leverage up or down the portfolio P so as to bring its volatility in line with B. The risk-adjusted performance (RAP) is then[4]

$$\text{RAP}_P = R_F + \frac{\sigma_B}{\sigma_P}[\mu(R_P) - R_F] \tag{16.8}$$

This is illustrated in Figure 16.1. The average return on portfolio P is greater than that of B. Its volatility, however, is much higher. The straight line going from R_F to R_P represents portfolios that mix the risk-free asset with P. For example, an investment of 50% in each will give an average return that is the mean of R_F and $\mu(R_P)$, and a volatility that is half that of P. The slope of this line represents the Sharpe ratio, given by Equation (16.5).

Portfolio P^* has the same level of risk as B; its performance is given by Equation (16.8). We can then compare directly RAP$_P$ and $\mu(R_B)$. In this case, portfolio P underperforms B on a risk-adjusted basis. We obtain the same ranking between P and B, however, using the Sharpe ratio.

TABLE 16.2   Absolute and Relative Performance

	Average	Volatility	Performance
Cash	3%	0%	
Portfolio P	−6%	25%	$SR = -0.36$
Benchmark B	−10%	20%	$SR = -0.65$
Deviation e	4%	8%	$IR = 0.50$

[3] See Grinold and Kahn (2000), *Active Portfolio Management*, McGraw-Hill, New York.
[4] This performance measure is sometimes called M-square.
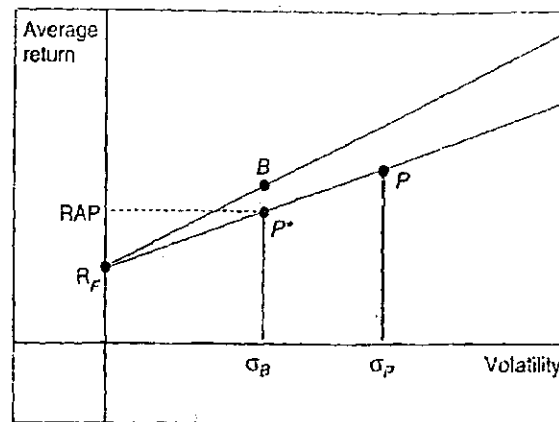
FIGURE 16.1 Risk-Adjusted Performance

EXAMPLE 16.4: SHARPE AND INFORMATION RATIOS

A portfolio manager returns 10% with a volatility of 20%. The benchmark returns 8% with risk of 14%. The correlation between the two is 0.98. The risk-free rate is 3%. Which of the following statement is *correct*?

a. The portfolio has higher SR than the benchmark.

b. The portfolio has negative IR.

c. The IR is 0.35.

d. The IR is 0.29.

16.2.4 Performance Attribution

So far, we have implemented a simple adjustment for risk that takes into account a volatility measure. To evaluate the performance of investment managers, however, it is crucial to decompose the total return into a component due to market risk premia and to other factors. Exposure to the stock market is widely believed to reward investors with a long-term premium, called the equity premium. Assume that this premium is EP = 4% annually. This is the expected return in excess of the risk-free rate. For simplicity, it is usually assumed that the same rate applies to lending and borrowing.

Now take the example of an investment fund of $1 million. A long position of $1.5 million, or 150% in passive equities financed by 50% cash borrowing should have an *excess return* composed of the total return on the 150% equity position, minus the cost of borrowing 50%, minus the risk-free rate. This gives

$$[150\% \times (EP + R_F) - 50\% R_F] - R_F = 1.5 \times EP = 6\%$$

This could be also achieved by taking a notional position of \$1.5 million in stock index futures and parking the investment in cash, including the margin. So, an investment manager who returns 6% in excess of the risk-free rate in this way is not really delivering any value added because this extra amount is simply due to exposure to the market. Therefore, it is crucial to account for factors that are known to generate risk premia.

Define $R_{M,t}$ as the rate of return in period t on the stock market, say the S&P 500 for U.S. equities, $R_{F,t}$ as the risk-free rate, and $R_{P,t}$ is the return on the portfolio. The general specification for this adjustment consists of estimating the regression

$$R_{P,t} - R_{F,t} = \alpha_P + \beta_P[R_{M,t} - R_{F,t}] + \epsilon_{P,t}, \quad t = 1, \cdots, T \tag{16.9}$$

where β_P is the exposure of portfolio P to the market factor, or systematic risk, and α_P is the abnormal performance after taking into account the exposure to the market.

"Abnormal" can only be defined in terms of a "normal" performance. One such definition is the capital asset pricing model (CAPM), developed by Professor William Sharpe. Under some conditions, he demonstrated that equilibrium in capital markets implies that the market portfolio is mean-variance efficient. In other words, it has the highest Sharpe ratio of any feasible portfolio. In Figure 16.1, taking B as the market M, the line passing through B is also known as the capital market line. It can be shown that the efficiency of the market implies a linear relationship between expected excess returns and systematic risk. For stock or portfolio i, we must have

$$E(R_i) - R_F = 0 + \beta_i[E(R_M) - R_F] \tag{16.10}$$

Comparing with Equation (16.9), this requires all α's to be zero in equilibrium.[5] A related measure is the **Treynor ratio**, which is

$$TR = \frac{[\mu(R_P) - R_F]}{\beta_P} \tag{16.11}$$

Taking the average of the two sides of Equation (16.9) shows that this represents α_P/β_P, plus the excess return on the market. Instead of focusing on total risk, as in the Sharpe ratio, the Treynor ratio focuses on systematic risk. Thus, this measure is appropriate for well-diversified portfolios.

This specification can be generalized to multiple factors. Assume we believe that in addition to the market premium, a premium is earned for *value* (or for low price-to-book companies) and *size* (of for small firms). We need to take this

---

[5] The CAPM is based on equilibrium in capital markets, which requires that the demand for securities from risk-averse investors matches the available supply. It also assumes that asset returns have a normal distribution. A major problem with this theory is that it may not be testable unless the "market" is exactly identified.

information into account in evaluating the manager, otherwise he or she may load up on factors that are priced but not recorded in the performance attribution system.

With K factors, Equation (16.9) can be generalized to

$$R_i = \alpha_i + \beta_{i1} y_1 + \cdots + \beta_{iK} y_K + \epsilon_i \qquad (16.12)$$

As in the case of the CAPM, the arbitrage pricing theory (APT), developed by Professor Stephen Ross, shows that there is a a relationship between α_i and the factor exposures.[6]

In Equation (16.9) or (16.12), the intercept is also known as Jensen's alpha. This term is widely used in the investment management industry to describe the performance adjusted for market factors.

This decomposition is also useful to detect timing ability, which consists of adding value by changing exposures on risk factors.[7] A manager could, for example, move into stocks with higher betas in anticipation of the market going up. Timing ability can be detected by adding a quadratic term to Equation (16.9)

$$R_{P,t} - R_{F,t} = \alpha_P + \beta_P[R_{M,t} - R_{F,t}] + \delta_P[R_{M,t} - R_{F,t}]^2 + \epsilon_{P,t} \qquad (16.13)$$

A positive coefficient δ_P indicates that the manager has added value from market timing, implying that beta is positively correlated with the market.

Return to the estimation of Equation (16.9). Denoting $\overline{R} = (1/T)\sum_{t=1}^{T}(R_t - R_{F,t})$ as the average over the sample period, the estimated alpha is

$$\widehat{\alpha} = \overline{R} - \widehat{\beta}\,\overline{R}_M \qquad (16.14)$$

If there is no exposure to the market ($\beta = 0$), Equation (16.14) shows that alpha is the sample average of the investment returns. More generally, Equation (16.14) properly accounts for the exposure to the systematic risk factor. In the case of our investment fund, we have $\overline{R} = 6\%$, and $\beta = 1.5\%$. So, the alpha is

$$\widehat{\alpha} = 6\% - 1.5 \times 4\% = 0$$

which correctly indicates that there is no value added.

---

[6] The theory does not rely on equilibrium but simply on the assumption that there should be no arbitrage opportunities in capital markets, a much weaker requirement. It does not even need the factor model to hold strictly. Instead, it requires only that the residual risk is very small. This must be the case if a sufficient number of common factors is identified and in a well-diversified portfolio. The APT model does not require the market to be identified, which is an advantage. Like the CAPM, however, tests of this model are ambiguous since the theory provides no guidance as to what the factors should be.

[7] See Treynor and Mazuy (1966), Can Mutual Funds Outguess the Market? *Harvard Business Review* 44, 131–136.

KEY CONCEPT

Performance evaluation must take into account the component of returns that can be attributed to exposures on general market factors (or risk premia). An investment manager only adds value if the residual return, called alpha, is positive.

EXAMPLE 16.5: FRM EXAM 2007—QUESTION 132

Which of the following statements about the Sharpe ratio is *false?*

a. The Sharpe ratio considers both the systematic and unsystematic risks of a portfolio.

b. The Sharpe ratio is equal to the excess return of a portfolio over the risk-free rate divided by the total risk of the portfolio.

c. The Sharpe ratio cannot be used to evaluate relative performance of undiversified portfolios.

d. The Sharpe ratio is derived from the Capital Market Line.

EXAMPLE 16.6: PERFORMANCE EVALUATION

Assume that a hedge fund provides a large positive alpha. The fund can take leveraged long and short positions in stocks. The market went up over the period. Based on this information,

a. If the fund has net positive beta, all of the alpha must come from the market

b. If the fund has net negative beta, part of the alpha comes from the market

c. If the fund has net positive beta, part of the alpha comes from the market

d. If the fund has net negative beta, all of the alpha must come from the market

16.2.5 Performance Evaluation and Survivorship

Another key issue when evaluating the performance of a group of investment managers is survivorship. This occurs when funds are dropped from the investment universe for reasons related to poor performance and "survivors" only are considered. Commercial databases often give information on funds that are "alive" only, because clients are no longer interested in "dead" funds.

The problem is that the average performance of the group of funds under examination becomes subject to survivorship bias. In other words, the apparent performance of the existing funds is too high, or biased upward relative to the true performance of the underlying population, due to the omission of some poor performing funds.

The extent of this bias depends on the attrition rate of the funds and can be very severe. Mutual fund studies, for example, report an attrition rate of 3.6% per year. This represents the fraction of funds existing at the beginning of the year that becomes "dead" during the year. In this sample, the survivorship bias is estimated at approximately 0.70% per annum.[8] This represents the difference between the performance of the survived sample and that of the true population. This is a significant number because it is on the order of management fees, which are around 1% of assets per annum. Samples with higher attrition rates have larger biases. For example, Commodity Trading Advisors (CTA), a category of hedge funds, are reported to have an attrition rate of 16% per year, leading to survivorship biases on the order of 5.2% per annum, which is very high.[9]

Other sources of bias can be introduced, due to the inclusion criteria and the voluntary reporting of returns. A fund with excellent performance is more likely to be chosen for inclusion by the database vendor. Or, the investment manager of such a fund may be more inclined to submit the fund returns to the database. Consequently, there is a bias toward adding funds with better returns. Or, a fund may decide to stop reporting returns if its performance drops. This is called selection bias. This bias differs from the previous one because it also exists when dead funds are included in the sample.

Finally, another subtle bias arises when firms "incubate" different types of funds before making them available to outsiders. Say 10 different funds are started by the same company over a two-year period. Some will do well and others will not, partly due to chance. The best performing fund is then open to the public, with its performance instantly backfilled for the previous two years. The other funds are ignored or disbanded. As a result, the performance of the public fund is not representative of the entire sample. This is called instant-history bias. The difference between this bias and selection bias is that the fund was not open to investors during the reported period.

KEY CONCEPT

Performance evaluation can be overly optimistic if based on a sample of funds affected by survivorship, selection, or instant-history bias. The extent of survivorship bias increases with the attrition rate.

---

[8] Carhart, Mark, Jennifer Carpenter, Anthony Lynch, and David Musto (2002), Mutual Fund Survivorship, *Review of Financial Studies* 15, 1355–1381.

[9] CTAs are investment managers who trade futures and options. In the United States, they are regulated by the Commodity Futures Trading Commission (CTFC).

EXAMPLE 16.7: FRM EXAM 2005—QUESTION 103

A database of hedge fund returns is constructed as follows. The first year of the database is 1994. All funds existing as of the end of 1994 that were willing to report their verified returns for that year are included in that year.

The database was extended by asking the funds for verified returns before 1994. Subsequently, funds are added as they are willing to report verified returns to the database. If a fund stops reporting returns, its returns are deleted from the database, but the database has an agreement with funds that they will keep reporting verified returns even if they stop being open to new investors.

Consider the four following statements:

I. The database suffers from backfilling bias.
II. The database suffers from survivorship bias.
III. The database suffers from an errors-in-variables bias.
IV. The equally-weighted annual return average of fund returns will underestimate the performance one would expect from a hedge fund.

Which one of the following is correct?

a. All the above statements are correct.
b. Statements I and II are correct.
c. Statements I, II, and III are correct.
d. Statements II and IV are correct.

16.3 RISK BUDGETING

The revolution in risk management reflects the recognition that risk should be measured at the highest level—that is, firmwide or portfoliowide. This ability to measure total risk has led to a top-down allocation of risk, called risk budgeting. Risk budgeting is the process of parceling out the total risk of the fund, or risk budget, to various assets classes and managers.

This concept is being implemented by institutional investors as a follow-up to their asset allocation process. Asset allocation consists of finding the optimal allocation into major asset classes (i.e., the allocation that provides the best risk/return trade-off for the investor). This choice defines the total risk profile of the portfolio.

16.3.1 Illustration

Consider for instance an investor having to decide how much to invest in U.S. stocks, in U.S. bonds, and in non-U.S. bonds. Risk is measured in absolute terms,

assuming returns have a joint normal distribution. More generally, this could be extended to other distributions or to a historical simulation method. The allocation will depend on the expected return and volatility of each asset class, as well as well as their correlations. Table 16.3 illustrates these data, which are based on historical dollar returns measured over the period 1978 to 2003.

Say the investor decides that the portfolio with the best risk/return trade-off has an expected return of 12.0% with total risk of 10.3%. Table 16.3 shows a portfolio allocation of 60.0%, 7.7%, and 32.3% to U.S. stocks, U.S. bonds, and non-U.S. bonds, respectively.

The volatility can be measured in terms of a 95% annual VAR. This defines a total risk budget of VAR $= \alpha\sigma W = 1.645 \times 10.3\% \times \$100 = \$16.9$ million. This VAR budget can then be parceled out to various asset classes and active managers within asset classes.

Risk budgeting is the process by which these efficient portfolio allocations are transformed into VAR assignments. At the asset class level, the individual VARs are $15.3, $0.9, and $5.9 million, respectively. For instance, the VAR budget for U.S. stocks is $60.0\% \times (1.645 \times 15.50\% \times \$100) = \$15.3$ million. Note that the sum of individual VARs is $22.1 million, which is greater than the portfolio VAR of $16.9 million due to diversification effects.

The process can be repeated at the next level. The fund has a risk budget of $15.3 million devoted to U.S. equities, with an allocation of $60 million. This allocation could be split equally between two active equity managers. Assume that the two managers are equally good, with a correlation of returns of 0.5. The optimal risk budget for each is then $8.83 million. We can verify that the total risk budget is

$$\sqrt{8.83^2 + 8.83^2 + 2 \times 0.5 \times 8.83 \times 8.83} = \sqrt{233.91} = \$15.3$$

Note that, as in the previous step, the sum of the risk budgets, which is $8.83 + $8.83 = $17.66 million, is greater than the total risk budget of $15.3 million. This is because the latter takes into account diversification effects. If the two managers were perfectly correlated with each other, the risk budget would have to be $15.3/2 = $7.65 million for each. This higher risk budget is beneficial for the investor because it creates more opportunities to take advantage of the managers' positive alphas.

TABLE 16.3 Risk Budgeting

Asset		Expected Return	Volatility	Correlations			Percentage Allocation	VAR
				1	2	3		
U.S. stocks	1	13.80%	15.50%	1.00			60.0	$15.3
U.S. bonds	2	8.40%	7.40%	0.20	1.00		7.7	$0.9
Non-U.S. bonds	3	9.60%	11.10%	0.04	0.40	1.00	32.3	$5.9
Portfolio		12.00%	10.30%				100.0	$16.9

The risk budgeting process highlights the importance of correlations across managers. To control their risk better, institutional investors often choose equity managers that follow different market segments or strategies. For example, the first manager could invest in small growth stocks, the second in medium-size value stocks. Or the first manager could follow momentum-based strategies, the second value-based strategies. The first type tends to buy more of a stock after its price has gone up, and the second after the price has become more attractive (i.e., low). Different styles lead to low correlations across managers. For a given total risk budget, low correlations mean that each manager can be assigned a higher risk budget, leading to a greater value added for the fund.

These low correlations explain why investors much watch for style drift, which refers to a situation where an investment manager changes investment style. This is a problem for the investor because it can change the total portfolio risk. If all the managers, for instance, drift into the small-growth category, the total risk of the fund will increase. Style drift is controlled by the choice of benchmarks with different characteristics, such as small-growth and medium-value indices, and by controls on the tracking error volatility for each manager.

In conclusion, this risk budgeting approach is spreading rapidly to the field of investment management. This approach provides a consistent measure of risk across all subportfolios. It forces managers and investors to confront squarely the amount of risk they are willing to assume. It gives them tools to monitor their risk in real time.

16.3.2 Marginal Risk and Contribution to Risk

A well-designed risk system should also provide tools to understand how to manage risk. A risk report should display measures of marginal risk. This represents the change in risk due to a small increase in one of the allocations. Using the volatility of returns are the risk measure, this is

$$\text{MRISK} = \frac{\partial \sigma_P}{\partial w_i} = \frac{\text{cov}(R_i, R_p)}{\sigma_P} = \beta_{i,P} \sigma_P \qquad (16.15)$$

Thus, beta represents the marginal contribution to the risk of the total portfolio P. A large value for β indicates that a small addition to this position will have a relatively large effect on the portfolio risk. Conversely, positions with large betas should be cut first because they will lead to the greatest reduction in risk.

This can be expanded to measure contributions to the portfolio risk. The risk contribution, or risk allocation, is obtained by multiplying the marginal risk for position i by its weight i

$$\text{CRISK} = w_i \beta_{i,P} \sigma_P \qquad (16.16)$$

Because the beta of a portfolio with itself is one, the sum of $w_i \beta_{i,P}$ is guaranteed to be one. Hence, the sum of the risk contributions adds up exactly to the total

TABLE 16.4 Risk Analysis

Asset	Volatility	Market Allocation	Marginal Risk	Risk Allocation
U.S. stocks	15.50%	60.0%	0.1438	8.63%
U.S. bonds	7.40%	7.7%	0.0278	0.21%
Non-U.S. bonds	11.10%	32.3%	0.0451	1.46%
Portfolio	10.30%	100.0%		10.30%

portfolio risk, σ_P. When risk is expressed in terms of VAR, this measure is called component VAR.

Table 16.4 gives an example, expanding on the previous table. The marginal risk column shows that U.S. stocks are the asset class with the greatest marginal contribution to the risk of the portfolio. As an example, increasing the allocation from 60% to 61% increases the portfolio risk from 10.30% to 10.44%, which is an increase of 0.14%. This is precisely the marginal risk number of 0.14 multiplied by the 1% weight increase.

The last column shows the risk contribution, or allocation. Out of a total portfolio risk of 10.30%, 8.63% is attributed to U.S. stocks. This high number reflects the high volatility of this asset class, its high weight in the portfolio, as well as correlations. Reporting systems should therefore display not only the conventional weights, or market allocations, but also risk allocations.

Such analysis provides useful insights into the structure of the portfolio. Given a scarce risk budget, high risk allocations can only be justified by expected returns that are high relative to other assets. In fact, an exact relationship holds for portfolios that are mean-variance efficient, i.e., maximize the Sharpe ratio. If this is the case with portfolio P, then the ratio of excess returns on all assets to their marginal risk, which is also proportional to the Treynor ratio, must be the same. On the other hand, if P is not efficient, then we should be able to improve its performance by tilting toward assets that provide a greater ratio of expected return to their contribution to risk. Thus, this top-down analysis of portfolio risk can help investors improve the performance of their portfolios, given a set of risk measures and asset class forecasts.

EXAMPLE 16.8: PENSION FUND RISK

The AT&T pension fund reports total assets worth $19.6 billion and liabilities of $17.4 billion. Assume the surplus has a normal distribution and volatility of 10% per annum. The 95% Surplus at Risk over the next year is

a. $360 million

b. $513 million

c. $2,860 million

d. $3,220 million

EXAMPLE 16.9: RISK BUDGETING

The AT&T pension fund has 68%, or about $13 billion invested in equities. Assume a normal distribution and volatility of 15% per annum. The fund measures absolute risk with a 95%, one-year VAR, which gives $3.2 billion. The pension plan wants to allocate this risk to two equity managers, each with the same VAR budget. Given that the correlation between managers is 0.5, the VAR budget for each should be

a. $3.2 billion
b. $2.4 billion
c. $1.9 billion
d. $1.6 billion

EXAMPLE 16.10: FRM EXAM 2005—QUESTION 140

Suppose a portfolio consists of four assets. The risk contribution of each asset is as follows: UK Large Cap, 3.9%; UK Small Cap, 4.2%; UK Bonds, 0.9%; Non-UK Bonds, 1.1%. Which of the following explanations would not be a possible explanation for the relatively high risk contribution values for UK equities?

a. High expected returns on UK equities
b. High weights on UK equities
c. High volatilities of UK equities
d. High correlation of UK equities with all other assets in the portfolio

16.4 IMPORTANT FORMULAS

Absolute risk: $\sigma(\Delta P) = \sigma(\Delta P / P) \times P = \sigma(R_P) \times P$

Relative risk: $\sigma(e)P = [\sigma(R_P - R_B)] \times P = [\sigma(\Delta P / P - \Delta B / B)] \times P = \omega \times P$

Tracking error volatility (TEV): $\omega = \sigma(\Delta P / P - \Delta B / B)$

Sharpe ratio (SR): $SR = \frac{[\mu(R_P) - R_F]}{\sigma(R_P)}$

Risk-adjusted performance (RAP): $RAP_P = R_F + \frac{\sigma_B}{\sigma_P}[\mu(R_P) - R_F]$

Information ratio (IR): $IR = \frac{[\mu(R_P) - \mu(R_B)]}{\omega}$

Alpha, from the intercept in: $R_{P,t} - R_{F,t} = \alpha + \beta_P[R_{M,t} - R_{F,t}] + \epsilon_{P,t}$

Treynor ratio (TR): $TR = \frac{[\mu(R_P) - R_F]}{\beta_P}$

Market timing skill, positive δ in:

$$R_{P,t} - R_{F,t} = \alpha_P + \beta_P[R_{M,t} - R_{F,t}] + \delta_P[R_{M,t} - R_{F,t}]^2 + \epsilon_{P,t}$$

Marginal risk: The change in total portfolio risk due to a small change in position i

$$\text{MRISK} = \frac{\partial \sigma_P}{\partial w_i} = \frac{\text{cov}(R_i, R_p)}{\sigma_P} = \beta_{i,P}\sigma_P$$

Risk contribution: A component of total portfolio risk due to one position

$$\text{CRISK} = w_i \beta_{i,P}\sigma_P$$

16.5 ANSWERS TO CHAPTER EXAMPLES

Example 16.1: Absolute and Relative Risk

c. This is an example of risk measured in terms of deviations of the active portfolio relative to the benchmark. Answers a. and b. are incorrect because they refer to absolute risk. Answer d. is also incorrect because it refers to the absolute risk of the benchmark.

Example 16.2: Pension Liabilities

d. We can compute the modified duration of the liabilities as $D^* = -(\Delta P / P)/\Delta y = -(0.8/17.4)/0.0005 = 9.2$ years. So, the liabilities behave like a short position in a bond with a duration around 9 years. Answers a. and b. are incorrect because the liabilities have fixed future payoffs, which do not resemble cash flow patterns on equities nor cash. Answer c. is incorrect because the duration of a bond with a nine-year maturity is less than nine years. For example, the duration of a 6% coupon par bond with nine-year maturity is seven years only.

Example 16.3: FRM Exam 2006—Question 25

c. Answer I. is incorrect because this liability is similar to a *short* (not long) position in a bond. Answer II. is correct because surplus at risk is a relative risk measure, assets minus liabilities. Answer III. is incorrect because it needs to invest $8.8 billion, not $8 billion.

Example 16.4: Sharpe and Information Ratios

d. The Sharpe ratios of the portfolio and benchmark are $(10\% - 3\%)/20\% = 0.35$, and $(8\% - 3\%)/14\% = 0.36$, respectively. So, the SR of the portfolio is lower than that of the benchmark. Answer a. is incorrect. The TEV is the square root of $20\%^2 + 14\%^2 - 2 \times 0.98 \times 20\% \times 14\%$, which is $\sqrt{0.00472} = 6.87\%$. So, the IR of the portfolio is $(10\% - 8\%)/6.87\% = 0.29$. This is positive, so answer b. is incorrect. Answer c. is the SR of the portfolio, not the IR, so it is incorrect.

Example 16.5: FRM Exam 2007—Question 132

c. The SR considers total risk, which includes systematic and unsystematic risks, so a. and b. are correct statements, and incorrect answers. Similarly, the SR is derived

from the CML, which states that the market is mean-variance efficient and hence has the highest Sharpe ratio of any feasible portfolio. Finally, the SR can be used to evaluate undiversified portfolios, precisely because it includes idiosyncratic risk.

Example 16.6: Performance Evaluation

c. Because the market went up, a portfolio with positive beta will have part of its positive performance due to the market effect. A portfolio with negative beta will have in part a negative performance due to the market. Answer a. is incorrect because the fund manager could still have generated some of its alpha through judicious stock-picking. Answers b. and d. are incorrect because a negative beta combined with a market going up should lead to a decrease, not an increase, in the alpha.

Example 16.7: FRM Exam 2005—Question 103

b. The database includes histories before 1994 and therefore suffers from backfill bias. Next, funds that stop reporting are deleted from the database, so this has survival bias. Errors-in-variables biases arise in other contexts, such as regression. Finally, the average of fund returns will be too high (not too low) because of these two biases. Hence, I. and II. are correct.

Example 16.8: Pension Fund Risk

a. The fund's surplus is the excess of assets over liabilities, which $19.6 - $17.4 = $2.2 billion. The Surplus at Risk at the 95% level over one year is, assuming a normal distribution, $1.645 \times 10\% \times \$2,200 = \$360$ million. Answer b. is incorrect because it uses a 99% confidence level. Answers c. and d. are incorrect because they apply the risk to the liabilities and assets instead of the surplus.

Example 16.9: Risk Budgeting

c. Call x the risk budget allocation to each manager. This should be such that $x^2 + x^2 + 2\rho\ xx = \$3.2^2$. Solving for $x\sqrt{1 + 1 + 2\rho} = x\sqrt{3} = \$3.2$, we find $x = \$1.85$ billion. Answer a. is incorrect because it refers to the total VAR. Answer b. is incorrect because it assumes a correlation of zero. Answer d. is incorrect because it simply divides the $3.2 billion VAR by 2, which ignores diversification effects.

Example 16.10: FRM Exam 2005—Question 140

a. The risk contribution is proportional to the weight times the beta. The latter involves the correlation between the asset and the portfolio, as well as the volatility of the asset. Higher weight, correlation, and volatility would create higher risk contribution. On the other hand, high expected returns would explain a high weight, but not risk contribution.

Hedge Fund Risk Management

The first hedge fund was started by A.W. Jones in 1949. Unlike the typical equity mutual fund, the fund took long *and* short positions in equities. Over the last decades, the hedge fund industry has undergone exponential growth. As of December 2007, hedge funds accounted for more than $1,900 billion in equity capital, called assets under management (AUM). Hedge funds are private partnership funds that have very few limitations on investment strategy. As a result, they can take long and short positions in various markets and also allow the use of leverage. Due to this leverage, the assets they control are greater than their AUM. Hedge funds have become an important force in financial markets, accounting for the bulk of trading in some markets.

Unlike mutual funds, who are open to any investor, hedge funds are accessible only to accredited investors. This is because of their perceived risk, which can be traced to their use of leverage and short positions. To control their risk, most hedge funds have also adopted risk controls using position-based, VAR-type techniques. Because some types of hedge fund strategies are very similar to those of proprietary trading desks of commercial banks, it was only natural for hedge funds to adopt similar risk management tools.

The purpose of this chapter is to provide an overview of risk management for the hedge fund industry. Section 17.1 gives an introduction to the hedge fund industry. Section 17.2 presents the mechanics of shorting and various measures of leverage. Section 17.3 then analyzes commonly used strategies for hedge funds and shows how to identify and measure their risk. The risk factors that are largely specific to hedge funds are presented in Section 17.4. Section 17.5 shows how to deal with hedge fund risk. Finally, Section 17.6 discusses the general role of hedge funds in financial markets.

17.1 THE HEDGE FUND INDUSTRY

The growth of the hedge fund industry is described in Figure 17.1. By now, there are close to 10,000 hedge fund managers controlling close to $1,900 billion in equity capital, also called net assets, up from $30 billion in 1990. This represents
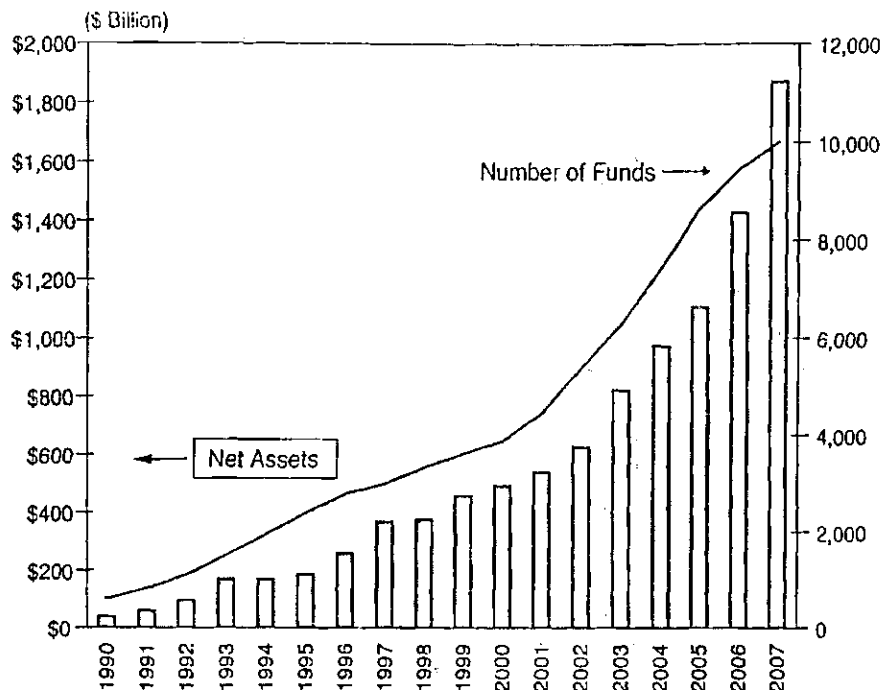
FIGURE 17.1 Growth of Hedge Fund Industry
Source: Hedge Fund Research. Data as of December of each year.

an annualized rate of growth of 25%. In comparison, U.S. mutual funds currently manage $12 trillion, up from $1.1 trillion in 1991. This represents an annualized rate of growth of 15%. Thus, hedge funds have grown at twice the rate of mutual funds over the same period. In 2008, however, the credit crisis will cause a contraction of the entire asset management industry.

The growth of this industry is due to a number of factors. On the investor side, the performance of hedge funds has been attractive, especially compared to the poor record of stock markets during the 2000–2002 period. Hedge funds also claim to have low beta, which makes them useful as diversifiers.

On the manager side, hedge funds provide greater remuneration than traditional investment funds. Typical investment management fees for mutual funds range from a fixed 0.5% to 2% of AUM. In contrast, hedge funds commonly charge a fixed management fee of 1% to 2% of assets plus an incentive fee of 20% of profits.

Hedge funds also typically have fewer restrictions on their investment strategy and are less regulated, giving more leeway to portfolio managers. More flexible investment opportunities include the ability to short securities, to leverage the portfolio, to invest in derivatives, and generally to invest a cross a broader pool of assets. The lighter regulatory environment creates an ability to set performance fees, lockups periods, or other forms of managerial discretion.

The unprecedented turbulence of 2008, however, has hit the hedge fund industry hard. Many funds posted poor performance and suffered widespread investor redemptions, leading to many hedge fund closures. Even so, hedge funds in general suffered only half the loss of equities.

> ### Example: Computing Fees
>
> The most common fee structure consists of (1) a management fee that is a fixed proportions of assets under management, typically 2%, and (2) an incentive fee, which is typically 20% of the profits over a year when positive. Sometimes, the incentive fee is paid after the performance exceeds a hurdle rate, such as LIBOR.
>
> As an example, assume that the Net Asset Value (NAV) goes from 100 to 120 over the year before fees (gross) and that LIBOR is at 5%. Fees are at the usual 2 and 20. The total fee is $100 \times 2\% + [(120 - 102) - 0.05 \times 100] \times 20\%$. In this case, the gross return is 20%, the total fees to the manager are 4.6%, and the net return is 15.4%. Without a hurdle, the net return would be 14.4%.

17.2 LEVERAGE, LONG, AND SHORT POSITIONS

Hedge funds can achieve leverage and implement short-sales through their prime broker (PB). PBs provide various back-office services to hedge funds, including trade reconciliation (clearing and settlement), custody services, and risk management, as well as record keeping. In addition, they provide credit lines for financing leverage and short selling capabilities.

To understand the mechanics of hedge funds, we need to describe how stock borrowing and margins work. In typical corporate balance sheet analysis, balance sheet leverage is defined as the ratio of balance sheet assets over equity. This simplistic measure, however, assumes that all the risk is coming from the assets, or that the future value of liabilities is known. Such definition is not adequate for hedge funds, or most financial institutions, for that matter. In these cases, both assets and liabilities, long and short positions, are risky.

In what follows, we illustrate the use of long and short positions in stocks. This analysis, however, can be extended to any asset that can be shorted, subject to its own specific margin requirements.

17.2.1 Long Position

Let us start with the simplest case, which is a long position in a risky asset. Consider an investor with $100 (say millions) invested in one stock. This can be achieved with $100 of investor equity. Or the investor can borrow. Suppose the broker requires a 50% margin deposit, which is the minimum requirement under Regulation T in the United States. The investor needs to invest only $50 and the remainder is provided by the broker, who gives a $50 loan. The balance sheet of the position is described below, with the risky entry in bold. Defining leverage L as the ratio of assets over equity, the leverage of this position is 2 to 1.

Assets	Liabilities
$100 Long stock	$50 Broker loan
	$50 Equity

The risk is that of a *decrease* in the value of the stock. For instance, a loss of $1, which is 1% of the value of the stock, translates into a $1 loss in the value of equity, which is a 2% loss in relative terms. Thus, movements in the asset value are magnified by the leverage factor L. If there is no leverage $(L = 1)$, the worst loss occurs when the stock price goes to zero.

The rate of return on the equity is the summation of L times the rate of return on the long stock position R_S minus $(L - 1)$ times the cost of the loan, R_F:

$$R_E = LR_S - (L - 1)R_F = R_F + L(R_S - R_F) \tag{17.1}$$

Hence, the volatility of equity will be L times that of the stock position. Similarly, the beta and idiosyncratic volatility are multiplied by the leverage:

$$\beta_E = L\beta_S \tag{17.2}$$

Leverage amplifies returns but also creates more risk.

Note that leverage can also be obtained by using derivatives, instead of cash instruments. This includes single stock futures, contracts for differences, or equity swaps. If a stock futures position can be entered with a margin of only 10%, the economically equivalent liabilities would consist of a loan of $90 plus equity of $10. The dollar exposure is still the same, at $100, but the leverage is now much higher than before, at 10 to 1.

17.2.2 Short Position

Consider next a situation where the investor wants to short the stock instead. Under a stock loan agreement, the owner of a stock lends the stock to our investor in exchange for cash and a future demand to get the stock back. In the meantime, the investor must pass along any cash flow on the stock, such as dividends, to the original owner.[1] When the operation is reversed, the stock lender returns the cash plus the short-term interest rate minus a stock loan fee. This is typically 20 basis points (bp) for most stocks but can reach 400bp for stocks that are hard to borrow (said to be "on special"). In the meantime, the stock lender will have invested the cash, thus earning a net fee of 20bp. From the viewpoint of the stock lender, this is an easy way to increase the return on the stock by a modest amount.

The stock borrower will now sell the stock in anticipation of a fall in the price. The sale, however, will go through a broker, who will not allow the seller to have full access to the sales proceeds. In the United States, under Regulation T, the broker keeps 50% of the sales proceeds. This margin, which can be posted as any security owned clear by the investor, imposes a limit on the investor's leverage.

So, the investor receives $100 worth of stocks, sells it, and keeps at least $50 as margin with the broker. The hope is for a fall in the stock price, so that the stock can be repurchased later at a lower price.

---

[1] Traditional stock loans are made on a day-to-day basis. The lender can demand the return of the stock at any time, with a three-day period for delivery.

All of the cash flows are arranged at the same time. The investor needs to send $100 to the stock lender, half of which will come from the remaining proceeds and the other half from the equity invested, or own funds. The balance sheet for the short position is described below, with the risky entry in bold. Here, leverage can be defined as the ratio of the absolute value of the short position to the equity, which is 2 to 1. As in the previous long-only case, we have a position of $50 in equity leveraged into a position of $100 in stocks. Regulation T imposes a maximum leverage ratio of 2, which is the inverse of the 50% of the short-sales proceeds kept by the broker.

Assets	Liabilities
$100 Cash lent to stock owner	$100 Short stock
$50 Margin at the broker	$50 Equity

Here, the risk is that of an *increase* in the value of the stock. If the stock price goes up by $1, or 1%, the equity loses $1, which is 2% in relative terms. This ratio equal the leverage of 2. The beta of the equity is now negatively related to that of the stock:

$$\beta_E = -L\beta_S \qquad (17.3)$$

Short positions are intrinsically more risky than long positions, however. This is because the distribution of prices is asymmetrical. The price has a lower bound of zero but has unlimited upper values, albeit with decreasing probabilities. With a long position, the most that could be lost is $100 million. With a short position, the price could go from $100 to $200 or even higher, in which case the dollar loss would exceed $100 million.

17.2.3 Long and Short Positions

Consider now a typical hedge fund, which has both long and short positions. Say the initial capital is $100. This is the equity, or net asset value (NAV). The fund could buy $100 worth of stocks and short $100 worth of stocks as above. Part of the long stock position can be used to satisfy the broker's minimum margin requirement of $50 for shorting the stock. The balance sheet for the long and short positions is described below, with the risky entries in bold.

Assets	Liabilities
$100 Long stock	
$100 Cash lent to stock owner	$100 Short stock
	$100 Equity

Let us now turn to traditional risk measures. Define V_L, V_S, and V_F as the (absolute) dollar values of the long stock positions, short stock positions, and equity, respectively. V_A is the value of total assets. If β_L and β_S are the betas of

the long and short stock positions, the total dollar beta is

$$(\beta_L \, V_L - \beta_S \, V_S) = \beta_E \, V_E \tag{17.4}$$

which defines the net beta of equity, or β_E. This net measure of systematic risk, however, ignores idiosyncratic risk.

Traditional leverage is commonly used as a risk measure

$$\text{Leverage} = \frac{V_A}{V_E} = \frac{\text{Long stock positions plus cash}}{\text{Equity}} \tag{17.5}$$

In our example, this is ($100 + $100)/($100) = 2. If cash is ignored, long leverage is 1. This, however, ignores the hedging effect of short stock positions, so it is inadequate.

Using gross amounts, gross leverage is

$$\text{Gross Leverage} = \frac{V_L + V_S}{V_E}$$

$$= \frac{\text{Long positions plus absolute value of short positions}}{\text{Equity}} \tag{17.6}$$

In our example, this is ($100 + $100)/($100) = 2.

Gross leverage is often used as a rough measure of hedge fund risk. This measure, however, fails to capture the systematic risk of the equity position adequately. If the long and short positions have the same value and market beta, the net beta is zero, so there is no directional market risk. In the limit (even though there would be no reason to do so), if the long and short positions are invested in the same stock, there is no risk. Yet, gross leverage is high.

Another definition sometimes used is net leverage, which is

$$\text{Net Leverage} = \frac{V_L - V_S}{V_E}$$

$$= \frac{\text{Long positions minus absolute value of short positions}}{\text{Equity}} \tag{17.7}$$

In our example, this is ($100 − $100)/($100) = 0.

Net leverage is also inadequate as a risk measure. Although it roughly accounts for systematic risk, it fails to take into account potential divergences in the value of the long and short positions. It is only appropriate under restrictive assumptions. For example, if the betas of the long and short positions are the same, then the equity beta is

$$\beta_E = \frac{\beta_L(V_L - V_S)}{V_E} = \beta_L \times \text{Net Leverage} \tag{17.8}$$

so this net leverage term measures the multiplier applied to the beta of the long position. This totally ignores idiosyncratic risk, however, which is precisely the type of risk that the hedge fund manager should take views on. In conclusion, these leverage measures should only be viewed as rough indicators of risk. They are robust and easy to compute, however.

This is why the industry has moved to more comprehensive position-based risk measures. VAR, for example, accounts for the size of positions, volatilities, as well as correlations between assets and liabilities. As such, it is a superior measure of the risk of loss.

EXAMPLE 17.1: FRM EXAM 2006—QUESTION 41

A hedge fund is long $315 million in certain stocks and short $225 million in other stocks. The hedge fund's equity is $185 million. The fund's overall beta is 0.75. Calculate the gross and net leverage.

a. 2.91 and 0.48
b. 2.18 and 0.36
c. 2.91 and 0.36
d. 2.18 and 0.48

EXAMPLE 17.2: HEDGING AND RETURNS

Continuing with the previous question, assume the stock market went up by 20% last year. Ignore the risk-free rate and idiosyncratic risk, and assume the average beta of both long and short positions is one. Over the same period, the return on the fund should be about

a. 20%
b. 15%
c. 10%
d. 5%

EXAMPLE 17.3: FRM EXAM 2004—QUESTION 2

A relative value hedge fund manager holds a long position in Asset A and a short position in Asset B of roughly equal principal amounts. Asset A currently has a correlation with Asset B of 0.97. The risk manager decides to overwrite this correlation assumption in the variance-covariance based VAR model to a level of 0.30. What effect will this change have on the resulting VAR measure?

a. It increases VAR.
b. It decreases VAR.
c. It has no effect on VAR, but changes profit or loss of strategy.
d. Do not have enough information to answer.

17.3 HEDGE FUNDS: MARKET RISKS

17.3.1 Types of Market Risks

Hedge funds are a much more heterogeneous group of investment managers than others. They follow a great variety of strategies, which can be classified into different styles. More generally, they can be categorized into taking directional or nondirectional risks.

- Directional risks involve exposures to the direction of movements in major financial market variables. These directional exposures are measured by first-order or linear approximations such as
  - Beta for exposure to general stock market movements
  - Duration for exposure to the level of interest rates
  - Spread duration for exposure to movements in credit spreads
  - Delta for exposure of options to the price of the underlying asset
- Nondirectional risks involve other remaining exposures, such as nonlinear exposures, exposures to hedged positions, and exposures to volatilities. These nondirectional exposures are measured by exposures to differences in price movements, or quadratic exposures such as
  - Basis risk, which involves differences in prices of related assets
  - Convexity risk, which involves quadratic effects for interest rates
  - Gamma risk, which involves quadratic effects with options
  - Volatility risk, which involves movements in the volatility

Directional trades can take long or short positions on the major risk factors, such as equities, currencies, fixed-income instruments, and commodities. As a result, directional positions have greater volatilities than nondirectional ones. For funds that take directional risks, total portfolio risk is controlled through diversification across sources of risks, across trading strategies, and with risk limits.

Many categories of hedge funds are hedged against directional risks. As a result, they are exposed to nondirectional risks. Such strategies need to take long *and* short positions in directional trades. The example we gave in the previous section was long $100 in a stock offset by a short position worth $100 in another stock. Such strategy has little directional risk to the stock market, but is exposed to changes in the relative value of the two stocks. Limiting risk also limits rewards, however. As a result, nondirectional strategies are often highly leveraged in order to multiply gains from taking nondirectional bets.

17.3.2 Hedge Fund Styles

Hedge funds can be classified into various styles, reflecting the type of trading and markets they are exposed to. Table 17.1 lists various hedge fund styles. To some extent, this classification is arbitrary. Definitions of categories vary within the industry. Different hedge fund index providers, for example, use different

TABLE 17.1   Hedge Fund Styles

Style	AUM ($b)	Nb. of Funds	Risk (%pa)	Description
Directional Strategies				
Long/short equity	424	975	11%	Combination of long and short equity positions with net long bias
Emerging markets	115	205	12%	Equity and bond positions in emerging countries, with net long bias
Global macro	122	148	11%	Long and/or short positions across all asset classes
Nondirectional Strategies				
Relative value:				
Equity market neutral	71	183	6%	Combination of long and short equity positions with net beta close to zero
Fixed-income arbitrage	59	138	6%	Offsetting long and short positions in fixed-income securities
Convertible arbitrage	39	79	5%	Long positions in convertible bonds hedged for stock risk and interest risk
Event driven:	318	289	6%	
Merger arbitrage, Distressed securities, Credit hedging				Positions driven by corporate events such as mergers, reorganizations, and bankruptcy proceedings
Fund Structure				
Managed futures	57	235	17%	Positions in futures and option contracts (includes CTAs)
Multistrategy	181	262	9%	Combinations of hedge fund strategies in the same fund
Funds of funds		872	6%	Diversified portfolios of hedge funds

Source: TASS database, sample of live funds reporting in U.S. dollars as of December 2007. Risk is cross-sectional average of annualized volatility over the last four years.

classifications, even though the underlying pool of hedge funds is similar. Classifications can also lose meaning if hedge funds managers change strategies over time.

The table also reports the number of existing funds in each group, as well as their typical risk, measured as the annual standard deviation averaged across all funds.[2] Styles are listed in order of decreasing risk.

Table 17.2 presents the performance of typical hedge fund indices, measured over the period 1994 to 2007. CSFB builds each sector index as a value-weighted average of eligible funds. The table shows the compound growth, volatility, beta to the S&P stock index, return in excess of cash, and the alpha from a regression on equities (as discussed in the previous chapter).

The table shows that the overall hedge fund index returned 10.3% over this period, which is above the 9.1% return of the S&P index, with much lower

---

[2] Note that the risk measures are for live funds only. Hence, the data are subject to survivorship bias. The risk of existing funds is less than that of dead funds.

TABLE 17.2   Hedge Fund Performance: CSFB Indices, 1994 to 2007

	Growth	Volatility	Beta	Exc. Ret.	Alpha
Overall Index	10.3%	7.5%	0.26	6.0%	4.7%
Sectors:					
Long/short equity	11.2%	9.8%	0.42	6.9%	4.9%
Short biased	−1.1%	16.9%	−0.91	−5.4%	−0.9%
Emerging markets	8.7%	15.4%	0.52	4.5%	2.0%
Global macro	13.5%	10.4%	0.15	9.2%	8.5%
Equity market neutral	9.7%	2.8%	0.07	5.4%	5.1%
Fixed-income arbitrage	5.6%	4.1%	0.02	1.3%	1.2%
Convertible arbitrage	7.7%	4.9%	0.06	3.4%	3.1%
ED-Merger arbitrage	7.6%	4.1%	0.13	3.3%	2.7%
ED-Distressed securities	12.3%	6.2%	0.24	8.0%	6.8%
ED-Credit hedging	10.4%	6.1%	0.21	6.1%	5.1%
Managed futures	6.7%	12.0%	−0.11	2.4%	2.9%
Multistrategy	9.1%	4.4%	0.04	4.8%	4.6%
Benchmarks:					
Cash	4.3%	0.0%	0.00	0.0%	0.0%
S&P Index	9.1%	14.1%	1.00	4.8%	0.0%
Treasury Index	6.1%	4.5%	−0.04	1.8%	
High Yield Index	6.4%	6.7%	0.26	2.1%	

Notes: Excess returns are measured relative to one-month London Interbank Deposit (LIBID). Alpha is measured from a market model regression on the S&P index.

volatility. The index, however, has slightly positive beta of 0.26, so part if its performance is due to the equity premium. The last column shows an alpha of 4.7%, which is significant.

Also note that the volatility in this table is not directly comparable to that in Table 17.1, because this is the volatility of a portfolio, instead of the average fund volatility. In contrast, the beta of a portfolio is a weighted average of the fund betas, so the portfolio beta gives a good indication of the typical beta of individual funds.[3]

Long/Short Equity   The first category consists of directional strategies. These include long/short equity funds, which as Table 17.1 shows, is the most prevalent strategy. These funds are not market neutral. Most have a long bias, e.g., 100% of NAV in long positions, and 50% in short positions. Table 17.2 indeed shows a beta of 0.42.

A related category consists of short biased funds, which are net short. Table 17.2 shows a negative beta, close to −1. Another category is emerging markets, which consists of equity and bond positions in emerging countries, such as Brazil, Russia, India, and China.

---

[3] In addition, the numbers are not directly comparable across the two tables because they are measured over different periods.

These funds are exposed to the general market risk factor, in addition to sector and idiosyncratic risks. Because of leverage, volatility is high, at 11% on average across all such funds. This is on the order of the volatility of an unleveraged position in the S&P 500.

Global Macros Next are global macros funds, which take directional, leveraged bets on global asset classes, equities, fixed-income, currencies, and commodities. Because they span so many markets, these funds do not have a homogeneous risk profile. An example is the Soros fund that shorted the British pound against the German mark just before its devaluation, leading to a reported gain of $1 billion for the hedge fund. This group is close to global tactical asset allocation (GTAA), which is a traditional investment manager category. GTAA managers take positions across national stock markets, fixed-income markets, and currencies to take advantage of short-term views, often through derivatives.

These funds are exposed to a number of general market risk factors, in addition to sector and idiosyncratic risks. The average volatility is 11%. This is less than the previous category because these funds also invest in other markets, which are less volatile than equities.

We now turn to nondirectional strategies. The first three categories are sometimes called relative value funds, because they rely on comparisons of securities with similar characteristics, buying the cheap ones while selling the expensive ones in the hope of future convergence.

Equity Market Neutral The first group is equity market neutral funds, which attempt to maintain zero beta through balanced long and short positions in equity markets. These funds may or may not be neutral across other risk factors, including industries, styles, and countries.

So, these funds are exposed to these other risk factors (industries, styles, countries) in addition to idiosyncratic, stock-specific risks. Balance sheet leverage is typically three times on each side, i.e., both longs and shorts add up to 300% of equity. The average volatility is 6%, which is much less than that of equity indices, due to the hedging effect of the short positions.

Fixed-Income Arbitrage The next group is fixed-income arbitrage funds. This is a generic term for a number of strategies that involve fixed-income securities and derivatives. The hedge fund manager assesses the relative value of various fixed-income instruments. For instance, if the on-the-run bond is expensive relative to the off-the-run bond, the fund would buy the undervalued security and sell the expensive one. This position has a net duration close to zero but is exposed to the spread between the two securities. Other examples include taking positions in swap spreads, or in asset-backed securities when their option-adjusted spread is high. This group includes mortgage arbitrage.

These funds avoid directional exposures to interest rates but are exposed to other nondirectional risks, such as spread risk. Due to the small expected profit of

each trade, fixed-income arbitrage funds are highly leveraged, with leverage ratios ranging from 10 to 25.

Example: LTCM's Bet

Long-Term Capital Management (LTCM) started as a fixed-income arbitrage fund, taking positions in relative value trades, such as duration-matched positions in long swap, short Treasuries. It started the year 1998 with $4.7 billion in equity capital.

On August 21, 1998, the 10-year Treasury yield dropped from 5.38% to 5.32%. The swap rate, in contrast, increased from 6.01% to 6.05%. This divergence was highly unusual. Assuming a notional position of $50 billion and modified duration of eight years, this leads to a value change of $-8 \times (5.32 - 5.38)/100 \times \$50,000 = +\$240$ million on a long Treasury position and $-8 \times (6.05 - 6.01)/100 \times \$50,000 = -\$160$ million on a long swap position. As the spread position is long the swap and short Treasury, this leads to a total loss of $400 million, close to 10% of capital.

LTCM also took positions in option markets, selling options when they were considered expensive and dynamically hedging to maintain a net delta of zero. Implied volatilities went up sharply on August 21, leading to further losses on the option positions. On that day, LTCM's reported loss was $550 million.

The average volatility of this group is 6%. The distribution of payoffs is typically asymmetric, however. Swap spreads, for example, cannot narrow below zero but can increase to very large values, and have done so. This asymmetry in the distribution of spreads is reflected in that of profits. Such hedge funds have negatively skewed distribution. When they lose money, they can lose large amounts.

Convertible Arbitrage   The last group in the relative value category is convertible arbitrage funds. The hedge fund manager assesses the relative value of convertible bonds using proprietary option pricing models. If the convertible bond is cheap, the hedge fund buys the bond while hedging the major risks.

Because a convertible bond involves a long call option position, it has positive delta with respect to the underlying stock. Therefore, the manager should short the stock to bring the net delta of the position close to zero. Typically, interest rate risk is hedged by shorting Treasury bonds, or T-bond futures. Sometimes, credit spread risk is hedged by buying credit default swaps.

These funds avoid directional exposures to interest rates but are exposed to other nondirectional risks, such as spread risk. Being typically long convertible bonds, the long option position creates positive gamma and vega (long implied volatility). The bond position creates positive convexity, unless the bond is callable. This strategy is also exposed to corporate event risk, such as default (if not hedged) and takeover. Leverage is moderate. Typically, the long convertible

bond position is no more than three times equity. The average volatility of this group is 5%, which is fairly low, in part because of illiquidity.

Event Driven   The next group includes event-driven funds, which attempt to capitalize on the occurrence of specific corporate events. This group includes merger arbitrage funds and distressed securities funds.

Let us focus first on merger arbitrage funds, also known as risk arbitrage funds. Mergers and acquisitions are transactions that combine two firms into one new firm.[4] The parties can be classified as the acquiring firm, or bidder, which initiates the offer and the target firm, or acquired firm, which receives the offer. The bidder offers to buy the target at a takeover premium, which is the difference between the offer price and the target's stock price before the bid. This premium is typically high, averaging 50% of the initial share price.

Upon the announcement of the merger, the price of the target firm reacts strongly, increasing by, say, 40%. This still falls short of the takeover price, due to the uncertainty as to whether the transaction will occur. The completion rate is 83% on average, so there is always a possibility the transaction could fail. When this happens, the target firm typically suffers a large price drop. As a result, it is important to diversify by spreading the portfolio over many deals.

Offers can take the form of cash or stock of the bidding company. For a cash deal, the risk arbitrage position simply consists of buying the target's stock, and hoping the price will eventually move to the takeover price. For a stock deal, the bidder offers to exchange each target share for Δ shares of the bidder. The risk arbitrage position then consists of a long position in the target offset by a short position of Δ in the bidder's stock. These positions generate an average annualized excess return of 10%.[5]

The volatility of this group is relatively low. Because the stochastic process for the target's stock price changes after the announcement, traditional position-based risk measures are not appropriate measures of risk.[6]

Example: Exxon-Mobil Merger

On December 1, 1998, Exxon confirmed that it had agreed to buy Mobil, another major oil company, in a transaction valued at $85 billion, which was the biggest acquisition ever. The deal created the world's largest traded oil company, with a market capitalization of $250 billion. Under the terms of the agreement, each shareholder of Mobil would receive $\Delta = 1.32015$ shares of Exxon in exchange.

---

[4] These are sometimes called takeovers. Takeovers can take the form of mergers or tender offers. Mergers are negotiated directly with the target manager's, approved by the board of directors, and then by shareholder vote. Tender offers are offers to buy shares made directly to target shareholders.
[5] This is a risk-adjusted excess return. These profits, however, seem to be related to limits to arbitrage, as there are lower for firms that are large and have low idiosyncratic risk. See Baker, Malcolm and Serkan Savasoglu, 2002, Limited Arbitrage in Mergers and Acquisitions, *Journal of Financial Economics* 64, 91–115.
[6] See Jorion, P., 2008, Risk Management for Event-Driven Funds, *Financial Analysts Journal* 64, 61–73.

Before the announcement, the initial prices of Mobil and Exxon were $78.4 and $72.7, respectively, which implies a modest premium of (1.32016 × $72.7)/$78.4 − 1 = 22%. Over the three days around the announcement, Mobil's stock price went up by +6.9% to $84.2 and Exxon's price went down by −1.5% to $71.6. This stock price reaction is typical of acquisition announcements.

The exchange was consummated on November 30, 1999, after regulatory and shareholder approval. On that day, the respective stock prices for Mobil and Exxon were $104.4 and $79.3. Multiplying the latter by 1.32015, we get $104.7, which is close to the final stock price for Mobil. So, the two prices converged to the same converted value. The profit from the risk arbitrage trade was ($104.4 − $84.2) − 1.32016($79.3 − $71.6) = $10.0 per share.

Event-driven funds also include distressed securities funds, which take positions in securities, debt or equity, of firms in financial difficulty. In such situations, the hedge fund manager needs to assess the effect of restructuring or the bankruptcy process on the market price of the securities. This requires an evaluation of the financial situation of the firm, as well as a good understanding of legal issues involved. If, for instance, the debt of a bankrupt company trades at 40 cents on the dollar, the hedge fund would benefit if the total payment after reorganization is 50 cents. Such funds are also actively involved in the bankruptcy processes and the reorganization plans.

These funds are exposed to event risks, that is, that the takeover or reorganization fails. They may also be exposed to equity market risk and interest rate risk if these exposures are not hedged. Because distressed securities do not trade actively, there is also liquidity risk. Leverage for event-driven funds is low to moderate, no more than two times.

The average volatility for event-driven funds is 6%, which is fairly low. This, however, hides the fact that the distribution of payoffs is asymmetric. Typically, the upside is more limited than the downside, should the takeover or reorganization fail. So, these funds are short volatility, or exposed to rare events. Because of the unusual nature of the event, measures of risk based on historical returns can be inaccurate for forecasting risk.

Managed Futures Funds   The next category of hedge funds differs from others on the basis of the fund structure. Managed futures funds consist of managers who use commodity and financial futures and options traded on organized exchanges. Trading strategies often involve technical trading, where positions depend on patterns in price histories. Leverage is high, leading to high volatility.

These funds have directional exposures to all the markets that have listed futures contracts. Their risk factors overlap with global macro funds. GTAA strategies, for instance, often involve stock index and currency futures. The average volatility of this group is 17%, which is fairly high.

Multistrategy Funds   Next, multistrategy funds are hedge funds that cover combinations of previously described fund strategies. One advantage of such funds

is that they can reallocate capital quickly from strategy to another. They also provide automatic diversification across strategies. The average volatility of this group is 9%.

On the other hand, multistrategy funds trend to be more concentrated in one type of strategy, and do not provide as much diversification as funds of funds, described below. Amaranth, for example, initially started as a multistrategy fund focused on convertible bond arbitrage, then morphed into a natural gas trading operation, and eventually blew up. Because all strategies are run within the same fund, a large loss in one strategy may affect the capital of other strategies. In other words, strategies are not firewalled, unlike in the fund of funds structure.

Funds of Funds   Finally, funds of funds, also called multimanager funds, are portfolios of hedge funds. These add value by careful selection of styles and investment managers. They also perform essential functions, such as the due diligence process when evaluating new managers and their continuous monitoring. Funds of funds can take views on strategies, increasing allocation to strategies that are expected to perform better.

Funds of funds charge additional management fees on top of those levied by the underlying funds, typically around 1%. On the other hand, because of their size, funds of funds can negotiate lower fees from the hedge fund managers.

Relative to multistrategy funds, funds of funds have higher fees. This cost difference, however, is offset by the fact that a fund of funds has access to the best managers, who generally want to run their own fund, thus creating better performance.

Also, funds of funds have lower risk of losses due to blowups than multistrategy funds, where the entire investment can be lost, as in the case of Amaranth. This is for two reasons. First, funds of funds are generally better diversified across strategies than multistrategy funds. Second, the hedge funds in a fund of fund pool are legally separate from each other. As a result, a blowup in one fund will not contaminate the rest of the portfolio, as in the case of a multistrategy fund.

Funds of funds provide convenient access to a diversified portfolio of hedge funds. Because hedge funds have minimum investments amounts, such diversification is difficult to achieve for small mandates allocated to hedge funds. For example, a $100 million allocation to hedge funds can be realistically invested to at most 10 hedge funds. A typical fund of funds, in contrast, will invest in 50 funds. Funds of funds provide economies of scale in the due diligence and risk monitoring processes. Funds of funds can also negotiate greater capacity and better liquidity than other investors.

Table 17.1 shows that the average volatility of this group is 6%. This low number reflects effective diversification across managers and styles.

A related category is hedge fund indices. These are unmanaged, passive baskets of hedge funds. In practice, however, their fees are similar to those charged by funds of funds, unlike fees for indexed mutual funds, which are much lower, around 0.10%. Also relatively new are collateralized fund obligations (CFOs), which are pools of hedge funds whose total payoff is sliced into various tranches, much like CMOs.

This list makes it clear that hedge funds are a very heterogeneous group. They are exposed to a wide variety of risk factors, follow different trading rules, and have varying levels of leverage and risk. The common element, however, is the need to manage risk.

EXAMPLE 17.4: RISKS IN FIXED-INCOME ARBITRAGE

Identify the risks in a fixed-income arbitrage strategy that takes long positions in interest rate swaps hedged with short positions in Treasuries.

a. The strategy could lose from decreases in the swap-Treasury spread.
b. The strategy could lose from increases in the Treasury rate, all else fixed.
c. The payoff in the strategy has negative skewness.
d. The payoff in the strategy has positive skewness.

EXAMPLE 17.5: RISKS IN CONVERTIBLE ARBITRAGE

Identify the risks in a convertible arbitrage strategy that takes long positions in convertible bonds hedged with short positions in Treasuries and the underlying stock.

a. Short implied volatility
b. Long duration
c. Long stock delta
d. Positive gamma

EXAMPLE 17.6: RISKS IN MERGER ARBITRAGE—I

A major acquisition has just been announced, targeting Company B. The bid from Company A is an exchange offer with a ratio of 2. Just after the announcement, the prices of A and B are $50 and $90, respectively. A hedge fund takes a long position in Company B hedged with A's stock. After the acquisition goes through, the prices move to $60 and $120. For each share of B, the gain is

a. $30
b. $20
c. $10
d. $0 since the acquisition is successful

EXAMPLE 17.7: RISKS IN MERGER ARBITRAGE—II

Suppose the payoff from a merger arbitrage operation is $5 million if successful, −$20 million if not. The probability of success is 83%. The expected payoff on on the operation is

  a. $5 million
  b. $0.75 million
  c. $0 since markets are efficient
  d. Symmetrically distributed

EXAMPLE 17.8: FRM EXAM 2005—QUESTION 47

The Big Bucks Hedge fund has the following description of its activities. It uses simultaneous long and short positions in equity with a net beta close to zero. Which of the following statements about Big Bucks are correct?

I. It uses a directional strategy.
II. It is a relative value hedge fund.
III. This fund is exposed to idiosyncratic risks.
  a. I and II
  b. II and III
  c. I and III
  d. II only

17.4 HEDGE FUNDS: SPECIFIC RISKS

17.4.1 Agency Risk

Hedge fund managers act as agents for investors. This can cause misalignment of incentives, however. Incentive fees make a payment that is a fraction of the profits, if positive. As a result, the hedge fund manager is long an option. Because the value of an option increases with the volatility, the fund manager may have an incentive to increase the risk of the fund.

Another potential problem is that of style drift, which occurs when managers change their investment patterns or stray into new markets.

This type of behavior can be minimized in a number of ways. Most importantly, hedge fund managers should invest a large fraction of their personal wealth in the fund they manage. This lessens the incentive to take on too much risk.

Some risk monitoring occurs at the level of the prime broker (PB). Because the PB is primarily concerned about the risk of loss from lending to the hedge fund,

however, its interests do not align with those of fund investors. For example, a lender may use margin calls to force liquidation of the fund asset at distressed prices. As long as there is excess collateral, the lender would be protected, but at the expense of investors.

The incentive to take risk is also lessened with high water marks, also known as loss carryforward provisions.[7]

The manager only receives performance fees to the extent that the current NAV exceeds the highest NAV previously achieved. Suppose, for instance, that the NAV changes from $100 to $130 to $120 to $140 in four consecutive years. The first year, the performance fee would apply to $30. The second year, there is no performance fee because the fund lost money. The third year, the performance fee only applies to the portion of $140 in excess of $130, which is the highest previous NAV. This mechanism, however, may not provide complete protection if the watermark is too high. In this case, the fund manager may choose to close the fund and to start a new one (if investors can be found).

17.4.2 Liquidity and Leverage Risk

Hedge funds take leveraged positions to increase returns, especially with nondirectional trades such as fixed-income arbitrage, where the expected return on individual trades is generally low.

Perversely, this creates other types of risks, including liquidity risk. This strategy indeed failed for Long-Term Capital Management (LTCM), a highly leveraged hedge fund that purported to avoid directional risks. LTCM had a leverage ratio of 25 to 1. It had grown to $125 billion in assets, four times the asset size of the next larger hedge fund. Once the fund started to accumulate losses, it became difficult to cut positions given its size. LTCM also had to meet margin calls from brokers. The fund ended up losing $4.4 billion, or 92% of its equity.

Table 17.3 links sources of liquidity risk to a hedge fund balance sheet. Liquidity risk arises on the asset side and is a function of the size of the positions as well as of the price impact of a trade. On the liabilities side, funding risk arises when the hedge fund cannot rollover funding from its broker, or when losses in marked-to-market positions or increases in haircuts lead to cash outflows. This is often a major source of risk for hedge funds, as the failure to meet margin calls can lead the lender to seize the collateral, forcing liquidation of the fund. Finally, funding risk also arises when the fund faces investor redemptions.

TABLE 17.3 Sources of Liquidity Risk

Assets	Liabilities
Size of position	Funding
Price impact	Mark-to-market, haircuts
	Equity
	Investor redemptions

[7] Sometimes, a claw-back provision is included, which requires the fund manager to pay back performance fees when the value of the fund drops.

The price impact function is instrument-specific. For example, major currencies, large stocks, Treasury bills, and Treasury bonds are very liquid, meaning that large amounts can be transacted without too much effect on the price. Other markets are by nature less liquid. For instance, minor currencies, small stocks, and most corporate debt instruments are generally illiquid.

LTCM dealt with mostly liquid instruments but was exposed to liquidity risk due to the sheer size of its positions. This is why hedge funds often say they have a maximum capacity. Beyond that optimal size, trading becomes difficult due to market impact.

Another type of risk that is exacerbated by leverage is model risk. This occurs when the investment strategy relies on valuation or risk models that are flawed. Due to leverage, small errors in the model can create big errors in the risk measure. Indeed, LTCM's risk measurement system was deficient, leading to an fatal underestimation of the amount of capital required to support its positions.

Some categories of hedge funds have intrinsic liquidity risk because the instruments are thinly traded, implying a large price impact for most trades. This is the case with convertible bonds and especially so with distressed securities. Because these funds invest in thinly traded securities, liquidity risk arises even for small funds.

Typically, funds with greater liquidity risk impose a longer lockup period and redemption notice period. The former refers to the minimum time period during which investor money is to be held in the fund. The latter refers to the period required to notify the fund of an intended redemption. Lockup periods average three months, and can extend to five years. Advance notice periods average 30 days. Funds also often have gates, which limit the amount of withdrawals to a fraction of the net assets. In the extreme, funds might be able to impose an outright suspension of redemptions.

Instrument liquidity risk creates a major problem for the measurement of risk. Typically, funds report their net asset value at the end of each month. If transaction prices are not observed at the end of the month, the valuation may be using a price from a trade that occurred in the middle of the month. This price is called a stale prices because it is "old" and does not reflect a market-clearing trade on the day of reporting. Unfortunately, this will distort the reported NAVs as well as the risk measures.

The first effect is that the reported monthly volatility will be less than the true volatility. This is because prices are based on trades during the month, which is similar to an averaging process. Averages are less volatile than end-of-period values.

The second effect is that monthly changes will display positive autocorrelation. A movement in one direction will be only partially captured using prices measured during the month. The following month, part of the same movement will show up in the return. This positive autocorrelation substantially increases the risk over longer horizons. Consider for instance the extrapolation of a one-month volatility to two months. The usual adjustment factor is $\sqrt{T} = \sqrt{2} = 1.41$. With autocorrelation of $\rho = 0.5$, this adjustment factor is instead $\sqrt{(1 + 1 + 2\rho)} = \sqrt{2(1 + 0.5)} = 1.73$. The true risk is understated by $(1.73 - 1.41)/1.73 = 18\%$.

This effect increases with the length of the horizon. As a result, the annualized volatility presented in Table 17.1, which extrapolates monthly volatility using the square root of time, may understate the true annual risk. Long-term measures of risk must specifically account for the observed autocorrelation.

A third, related effect is that measures of systematic risk will be systematically biased downward. If the market goes up during a month, only a fraction of this increase will be reflected in the NAV, leading to beta measures that are too low. Corrections to the beta involve measuring the portfolio's beta with the contemporaneous market return plus the beta with respect to the one-month lagged return plus the beta with respect to the one-month future return. With thin trading, the sum of these three betas should be higher than the contemporaneous beta, and also closer to the true systematic risk.[8]

Leverage can create other problems, which can be classified as crowded trade risk. This arises when many leveraged investors are on the same side of a trade.[9] A loss in their portfolio may require them to post additional margin, which may be satisfied by several funds selling similar assets at the same time, which can create disruptions in markets. Apparently, this explains why many quant funds, which are generally equity market neutral (EMN) strategies driven by quantitative models, suffered heavy losses in August 2007. The story is that a large multistrategy fund lost money on credit trades and then liquidated its equity positions, because these are more liquid than others. This caused large losses in EMN portfolios, both on the long and short sides, which was highly unusual.

This is not just a problem with leverage, however. Any mechanistic trading rule that involves cutting positions after a loss is incurred may have similar effects if there is an imbalance between demand and supply. This includes, for instance, stop losses, which are equivalent to synthetically replicating long positions in options.

17.4.3 Leverage and Counterparty Risk

Leverage also create another type of risk, which is counterparty risk. Hedge funds that use leverage give collateral to prime brokers. Hypothecation is the pledge of client-owned securities in a margin account to secure a loan. The broker then has the right to rehypothecate the securities to another party. If the broker goes bankrupt, however, the rehypothecated assets become part of the claims against the broker.

In the case of the Lehman failure, $22 billion of out the $40 billion held by Lehman's European prime brokerage had been rehypothecated. As a results,

---

[8] This correction is called the *Dimson beta*. See Dimson, E. (1979), Risk Measurement When Shares are Subject to Infrequent Trading, *Journal of Financial Economics* 7, 197–226.

[9] Of course, other investors must be on the other side of the trade. This classification supposes that the other side is not so leveraged, otherwise trades could be crossed without much effect on prices. In practice, positions are confidential, and it is impossible to know who is on which side of a trade, except anecdotally.

hedge funds trying to reclaim these assets found themselves in the line of general creditors with claims against Lehman.

17.4.4 Fraud Risk

A last issue, especially with complex or illiquid assets, is improper valuation of assets. This problem arises when assets do not have market-clearing prices at the end of the reporting period and when fund managers calculate the NAV themselves. As a result, some unscrupulous hedge fund managers have succumbed to the temptation to misreport the value of the fund's assets in order to hide their trading losses.[10] Others have even stolen investors' assets.

Indeed, a recent study has shown that valuation problems played a role in 35% of hedge fund failures, and that 57% of those valuation problems were caused by fraud or misrepresentation.[11] The growth of the hedge fund industry, along with the increasing occurrence of fraud, explains why the Securities and Exchange Commission (SEC) has issued a new rule in December 2004 that requires hedge funds to *register* as investment advisors.[12] Registration gives the SEC the authority to conduct examinations of hedge fund activities. The goal is to help to identify compliance problems at an early stage and to provide a deterrent to fraud. Registration also requires the hedge fund to designate a chief compliance officer. In June 2006, this registration requirement was annulled by the U.S. Court of Appeals. In practice, however, the majority of U.S.-based hedge funds have voluntarily registered as investment advisors. Registration is often required by investors as a precondition for investing.

The possibility of fraud can be lessened when a fund has an independent administrator. Administrators perform day-to-day administrative duties associated with running a fund, in particular financial and tax reporting. They calculate net asset values, maintain the statutory books and records, and provide shareholder services. In addition, an outside auditor provides important additional information. Auditors issue a written opinion upon the fair presentation of the fund's financial statements, typically on an annual basis. To protect against theft, investors should insist on an external custodian, which is a financial institution such as a bank or trust company that holds the fund's assets. Usually a fund's prime broker will perform the role of custodian.

Example: Ponzi Scheme

The term Ponzi scheme is attributed to Carlo Ponzi, who in 1919 established an inventive pyramid scheme using new investor funds to repay earlier investors. The

---

[10] A 2003 report by the SEC, however, notes that there is no evidence that hedge fund advisors engage *disproportionately* in fraudulent activities.

[11] See Kundro, C. and S. Feffer (2003), Valuation Issues and Operational Risk in Hedge Funds, *Capco White Paper*.

[12] This rule applies to U.S.-based hedge funds, and to non-U.S. funds that have at least fourteen U.S. investors. Funds with less than $25 million under management do not have to comply. The rule took effect in February 2006.

investment was based on a relative-value trade, in which postal coupons were bought overseas for the equivalent of one U.S. cent and resold for six American one-cent stamps. After transaction costs were factored in, however, the trade was unprofitable. Nevertheless, thousands of people invested with him, lured by a promise of 50% return in 90 days. Ultimately, he lost $140 million of investor funds, in today's dollars, and was jailed for fraud.

The most famous case of a Ponzi scheme was perpetrated by Bernard Madoff, who was arrested in December 2008 after admitting to defrauding investors of perhaps $50 billion through his brokerage firm, Bernard Madoff Investment Securities (BMIS). BMIS was established in 1960 and by the end of 2007 was managing about $17 billion in hedge funds investments. Returns to initial investors were paid using new investments. The scheme collapsed when investors sought about $7 billion in redemptions during 2008, which could not be met. Ponzi schemes only work as long as new money flows in.

Investors had been attracted by the high and steady returns from Madoff's funds, which turned out to have been fabricated. Many leading funds of funds, however, had performed due diligence on Madoff and had spotted enough warning signs to stay clear of him. BMIS acted as custodian of investor assets, had no external administrator, and relied on an unknown three-person audit firm. In addition, some people who had analyzed the options-trading strategy had concluded that it was infeasible because it would have implied trading volumes far in excess of exchange-traded volumes. This illustrates the value of a thorough due diligence process.

17.4.5 Regulatory Risks

Finally, hedge funds are also subject to regulatory risks. This is the risk of loss due to regulatory changes. In September 2008, for example, many countries put in place outright bans on short-selling equities.

This ban created havoc with strategies that rely on short sales to hedge. For example, the convertible bond arbitrage sector suffered extreme losses during these months. Statistical arbitrage funds, which take long and short positions in stocks, had to withdraw from many markets.

There is broad consensus, however, that such bans are ineffective at stopping prices from falling.[13] Such bans also have far-reaching consequences. In the case of convertible bonds, for example, the market freeze prevented many institutions, including banks, that traditionally issue convertible bond debt to raise new funds. In addition, such bans cause investors to withdraw from markets, sapping liquidity, which actually increases volatility. Bans on short-selling removed hedge funds' main mechanism for risk management, which led to an acceleration of withdrawal of capital when the market needed it most.

---

[13] Marsh, I. and N. Niemer (2008), The Impact of Short Sales Restrictions, *Working Paper*, Cass Business School, London.

EXAMPLE 17.9: LIQUIDITY RISKS

Asset liquidity risk is most pronounced for

a. A $10 million position in distressed securities
b. A $10 million position in Treasury bonds
c. A $100 million position in distressed securities
d. A $100 million position in Treasury bonds

EXAMPLE 17.10: FRM EXAM 2007—QUESTION 62

You are asked to estimate the exposure of a hedge fund to the S&P 500. Though the fund claims to mark to market weekly, it does not do so and marks to market once a month. The fund also does not tell investors that it simply holds an Exchange Traded Fund (ETF) indexed to the S&P 500. Because of the claims of the hedge fund, you decide to estimate the market exposure by regressing weekly returns of the fund on the weekly return of the S&P 500. Which of the following correctly describes a property of your regression estimates?

a. The intercept of your regression will be positive, showing that the fund has a positive alpha when estimated using an OLS regression.
b. The beta will be misestimated because hedge fund exposures are nonlinear.
c. The beta of your regression will be one because the fund holds the S&P 500.
d. The beta of your regression will be zero because the fund returns are not synchronous with the S&P 500 returns.

EXAMPLE 17.11: FRM EXAM 2006—QUESTION 112

For a portfolio of illiquid assets, hedge fund managers often have considerable discretion in portfolio valuation at the end of each month and may have incentives to smooth returns by marking values below actual in high-return months and above actual in low-return months. Which of the following is *not* a consequence of return smoothing over time?

a. Higher Sharpe ratio
b. Lower volatility
c. Higher serial correlation
d. Higher market beta

17.5 DEALING WITH HEDGE FUND RISKS

Because of these risks, hedge funds need to be monitored closely. This starts with due diligence, which is the process of systematically investigating the fund before investing. On the operational side, this involves an analysis of the fund documents, of the key personnel (including background checks), of the fund service providers (administrator, prime broker, legal counsel, auditor), of the regulatory registration, and of the operations and valuation procedures. On the investment side, this involves an analysis of the investment strategy, of the risk factors, and of the risk control systems. Once a hedge fund manager is hired, some components of this due diligence process need to be verified periodically.

Without information about the positions, however, this process is rather incomplete. It is very difficult to detect style drift, for example, from historical returns. Because returns are typically provided at monthly intervals, structural breaks can generally be identified only after a few years.

17.5.1 Hedge Fund Transparency Issues

Hedge funds, however, are generally reluctant to reveal information about their positions. This lack of transparency has serious disadvantages for investors, however.

Disclosure allows *risk monitoring* of the hedge fund, which is especially useful with active trading. This can help to avoid situations where the hedge fund manager unexpectedly increases leverage or changes style. Closer monitoring of the fund can also decrease the probability of fraud or misvaluation of assets.

Disclosure is also important for *risk aggregation*. The investor should know how the hedge fund interacts with other assets in the portfolio. Whether the hedge fund has a positive or negative correlation with the rest of the portfolio affects the total portfolio risk.

Example: Why Risk Aggregation?

Aggregation of positions is important to identify potential concentrations to individual names, or companies. The story is that of a large pension fund, which had allocated assets to outside managers investing in corporate bonds, growth equities, and value equities. In 2000, Enron was rated investment-grade and viewed as a growth stock, reflecting its high stock price of $90. The fund had positions in Enron corporate bonds and in Enron stock, through the growth manager.

As negative news unfolded the following year, the stock price fell to $15 by October 2001. Many saw this decline as a great opportunity to buy the stock. The pension fund's value managers started to buy the stock. As the same time, its other managers did not have the discipline to sell. By December 2001, the stock price had fallen to $0.03. The pension fund ended up with large holdings of Enron stock and bonds, and a huge loss due to its failure to identify this concentration risk.

Greater disclosure is often resisted on the grounds that it would disclose *pro- prietary information*, leading to the possibility of a third party trading against the hedge fund. This threat, however, comes from the broker-dealer community, generally not from investors. If this is an issue, confidentiality agreements should prevent leakages of sensitive information. Hedge funds generally prefer to release such information to investors with no trading operations, whether directly or through affiliates, who would not be able to profit from this information.

Another argument sometimes advanced is the *lack of investor sophistication.* In other words, disclosing positions would give too much information to investors who might not able to use it.

17.5.2   Solutions for Transparency

These arguments can be addressed with a number of solutions. The first consists of external risk measurement services. These firms have access to the individual positions of hedge funds, with the proper confidentiality agreements, and provide aggregate risk measures to investors. They release only exposures to major risk factors, such as net duration, net systematic risk, and so on. This solution neatly solves the problems of risk aggregation and managers' widespread reluctance to disclose detailed information about their positions. Another solution is to go through a fund of funds that has position-level information. Although this is still fairly rare, such fund of funds can perform the risk measurement and monitoring function for the investor, thereby justifying some of its added fee.

As will be seen in the chapter on operational risk, sound risk management relies heavily on the principle of independence. Some hedge funds are now using external companies that provide independent valuation services. Figure 17.2 describes a best-practice architecture for valuation and risk measurement in hedge funds.

Portfolio positions are recorded by the prime broker, who can provide port-folio valuation. The hedge fund, however, may also use external service providers for valuation and risk measurement, both of which are based on the portfolio po-sitions. All of this information is then fed into the portfolio and risk measurement system for the hedge fund manager.
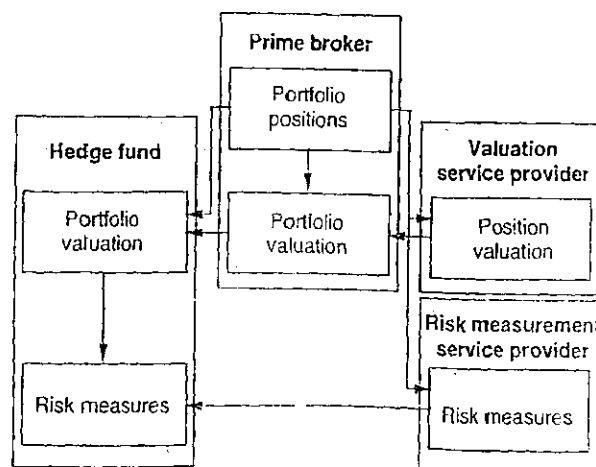
FIGURE 17.2   Architecture of Hedge Fund Risk Systems

EXAMPLE 17.12: TRANSPARENCY

Investors should insist on learning about the positions of hedge funds because

a. They want to trade ahead of the hedge fund.
b. They do not understand the trading strategies behind the positions.
c. They want to aggregate the risk of hedge funds with the rest of their portfolio.
d. They receive the information from the prime broker, anyway.

EXAMPLE 17.13: FRM EXAM 2003—QUESTION 51

As one of your duties as the chief risk officer for a fund of funds, you evaluate the risk management of candidate hedge funds. In your evaluation of a newly organized two-person hedge fund, which is your primary consideration?

a. Risk reporting structure
b. Investment style
c. Assets under management
d. Last month's return

17.6 IMPORTANT FORMULAS

Net Beta: $(\beta_L V_L - \beta_S V_S) = \beta_E V_E$

Leverage: $\frac{V_A}{V_E} = \dfrac{\text{Long stock positions plus cash}}{\text{Equity}}$

Gross Leverage: $\frac{V_L + V_S}{V_E} = \dfrac{\text{Long positions plus absolute value of short positions}}{\text{Equity}}$

Net Leverage: $\frac{V_L - V_S}{V_E} = \dfrac{\text{Long positions minus absolute value of short positions}}{\text{Equity}}$

17.7 ANSWERS TO CHAPTER EXAMPLES

Example 17.1: FRM Exam 2006—Question 41

c. The gross leverage is $(315 + 225)/185 = 2.9$. The net leverage is $(315 - 225)/185 = 0.5$. Note that beta is not needed for this calculation.

Example 17.2: Leverage and Returns

c. The net return on the stock portfolio is $(\beta_L \$315 - \beta_S \$225) \times 20\%$. With betas of 1, this is $18 million. Given that the equity is $185 million, the rate of return is about 10%. The rate of return is less than that on the market because most of the exposure to the market is hedged.

Example 17.3: FRM Exam 2004—Question 2

a. Because the position is both long and short, high correlation implies low risk. Conversely, lowering correlation increases risk.

Example 17.4: Risks in Fixed-Income Arbitrage

c. The strategy has no exposure to the level of rates but is exposed to a widening of the swap-Treasury spread. Assume for instance that the swap and Treasury rates are initially 5.5% and 5%. If these rates change to 5.3% and 4.5%, for example, both values for the swap and the Treasury bond would increase. Because the drop in the Treasury rate is larger, however, the price of the Treasury bond would fall more than the swap, leading to a net loss on the position. The strategy should *gain* from decreases in the swap-Treasury spread, so a. is wrong. The strategy should *gain* from increases in the Treasury rate, all else equal, so b. is wrong. Finally, the distribution of the payoff depends on the distribution of the swap-Treasury spread. Because this cannot go below zero, there is a limit on the upside. The position has negative skewness, so c. is correct.

Example 17.5: Risks in Convertible Arbitrage

d. This position is hedged against interest rate risk, so b. is wrong. It is also hedged against directional movements in the stock, so c. is wrong. The position is long an option (the option to convert the bond into the stock) so is long implied volatility, so a. is wrong. Long options positions have positive gamma.

Example 17.6: Risks in Merger Arbitrage—I

c. The position is long one share of Company B offset by a short position in two shares of Company A. The payoff is $(\$120 - \$90) - 2(\$60 - \$50) = \$30 - \$20 = \$10$.

Example 17.7: Risks in Merger Arbitrage—II

b. The expected payoff is the sum of probabilities times the payoff in each state of the world, or $83\% \times \$5 + 17\% \times (-\$20) = \$4.15 - \$3.40 = \$0.75$. Note that the distribution is highly asymmetric, with a small probability of a large loss.

Example 17.8: FRM Exam 2005—Question 47

b. This fund has zero beta, so is a relative value fund. It is, however, exposed to idiosyncratic, stock-specific risk.

Example 17.9: Liquidity Risks

c. Asset liquidity risk is a function of the size of the position and the intrinsic liquidity of the instrument. Distressed securities trade much less than Treasury bonds, so have more liquidity risk. A $100 million is more illiquid than a $10 million in the same instrument.

Example 17.10: FRM Exam 2007—Question 62

d. The weekly returns are not synchronized with those of the S&P. As a result, the estimate of beta from weekly data will be too low.

Example 17.11: FRM Exam 2006—Question 112

d. Illiquidity creates an *understatement* of the total risk measure. As a result, the Sharpe ratio will be artificially higher. Illiquidity creates trends in returns (higher serial correlation), as market shocks during a month will be partially recorded in two consecutive months. Illiquidity, however, biases down the market beta.

Example 17.12: Transparency

c. Risk aggregation is an important reason for investor to learn about the positions of their investment in hedge funds. Answer a. is incorrect because front-running the hedge fund would be a reason *not* to disclose position information. Answer b. is incorrect because misunderstanding the trading strategies would be a reason *not* to require position information. Answer d. is incorrect because they do not receive position information from the prime broker.

Example 17.13: FRM Exam 2003—Question 51

b. The market risk of the fund is primarily driven by the investment style. Answer a. about the risk reporting structure, is not too important for this fund because there will not be much separation of duties, anyway, for such a small team.

Credit Risk Management
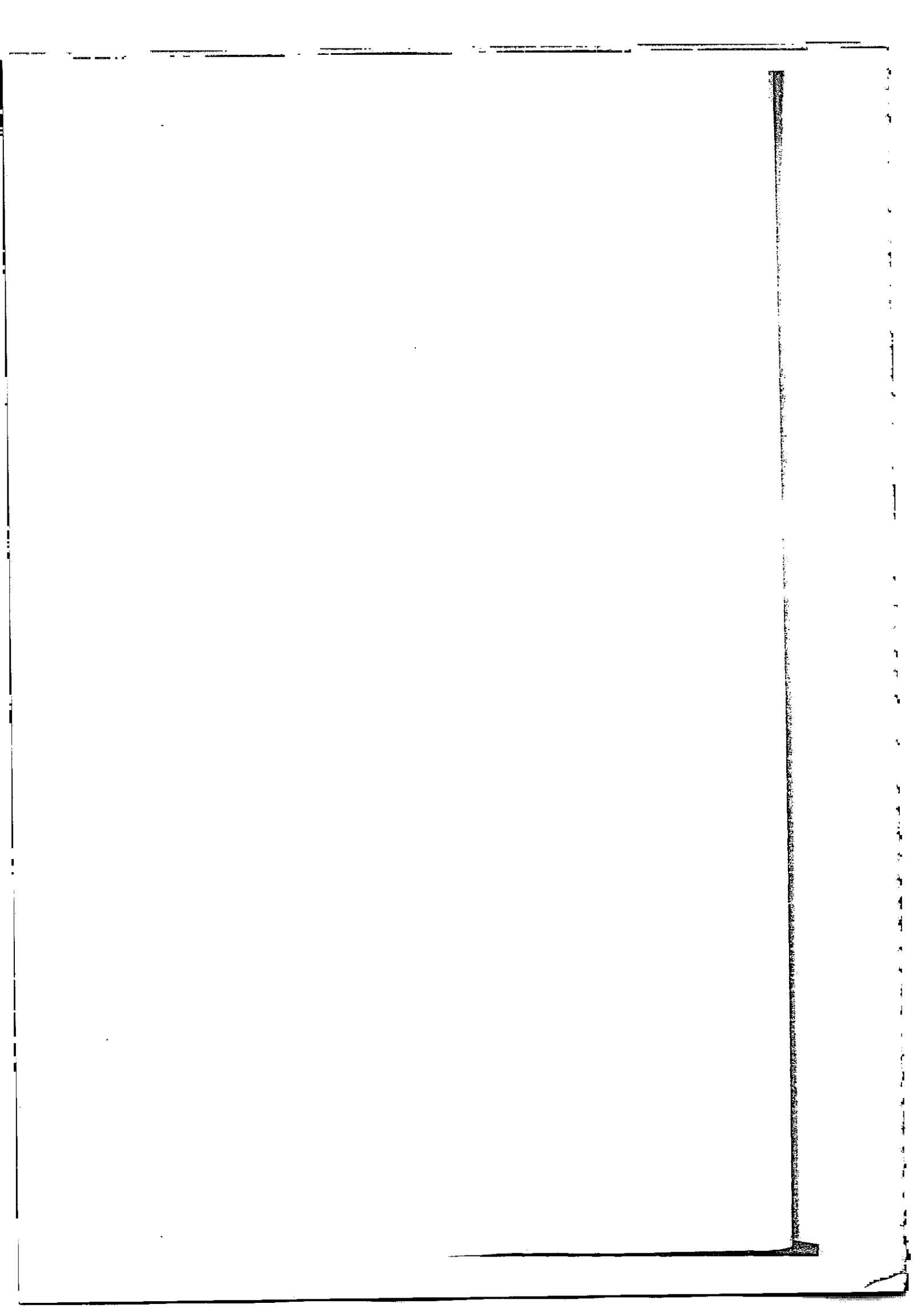
Introduction to Credit Risk

Credit risk is the risk of an economic loss from the failure of a counterparty to fulfill its contractual obligations. Its effect is measured by the cost of replacing cash flows if the other party defaults.

This chapter provides an introduction to the measurement of credit risk. The study of credit risk has undergone vast developments in the last few years. Fuelled by advances in the measurement of market risk, institutions are now, for the first time, attempting to quantify credit risk on a portfolio basis.

Credit risk, however, offers unique challenges. It requires constructing the distribution of default probabilities, of loss given default, and of credit exposures, all of which contribute to credit losses and should be measured in a portfolio context. In comparison, the measurement of market risk using value at risk is a simple affair. These challenges explain why many of these models performed poorly during the credit crisis that started in 2007.

For most institutions, however, market risk pales in significance compared with credit risk. Indeed, the amount of risk-based capital for the banking system reserved for credit risk is vastly greater than that for market risk. The history of financial institutions has also shown that the biggest banking failures were due to credit risk.

Credit risk involves the possibility of nonpayment, either on a future obligation or during a transaction. Section 18.1 introduces settlement risk, which arises from the exchange of principals in different currencies during a short window, typically a day. We discuss exposure to settlement risk and methods to deal with it.

Traditionally, however, credit risk is viewed as presettlement risk, which arises during the life of the obligation. Section 18.2 analyzes the components of a credit risk system and the evolution of credit risk measurement systems. Section 18.3 then shows how to construct the distribution of credit losses for a portfolio given default probabilities for the various credits in the portfolio.

The key drivers of portfolio credit risk are the correlations between defaults. Section 18.4 takes a fixed $100 million portfolio with an increasing number of obligors and shows how the distribution of losses is dramatically affected by correlations.

18.1 SETTLEMENT RISK

18.1.1 Presettlement versus Settlement Risk

Counterparty credit risk consists of both presettlement and settlement risk. Presettlement risk is the risk of loss due to the counterparty's failure to perform on an obligation during the life of the transaction. This includes default on a loan or bond or failure to make the required payment on a derivative transaction. Presettlement risk exists over long periods—years—starting from the time it is contracted until settlement.

In contrast, settlement risk is due to the *exchange* of cash flows and is of a much shorter-term nature. This risk arises as soon as an institution makes the required payment and exists until the offsetting payment is received. This risk is greatest when payments occur in different time zones, especially for foreign exchange transactions where notionals are exchanged in different currencies. Failure to perform on settlement can be caused by counterparty default, liquidity constraints, or operational problems.

Most of the time, settlement failure due to operational problems leads to minor economic losses, such as additional interest payments. In some cases, however, the loss can be quite large, extending to the full amount of the transferred payment. An example of major settlement risk is the 1974 failure of Herstatt Bank. The day the bank went bankrupt, it had received payments from a number of counterparties but defaulted before payments were made on the other legs of the transactions.

18.1.2 Managing Settlement Risk

In March 1996, the Bank for International Settlements (BIS) issued a report warning that the private sector should find ways to reduce settlement risk in the $1.2 trillion-a-day global foreign exchange market.[1] The report noted that central banks had "significant concerns regarding the risk stemming from the current arrangements for settling FX trades." It explained that "the amount at risk to even a single counterparty could exceed a bank's capital," which creates systemic risk. The threat of regulatory action led to a reexamination of settlement risk.

Examining the various stages of a trade, its status can be classified into five categories:

1. *Revocable:* When the institution can still cancel the transfer without the consent of the counterparty
2. *Irrevocable:* After the payment has been sent and before payment from the other party is due
3. *Uncertain:* After the payment from the other party is due but before it is actually received

---

[1] Committee on Payment and Settlement Systems (1996), *Settlement Risk in Foreign Exchange Transactions*, BIS (online), available at http://www.bis.org/publ/cpss17.pdf.

4. *Settled:* After the counterparty payment has been received
5. *Failed:* After it has been established that the counterparty has not made the payment

Settlement risk occurs during the periods of irrevocable and uncertain status, which can take from one to three days.

While this type of credit risk can lead to substantial economic losses, the short-term nature of settlement risk makes it fundamentally different from presettlement risk. Managing settlement risk requires unique tools, such as real-time gross settlement (RTGS) systems. These systems aim at reducing the time interval between the time an institution can no longer stop a payment and the receipt of the funds from the counterparty.

Settlement risk can be further managed with netting agreements. One such form is bilateral netting, which involves two banks. Instead of making payments of gross amounts to each other, the banks tot up the balance and settle only the net balance outstanding in each currency. At the level of instruments, netting also occurs with contracts for differences (CFD). Instead of exchanging principals in different currencies, the contracts are settled in dollars at the end of the contract term.[2]

The next step up is a multilateral netting system, also called continuous-linked settlements, where payments are netted for a group of banks that belong to the system. This idea became reality when the CLS Bank, established in 1998 with 60 bank participants, became operational on September 9, 2002. Every evening, CLS Bank provides a schedule of payments for the member banks to follow during the next day. Payments are not released until funds are received and all transaction confirmed. The risk now has been reduced to that of the netting institution. In addition to reducing settlement risk, the netting system has the advantage of reducing the number of trades between participants, by up to 90%, which lowers transaction costs.

EXAMPLE 18.1: FRM EXAM 2000—QUESTION 36

Settlement risk in foreign exchange is generally due to

a. Notionals being exchanged
b. Net value being exchanged
c. Multiple currencies and countries involved
d. High volatility of exchange rates

---

[2] These are similar to nondeliverable forwards, which are used to trade emerging-market currencies outside the jurisdiction of the emerging-market regime and are also settled in dollars.

EXAMPLE 18.2: FRM EXAM 2000—QUESTION 85

Which one of the following statements about multilateral netting systems is *not* accurate?

a. Systemic risks can actually increase because they concentrate risks on the central counterparty, the failure of which exposes all participants to risk.

b. The concentration of risks on the central counterparty eliminates risk because of the high quality of the central counterparty.

c. By altering settlement costs and credit exposures, multilateral netting systems for foreign exchange contracts could alter the structure of credit relations and affect competition in the foreign exchange markets.

d. In payment netting systems, participants with net-debit positions will be obligated to make a net settlement payment to the central counterparty that, in turn, is obligated to pay those participants with net-credit positions.

18.2 OVERVIEW OF CREDIT RISK

18.2.1 Drivers of Credit Risk

We now examine the drivers of credit risk, traditionally defined as presettlement risk. Credit risk measurement systems attempt to quantify the risk of losses due to counterparty default. The distribution of credit risk can be viewed as a compound process driven by these variables:

- Default, which is a discrete state for the counterparty—either the counterparty is in default or not. This occurs with some probability of default (PD).
- Credit exposure (CE), which is the economic or market value of the claim on the counterparty. It is also called exposure at default (EAD) at the time of default.
- Loss given default (LGD), which represents the fractional loss due to default. As an example, take a situation where default results in a fractional recovery rate of 30% only. LGD is then 70% of the exposure.

Traditionally, credit risk has been measured in the context of loans or bonds for which the exposure, or economic value, of the asset is close to its notional, or face value. This is an acceptable approximation for bonds but certainly not for derivatives, which can have positive or negative value. Credit exposure is defined as the positive value of the asset:

$$CE_t = Max(V_t, 0) \tag{18.1}$$

This is so because if the counterparty defaults with money owed to it, the full amount has to be paid.[3] In contrast, if the counterparty owes money, only a fraction may be recovered. Thus, presettlement risk only arises when the contract's replacement cost has a positive value to the institution (i.e., is in-the-money).

18.2.2 Measurement of Credit Risk

The evolution of credit risk measurement tools has gone through these steps:

Notional amounts, adding up simple exposures

Risk-weighted amounts, adding up exposures with a rough adjustment for risk

Notional amounts combined with credit ratings, adding up exposures adjusted for default probabilities

Internal portfolio credit models, integrating all dimensions of credit risk

Initially, risk measures were based on the total notional amount. A multiplier, say 8%, was applied to this amount to establish the amount of required capital to hold as a reserve against credit risk.

The problem with this approach is that it ignores variations in the probability of default. In 1988, the Basel Committee instituted a rough categorization of credit risk by *risk class*, providing risk weights to multiply notional amounts. This was the first attempt to force banks to carry capital in relation to the risks they were taking. The Basel rules are explained in more detail in Chapter 29.

These risk weights proved to be too simplistic, however, creating incentives for banks to alter their portfolios in order to maximize their shareholder returns subject to the Basel capital requirements. This had the perverse effect of introducing more risk into the balance sheets of commercial banks, which was certainly not the intended purpose of the 1988 rules. As an example, there was no differentiation between AAA-rated and C-rated corporate credits. Because loans to lower quality credits are generally more profitable than those to high quality credits but require the same amount of regulatory capital, some banks have responded by shifting its loan mix toward lower-rated credits. This led to the Basel II rules, which allow banks to use their own internal or external credit ratings. These credit ratings provide more refined representation of credit risk.

Even with these improvements, the credit risk charges are computed separately for all exposures and then added up. This approach may not properly account for diversification effects, unlike internal portfolio credit models. Such models, however, create special challenges and as a result, are still not accepted by the Basel Committee for capital adequacy requirements.

---

[3] This is due to *no walk-away clauses*, explained in Chapter 27.

TABLE 18.1   Comparison of Market Risk and Credit Risk

Item	Market Risk	Credit Risk
Sources of risk	Market risk only	Default risk, recovery risk, market risk
Distributions	Mainly symmetric, perhaps fat tails	Skewed to the left
Time horizon	Short term (days)	Long term (years)
Aggregation	Business/trading unit	Whole firm vs. counterparty
Legal issues	Not applicable	Very important

18.2.3 Credit Risk versus Market Risk

The tools recently developed to measure market risk have proved invaluable in assessing credit risk. Even so, there are a number of major differences between market and credit risks, which are listed in Table 18.1.

As mentioned previously, credit risk results from a compound process with three types of risk. The nature of this risk creates a distribution that is strongly skewed to the left, unlike most market risk factors. This is because credit risk is akin to short positions in options. At best, the counterparty makes the required payment and there is no loss. At worst, the entire amount due is lost.

The time horizon is also different. Whereas the time required for corrective action is relatively short in the case of market risk, it is much longer for credit risk. Positions also turn over much more slowly for credit risk than for market risk, although the advent of credit derivatives has made it easier to hedge credit risk.

Finally, the level of aggregation is different. Limits on market risk may apply at the level of a trading desk, business units, and eventually the whole firm. In contrast, limits on credit risk must be defined at the counterparty level, for all positions taken by the institution.

Legal issues are also very important for credit risk. Exposures can be controlled by netting contracts, for example. Finally, an event of default invariably puts in motion a level process.

18.3 MEASURING CREDIT RISK

18.3.1 Credit Losses

To develop the intuition of credit models, let us start with a simple case where losses are due to the effect of defaults only. This is what is called default mode. In other words, there is no intermediate marking-to-market. This example would

be typical of loans held in a banking book as opposed to bonds held in a trading account.

The distribution of credit losses from a portfolio of N instruments issued by different obligors can be described as

$$\text{Credit loss} = \sum_{i=1}^{N} b_i \times \text{CE}_i \times (1 - f_i) \tag{18.2}$$

where

- b_i is a (Bernoulli) random variable that takes the value of 1 if default occurs and 0 otherwise, with probability p_i, such that $E[b_i] = p_i$
- CE_i is the credit exposure at the time of default
- f_i is the recovery rate, or $(1 - f_i)$ the loss given default

18.3.2 Joint Events

In theory, all of these could be random variables. For what follows, we will assume that the only random variable is the default indicator b.

Equation (18.2) then shows that the expected credit loss is

$$E[\text{CL}] = \sum_{i=1}^{N} E[b_i] \times \text{CE}_i \times (1 - f_i) = \sum_{i=1}^{N} p_i \times \text{CE}_i \times (1 - f_i) \tag{18.3}$$

The dispersion in credit losses, however, critically depends on the correlations between the default events.

Closed-form solutions can be easily derived when the default events are statistically independent. This simplifies the analysis considerably, as the probability of any joint event is simply the product of the individual event probabilities:

$$p(A \text{ and } B) = p(A) \times p(B) \tag{18.4}$$

At the other extreme, if the two events are perfectly correlated, that is, if B always defaults when A defaults, we have

$$p(A \text{ and } B) = p(B \mid A) \times p(A) = 1 \times p(A) = p(A) \tag{18.5}$$

when the marginal probabilities are equal, $p(A) = p(B)$.

Suppose, for instance, that the marginal probabilities are $p(A) = p(B) = 1\%$. Then the probability of the joint event is 0.01% in the independence case and 1% in the perfect-correlation case.

More generally, one can show that the probability of a joint default depends on the marginal probabilities and the correlations. As we have seen in Chapter 2,

TABLE 18.2   Joint Probabilities (Default Correlation = 0.5)

A	B Default	No Default	Marginal
Default	0.00505	0.00495	0.01
No default	0.00495	0.98505	0.99
Marginal	0.01	0.99	

the expectation of the product can be related to the covariance

$$E[b_A \times b_B] = \text{Cov}[b_A, b_B] + E[b_A]E[b_B] = \rho \sigma_A \sigma_B + p(A)p(B) \qquad (18.6)$$

Given that b_A is a Bernoulli variable, Chapter 2 has shown that its standard deviation is $\sigma_A = \sqrt{p(A)[1 - p(A)]}$ and similarly for b_B. We then have

$$p(A \text{ and } B) = \text{Corr}(A, B)\sqrt{p(A)[1 - p(A)]}\sqrt{p(B)[1 - p(B)]} + p(A)p(B) \quad (18.7)$$

For example, if the correlation is unity and $p(A) = p(B) = p$, we have

$$p(A \text{ and } B) = 1 \times [p(1 - p)]^{1/2} \times [p(1 - p)]^{1/2} + p^2 = [p(1 - p)] + p^2 = p$$

as shown in Equation (18.5).

If the correlation is 0.5 and $p(A) = p(B) = 0.01$, however, then we have $p(A \text{ and } B) = 0.00505$, which is only half of the marginal probabilities. This example is illustrated in Table 18.2, which lays out the full joint distribution. Note how the probabilities in each row and column sum to the marginal probability. From this information, we can infer all remaining probabilities. For example, the probability of B not defaulting when A is in default is $0.01 - 0.00505 = 0.00495$. This allows us to fill the entire table with joint probabilities.

18.3.3 An Example

As an example of credit loss distribution, consider a portfolio of $100 million with three bonds, A, B, and C, with various probabilities of default. To simplify, we assume (1) that the exposures are constant, (2) that the recovery in case of default is zero, and (3) that default events are independent across the three issuers.

Table 18.3 displays the exposures and default probabilities. We can easily compute the expected loss as $E[CL] = \sum p_i \times CE_i = 0.05 \times 25 + 0.10 \times 30 + 0.20 \times 45$, or

$$E[CL] = \sum p_i \times CE_i = \$13.5 \text{ million}$$

This is the average credit loss over many repeated, hypothetical "samples." This computation is very easy and does not require any information about the

TABLE 18.3   Portfolio Exposures, Default Risk, and Credit Losses

	Issuer	Exposure	Probability
	A	$25	0.05
	B	$30	0.10
	C	$45	0.20

Default i	Loss L_i	Probability $p(L_i)$	Cumulative Probability	Expected $L_i p(L_i)$	Variance $(L_i - EL_i)^2 p(L_i)$
None	$0	0.6840	0.6840	0.000	120.08
A	$25	0.0360	0.7200	0.900	4.97
B	$30	0.0760	0.7960	2.280	21.32
C	$45	0.1710	0.9670	7.695	172.38
A, B	$55	0.0040	0.9710	0.220	6.97
A, C	$70	0.0090	0.9800	0.630	28.99
B, C	$75	0.0190	0.9990	1.425	72.45
A, B, C	$100	0.0010	1.0000	0.100	7.53
Sum				$13.25	434.7

distribution other than default probabilities. Also note that we define the loss as a positive number, which is the usual convention.

On the other hand, we do need to described the full distribution to derive a worst-loss measure. This is done in the second panel, which lists all possible states. In state 1, there is no default, which has a probability of $(1 - p_1)(1 - p_2)(1 - p_3) = (1 - 0.05)(1 - 0.10)(1 - 0.20) = 0.684$, given independence. In state 2, bond A defaults and the others do not, with probability $p_1(1 - p_2)(1 - p_3) = 0.05(1 - 0.10)(1 - 0.20) = 0.036$ (and so on for the other states).

Figure 18.1 plots the frequency distribution of credit losses. The table also shows how to compute the variance as

$$V[CL] = \sum_{i=1}^{N} (L_i - E[CL_i])^2 p(L_i) = 434.7$$

which gives a standard deviation of $\sigma(CL) = \sqrt{434.7} = \$20.9$ million.

Alternatively, we can express the range of losses with a 95% quantile, which is the lowest number CL_i such that

$$P(CL \leq CL_i) \geq 95\% \qquad (18.8)$$

Because losses are recorded as positive numbers, the quantile is the lowest loss such that the cumulative probability of a lower loss is at or just above 95%.

From Table 18.3, we see that the fourth row has a cumulative probability of 0.9670, just above 0.95 and corresponds to a loss of $45 million. Figure 18.2 plots the cumulative distribution function and shows that the 95% quantile is $45 million.
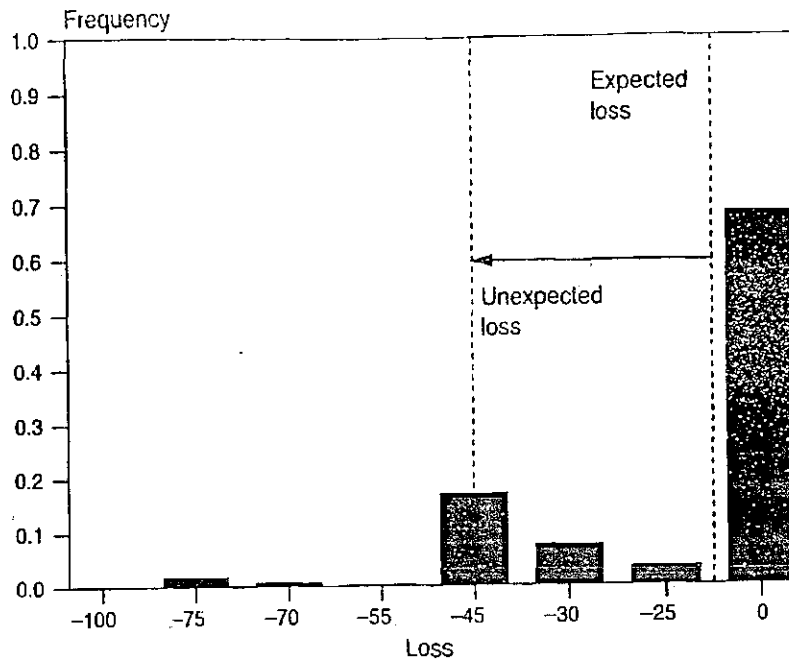
FIGURE 18.1   Distribution of Credit Losses

In terms of deviations from the mean, this gives an unexpected loss of 45 − 13.2 = $32 million. This is a measure of credit VAR. Sometimes, however, credit VAR is measured as the total loss, or $45 million in this case.

This very simple three-bond portfolio provides a useful example of the measurement of the distribution of credit risk. It shows that the distribution is skewed to the left. In addition, the distribution has irregular "bumps" that correspond to
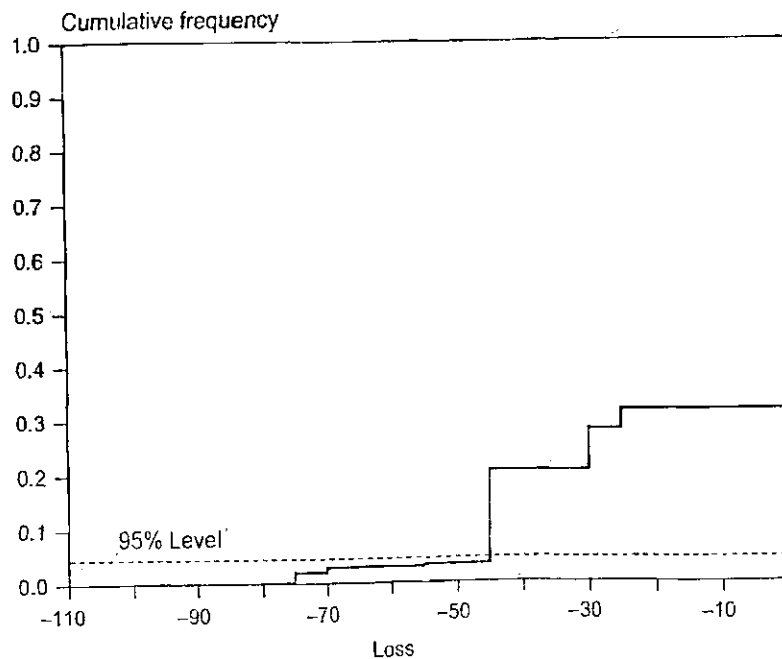
FIGURE 18.2   Cumulative Distribution of Credit Losses

the default events. The chapter on managing credit risk will further elaborate on this point.

KEY CONCEPT

The expected credit loss depends on default probabilities but not on default correlation. In contrast, the higher the default correlation, the higher the unexpected credit loss.

EXAMPLE 18.3: FRM EXAM 2002—QUESTION 130

You have granted an unsecured loan to a company. This loan will be paid off by a single payment of $50 million. The company has a 3% chance of defaulting over the life of the transaction and your calculations indicate that if they default you would recover 70% of your loan from the bankruptcy courts. If you are required to hold a credit reserve equal to your expected credit loss, how great a reserve should you hold?

a. $450,000
b. $750,000
c. $1,050,000
d. $1,500,000

EXAMPLE 18.4: FRM EXAM 2003—QUESTION 17

An investor holds a portfolio of $100 million. This portfolio consists of A-rated bonds ($40 million) and BBB-rated bonds ($60 million). Assume that the one-year probabilities of default for A-rated and BBB-rated bonds are 3% and 5%, respectively, and that they are independent. If the recovery value for A-rated bonds in the event of default is 70% and the recovery value for BBB-rated bonds is 45%, what is the one-year expected credit loss from this portfolio?

a. $1,672,000
b. $1,842,000
c. $2,010,000
d. $2,218,000

EXAMPLE 18.5: FRM EXAM 2007—QUESTION 73

A portfolio consists of two bonds. The credit-VAR is defined as the maximum
loss due to defaults at a confidence level of 98% over a one-year horizon.
The probability of joint default of the two bonds is 1.27%, and the default
correlation is 30%. The bond value, default probability, and recovery rate
are USD 1,000,000, 3%, and 60% for one bond, and USD 600,000, 5%,
and 40% for the other. What is the expected credit loss of the portfolio?

a. USD 0
b. USD 9,652
c. USD 20,348
d. USD 30,000

EXAMPLE 18.6: FRM EXAM 2007—QUESTION 74

Continuing with the previous question, what is the best estimate of the un-
expected credit loss (away from the ECL), or credit VAR, for this portfolio?

a. USD 570,000
b. USD 400,000
c. USD 360,000
d. USD 370,000

EXAMPLE 18.7: FRM EXAM 2007—QUESTION 102

Suppose Bank Z lends EUR 1 million to X and EUR 5 million to Y. Over the
next year, the PD for X is 0.2 and for Y is 0.3. The PD of joint default is 0.1.
The loss given default is 40% for X and 60% for Y. What is the expected
loss of default in one year for the bank?

a. EUR 0.72 million
b. EUR 0.98 million
c. EUR 0.46 million
d. EUR 0.64 million

EXAMPLE 18.8: FRM EXAM 2004—QUESTION 46

Consider an A-rated bond and a BBB-rated bond. Assume that the one-year probabilities of default for the A- and BBB-rated bonds are 2% and 4%, respectively, and that the joint probability of default of the two bonds is 0.15%. What is the default correlation between the two bonds?

a. 0.07%
b. 2.6%
c. 93.0%
d. The default correlation cannot be calculated with the information provided.

18.4 CREDIT RISK DIVERSIFICATION

Modern banking was built on the sensible notion that a portfolio of loans is less risky than are single loans. As with market risk, the most important feature of credit risk management is the ability to diversify across defaults.

To illustrate this point, Figures 18.3a through 3c presents the distribution of losses for a $100 million loan portfolio. The probability of default is fixed at 1%. If default occurs, recovery is zero.

1 credit of $100 million
N=1, E(Loss)=$1 million, SD(Loss)=$10 million

10 independent credits of $10 million
N=10, E(Loss)=$1 million, SD(Loss)=$3 million
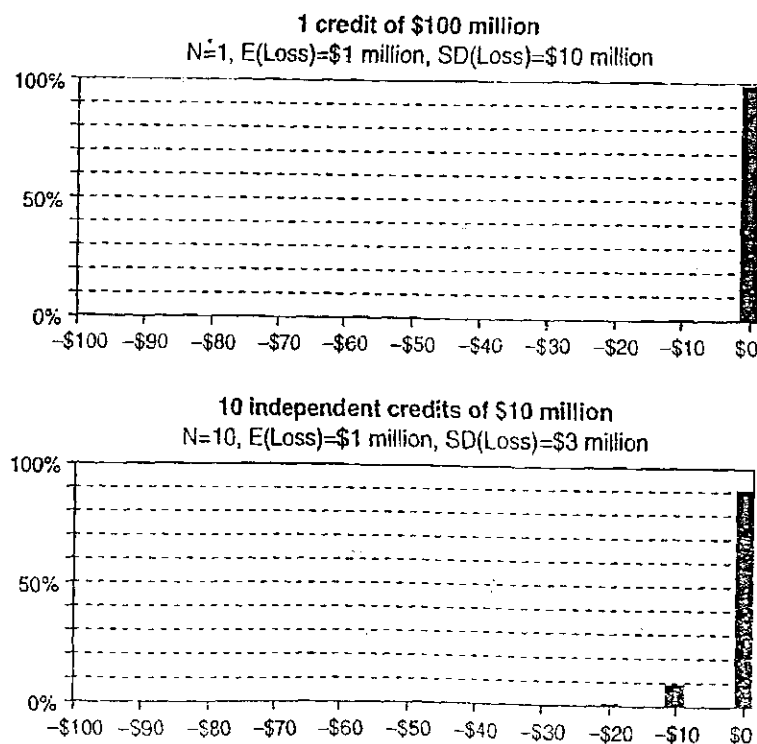
FIGURE 18.3a Distribution of Credit Losses

100 independent credits of $1 million
N=100, E(Loss)=$1 million, SD(Loss)=$1 million

1,000 independent credits of $100,000
N=1,000, E(Loss)=$1 million, SD(Loss)=$0.3 million
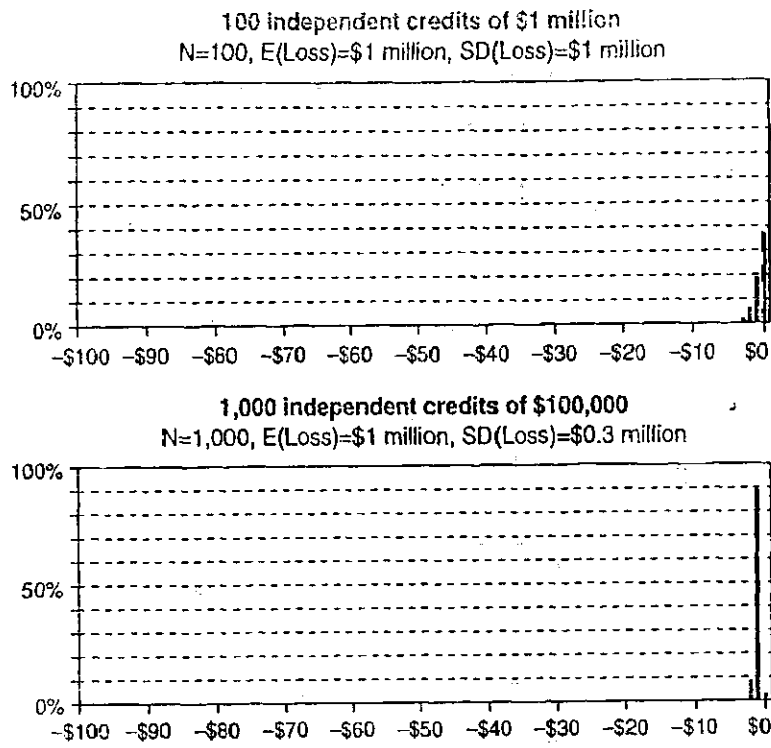
FIGURE 18.3b   Distribution of Credit Losses

1,000 independent credits of $100,000
N=1,000, E(Loss)=$1 million, SD(Loss)=$0.3 million
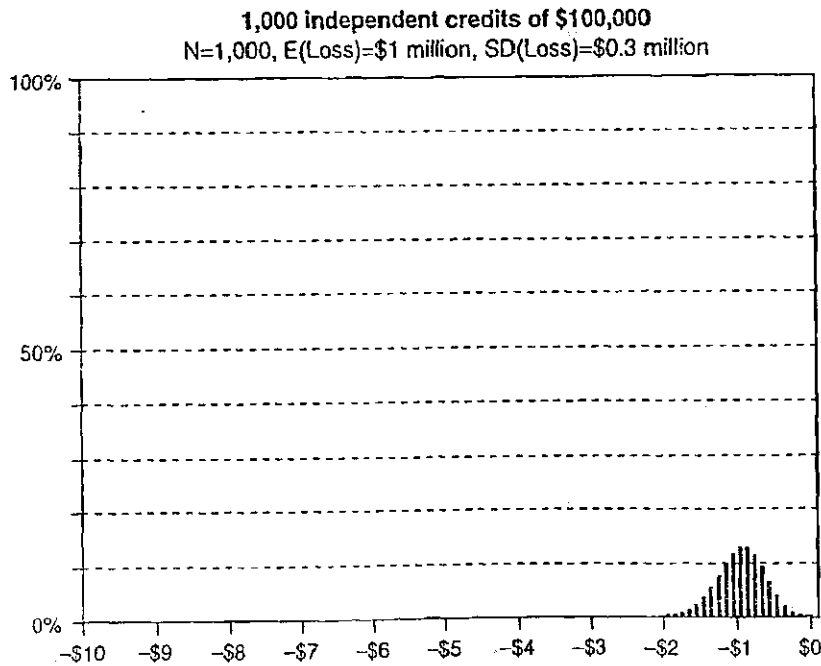
FIGURE 18.3c   Distribution of Credit Losses

In the first panel, we have one loan only. We can either have no default, with probability 99%, or a loss of $100 million with probability 1%. The expected loss is

$$EL = 0.01 \times \$100 + 0.99 \times 0 = \$1 \text{ million}$$

The problem, of course, is that, if default occurs, it will be a big hit to the bottom line, possibly bankrupting the lending bank.

Basically, this is what happened to Peregrine Investments Holdings, one of Hong Kong's leading investment banks which shut down due to the Asian crisis of 1997. The bank failed in large part from a single loan to PT Steady Safe, an Indonesian taxi-cab operator, that amounted to $235 million, a quarter of the bank's equity capital.

In the case of our single loan, the spread of the distribution is quite large, with a variance of 99, which implies a standard deviation (SD) of about $10 million. Simply focusing on the standard deviation, however, ignores the severe skewness in the distribution.

In the second panel, we consider 10 loans, each for $10 million. The total notional is the same as before. We assume that defaults are independent. The expected loss is still $1 million, or $10 \times 0.01 \times \$10$ million. The SD, however, is now $3 million, much less than before.

Next, the third panel considers 100 loans of $1 million each. The expected loss is still $1 million, but the SD is now $1 million, even lower. Finally, the fourth panel considers 1,000 loans of $100,000, which create a SD of $0.3 million.

For comparability, all these graphs use the same vertical and horizontal scale. This, however, does not reveal the distributions fully. This is why the fifth panel expands the distribution with 1,000 counterparties, which looks similar to a normal distribution. This reflects the central limit theorem, which states that the distribution of the sum of *independent* variables tends to a normal distribution. Remarkably, even starting from a highly skewed distribution, we end up with a normal distribution due to diversification effects.

In addition, the spread of the distribution becomes very small. This explains why portfolios of consumer loans, which are spread over a large number of credits, are less risky than typical portfolios of corporate loans.

With N events that occur with the same probability p, define the variable $X = \sum_{i=1}^{N} b_i$ as the number of defaults (where $b_i = 1$ when default occurs). The expected credit loss on our portfolio is then

$$E[CL] = E[X] \times \$100/N = pN \times \$100/N = p \times \$100 \qquad (18.9)$$

which depends not on N but rather on the average probability of default and total exposure, $100 million. When the events are independent, the variance of this variable is, using the results from a binomial distribution,

$$V[CL] = V[X] \times (\$100/N)^2 = p(1-p)N \times (\$100/N)^2 \qquad (18.10)$$

which gives a standard deviation of

$$SD[CL] = \sqrt{p(1-p)} \times \$100/\sqrt{N} \qquad (18.11)$$

For a constant total notional, this shrinks to zero as N increases.

We should note the crucial assumption that the credits are independent. When this is not the case, the distribution will lose its asymmetry more slowly. Even with a very large number of consumer loans, the dispersion will not tend to zero because the general state of the economy is a common factor behind consumer credits. Indeed, many more defaults occur in a recession than in an expansion. This is one of the reasons for being suspicious of credit risk models that have been calibrated over periods of expansion.

Institutions loosely attempt to achieve diversification by concentration limits. In other words, they limit the extent of exposure, say loans, to a particular industrial or geographical sector. The rationale behind this is that defaults are more highly correlated within sectors than across sectors. Conversely, concentration risk is the risk that too many defaults could occur at the same time.

The distributions in this section were derived using closed-form solutions, which assume *homogeneity*, or the same probability of default p, and independence. In this situation, all the possible states of the world can be described by using the binomial expansion. This general theorem states that

$$(x+y)^N = a_0 x^N + a_1 x^{N-1} y^1 + a_2 x^{N-2} y^2 + \ldots + a_{N-1} x^1 y^{N-1} + a_N y^N$$
$$(18.12)$$

where the coefficients a_i in this expansion are the number of combinations of N things taken i at a time, or

$$a_i = \binom{N}{i} = \frac{N!}{i!(N-i)!} \qquad (18.13)$$

If we define $x = p$ and $y = 1 - p$, the expansion must sum to unity, and each term gives the probability of this particular combination of events.

$$1 = p^N + Np^{N-1}(1-p)^1 + \tfrac{N(N-1)}{2}p^{N-2}(1-p)^2 + \ldots + Np^1(1-p)^{N-1}$$
$$+ (1-p)^N$$
$$(18.14)$$

As an example, with $N = 3$, we have

$$1 = p^3 + 3p^2(1-p) + 3p(1-p)^2 + (1-p)^3 \qquad (18.15)$$

The first term gives the probability of three defaults. The second is the probability of exactly two defaults, which involves three events.[4] The last term gives the

---

[4] The events are (1) bond 1 in no default with bonds 2 and 3 in default, (2) bond 2 in no default with others in default, and (3) bond 3 in no default with others in default.

probability of no defaults. This is a convenient decomposition but is only valid when defaults are independent.

EXAMPLE 18.9: FRM EXAM 2002—QUESTION 82

A portfolio of bonds consists of five bonds whose default correlation is zero. The one-year probabilities of default of the bonds are: 1%, 2%, 5%, 10%, and 15%. What is the one-year probability of no default within the portfolio?

a. 71%
b. 67%
c. 85%
d. 99%

EXAMPLE 18.10: FRM EXAM 2004—QUESTION 15

There are 10 bonds in a credit default swap basket. The probability of default for each of the bonds is 5%. The probability of any one bond defaulting is completely independent of what happens to the other bonds in the basket. What is the probability that exactly one bond defaults?

a. 5%
b. 50%
c. 32%
d. 3%

18.5   IMPORTANT FORMULAS

Credit loss $= \sum_{i=1}^{N} b_i \times CE_i \times (1 - f_i)$

Joint probability with independence: $p(A \text{ and } B) = p(A) \times p(B)$

Joint probability:

$p(A \text{ and } B) = \text{Corr}(A, B)\sqrt{p(A)[1 - p(A)]}\sqrt{p(B)[1 - p(B)]} + p(A)p(B)$,

using $E[b_A \times b_B] = \text{Cov}[b_A, b_B] + E[b_A]E[b_B]$

Binomial expansion:

$1 = p^N + Np^{N-1}(1 - p)^1 + \frac{N(N-1)}{2}p^{N-2}(1 - p)^2 + \ldots + Np^1(1 - p)^{N-1} + (1 - p)^N$

18.6 ANSWERS TO CHAPTER EXAMPLES

Example 18.1: FRM Exam 2000—Question 36

a. Settlement risk is due to the exchange of notional principal in different currencies at different points in time, which exposes one counterparty to default after it has made payment. There would be less risk with netted payments.

Example 18.2: FRM Exam 2000—Question 85

b. Answers c. and d. are both correct. Answers a. and b. are contradictory. A multilateral netting system concentrates the credit risk into one institution. This could potentially create much damage if this institution fails.

Example 18.3: FRM Exam 2002—Question 130

a. The Expected Credit Loss (ECL) is the notional amount times the probability of default times the loss given default. This is $50,000,000 \times 0.03 \times (1 - 70\%)$ $= \$450,000$.

Example 18.4: FRM Exam 2003—Question 17

c. The expected loss is $\sum_i p_i \times CE_i \times (1 - f_i) = \$40,000,000 \times 0.03(1 - 0.70) +$ $\$60,000,000 \times 0.05(1 - 0.45) = \$2,010,000$.

Example 18.5: FRM Exam 2007—Question 73

d. The ECL is for the first bond $1,000,000 \times 3\% \times (1 - 60\%) = 12,000$, and for the second $600,000 \times 5\% \times (1 - 40\%) = 18,000$. This adds up to $30,000. Note that this number does not depend on the default correlation.

Example 18.6: FRM Exam 2007—Question 74

d. Here, the joint default probability matters. If the two bonds default, the loss is $\$1,000,000 \times (1 - 40\%) + \$600,000 \times (1 - 60\%) = \$400,000 + \$360,000 = \$760,000$. This will happen with probability 1.27%. The next bigger loss is $400,000, which has probability of $3.00 - 1.27 = 1.73\%$. Its cumulative probability must be $100.00 - 1.17 = 98.73\%$. This is slightly above 98%, so $400,000 is the quantile at the 98% level of confidence or higher. Subtracting the mean gives $370,000.

Example 18.7: FRM Exam 2007—Question 102

b. The joint PD does not matter for the ECL. This is $ECL = 1 \times 20\% \times 40\% + 5 \times 30\% \times 60\%, = 0.08 + 0.90$, or EUR 0.98 million.

Example 18.8: FRM Exam 2004—Question 46

b. From Equation (18.7), the default correlation is $\text{Corr}(A, B) = [p(A \text{ and } B) - p(A)p(B)]/\{\sqrt{p(A)[1 - p(A)]}\sqrt{p(B)[1 - p(B)]}\} = [0.0015 - 0.02 \times 0.04]/\{\sqrt{0.02[1 - 0.02]}\sqrt{0.04[1 - 0.04]}\} = 0.025516.$

Example 18.9: FRM Exam 2002—Question 92

a. Because the events are independent, the joint probability is given by the product $(1 - p_1)(1 - p_2)(1 - p_3)(1 - p_4)(1 - p_5) = (1 - 1\%)(1 - 2\%)(1 - 5\%)(1 - 10\%)(1 - 20\%) = 70.51\%.$

Example 18.10: FRM Exam 2004—Question 15

c. Using the second term in Equation (18.14), we have $a_1 = \binom{10}{1} = 10$, and the probability is $10p^1(1 - p)^9 = 10 \times 0.05 \times (1 - 0.05)^9 = 0.315.$
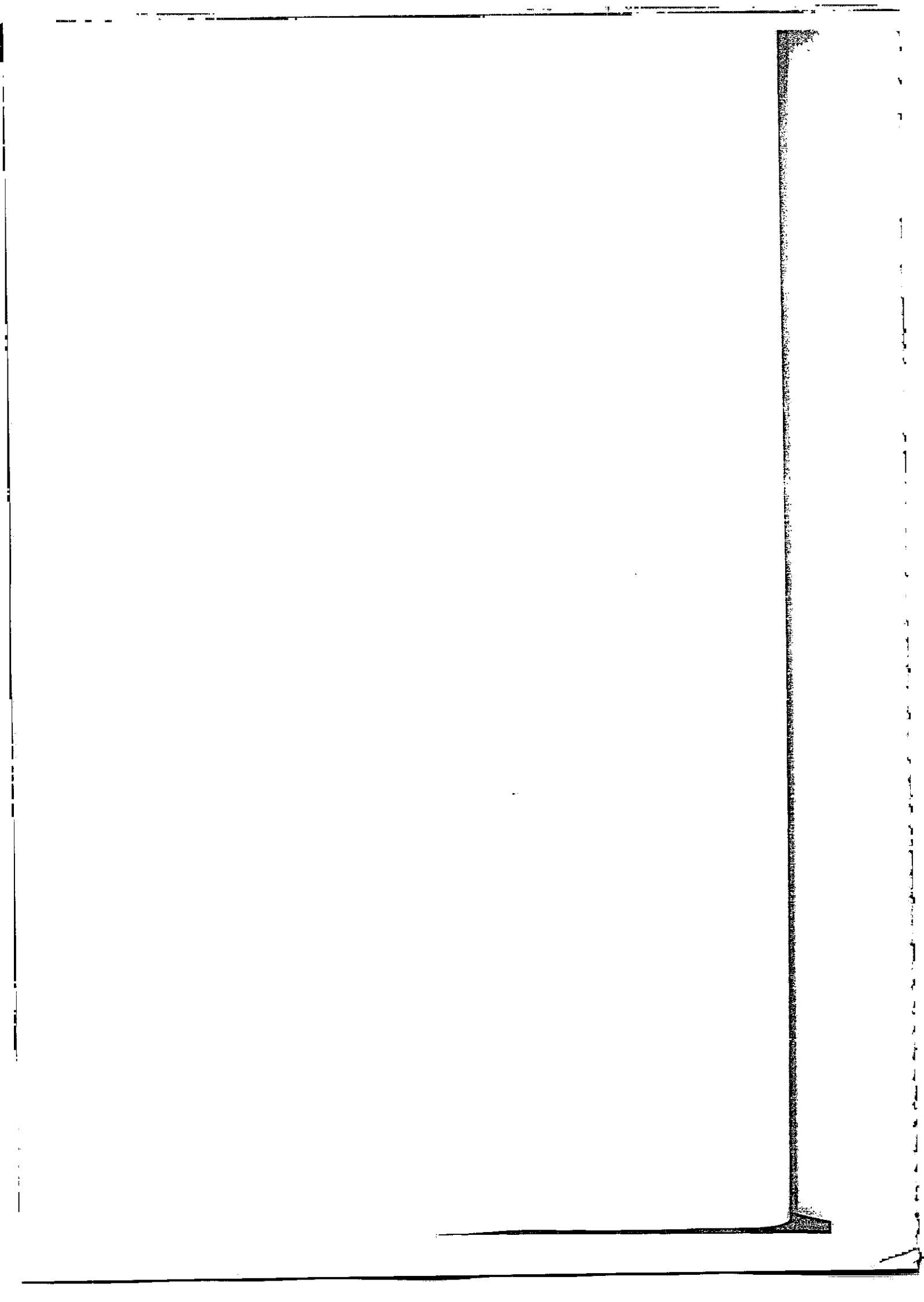
Measuring Actuarial Default Risk

Default risk is the primary driver of credit risk. It is represented by the probability of default (PD). When default occurs, the actual loss is the combination of exposure at default (EAD) and loss given default (LGD).

Default risk can be measured using two approaches: (1) actuarial methods, which provide "objective" measures of default rates, usually based on historical default data; and (2) market-price methods, which infer from traded prices of debt, equity, or credit derivatives "risk-neutral" measures of default risk.

Risk-neutral measures were introduced earlier in relation to option prices. They provide a useful shortcut for asset *pricing*. A major benefit of risk-neutral measures is that they are forward looking and responsive to the latest news because they are based on current market prices.

For *risk management* purposes, however, they are contaminated by the effect of risk premiums and contain measures of loss given default. As a result, they do not directly measure default probabilities. In contrast, objective measures describe the "actual" or "natural" probability of default.

Actuarial measures of default probabilities are provided by credit rating agencies, which classify borrowers by credit ratings that are supposed to quantify default risk. Such ratings are external to the firm. Similar techniques can be used to develop internal ratings.

Ratings usually start with accounting variables models, which relate the occurrence of default to a list of accounting variables, or more generally firm characteristics. Statistical techniques such as discriminant analysis can be used to examine how these variables are related to the occurrence or nonoccurrence of default. Accounting variables are augmented by information from financial markets and about the economic environment. Rating agencies also have access to management and private information about the firm.

This chapter focuses on actuarial measures of default risk. Market-based measures of default risk will be examined in the next chapter. Section 19.1 examines first the definition of a credit event. Section 19.2 then examines credit ratings, describing how historical default rates can be used to infer default probabilities. Recovery rates are analyzed in Section 19.3. Finally, Section 19.4 broadly discusses corporate and sovereign credit risk, as well as the role of credit rating agencies.

451

19.1 CREDIT EVENT

A credit event is a discrete state of the world. Either it happens or does not, depending on the definition of the event, which must be framed as precisely as possible.

The definition of default for a bond obligation is very narrow. Default on a bond occurs when payment on that bond is missed. The state of default is defined by Standard & Poor's (S&P), a credit rating agency, as

> *The first occurrence of a payment default on any financial obligation, rated or un-rated, other than a financial obligation subject to a bona fide commercial dispute; an exception occurs when an interest payment missed on the due date is made within the grace period.*

Default on a particular bond, however, generally reflects the creditor's financial distress and is typically accompanied by default on other obligations. This is why rating agencies give a credit rating for the *issuer* in addition to a rating for specific bonds. The rating for specific bonds can be higher or lower than this issuer rating, depending on their relative priority.

This definition, however, needs to be defined more precisely for credit derivatives, whose payoffs are directly related to credit events. We will cover credit derivatives in Chapter 22. The definition of a credit event has been formalized by the International Swaps and Derivatives Association (ISDA), an industry group, which lists these events:

- Bankruptcy, which is defined as a situation involving either of
  - The *dissolution* of the obligor (other than merger)
  - The *insolvency*, or inability to pay its debt
  - The *assignment* of claims
  - The *institution of bankruptcy proceeding*
  - The *appointment of receivership*
  - The *attachment of substantially all assets by a third party*
- Failure to pay, which means failure of the creditor to make due payment; this is usually triggered after an agreed-upon grace period and when the payment due is above a certain amount
- Obligation/cross default, which means the occurrence of a default (other than failure to make a payment) on any other similar obligation
- Obligation/cross acceleration, which means the occurrence of a default (other than failure to make a payment) on any other similar obligation, resulting in that obligation becoming due immediately
- Repudiation/moratorium, which means that the counterparty is rejecting, or challenging, the validity of the obligation
- Restructuring, which means a waiver, deferral, or rescheduling of the obligation with the effect that the terms are less favorable than before

Notably, credit events occurred in 2008 for Fannie Mae and Freddie Mac (receivership), Lehman Brothers Holdings and Washington Mutual (bankruptcy),

and Ecuador (failure to pay). In addition, other events sometimes included are

- Downgrading, which means the credit rating is lower than previously, or is withdrawn
- Currency inconvertibility, which means the imposition of exchange controls or other currency restrictions by a governmental or associated authority
- Governmental action, which means either (1) declarations or actions by a government or regulatory authority that impair the validity of the obligation, or (2) the occurrence of war or other armed conflict that impairs the functioning of the government or banking activities

Ideally, the industry should agree on a common set of factors defining a credit event, to minimize the possibility of disagreements and costly legal battles that create uncertainty for everybody. The ISDA definitions are designed to minimize legal risk by precisely wording the definition of credit event.

Even so, unforeseen situations sometimes develop. For example, there have been differences of opinion as to whether a bank debt restructuring constitutes a credit event, as in the recent cases of Conseco, Xerox, and Marconi.

Another notable situation is that of Argentina, which represents the largest sovereign default recorded so far, in terms of external debt. Argentina announced in November 2001 a restructuring of its local debt that was more favorable to itself. Some holders of credit default swaps argued that this was a "credit event," since the exchange was coerced, and that they were entitled to payment. Swap sellers disagreed. This became an unambiguous default; however, when Argentina announced in December it would stop paying interest on its $135 billion foreign debt. Nonetheless, the situation was unresolved for holders of credit swaps that expired just before the official default.

EXAMPLE 19.1: DEFINITION OF A CREDIT EVENT

Which of the following events is *not* a "credit event"?

a. Bankruptcy
b. Calling back a bond
c. Downgrading
d. Default on payments

19.2 DEFAULT RATES

19.2.1 Credit Ratings

A credit rating is an "evaluation of creditworthiness" issued by a credit rating agency (CRA). The major U.S. bond rating agencies are Moody's Investors

Services, Standard & Poor's (S&P), and Fitch. More technically, a credit rating is defined by Moody's as an

> *opinion of the future ability, legal obligation, and willingness of a bond issuer or other obligor to make full and timely payments on principal and interest due to investors.*

Table 19.1 presents the interpretation of various credit ratings issued by Moody's and S&P. These ratings correspond to long-term debt; other ratings apply to short-term debt. Generally, the two agencies provide similar ratings for the same issuer.

Ratings are divided into the following:

- Investment grade, that is, at and above BBB for S&P and Baa for Moody's
- Speculative grade, or below investment grade, for the rest

Each letter is known as a class. In addition, the CRA use modifiers, also called notches. For instance, S&P subdivides the BBB category into BBB+, BBB, and BBB−. For Moody's, the equivalent ratings are Baa1, Baa2, and Baa3.

These ratings represent objective (or actuarial) probabilities of default.[1] Indeed, the agencies have published studies that track the frequency of bond defaults, classified by initial ratings for different horizons. These frequencies can be used to convert ratings to default probabilities.

The agencies use a number of criteria to decide on the credit rating, including various accounting ratios. Table 19.2 presents median value for selected accounting ratios for U.S. industrial corporations. The first column (under "leverage") shows that the ratio of total debt to total capital (debt plus book equity, or assets)

TABLE 19.1 Classification by Credit Ratings

Explanation	Standard & Poor's	Moody's Services
Investment grade:		
Highest grade	AAA	Aaa
High grade	AA	Aa
Upper medium grade	A	A
Medium grade	BBB	Baa
Speculative grade:		
Lower medium grade	BB	Ba
Speculative	B	B
Poor standing	CCC	Caa
Highly speculative	CC	Ca
Lowest quality, no interest	C	C
In default	D	

Modifiers: A+, A, A−, and A1, A2, A3

---

[1] In fact, the ratings measure the probability of default (PD) for S&P and the joint effect of PD×LGD for Moody's, where LGD is the proportional loss given default.

TABLE 19.2 S&P's Industrial Financial Ratios Across Ratings (2005 to 2007 Averages)

Rating	Leverage: (Percent) Total Debt/Capital	Cash Flow Coverage: (Multiplier)	
		EBITDA/Interest	EBIT/Interest
AAA	12	32.0	26.2
AA	35	19.5	16.4
A	37	13.5	11.2
BBB	45	7.8	5.8
BB	53	4.8	3.4
B	73	2.3	1.4
CCC	99	1.1	0.4

varies systematically across ratings. Highly rated companies have low leverage ratios, 12% for AAA firms. In contrast, BB-rated (just below investment grade) companies have a leverage ratio of 53%. This implies a capital-to-equity leverage ratio of 2.1 to 1.[2]

The right-hand panel (under "cash flow") also shows systematic variations in a measure of free cash flow divided by interest payments. This represents the number of times the cash flow can cover interest payments. Focusing on earnings before interest and taxes (EBIT), AAA-rated companies have a safe cushion of 26.2, whereas BB-rated companies have coverage of only 3.4.

A related model for bankruptcy prediction is the multiple discriminant analysis (MDA), such as the z-score model developed by Altman.[3] MDA constructs a linear combination of accounting data that provides the best fit with the observed states of default and non-default for the sample firms.

The variables used in the z-score are (1) working capital over total assets, (2) retained earnings over total assets, (3) EBIT over total assets, (4) market value of equity over total liabilities, and (5) net sales over total assets. Lower scores indicate a higher likelihood of default. Each variable enters with a positive sign, meaning that an increase in each of these variables decreases the probability of bankruptcy.

EXAMPLE 19.2: FRM EXAM 2003—QUESTION 100

What is the lowest tier of an investment grade credit rating by Moody's?

a. Baa1
b. Ba1
c. Baa3
d. Ba3

---

[2] Defining D, E as debt and equity, the debt-to-asset ratio is $D/(D + E) = 53\%$. We then have an asset-to-equity ratio of $(D + E)/E = [D/E] + 1 = [D/(D + E)]/[E/(D + E)] + 1 = 53\%/(1 - 53\%) + 1 = 2.1$

[3] Altman, E. (1968), Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance* 23, 589–609.

EXAMPLE 19.3: FRM EXAM 2005—QUESTION 86

You are considering an investment in one of three different bonds. Your investment guidelines require that any bond you invest in carry an investment-grade rating from at least two recognized bond rating agencies. Which, if any, of the bonds listed below would meet your investment guidelines?

a. Bond A carries an S&P rating of BB and a Moody's rating of Baa.
b. Bond B carries an S&P rating of BBB and a Moody's rating of Ba.
c. Bond C carries an S&P rating of BBB and a Moody's rating of Baa.
d. None of the above.

EXAMPLE 19.4: FRM EXAM 2002—QUESTION 110

If Moody's and S&P are equally good at rating bonds, the average default rate on BB bonds by S&P will be lower than the average default rate on bonds rated by Moody's as

a. Baa3
b. Ba1
c. Ba
d. Ba3

19.2.2 Historical Default Rates

Tables 19.3 and 19.4 display historical default rates as reported by Moody's and Standard and Poor's, respectively. These describe the proportion of firms that default, $\bar{X}$, which is a statistical estimate of the true default probability:

$$E(\bar{X}) = p \tag{19.1}$$

For example, borrowers with an initial Moody's rating of Baa experienced an average default rate of 0.29% over the next year. Similar rates are obtained for S&P's BBB-rated credits, who experienced an average 0.23% default rate over the next year. On the other hand, A-rated firms experience a default rate around 0.07% over the next year. Firms at or below Caa have a default rate of 13.73%. Higher ratings are associated with lower default rates. As a result, this information could be used to derive estimates of default probability for an initial rating class.

In addition, the tables show that the cumulative default rate increases sharply with the horizon, for a given initial credit rating. The default rate for Baa firms

TABLE 19.3 Moody's Cumulative Default Rates (Percent), 1920–2007

					Year					
Rating	1	2	3	4	5	6	7	8	9	10
Aaa	0.00	0.00	0.02	0.08	0.16	0.26	0.37	0.53	0.70	0.90
Aa	0.06	0.18	0.29	0.45	0.70	1.01	1.34	1.65	1.95	2.29
A	0.07	0.24	0.50	0.81	1.12	1.45	1.80	2.13	2.50	2.90
Baa	0.29	0.85	1.56	2.34	3.14	3.94	4.71	5.48	6.28	7.06
Ba	1.34	3.20	5.32	7.49	9.59	11.56	13.36	15.11	16.73	18.44
B	4.05	8.79	13.49	17.72	21.43	24.66	27.59	30.04	32.15	33.93
Caa-C	13.73	22.46	29.03	33.92	37.64	40.58	42.87	44.92	47.00	48.98
Inv.	0.14	0.43	0.81	1.23	1.69	2.16	2.63	3.09	3.58	4.08
Spec.	3.59	7.24	10.75	13.92	16.71	19.18	21.37	23.34	25.11	26.83
All	1.41	2.88	4.32	5.63	6.80	7.85	8.80	9.67	10.48	11.28

					Year					
Rating	11	12	13	14	15	16	17	18	19	20
Aaa	1.07	1.21	1.36	1.41	1.45	1.53	1.62	1.68	1.77	1.83
Aa	2.68	3.10	3.51	3.93	4.25	4.49	4.68	4.88	5.09	5.27
A	3.34	3.77	4.15	4.50	4.92	5.28	5.56	5.81	6.08	6.33
Baa	7.80	8.54	9.24	9.89	10.44	11.01	11.53	12.00	12.44	12.91
Ba	20.00	21.52	23.04	24.34	25.51	26.64	27.81	28.91	29.85	30.78
B	35.64	37.26	38.69	40.08	41.40	42.68	43.73	44.52	45.07	45.38
Caa-C	50.99	53.07	55.05	57.11	59.12	60.98	62.63	64.20	65.68	67.13
Inv.	4.58	5.09	5.57	6.00	6.42	6.79	7.12	7.41	7.71	8.00
Spec.	28.44	30.00	31.50	32.87	34.13	35.35	36.52	37.57	38.46	39.28
All	12.05	12.80	13.50	14.15	14.74	15.29	15.79	16.25	16.66	17.06

increases from 0.29% over one year to 7.06% over the following 10 years. Because these are cumulative default rates, the number must necessarily increase with the horizon. For investment-grade credits, however, the increase is more than proportional with the horizon. The ratio is $7.06/0.29 = 24$ for Baa-rated credits, which is more than 10. In contrast, the ratio for B-rated credits is $33.93/4.05 = 8$. For speculative-grade credits, the increase is less than proportional with the horizon.

One problem with such historical information, however, is the relative paucity of data in some cells. There are few defaults for highly-rated firms over a one-year horizon. For example, the one-year default rate for Aa firms is 0.06%. This corresponds to very few defaults. From 1939 to 2007, only two companies initially rated Aa defaulted during the following year, which happened in 1989. Changing the sample period or having another number of defaults could have a substantial effect on this average default rate.

In addition, the sample size decreases as the horizon lengthens. For instance, S&P reports default rates for horizons up to 15 years using data from 1981 to 2007. The one-year default rate is an average using 27 periods—that is, 1981, 1982, and so on to 2007. For 15-year horizons, however, the first period is 1981

TABLE 13.4 S&P's Cumulative Global Default Rates (Percent), 1981–2007

Rating	Year														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AAA	0.00	0.00	0.09	0.18	0.28	0.41	0.48	0.59	0.63	0.67	0.67	0.67	0.67	0.73	0.79
AA	0.01	0.05	0.09	0.19	0.29	0.40	0.52	0.62	0.71	0.81	0.91	0.99	1.09	1.17	1.21
A	0.06	0.16	0.29	0.45	0.64	0.85	1.11	1.32	1.53	1.76	1.95	2.11	2.26	2.39	2.61
BBB	0.23	0.65	1.13	1.75	2.38	2.98	3.47	3.96	4.42	4.89	5.37	5.75	6.22	6.68	7.20
BB	1.00	2.93	5.19	7.36	9.30	11.19	12.72	14.05	15.27	16.24	17.13	17.87	18.51	18.96	19.43
B	4.57	10.06	14.72	18.39	21.08	23.19	24.94	26.37	27.55	28.74	29.80	30.70	31.61	32.47	33.26
CCC	25.59	34.06	39.04	41.86	44.50	45.62	46.67	47.25	48.86	49.76	50.50	51.26	51.87	52.50	52.50
Inv.	0.10	0.30	0.52	0.81	1.11	1.42	1.69	1.95	2.19	2.44	2.66	2.85	3.05	3.24	3.47
Spec.	4.11	8.11	11.66	14.57	16.90	18.84	20.45	21.79	23.01	24.08	25.05	25.87	26.64	27.30	27.90
All	1.45	2.91	4.21	5.33	6.26	7.06	7.73	8.30	8.81	9.29	9.72	10.08	10.44	10.76	11.09

to 1995, the second is 1982 to 1996, and so on until the last period of 1993 to 2007. The average therefore uses only 13 periods, which is a much shorter sample. The data is also overlapping and therefore not independent. So, omitting or adding a few borrowers can drastically alter the reported default rate.

This can lead to inconsistencies in the tables. For instance, the default rate for CCC obligors is the same, at 52.50%, from year 14 to 15. This implies that there is no further risk of default after 14 years, which is an unrealistic implication. Also, when the categories are further broken down into modifiers, default rates sometimes do not decrease monotonically as the rating increases, which is probably a small-sample effect.

We can try to assess the accuracy of these default rates by computing their standard error. Consider for instance the default rate over the first year for AA-rated credits, which averaged out to $\bar{X} = 0.01\%$ in this S&P sample. Assume that this was taken out of a total of about $N = 10,000$ independent observations. The variance of the average is, from the distribution of a binomial process,

$$V(\bar{X}) = \frac{p(1-p)}{N} \tag{19.2}$$

which gives a standard error of about 0.01%. This is on the same order as the average of 0.01%, indicating that there is substantial imprecision in this average default rate. So we could not really distinguish an AA credit from an AAA credit.

The problem is made worse with lower sample sizes, which is the case in non-U.S. markets or when the true p is changing over time. For instance, if we observe a 5% default rate among 100 observations, the standard error becomes 2.2%, which is very large. Therefore, a major issue with credit risk is that estimation of default rates for low-probability events can be very imprecise.

19.2.3 Cumulative and Marginal Default Rates

The default rates reported in Tables 19.3 and 19.4 are cumulative default rates for an initial credit rating; that is, they measure the total frequency of default *at any time* between the starting date and year T. It is also informative to measure the marginal default rate, which is the frequency of default *during* year T.

The default process is illustrated in Figure 19.1. Here, d_1 is the marginal default rate during year 1, and d_2 is the marginal default rate during year 2. To default during the second year, the firm must have survived the first year and defaulted in the second. Thus, the probability of defaulting in year 2 is given by $(1 - d_1)d_2$. The cumulative probability of defaulting up to year 2 is then $C_2 = d_1 + (1 - d_1)d_2$. Subtracting and adding one, this is also $C_2 = 1 - (1 - d_1)(1 - d_2)$, which perhaps has a more intuitive interpretation, as this is 1 minus the probability of surviving the entire period.

More formally,

- $m[t + N \mid R(t)]$ is the number of issuers rated R at the end of year t that default in year $T = t + N$
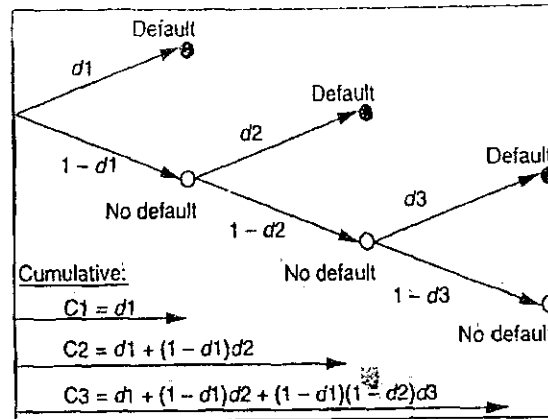
FIGURE 19.1   Sequential Default Process

- $n[t + N \mid R(t)]$ is the number of issuers rated R at the end of year t that have not defaulted by the beginning of year $t + N$

Marginal Default Rate during Year T   This is the proportion of issuers initially rated R at initial time t that default in year T, relative to the remaining number at the beginning of the same year T:

$$d_N(R) = \frac{m[t + N \mid R(t)]}{n[t + N \mid R(t)]}$$

Survival Rate   This is the proportion of issuers initially rated R that will not have defaulted by T:

$$S_N(R) = \Pi_{i=1}^{N}(1 - d_i(R)) \tag{19.3}$$

Marginal Default Rate from Start to Year T   This is the proportion of issuers initially rated R that defaulted in year T, relative to the initial number in year t. For this to happen, the issuer will have survived until year $t + N - 1$, and then default the next year. Hence, this is:

$$k_N(R) = S_{N-1}(R)d_N(R) \tag{19.4}$$

Cumulative Default Rate   This is the proportion of issuers rated R that defaulted at any point until year T:

$$C_N(R) = k_1(R) + k_2(R) + \cdots + k_N(R) = 1 - S_N(R) \tag{19.5}$$

Average Default Rate   We can express the total cumulative default rate as an average, per period default rate d, by setting

$$C_N = 1 - \Pi_{i=1}^{N}(1 - d_i) = 1 - (1 - d)^N \tag{19.6}$$

As we move from annual to semiannual and ultimately continuous compounding, the average default rate becomes

$$C_N = 1 - (1 - d^a)^N = 1 - (1 - d^s/2)^{2N} \to 1 - e^{-d^c N} \qquad (19.7)$$

where d^a, d^s, d^c are default rates using annual, semiannual, and continuous compounding. This is equivalent to the various definitions for the compounding of interest.

Example: Computing Cumulative Default Probabilities

Consider a B-rated firm that has default rates of $d_1 = 5\%$, $d_2 = 7\%$. Compute the cumulative default probabilities.

Answer

In the first year, $k_1 = d_1 = 5\%$. After one year, the survival rate is $S_1 = 0.95$. The probability of defaulting in year 2 is then $k_2 = S_1 \times d_2 = 0.95 \times 0.07 = 6.65\%$. After two years, the survival rate is $(1 - d_1)(1 - d_2) = 0.95 \times 0.93 = 0.8835$. Thus, the cumulative probability of defaulting in years 1 and 2 is $5\% + 6.65\% = 11.65\%$.

Based on this information, we can map these *forward*, or marginal, default rates from cumulative default rates for various credit ratings. Figure 19.2, for instance, displays cumulative default rates reported by Moody's in Table 19.3. The corresponding marginal default rates are plotted in Figure 19.3.

It is interesting to see that the marginal probability of default increases with maturity for initial high credit ratings, but decreases for initial low credit ratings. The increase is due to a mean reversion effect. The fortunes of an Aaa-rated firm can only stay the same at best, and often will deteriorate over time. In contrast, a B-rated firm that has survived the first few years must have a decreasing probability of defaulting as time goes by. This is a survival effect.
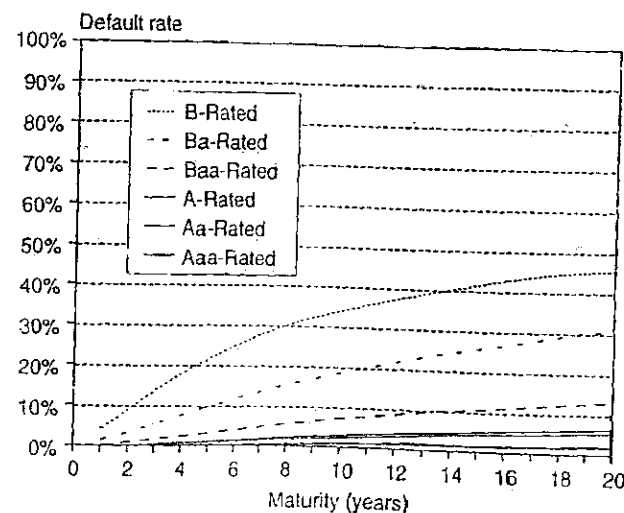
FIGURE 19.2 Moody's Cumulative Default Rates, 1920–2007
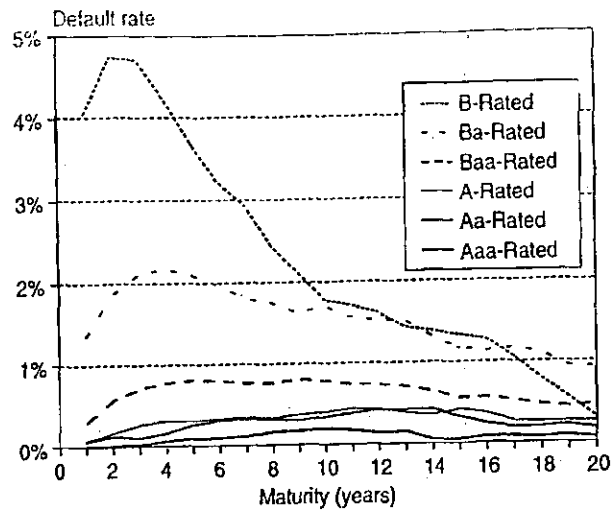
FIGURE 19.3   Moody's Marginal Default Rates, 1920–2007

EXAMPLE 19.5: FRM EXAM 2004—QUESTION 1

Company ABC was incorporated on January 1, 2004. It has an expected annual default rate of 10%. Assuming a constant quarterly default rate, what is the probability that company ABC will *not* have defaulted by April 1, 2004?

a. 2.40%

b. 2.50%

c. 97.40%

d. 97.50%

EXAMPLE 19.6: FRM EXAM 2002—QUESTION 77

If the default probability for an A-rated company over a three-year period is 0.30%, then the most likely probability of default for this company over a six-year period is:

a. 0.30%

b. Between 0.30% and 0.60%

c. 0.60%

d. Greater than 0.60%

EXAMPLE 19.7: FRM EXAM 2006—QUESTION 21

What is the survival rate at the end of three years if the annual default probabilities are 8%, 12%, and 15% in the first, second, and third years, respectively?

a. 68.8%
b. 39.1%
c. 99.9%
d. 65.0%

EXAMPLE 19.8: FRM EXAM 2004—QUESTION 14

A corporate bond will mature in three years. The marginal probability of default in year one is 3%. The marginal probability of default in year two is 4%. The marginal probability of default in year three is 6%. What is the cumulative probability that default will occur during the three-year period?

a. 12.47%
b. 12.76%
c. 13%
d. 13.55%

EXAMPLE 19.9: FRM EXAM 2000—QUESTION 34

What is the difference between the marginal default probability and the cumulative default probability?

a. Marginal default probability is the probability that a borrower will default in any given year, whereas the cumulative default probability is over a specified multiyear period.
b. Marginal default probability is the probability that a borrower will default due to a particular credit event, whereas the cumulative default probability is for all possible credit events.
c. Marginal default probability is the minimum probability that a borrower will default, whereas the cumulative default probability is the maximum probability.
d. Both a. and c. are correct.

19.2.4 Transition Probabilities

As we have seen, the measurement of long-term default rates can be problematic with small sample sizes. The computation of these default rates can be simplified by assuming a Markov process for the ratings migration, described by a transition matrix. Migration is a discrete process that consists of credit ratings changing from one period to the next.

The transition matrix gives the probability of moving to one rating conditional on the rating at the beginning of the period. The usual assumption is that these moves follow a Markov process, or that migrations across states are independent from one period to the next.[4] This type of process exhibits *no carry-over effect*. More formally, a Markov chain describes a stochastic process where the conditional distribution, given today's value, is constant over time. Only present values are relevant.

Table 19.5 gives an example of a simplified transition matrix for four states, A, B, C, D, where the last represents default. Consider a company in year 0 in the B category. The company could default

- In year 1, with probability $D[t_1 \mid B(t_0)] = P(D_1 \mid B_0) = 3\%$.
- In year 2, after going from B to A in the first year, then A to D in the second; or from B to B, then to D; or from B to C, then to D. The default probability is $P(D_2 \mid A_1)P(A_1) + P(D_2 \mid B_1)P(B_1) + P(D_2 \mid C_1)P(C_1) = 0.00 \times 0.02 + 0.03 \times 0.93 + 0.23 \times 0.02 = 3.25\%$.

The cumulative probability of default over the two years is then $3\% + 3.25\% = 6.25\%$. Figure 19.4 illustrates the various paths to default in years 1, 2, and 3.

The advantage of using this approach is that the resulting data are more robust and consistent. For instance, the 15-year cumulative default rate obtained this way will always be greater than the 14-year default rate.

TABLE 19.5 Credit Ratings Transition Probabilities

State Starting	Ending				Total Prob.
	A	B	C	D	
A	0.97	0.03	0.00	0.00	1.00
B	0.02	0.93	0.02	0.03	1.00
C	0.01	0.12	0.64	0.23	1.00
D	0	0	0	1.00	1.00

[4] There is some empirical evidence, however, that credit downgrades are not independent over time but instead display a momentum effect.

Paths to default:

B → D 0.03 = 0.0300 B → A → D 0.02 × 0.00 = 0.0000 B → A → A → D 0.02 × 0.97 × 0.00 = 0.0000
B → B → D 0.93 × 0.03 = 0.0279 B → A → B → D 0.02 × 0.03 × 0.03 = 0.0000
B → C → D 0.02 × 0.23 = 0.0046 B → A → C → D 0.02 × 0.00 × 0.23 = 0.0000
B → B → A → D 0.93 × 0.02 × 0.00 = 0.0000
B → B → B → D 0.93 × 0.93 × 0.03 = 0.0259
B → B → C → D 0.93 × 0.02 × 0.23 = 0.0043
B → C → A → D 0.02 × 0.00 × 0.00 = 0.0000
B → C → B → D 0.02 × 0.12 × 0.03 = 0.0001
B → C → C → D 0.02 × 0.64 × 0.23 = 0.0029

Default prob:	0.0300	0.0325	0.0333
Cumulative:	0.0300	0.0625	0.0958
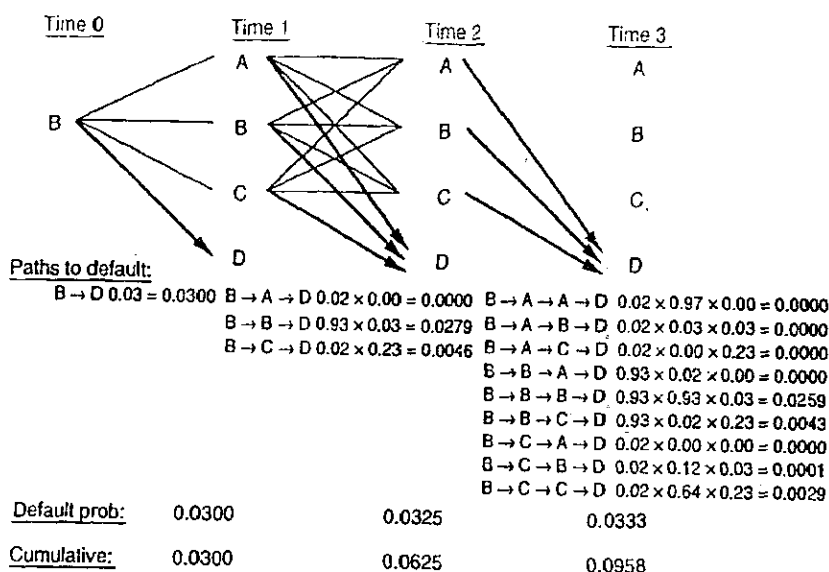
FIGURE 19.4 Paths to Default

EXAMPLE 19.10: FRM EXAM 2005—QUESTION 105

A rating transition table includes sufficient information to find all but the following item:

a. The likelihood that an AA-rated firm will fall to a BB rating over five years
b. The price of a bond that has been downgraded to BB from BBB
c. The probability of default on a B-rated bond
d. The probability that a high-yield bond will be upgraded to investment grade

---

EXAMPLE 19.11: FRM EXAM 2007—QUESTION 51

Fitch ratings provides a table indicating that the number of A-rated issuers (1) migrating to AAA is 2, (2) to AA is 5, (3) staying at A is 40, (4) migrating to BBB is 2, and (5) going into default is 3.

Based on this information, what is the probability that an issue with a rating of A at the beginning of the year will be downgraded by the end of the year?

a. 13.46%
b. 13.44%
c. 9.62%
d. 3.85%

EXAMPLE 19.12: FRM EXAM 2003—QUESTION 59

Given the following ratings transition matrix, calculate the two-period cumulative probability of default for a B credit.

Rating at beginning of period	Rating at end of period			
	A	B	C	Default
A	0.95	0.05	0.00	0.00
B	0.03	0.90	0.05	0.02
C	0.01	0.10	0.75	0.14
Default	0.00	0.00	0.00	1.00

a. 2.0%
b. 2.5%
c. 4.0%
d. 4.5%

19.2.5 Time Variation in Default Probabilities

Defaults are also correlated with economic activity. Moody's, for example, has compared the annual default rate to the level of industrial production since 1920. Moody's reports a marked increase in the default rate in the 1930s at the time of the Great Depression and during the recent recessions. These default rates, however, do not control for structural shifts in the credit quality. In recent years, many issuers came to the market with a lower initial credit rating than in the past. This should lead to more defaults even with a stable economic environment.

To control for this effect, Figure 19.5 plots the default rate for investment-grade and speculative credits over the years 1981 to 2006. As expected, the default rate of investment-grade bonds is very low. More interestingly, however, it displays minimal variation through time. We do observe, however, significant variation in the default rate of speculative-grade credits, which peaks during the recessions that started in 1981, 1990, and 2001. Thus, economic activity significantly affects the frequency of defaults. This effect is most marked for speculative-grade bonds.

19.3 RECOVERY RATES

Credit risk also depends on the loss given default (LGD). This can be measured as 1 minus the recovery rate, or fraction recovered after default.

19.3.1 The Bankruptcy Process

Normally, default is a state that affects all obligations of an issuer equally, especially when accompanied by a bankruptcy filing. In most countries, a formal

Default rate
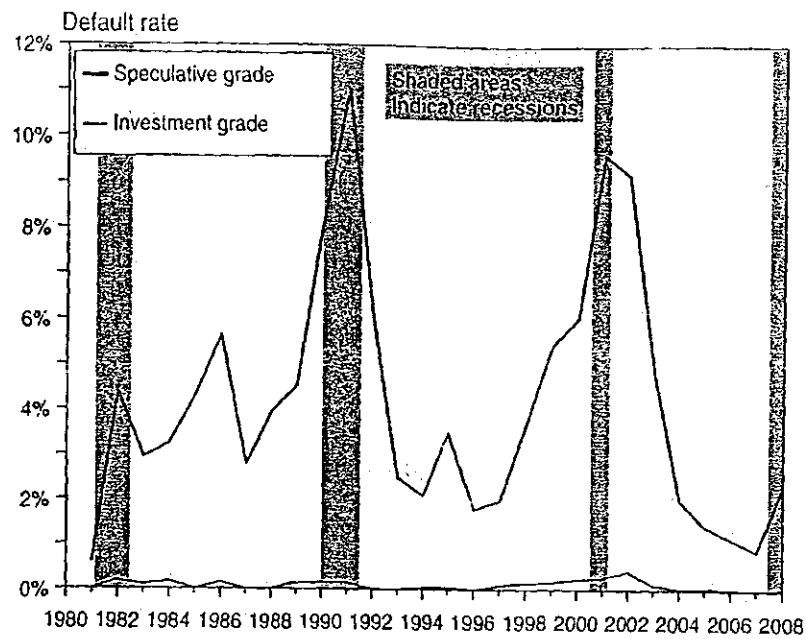
FIGURE 19.5 Time Variation in Defaults (from S&P)

bankruptcy process provides a centralized forum for resolving all the claims against the corporation. The bankruptcy process creates a pecking order for a company's creditors. This spells out the order in which creditors are paid, thereby creating differences in the recovery rate across creditors. Within each class, however, creditors should be treated equally.

In the United States, firms that are unable to make required payments can file for either Chapter 7 bankruptcy, which leads to the liquidation of the firm's assets, or Chapter 11 bankruptcy, which leads to a reorganization of the firm during which the firm continues to operate under court supervision.

Under Chapter 7, the proceeds from liquidation should be divided according to the absolute priority rule, which states that payments should be made first to claimants with the highest priority.

Table 19.6 describes the pecking order in bankruptcy proceedings. At the top of the list are secured creditors, who, because of their property right, are paid

TABLE 19.6 Pecking Order in U.S. Federal Bankruptcy Law

Seniority	Type of Creditor

Highest (paid first)	1. Secured creditors (up to the extent of secured collateral)
	2. Priority creditors
	  • Firms that lend money during bankruptcy period
	  • Providers of goods and services during bankruptcy period
	    (e.g., employees, lawyers, vendors)
	  • Taxes
	3. General creditors
	  • Unsecured creditors before bankruptcy

Lowest (paid last)	• Shareholders

to the fullest extent of the value of their collateral. Then come priority creditors, which consist mainly of post-bankruptcy creditors. Finally, general creditors can be paid if funds remain after distribution to others.

Similar rules apply under Chapter 11. In this situation, the firm must submit a reorganization plan, which specifies new financial claims to the firm's assets. The absolute priority rule, however, is often violated in Chapter 11 settlements. Junior debt holders and stockholders often receive some proceeds even though senior shareholders are not paid in full. This is allowed to facilitate timely resolution of the bankruptcy and to avoid future lawsuits. Even so, there remain sharp differences in recovery rates across seniority.

19.3.2 Estimates of Recovery Rates

Recovery rates are commonly estimated from the market prices of defaulted debt shortly after default. This is viewed as the best estimate of the future recovery and takes into account the value of the firm's assets, the estimated cost of the bankruptcy process, and various means of payment (e.g., using equity to pay bondholders), discounted into the present.

The recovery rate has been shown to depend on a number of factors:

- *The status or seniority of the debtor.* Claims with higher seniority have higher recovery rates. More generally, a greater debt cushion, or the percentage of total company debt below the instrument, also leads to higher recovery rates.
- *The state of the economy.* Recovery rates tend to be higher (lower) when the economy is in an expansion (recession).
- *The obligor's characteristics.* Recovery rates tend to be higher when the borrower's assets are tangible and when the previous rating was high. Utilities have more tangible assets, such as power-generating plants, than other industries and consequently have higher recovery rates. Also, companies with greater interest coverage, as measures by higher credit ratings, typically have higher recovery rates.
- *The type of default.* Distressed exchanges, as opposed to bankruptcy proceedings, usually lead to higher recovery rates. Unlike a bankruptcy proceeding, which causes all debts to go into default, a distressed exchange only involves the instruments that have defaulted.

Ratings can also include the loss given default. The same borrower may have various classes of debt, which may have different credit ratings due to the different level of protection. If so, debt with lower seniority should carry a lower rating.

Table 19.7 displays recovery rates for corporate debt, from Moody's. The average recovery rate for senior unsecured debt is around $f = 37\%$. Derivative instruments rank as senior unsecured creditors and should have the same recovery rates as senior unsecured debt.

Bank loans are usually secured and therefore have higher recovery rates, typically around 60%. As expected, subordinated bonds have the lowest recovery rates, typically around 20% to 30%.

TABLE 19.7 Moody's Recovery Rates for Global Corporate Debt (Percent)

Priority	Count	Mean	S.D.	Min.	10th.	Median	90th.	Max.
All Bank Loans	310	61.6	23.4	5.0	25.0	67.0	90.0	98.0
Equipment Trust	86	40.2	29.9	1.5	10.6	31.0	90.0	103.0
Sr. Secured	238	53.1	26.9	2.5	10.0	34.0	82.0	125.0
Sr. Unsecured	1,095	37.4	27.2	0.3	7.0	30.0	82.2	122.6
Sr. Subordinated	450	32.0	24.0	0.5	5.0	27.0	66.5	123.0
Subordinated	477	30.4	21.3	0.5	5.0	27.1	60.0	102.5
Jr. Subordinated	22	23.6	19.0	1.5	3.8	16.4	48.5	74.0
All Bonds	2,368	36.8	26.3	0.3	7.5	30.0	80.0	125.0

Source: Adapted from Moody's, based on 1982–2002 defaulted bond prices

There is, however, much variation around the average recovery rates. The table reports not only the average value but also the standard deviation, minimum, maximum, and 10th and 90th percentiles. Recovery rates vary widely. In addition, recovery rates are negatively related to default rates. During years with more bond defaults, prices after default are more depressed than usual. This correlation creates bigger losses, which extends the tail of the credit loss distribution. In practice, the distribution of recovery rates is often modeled with a beta distribution, which has an argument ranging from 0 to 1.

The legal environment is also a main driver of recovery rates. Differences across national jurisdictions cause differences among recovery rates. Table 19.8 compares mean recovery rates across Europe and North America. Recovery rates are significantly higher in the United States than in Europe.

Using trading prices of debt shortly after default as estimates of recovery is convenient because the bankruptcy process can be slow, often taking years. Computing the total value of payments to debt holders can also be complicated, and should take into account the time value of money.

The evidence, however, is that trading prices are on average lower than the discounted recovery rate, as shown in Table 19.9. The average discounted recovery

TABLE 19.8 Moody's Mean Recovery Rates (Percent): Europe and North America

Instrument	Europe	North America
Bank loans	47.6	61.7
Bonds		
Senior secured	52.2	52.7
Senior unsecured	25.6	37.5
Senior subordinated	24.3	32.1
Subordinated	13.9	31.3
Junior subordinated	NA	24.5
All bonds	28.4	35.3
Preferred stock	3.4	10.9
All instruments	27.6	35.9

Source: Adapted from Moody's, from 1982–2002 defaulted bond prices

TABLE 19.3   S&P's Recovery Rates for Corporate Debt (Percent)

Instrument	Trading Prices 15–45 days	Discounted Recovery
Bank loans	58.0	81.6
Senior secured bonds	48.6	67.0
Senior unsecured bonds	34.5	46.0
Senior subordinated bonds	28.4	32.4
Subordinated bonds	28.9	31.2

Source: Adapted from S&P, from 1988-2002 defaulted debt

rate is systematically higher than the indication given by trading prices. This could be due to different clienteles for the two markets, or to a risk premium in trading prices. In other words, trading prices may be artificially depressed because investors want to get risk of defaulted securities in their portfolio. If so, this creates an interesting trading opportunity. Buying the defaulted debt and working through the recovery process should create value. Indeed, this largely explains the existence of the hedge fund category called distressed securities funds. Such funds invest in selected distressed securities and benefit from their subsequent increase in value.

EXAMPLE 19.13: FRM EXAM 2005—QUESTION 74

Moody's estimates the average recovery rate for senior unsecured debt to be nearest to

a. 20%

b. 40%

c. 60%

d. 80%

EXAMPLE 19.14: FRM EXAM 2002—QUESTION 123

The recovery rate on credit instruments is defined as one minus the loss rate. The loss rate can be significantly influenced by the volatility of the value of a firm's assets before default. All other things being equal, in the event of a default, which type of company would we expect to have the highest recovery rate?

a. A trading company active in volatile markets

b. An Internet merchant of trendy consumer products

c. An asset-intensive manufacturing company

d. A highly leveraged hedge fund

19.4 ASSESSING CORPORATE AND SOVEREIGN RATING

19.4.1 Corporate Ratings

Rating agencies expend considerable effort and financial resources in coming up with publicly available credit ratings. As explained in Table 19.2, the primary inputs for the credit rating process are accounting variables such as balance sheet leverage and debt coverage. The weight assigned to these variables may change if their informativeness changes over time (i.e., if earnings management is suspected).

By nature, however, accounting information is backward looking. The economic prospects of a company are also crucial for assessing credit risk. These include growth potential, market competition, and exposure to financial risk factors. Rating agencies also have access to private information, including meetings with management during which they might be provided with confidential information.

Ratings agencies also need to account for structural differences across countries. These could arise because of a number of factors:

- *Differences in financial stability across countries.* Countries differ in terms of financial market structures and government policies. The mishandling of economic policy can turn, for instance, what should be a minor devaluation into a major problem, leading to a recession.
- *Differences in legal systems.* The protection accorded to creditors can vary widely across countries, some of which have not yet established a bankruptcy process.

In theory, ratings provided by credit rating agencies are supposed to be *consistent* across countries and industrial sectors. In other words, they should take into account such variations and represent the same probability of default.

Finally, credit ratings are supposed to look through the cycle, where a business cycle typically covers several years. This means that the rating should not depend on the current position in the business cycle. There is no point in assigning a high rating to a company enjoying peak prosperity if that performance is expected to be temporary, for instance due to high consumer demand that will revert soon to a long-run average. Figure 19.6 illustrates how corporate performance depends on the cycle. In theory, ratings should be constant through time.

In practice, a down cycle can have a lasting impact on credit quality, in extreme cases leading to default. As a result, actual ratings may be affected by the business cycle, in particular for speculative-grade firms.

Ratings less sensitive to cyclical factors should be more stable over time. This should reduce the procyclical effect of capital charges based on external credit ratings. Lower credit ratings during a recession would lead to higher capital charges, precisely after banks have suffered credit losses, which would force them to raise additional equity and to reduce lending, which would aggravate the recession. This effect is further discussed in Chapter 29. On the other hand, the practice of rating
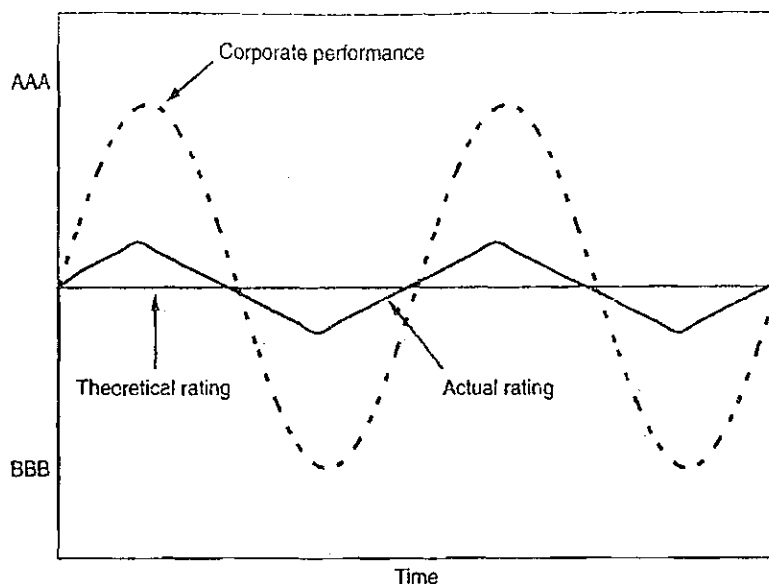
FIGURE 19.6   Business Cycle and Credit Ratings

through the cycle implies that credit ratings may underestimate the probabilities of default during a recession, and conversely during an expansion.

19.4.2 Sovereign Ratings

Rating agencies have only recently started to rate sovereign bonds. In 1975, S&P only rated seven countries, all of which were investment grade. By 1990, the pool had expanded to 31 countries, of which only nine were from emerging markets. Now, S&P rates approximately 120 countries. The history of default, however, is even more sparse. As a result, it is difficult to generalize from a very small sample.

Assessing credit risk for sovereign nations is significantly more complex than for corporates. When a corporate borrower defaults, legal action can be taken by the creditors. For instance, an unsecured creditor can file an action against a debtor and have the defendant's assets seized under a "writ of attachment." This creates a lien on its assets, or a claim on the assets as security for the payment of the debt. In contrast, it is impossible to attach the domestic assets of a sovereign nation. As a result, recovery rates on sovereign debt are usually lower than recovery rates on corporate debt. Thus, sovereign credit evaluation involves not only economic risk (the ability to repay debts when due), but also political risk (the willingness to pay). Some countries have unilaterally repudiated their debt, i.e., refused to make payments even when they had the ability to do so.

Sovereign credit ratings also differ depending on whether the debt is local currency debt or foreign currency debt. Table 19.10 displays the factors involved in local and foreign currency ratings.

Political risk factors (e.g., degree of political consensus, integration in global trade and financial system, and internal or external security risk) play an important part in sovereign credit risk. Countries with greater political stability have higher credit ratings.

TABLE 19.10 Credit Ratings Factors

Categories	Local Currency	Foreign Currency
Political risk	x	x
Income and economic structure	x	x
Economic growth prospects	x	x
Fiscal flexibility	x	x
Public debt burden	x	x
Contingent liabilities	x	x
Monetary flexibility	x	x
External liquidity		x
External debt burden		x

The second group of factors includes income and economic structure, and economic growth prospects. Countries that are richer or growing faster tend to have higher credit ratings.

The third group of factors includes fiscal flexibility, the public debt burden, and contingent liabilities. Countries with lower budget deficits and lower debt amounts in relation to the size of their economy tend to have higher credit ratings. The evaluation of debt, however, must also include public sector enterprises where a default would require government funding, as well as the financial sector, where the government may have to inject funds to ensure stability.

The fourth group includes monetary flexibility. High rates of inflation typically reflect economic mismanagement and are associated with political instability. Countries with high inflation rates tend to have low credit ratings.

All of the previous factors affect both *local currency debt* and *foreign currency debt*. Foreign currency debt is also affected by external liquidity and the external debt burden. External liquidity is assessed by balance of payment flows. Countries with large current account deficits, reflecting excess imports, tend to have low credit ratings, especially when financed by volatile capital inflows. In particular, the ratio of external interest payments to exports is closely watched. The maturity profile of flows is also important. In the case of the 1997 Asian crisis, rating agencies seem to have overlooked other important aspects of creditworthiness, such as the currency and maturity structure of national debt. Too many Asian creditors had borrowed short-term in dollars to invest in the local currency, which created a severe liquidity problem when their currency devalued. Finally, the external debt burden is assessed by the international investment position of a country (that is, public and private external debt) as well as the stock of foreign currency reserves.

Because local currency debt is backed by the taxation power of the government, local-currency debt is considered to have less credit risk than foreign-currency debt. Table 19.11 displays local- and foreign-currency debt ratings for a sample of countries. Ratings for foreign-currency debt are the same, or generally only one notch below, those of local currency debt. Similarly, sovereign debt is typically rated higher than corporate debt in the same country. Governments can repay foreign-currency debt, for instance, by controlling capital flows or seizing foreign-currency reserves.

TABLE 19.11 S&P's Sovereign Credit Ratings, December 2008

Issuer	Local Currency	Foreign Currency
Argentina	B−	B−
Australia	AAA	AAA
Belgium	AA+	AA+
Brazil	BBB+	BBB−
Canada	AAA	.AAA
China	A+	A+
France	AAA	AAA
Germany	AAA	AAA
Hong Kong	AA+	AA+
India	BBB−	BBB−
Italy	A+	A+
Japan	AA	AA
South Korea	A+	A
Mexico	A+	BBB+
Netherlands	AAA	AAA
Russia	BBB+	BBB
South Africa	A+	BBB+
Spain	AAA	AAA
Switzerland	AAA	AAA
Taiwan	AA−	AA−
Thailand	A	BBB+
Turkey	BB	BB−
United Kingdom	AAA	AAA
United States	AAA	AAA

Overall, sovereign debt ratings are considered less reliable than corporate ratings. Indeed, bond spreads are generally greater for sovereigns than corporate issuers. There are also greater differences in sovereign ratings across agencies than for corporates. Thus, the evaluation of sovereign credit risk is a much more subjective process than for corporates.

EXAMPLE 19.15: FRM EXAM 2005—QUESTION 79

In the context of evaluating sovereign risk, which of the following statements is *incorrect*?

a. Bankruptcy law does not typically protect investors from sovereign risk.

b. Debt repudiation is a postponement of all current and future foreign debt obligations of a borrower.

c. Debt rescheduling occurs when a group of creditors declares a moratorium on debt obligations and seeks to reschedule terms.

d. Sovereign risk can be a cause of default in a non-governmental borrower of high credit quality.

EXAMPLE 19.16: FRM EXAM 2006—QUESTION 126

The 1997 Thai default was unusual compared to past sovereign defaults because:

a. The country repudiated its debt, whereas most defaults are reschedulings.

b. The country had a low inflation level, whereas most previous defaults had high inflation, largely as the result of fiscal deficits.

c. The country had a strong banking system, whereas most previous defaults arose from weakness in the financial intermediation arena.

d. The country was a strong exporter prior to the crisis, whereas most defaulting countries were net importers.

19.4.3 The Role of Credit Rating Agencies

Credit rating agencies play an important role in financial markets. They provide widely available summary information about the credit risk of a wide range of instruments and obligors. They help mitigate the asymmetry of information between borrowers and investors. This is especially useful for small investors, who do not have the time nor the resources for detailed credit analyses.

Investors use these credit ratings to assess credit risk and to comply with investment guidelines and regulations. Sell-side firms such as broker–dealers use them to determine the amount of collateral to hold against credit exposure.

The role of credit ratings is also officially recognized. Indeed, credit rating agencies registered with the Securities and Exchange Commission (SEC) are known as Nationally Recognized Statistical Rating Organizations (NSRO). Several regulations at the federal and state levels, for example, explicitly use ratings from NSROs. Most recently, the Basel II rules for commercial banks, which will be explained in more detail in Chapter 29, include capital charges that depend on external credit ratings.

Credit rating agencies, however, are beset by conflicts of interest. Even though their rating is viewed as an "opinion" to the reader, they are paid directly by the firms that they rate. Thus, on the one hand, they may have an incentive to *maximize profits* by providing easy ratings to many client firms. On the other hand, they have a countervailing incentive to *protect their reputation* as a neutral third party that provides independent and objective advice. If their ratings became worthless, they would quickly lose their business franchise.

Credit rating agencies argue that there is no strong incentive to accommodate the preferences of bond issuers because each single issuer represents a small fraction of revenues. For a long time, these countervailing incentives were delicately balanced and the system seemed to work well.

In recent years, however, credit rating agencies have expanded into rating the different tranches of structured credit products. This was a booming market, with

complex products that will be described in Chapter 22. The business of rating structured products was highly profitable, much more so than before. Credit rating agencies were also involved in the design of the products, for instance offering advice as to the width of various tranches to achieve the desired rating. This was quite a step from giving an opinion on the default risk of a bond issued by a corporate entity. The banks that designed these structured products could also "shop around" the three major credit rating agencies, giving their business to the most accommodating one. These pressures seem to have led to a marked relaxation of credit standards. Credit rating agencies gave very high grades to securities that quickly went bad when subprime-backed debt started to default in 2007.

Investors, burned by these losses, lost faith in the quality of credit ratings and withdrew from credit markets. This put in motion the sequence of events that led to the recent credit crisis.

The credit rating agencies are widely viewed as having played a role in this crisis. As a result, the SEC strengthened the regulation of CRAs. It now prohibits them from providing both ratings and advice on how to structure securities. Agencies are also now required to provide information on the history of upgrades and downgrades over time for each asset class. In addition, EU regulators made it clear that self-regulation was no longer viewed as adequate and that further regulation was necessary.

19.5 IMPORTANT FORMULAS

Credit ratings by Standard & Poor's (modifiers, +, −):
  AAA, AA, A, BBB (investment grade); BB, B, CCC and below (speculative grade)

Credit ratings by Moody's (modifiers, 1, 2, 3):
  Aaa, Aa, A, Baa (investment grade); Ba, B, Caa and below (speculative grade)

Default rate $\bar{X}$ mean and variance: $E(\bar{X}) = p$, $V(\bar{X}) = \frac{p(1-p)}{N}$

Marginal default rate for firm initially rated R during year $T = t + N$:
  $d_N(R) = \frac{m[t+N]}{n[t+N]}$

Survival rate for N years: $S_N(R) = \Pi_{i=1}^{N}(1 - d_i(R))$

Marginal default rate from start to year T: $k_N(R) = S_{N-1}(R)d_N(R)$

Cumulative default rate: $C_N(R) = k_1(R) + k_2(R) + \cdots + k_N(R) = 1 - S_N(R)$

Average default rate, d: $C_N = 1 - \Pi_{i=1}^{N}(1 - d_i) = 1 - (1 - d)^N$

19.6 ANSWERS TO CHAPTER EXAMPLES

Example 19.1: Definition of a Credit Event

b. Calling back a bond occurs when the borrower wants to refinance its debt at a lower cost, which is not a credit event.

Example 19.2: FRM Exam 2003—Question 100

c. Baa3 is the lowest investment-grade rating for Moody's.

Example 19.3: FRM Exam 2005—Question 86

c. The lowest investment-grade ratings are BBB and Baa.

Example 19.4: FRM Exam 2002—Question 110

d. The BB rating by S&P is similar to a Ba rating by Moody's. A BB bond will have lower default rate than a bond rated lower. Hence, the answer is the next lower rating category by Moody's.

Example 19.5: FRM Exam 2004—Question 1

c. The probability of survival for one year is $S_1 = (1 - d) = (1 - d^Q)^4$. This gives a probability of surviving the first quarter of $(1 - d^Q) = (1 - 0.10)^{1/4} = 0.974$.

Example 19.6: FRM Exam 2002—Question 77

d. The marginal default rate increases with maturity. So, this could be for example 0.50% over the last three years of the six-year period. This gives a cumulative default probability greater than 0.60%.

Example 19.7: FRM Exam 2006—Question 21

a. The survival rate is $S_3 = (1 - d_1)(1 - d_2)(1 - d_3) = (1 - 0.08)(1 - 0.12)(1 - 0.15) = 68.8\%$.

Example 19.8: FRM Exam 2004—Question 14

b. This is one minus the survival rate over three years: $S_3(R) = (1 - d_1)(1 - d_2)(1 - d_3) = (1 - 0.03)(1 - 0.04)(1 - 0.06) = 0.8753$. Hence, the cumulative default rate is 0.1247.

Example 19.9: FRM Exam 2000—Question 34

a. The marginal default rate is the probability of defaulting over the next year, conditional on having survived to the beginning of the year.

Example 19.10: FRM Exam 2005—Question 105

b. A rating transition table has probabilities of changing from one rating to another over one year, which can be extrapolated over several years. Hence, statement a. is correct. This also include default, hence statement c. is correct. The probabilities

can be used to group ratings, hence, statement d. is correct. Transitions matrices have no information about prices, so answer b. is the correct one.

Example 19.11: FRM Exam 2007—Question 51

c. This is given by the ratio of entries to BBB and D, which is 2 + 3 over the total of 52, which is 0.096.

Example 19.12: FRM Exam 2003—Question 59

d. B can go into default the first year, with probability of 0.02. Or it could go to A then D, with probability of $0.03 \times 0.00 = 0$. Or it could go to B then D, with probability of $0.90 \times 0.02 = 0.018$. Or it could go to C then D, with probability of $0.05 \times 0.14 = 0.007$. The total is 0.045.

Example 19.13: FRM Exam 2000—Question 58

a. The recovery rate on loans is typically higher than that on bonds. Hence, the credit rating, if it involves both probability of default and recovery, should be higher for loans than for bonds.

Example 19.14: FRM Exam 2002—Question 123

c. The recovery rate is higher when the assets of the firm in default consist of tangible assets that can be resold easily. More volatile assets mean that there is a greater probability of a fall in market value upon liquidation. So, the tangible assets of a manufacturing company is the best answer.

Example 19.15: FRM Exam 2005—Question 79

b. Statements a., c., and d. are all correct. Debt repudiation is a cancellation, not a postponement, so b. is incorrect.

Example 19.16: FRM Exam 2006—Question 126

b. Thailand did not repudiate its debt, so a. is false. Thailand had a weak banking system and suffered from a poor allocation of credit, so c. is false. Thailand had strong imports, so d. is false. Thailand had low inflation but a very large current account deficit finance by capital inflows that eventually led to a credit crisis.

20.

Measuring Default Risk from Market Prices

The previous chapter discussed how to quantify credit risk from categorization into credit risk ratings. Based on these external ratings, we can forecast credit losses from historical default rates and recovery rates.

Credit risk can also be assessed from market prices of securities whose values are affected by default. These include corporate bonds, equities, and credit derivatives. In principle, these should provide more up-to-date and accurate measures of credit risk because financial markets have access to a very large amount of information and are forward-looking. Agents also have very strong financial incentives to impound this information in trading prices. This chapter shows how to infer default risk from market prices.

Section 20.1 will show how to use information about the market prices of credit-sensitive bonds to infer default risk. In this chapter, we will call defaultable debt interchangeably "credit-sensitive," "corporate," and "risky" debt. Here *risky* refers to credit risk and not market risk. We show how to break down the yield on a corporate bond into a default probability, a recovery rate, and a risk-free yield.

Section 20.2 turns to equity prices. The advantage of using equity prices is that they are much more widely available and of much better quality than corporate bond prices. We show how equity can be viewed as a call option on the value of the firm and how a default probability can be inferred from the value of this option. This approach also explains why credit positions are akin to short positions in options and are characterized by distributions that are skewed to the left. Chapter 22 will discuss credit derivatives, which can also be used to infer default risk.

20.1 CORPORATE BOND PRICES

To assess the credit risk of a transaction with a counterparty, consider credit-sensitive bonds issued by the same counterparty. We assume that default is a state that affects all obligations equally.

20.1.1 Spreads and Default Risk

Assume for simplicity that the bond makes only one payment of $100 in one period. We can compute a market-determined yield y^* from the price P^* as

$$P^* = \frac{\$100}{(1 + y^*)} \tag{20.1}$$

This can be compared with the risk-free yield over the same period y.

The payoffs on the bond can be described by a simplified default process, which is illustrated in Figure 20.1. At maturity, the bond can be in default or not. Its value is $100 if there is no default and $f \times \$100$ if default occurs, where f is the fractional recovery. We define π as the default rate over the period. How can we value this bond?

Using risk-neutral pricing, the current price must be the mathematical expectation of the values in the two states, discounting the payoffs at the risk-free rate. Hence,

$$P^* = \frac{\$100}{(1 + y^*)} = \left[\frac{\$100}{(1 + y)}\right] \times (1 - \pi) + \left[\frac{f \times \$100}{(1 + y)}\right] \times \pi \tag{20.2}$$

Note that the discounting uses the risk-free rate y because there is no risk premium with risk-neutral valuation. After rearranging terms,

$$(1 + y) = (1 + y^*)[1 - \pi(1 - f)] \tag{20.3}$$

which implies a default probability of

$$\pi = \frac{1}{(1 - f)}\left[1 - \frac{(1 + y)}{(1 + y^*)}\right] \tag{20.4}$$
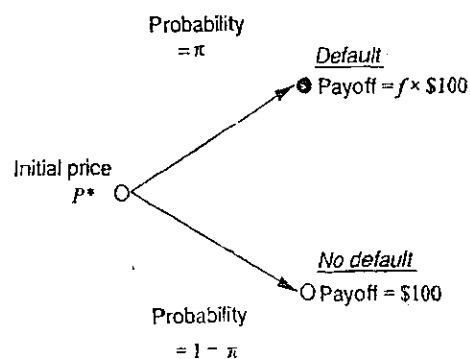
FIGURE 20.1   A Simplified Bond Default Process

Assuming that yields and default probabilities are small, and dropping second-order terms, this simplifies to

$$y^* \approx y + \pi(1 - f) \tag{20.5}$$

This equation shows that the credit spread $y^* - y$ measures credit risk. More specifically, it includes the probability of default, π, times the loss given default, $(1 - f)$. This makes sense because there is no potential credit loss either if the default probability is zero or if the loss given default is zero.

Let us now consider multiple periods, which number T. We compound interest rates and default rates over each period. In other words, π^a is now the *average* annual default rate. Assuming one payment only, the present value is

$$P^* = \frac{\$100}{(1 + y^*)^T} = \left[\frac{\$100}{(1 + y)^T}\right] \times (1 - \pi^a)^T + \left[\frac{f \times \$100}{(1 + y)^T}\right] \times [1 - (1 - \pi^a)^T] \tag{20.6}$$

which can be written as

$$(1 + y)^T = (1 + y^*)^T \{(1 - \pi^a)^T + f[1 - (1 - \pi^a)^T]\} \tag{20.7}$$

Unfortunately, this does not simplify easily. Alternatively, using the cumulative default probability,

$$\frac{1}{(1 + y^*)^T} = \left[\frac{1}{(1 + y)^T}\right] \times (1 - \pi) + \left[\frac{f \times 1}{(1 + y)^T}\right] \times [1 - (1 - \pi)] \tag{20.8}$$

or

$$\frac{1}{(1 + y^*)^T} = \frac{1}{(1 + y)^T} \times [1 - \pi(1 - f)] \tag{20.9}$$

for which a very rough approximation is

$$y^* \approx y + (\pi/T)(1 - f) \tag{20.10}$$

When we have risky bonds of various maturities, they can be used to compute default probabilities for different horizons. If we have two periods, for example, we could use Equation (20.3) to find the probability of defaulting over the first period, π_1, and Equation (20.7) to find the annualized, or average, probability of defaulting over the first two periods, π_2. As we saw in the previous chapter, the marginal probability of defaulting in the second period, d_2, is given by solving

$$(1 - \pi_2)^2 = (1 - \pi_1)(1 - d_2) \tag{20.11}$$

This enables us to recover a term structure of forward default probabilities from a sequence of zero-coupon bonds. In practice, if we have access to only

coupon-paying bonds, the computation becomes more complicated because we
need to consider the payments in each period with and without default.

20.1.2 Risk Premium

It is worth emphasizing that the preceding approach assumed risk-neutrality. As
in the methodology for pricing options, we assumed that the value of any asset
grows at the risk-free rate and can be discounted at the same risk-free rate. Thus
the probability measure π is a risk-neutral measure, which is not necessarily equal
to the objective, physical probability of default.

Defining this objective probability as π' and the discount rate as y', the current
price can be also expressed in terms of the true expected value discounted at the
risky rate y':

$$P^* = \frac{\$100}{(1+y^*)} = \left[\frac{\$100}{(1+y')}\right] \times (1-\pi') + \left[\frac{f \times \$100}{(1+y')}\right] \times \pi' \qquad (20.12)$$

Equation (20.4) allows us to recover a risk-neutral default probability only.
More generally, if investors require some compensation for bearing credit risk,
the credit spread will include a risk premium, rp:

$$y^* \approx y + \pi'(1 - f) + rp \qquad (20.13)$$

To be meaningful, this risk premium must be tied to some measure of bond riski-
ness as well as investor risk aversion. In addition, this premium may incorporate
a liquidity premium and tax effects.[1]

KEY CONCEPT

The yield spread between a corporate bond and an otherwise identical bond
with no credit risk reflects the expected actuarial loss, or annual default rate
times the loss given default, plus a risk premium.

Example: Deriving Default Probabilities

We wish to compare a 10-year U.S. Treasury strip and a 10-year zero issued by
International Business Machines (IBM), which is rated A by S&P and Moody's.
The respective yields are 6% and 7%, using semiannual compounding. Assuming

---

[1] For a decomposition of the yield spread into risk premium effects, see Elton, E., M. Gruber, D.
Agrawal, and C. Mann (2001), Explaining the Rate Spread on Corporate Bonds, *Journal of Finance*.
The authors find a high risk premium, which is related to common risk factors from the stock market.
Part of the risk premium is also due to tax effects. Because Treasury coupon payments are not taxable
at the state level (for example, New York state), investors are willing to accept a lower yield on
Treasury bonds, which increases the corporate yield spread.

that the recovery rate is 45% of the face value, what does the credit spread imply for the probability of default?

Equation (20.9) shows that $\pi(1-f) = 1 - (1 + y/200)^{20}/(1 + y^*/200)^{20} = 0.0923$. Hence, $\pi = 9.23\%/(1 - 45\%) = 16.8\%$. Therefore, the cumulative (risk-neutral) probability of defaulting during the next 10 years is 16.8%. This number is rather high compared with the historical record for this risk class. Table 19.3 shows that Moody's reports a historical 10-year default rate for A credits around 3% only.

If these historical default rates are used as the future probability of default, the implication is that a large part of the credit spread reflects a risk premium. For instance, assume that 80 basis points out of the 100-basis-point credit spread reflects a risk premium. We change the 7% yield to 6.2% and find a probability of default of 3.5%. This is more in line with the actual default experience of such issuers.

EXAMPLE 20.1: FRM EXAM 2007—QUESTION 77

The risk-free rate is 5% per year and a corporate bond yields 6% per year. Assuming a recovery rate of 75% on the corporate bond, what is the approximate market implied one-year probability of default of the corporate bond?

a. 1.33%

b. 4.00%

c. 8.00%

d. 1.60%

EXAMPLE 20.2: FRM EXAM 2007—QUESTION 48

The spread on a one-year BBB-rated bond relative to the risk-free Treasury of similar maturity is 2%. It is estimated that the contribution to this spread by all noncredit factors (e.g., liquidity risk, taxes) is 0.8%. Assuming the loss given default rate for the underlying credit is 60%, what is, approximately, the implied default probability for this bond?

a. 3.33%

b. 5.00%

c. 3.00%

d. 2.00%

EXAMPLE 20.3: FRM EXAM 2002—QUESTION 96

A loan of $10 million is made to a counterparty whose expected default rate is 2% per annum and whose expected recovery rate is 40%. Assuming an all-in cost of funds of LIBOR for the lender, what would be the fair price for the loan?

    a. LIBOR + 120 bp
    b. LIBOR + 240 bp
    c. LIBOR − 120 bp
    d. LIBOR + 160 bp

20.1.3 The Cross-Section of Yield Spreads

We now turn to actual market data. Figure 20.2 illustrates a set of par yield curves for various credits as of December 1998. For reference, the spreads are listed in Table 20.1. The curves are sorted by credit rating, from AAA to B, using S&P's ratings.

These curves bear a striking resemblance to the cumulative default rate curves reported in the previous chapter. They increase with maturity and with lower credit quality.

The lowest curve is the Treasury curve, which represents risk-free bonds. Spreads for AAA credits are low, starting at 46 bp at short maturities and
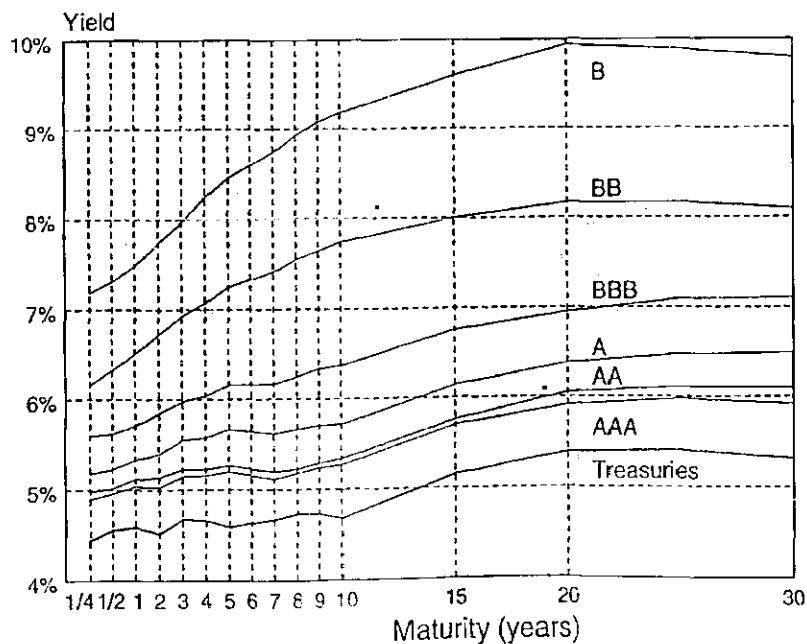
FIGURE 20.2 Yield Curves for Different Credits

TABLE 20.1 Credit Spreads

Maturity (Years)	Credit Rating					
	AAA	AA	A	BBB	BB	B
3M	46	54	74	116	172	275
6M	40	46	67	106	177	275
1	45	53	74	112	191	289
2	51	62	88	133	220	321
3	47	55	87	130	225	328
4	50	57	92	138	241	358
5	61	68	108	157	266	387
6	53	61	102	154	270	397
7	45	53	95	150	274	407
8	45	50	94	152	282	420
9	51	56	98	161	291	435
10	59	66	104	169	306	450
15	55	61	99	161	285	445
20	52	66	99	156	278	455
30	60	78	117	179	278	447

increasing to 60 bp at longer maturities. Spreads for B credits are much wider; they also increase faster, from 275 to 450. Finally, note how close together the AAA and AA spreads are, in spite of the fact that default probabilities approximately double from AAA to AA. The transition from Treasuries to AAA credits most likely reflects other factors, such as liquidity and tax effects, rather than actuarial credit risk.

The previous sections showed that we could use information in corporate bond yields to make inferences about credit risk. Indeed, bond prices represent the best assessment of traders, or real "bets," on credit risk. Thus, we would expect bond prices to be the best predictors of credit risk and to outperform credit ratings. To the extent that agencies use public information to form their credit ratings, this information should be subsumed into market prices. Bond prices are also revised more frequently than credit ratings. As a result, movements in corporate bond prices tend to *lead* changes in credit ratings.

EXAMPLE 20.4: FRM EXAM 2002—QUESTION 81

Which of the following is true?

a. Changes in bond spreads tend to lead changes in credit ratings.
b. Changes in bond spreads tend to lag changes in credit ratings.
c. Changes in bond spreads tend to occur at the exact same time as changes in credit ratings.
d. There is absolutely no perceived general relationship in the timing of changes in bond spreads and changes in credit ratings.

EXAMPLE 20.5: TERM STRUCTURE OF CREDIT SPREADS

Suppose XYZ Corporation has two bonds paying semiannually according to the following table:

Remaining Maturity	Coupon (sa 30/360)	Price	T-Bill Rate (Bank Discount)
6 months	8.0%	99	5.5%
1 year	9.0%	100	6.0%

The recovery rate for each in the event of default is 50%. For simplicity, assume that each bond will default only at the end of a coupon period. The market-implied risk-neutral probability of default for XYZ Corporation is

a. Greater in the first six-month period than in the second
b. Equal between the two coupon periods
c. Greater in the second six-month period than in the first
d. Cannot be determined from the information provided

20.1.4 Time Variation in Credit Spreads

Credit spreads reflect potential losses caused by default risk, and perhaps a risk premium. Some of this default risk is specific to the issuer and requires a detailed analysis of its prospective financial condition. Part of this risk, however, can be attributed to common credit risk factors. These common factors are particularly important, as they cannot be diversified away in a large portfolio of credit-sensitive bonds.

First among these factors are general economic conditions. Economic growth is negatively correlated with credit spreads. When the economy slows down, more companies are likely to have cash-flow problems and to default on their bonds.

Figure 20.3 compares the speculative-grade default rate from Moody's and the Baa-Treasury credit spread. Shaded areas indicate periods of recession. The graph shows that both default rates and credit spreads tend to increase around recessions. Because spreads are forward-looking, however, they tend to lead default rates, which peak after recessions. Also, the effect of the 2007–2008 credit crisis is apparent from the unprecedented widening of credit spreads. These reflect a combination of higher risk aversion and anticipation of very high default rates.

Volatility is also a factor. In a more volatile environment, investors may require larger risk premiums, thus increasing credit spreads. When this happens, liquidity may also dry up. Investors may then require a greater credit spread in order to hold increasingly illiquid securities.
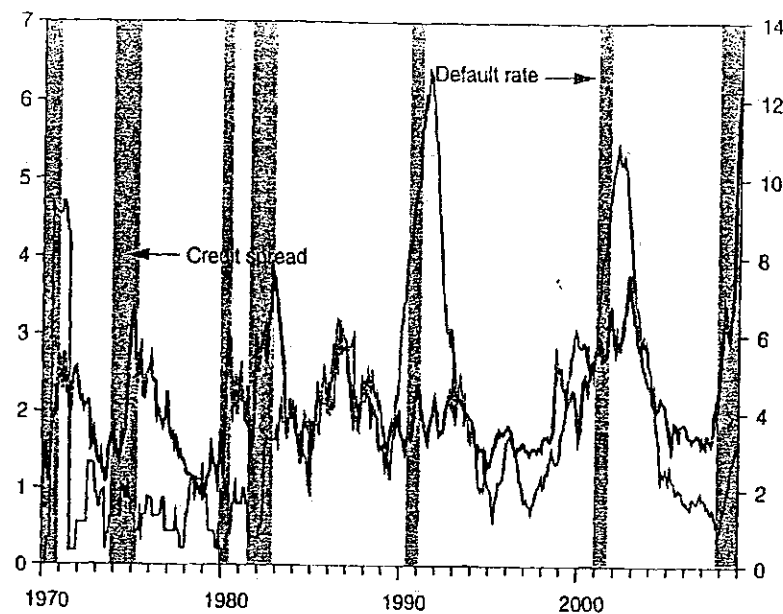
FIGURE 20.3 Default Rates and High Yield Spreads

Finally, volatility has another effect through an option channel. Corporate bond indices include many callable bonds, unlike Treasury indices. The buyer of a callable bond requires a higher yield in exchange for granting the call option. Higher volatility should increase the value of this option and therefore this yield, all else equal. Thus, credit spreads directly increase with volatility.

20.2 EQUITY PRICES

The credit spread approach, unfortunately, is only useful when there is good bond market data. The problem is that this is rarely the case, for a number of reasons.

- Many countries do not have a well-developed corporate bond market. As Table 7.2 has shown, the United States has by far the largest corporate bond market in the world. This means that other countries have much fewer outstanding bonds and a much less active market.
- The counterparty may not have an outstanding publicly traded bond or if so, the bond may contain other features such as a call that make it more difficult to interpret the yield.
- The bond may not trade actively and instead reported prices may simply be matrix prices, that is, interpolated from yields on other issuers.

An alternative is to turn to default risk models based on stock prices, because equity prices are available for a larger number of companies and because equities are more actively traded than corporate bonds. The Merton (1974) model views equity as akin to a call option on the assets of the firm, with an exercise price given by the face value of debt.

20.2.1 The Merton Model

To simplify to the extreme, consider a firm with total value V that has one bond due in one period with face value K. If the value of the firm exceeds the promised payment, the bond is repaid in full and stockholders receive the remainder. However, if V is less than K, the firm is in default and the bondholders receive V only. The value of equity goes to zero. We assume that there are no transaction costs and that the absolute-priority rule is followed. Hence, the value of the stock at expiration is

$$S_T = \text{Max}(V_T - K, 0) \tag{20.14}$$

Because the bond and equity add up to the firm value, the value of the bond must be

$$B_T = V_T - S_T = V_T - \text{Max}(V_T - K, 0) = \text{Min}(V_T, K) \tag{20.15}$$

The current stock price, therefore, embodies a forecast of default probability in the same way that an option embodies a forecast of being exercised. Figures 20.4 and 20.5 describe how the value of the firm can be split up into the bond and stock values.

Note that the bond value can also be described as

$$B_T = K - \text{Max}(K - V_T, 0) \tag{20.16}$$

In other words, a long position in a risky bond is equivalent to a long position in a risk-free bond plus a short put option, which is really a credit derivative.
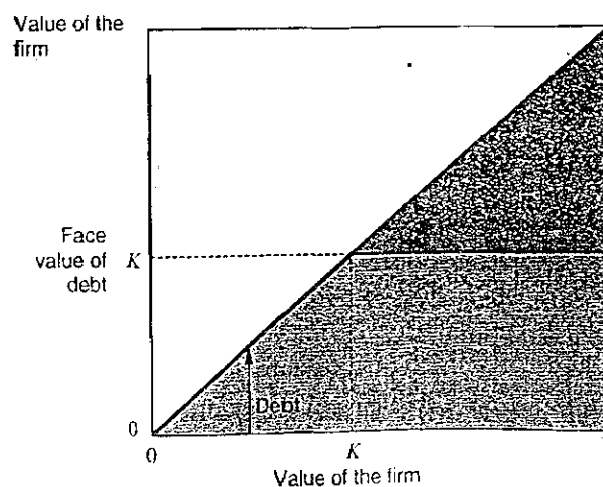
FIGURE 20.4   Equity as an Option on the Value of the Firm
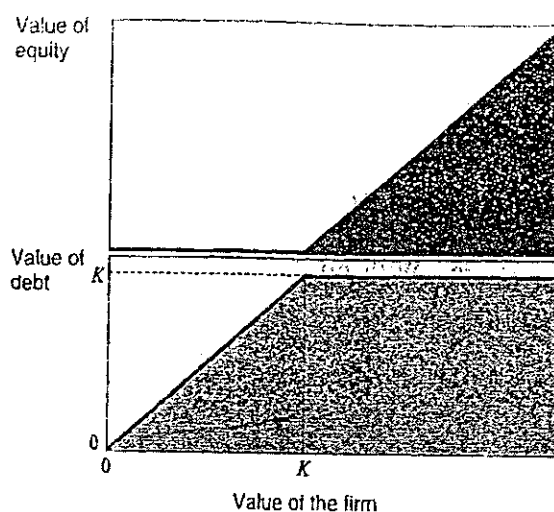
FIGURE 20.5 Components of the Value of the Firm

> **KEY CONCEPT**
>
> Equity can be viewed as a call option on the firm value with strike price equal to the face value of debt. Corporate debt can be viewed as risk-free debt minus a put option on the firm value.

This approach is particularly illuminating because it demonstrates that corporate debt has a payoff akin to a short position in an option, explaining the left skewness that is characteristic of credit losses. In contrast, equity is equivalent to a long position in an option due to its limited-liability feature. In other words, investors can lose no more than their equity investment.

20.2.2 Pricing Equity and Debt

To illustrate, we proceed along the lines of the usual Black–Scholes (BS) framework, assuming that the firm value follows the geometric Brownian motion process:

$$dV = \mu V\, dt + \sigma V\, dz \tag{20.17}$$

If we assume that markets are frictionless and that there are no bankruptcy costs, the value of the firm is simply the sum of the firm's equity and debt: $V = B + S$.

To price a claim on the value of the firm, we need to solve a partial differential equation with appropriate boundary conditions. The corporate bond price is obtained as

$$B = F(V, t), \quad F(V, T) = \mathrm{Min}[V, B_F] \tag{20.18}$$

where $B_F = K$ is the face value of the bond to be repaid at expiration, or the strike price.

Similarly, the equity value is

$$S = f(V, t), \qquad f(V, T) = \text{Max}[V - B_F, 0] \tag{20.19}$$

Stock Valuation   With no dividend, the value of the stock is given by the BS formula,

$$S = \text{Call} = V N(d_1) - Ke^{-r\tau} N(d_2) \tag{20.20}$$

where $N(d)$ is the cumulative distribution function for the standard normal distribution, and

$$d_1 = \frac{\ln(V/Ke^{-r\tau})}{\sigma\sqrt{\tau}} + \frac{\sigma\sqrt{\tau}}{2}, \qquad d_2 = d_1 - \sigma\sqrt{\tau}$$

where $\tau = T - t$ is the time to expiration, r the risk-free interest rate, and σ the volatility of *asset* value. The option value depends on two factors, $x = Ke^{-r\tau}/V$ and $\sigma\sqrt{\tau}$. The first factor is the debt/value ratio and is inversely related to leverage, which can be written as $l = V/(V - Ke^{-r\tau})$. The value of the stock increases as the volatility σ increases and as x decreases, or when leverage increases.

Firm Volatility   Note that this application is different from the BS model, where we plug in the value of V and of its volatility, $\sigma = \sigma_V$, and solve for the value of the call. Here we observe the market value of the firm S and the equity volatility σ_S and must infer the values of V and its volatility so that Equation (20.20) is satisfied. This can only be done iteratively. Defining $\Delta = N(d_1)$ as the hedge ratio, we have

$$dS = \frac{\partial S}{\partial V} dV = \Delta dV \tag{20.21}$$

Defining σ_S as the volatility of (dS/S), we have $(\sigma_S S) = \Delta(\sigma_V V)$ and

$$\sigma_V = (1/\Delta)\sigma_S(S/V) \tag{20.22}$$

Bond Valuation   Next, the value of the bond is given by $B = V - S$, or

$$B = Ke^{-r\tau} N(d_2) + V[1 - N(d_1)] \tag{20.23}$$

$$B/Ke^{-r\tau} = [N(d_2) + (V/Ke^{-r\tau})N(-d_1)] \tag{20.24}$$

Hence, the value of the bond is related to the firm by

$$dB = \frac{\partial B}{\partial V} dV = N(-d_1)dV \tag{20.25}$$

Equations (20.21) and (20.25) are sometimes used for capital arbitrage trades, which involve buying and selling different types of claims on the firm, using these hedge ratios to try to minimize risk—that is, assuming the model is correct.

The bond price can be expressed in terms of the annualized credit spread s

$$B = Ke^{-(r+s)\tau} \qquad (20.26)$$

which gives

$$s = -(1/\tau)[N(d_2) + (1/x)N(-d_1)] \qquad (20.27)$$

The value of the bond decreases when volatility increases and when leverage increases. The spread moves conversely. Equation (20.27) can create rich patterns in the term structure of credit spreads. For firms with low leverage and volatility, spreads are low for near maturities and increase with maturity. Firms with high leverage have high spreads, with a term structure that can have a negative slope.

Risk-Neutral Dynamics of Default   In the Black–Scholes model, $N(d_2)$ is also the probability of exercising the call, or that the bond will not default. Conversely, $1 - N(d_2) = N(-d_2)$ is the risk-neutral probability of default.

Pricing Credit Risk   At maturity, the credit loss is the value of the risk-free bond minus the corporate bond, $CL = B_F - B_T$. At initiation, the expected credit loss (ECL) is

$$
\begin{aligned}
B_F e^{-r\tau} - B &= Ke^{-r\tau} - \{Ke^{-r\tau}N(d_2) + V[1 - N(d_1)]\} \\
&= Ke^{-r\tau}[1 - N(d_2)] - V[1 - N(d_1)] \\
&= Ke^{-r\tau}N(-d_2) - VN(-d_1) \\
&= N(-d_2)[Ke^{-r\tau} - VN(-d_1)/N(-d_2)]
\end{aligned}
$$

This decomposition is quite informative. Multiplying by the future value factor $e^{r\tau}$ shows that the ECL at maturity is

$$\text{ECL}_T = N(-d_2)[K - Ve^{r\tau}N(-d_1)/N(-d_2)] = p \times [\text{Exposure} \times \text{LGD}] \qquad (20.28)$$

This involves two terms. The first is the probability of default, $N(-d_2)$. The second, between brackets, is the loss when there is default. This is obtained as the face value of the bond K minus the recovery value of the loan when in default, $Ve^{r\tau}N(-d_1)/N(-d_2)$, which is also the expected value of the firm in the state of default. Note that the recovery rate is endogenous here, as it depends on the value of the firm, time, and debt ratio.

Credit Option Valuation   This approach can also be used to value the put option component of the credit-sensitive bond. This option pays $K - B_T$ in case of default.

A portfolio with the bond plus the put is equivalent to a risk-free bond $Ke^{-rt} = B +$ put. Hence, using Equation (20.23), the credit put should be worth

$$\text{Put} = Ke^{-rt} - \{Ke^{-rt}N(d_2) + V[1 - N(d_1)]\} = -V[N(-d_1)] + Ke^{-rt}[N(-d_2)]$$

$$(20.29)$$

This will be used later in the chapter on credit derivatives.

EXAMPLE 20.6: FRM EXAM 2001—QUESTION 14

To what sort of option on the counterparty's assets can the current exposure of a credit-risky position better be compared?

a. A short call
b. A short put
c. A short knock-in call
d. A binary option

20.2.3 Applying the Merton Model

These valuation formulas can be used to recover, given the current value of equity and of nominal liabilities, the value of the firm and its probability of default. Figure 20.6 illustrates the evolution of the value of the firm. The firm defaults if this value falls below the liabilities at the horizon. We measure this risk-neutral probability by $N(-d_2)$.

In practice, default is much more complex than depicted here. We would have to collect information about all the liabilities of the company, as well as their
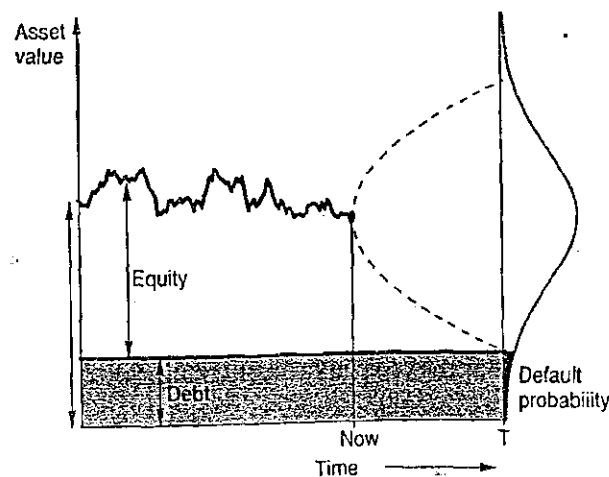
FIGURE 20.6 Default in the Merton Model

maturities. Default can also occur with coupon payments. So instead of default on a target date, we could measure default probability as a function of the distance relative to a moving floor that represents liabilities. This is essentially the approach undertaken by KMV Corporation, now part of Moody's, which sells estimated default frequencies (EDFs) for global firms. The approach will be explained in Chapter 23.

The Merton approach has many advantages. First, it relies on the prices of equities, which are more actively traded than bonds. Second, correlations between equity prices can generate correlations between defaults, which would be otherwise difficult to measure. Perhaps the most important advantage of this model is that it generates movements in EDFs that seem to *lead* changes in credit ratings.

Figure 20.7 displays movements in EDFs and credit rating for WorldCom, using the same vertical scale. WorldCom went bankrupt on July 21, 2002. With $104 billion in assets, this was America's largest bankruptcy ever. The agency rating was BBB until April 2002. It gave no warning of the impending default. In contrast, starting one year before the default, the EDF began to move up. In April, it reached 20%, presaging bankruptcy.

These models have disadvantages as well. The first limitation of the model is that it cannot be used to price sovereign credit risk, as countries obviously do not have a stock price. This is a problem for credit derivatives, where a large share of the market consists of sovereign risks.

A more fundamental drawback is that it relies on a static model of the firm's capital and risk structure. The debt level is assumed to be constant over the horizon. Similarly, the model cannot handle new injections of equity, which protect
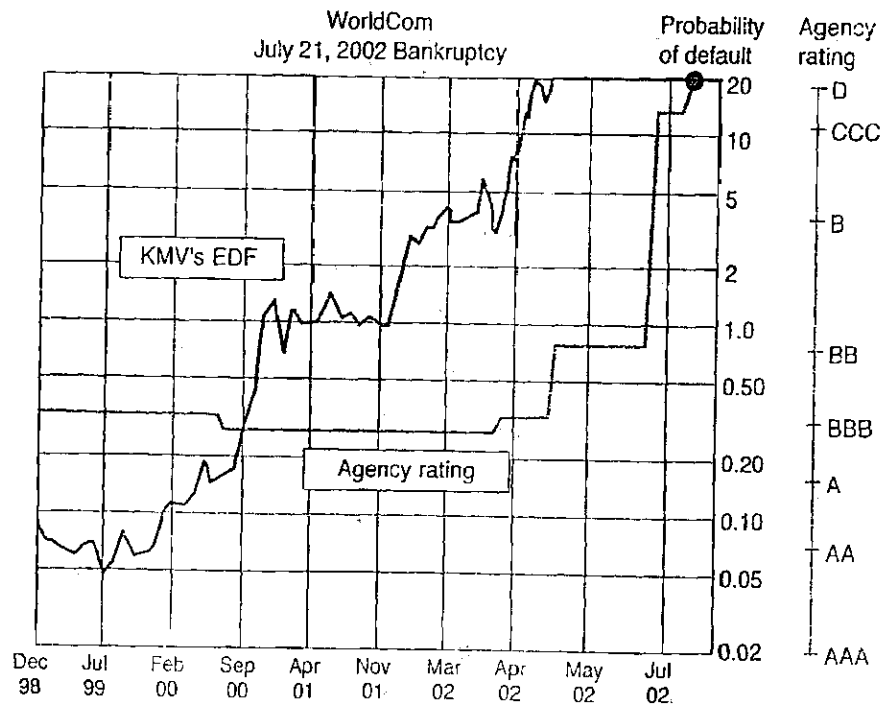
FIGURE 20.7 KMV's EDF and Credit Rating

existing debt holders. Also, the model needs to be expanded to a more realistic setting where debt matures at various points in time, which is not an obvious extension.

Another problem is that management could undertake new projects that increase not only the value of equity but also its volatility, thereby increasing the credit spread. This runs counter to the fundamental intuition of the Merton model, which is that, all else equal, a higher stock price reflects a lower probability of default and hence should be associated with a smaller credit spread.

Finally, this class of models fails to explain the magnitude of credit spreads we observe on credit-sensitive bonds. Recent work has attempted to add other sources of risk, such as interest rate risk, but still falls short of explaining these spreads. Thus these models are most useful in tracking *changes* in EDFs over time. Indeed, KMV calibrates the risk-neutral default probabilities to actual default data.

20.2.4 Example

It is instructive to work through a simplified example. Consider a firm with assets worth $V = \$100$ and with volatility $\sigma_V = 20\%$. In practice, one would have to start from the observed stock price and volatility and iterate to find σ_V.

The horizon is $\tau =$ one year. The risk-free rate is $r = 10\%$ using continuous compounding. We assume a leverage factor such that $x = 0.9$, which implies a face value of $K = \$99.46$ and a risk-free current value of $Ke^{-r\tau} = \$90$.

Working through the Merton analysis, one finds that the current stock price should be $S = \$13.59$. Hence, the current bond price is

$$B = V - S = \$100 - \$13.59 = \$86.41$$

which implies a yield of $\ln(K/B)/\tau = \ln(99.46/86.41) = 14.07\%$, or a yield spread of 4.07%. The current value of the credit put is then

$$P = Ke^{-r\tau} - B = \$90 - \$86.41 = \$3.59$$

The analysis also generates values for $N(d_2) = 0.6653$ and $N(d_1) = 0.7347$. Thus, the *risk-neutral* probability of default is EDF $= N(-d_2) = 1 - N(d_2) = 33.47\%$. Note that this could differ from the *actual* or *objective* probability of default since the stock could very well grow at a rate that is greater than the risk-free rate of 10%.

Finally, let us decompose the expected loss at expiration from Equation (20.28), which gives

$$N(-d_2)[K - Ve^{r\tau}N(-d_1)/N(-d_2)] = 0.3347 \times [\$99.46 - \$110.56 \times 0.2653/0.3347]$$
$$= 0.3347 \times [\$11.85] = \$3.96 \qquad (20.30)$$

This combines the probability of default with the expected loss upon default, which is $\$11.85$. This future expected credit loss of $\$3.96$ must also be the future value of the credit put, or $\$3.59e^{r\tau} = \$3.96$.

Note that the model needs very high leverage, here $x = 90\%$, to generate a reasonable credit spread of 4.07%. This implies a debt-to-equity ratio of $0.9/0.1 = 900\%$, which is unrealistically high for this type of spread.

With lower leverage, say $x = 0.7$, the credit spread shrinks rapidly, to 0.36%. At $x = 50\%$ or below, the predicted spread goes to zero. As this leverage would be considered normal, the model fails to reproduce the size of observed credit spreads. Perhaps it is most useful for tracking time variation in estimated default frequencies.

EXAMPLE 20.7: FRM EXAM 2002—QUESTION 97

Among the following variables, which one is the main driver of the probability of default in the KMV model?

a. Stock prices
b. Bond prices
c. Bond yield
d. Loan prices

EXAMPLE 20.8: FRM EXAM 2003—QUESTION 41

Which of the following is *not* a limitation of KMV's EDF model?

a. It is difficult to price sovereign credit risk since asset values and volatility are not directly observable.
b. EDFs are biased by periods of high or low defaults.
c. Takes a simplified view of the capital structure of a firm.
d. The model often fails to explain real-world credit spreads.

EXAMPLE 20.9: FRM EXAM 2005—QUESTION 108

The KMV model produces a measure called Expected Default Frequency. Which of the following statements about this variable is correct?

a. It decreases when the leverage of the firm falls.
b. It increases when the stock price of the firm has been rising.
c. It is the risk-neutral probability of default from Merton's model.
d. It tells investors how the default risk of a bond is correlated with the default risk of other bonds in the portfolio.

EXAMPLE 20.10: FRM EXAM 2007—QUESTION 82

Using the Merton model, the value of the debt increases if all other parameters are fixed and

I. The value of the firm decreases.
II. The riskless interest rate decreases.
III. Time to maturity increases.
IV. The volatility of the firm value decreases.
    a. I and II only
    b. I and IV only
    c. II and III only
    d. II and IV only

EXAMPLE 20.11: FRM EXAM 2005—QUESTION 134

You have a large position of bonds of firm XYZ. You hedge these bonds with equity using Merton's debt valuation model. The value of the debt falls unexpectedly, but the value of equity does not fall, so you make a loss. Consider the following statements:

I. Interest rates increased.
II. Volatility fell.
III. Volatility increased.
IV. A liquidity crisis increased the liquidity component of the credit spreads.

Which statements are possible explanations for why your hedge did not work out?

    a. I and II only
    b. I and III only
    c. I, III, and IV only
    d. III and IV only

20.3 IMPORTANT FORMULAS

Implied default probability, 1 period: $(1 + y) = (1 + y^*)[1 - \pi(1 - f)]$

Approximation of implied default probability: $y^* \approx y + \pi(1 - f)$

Implied default probability, T period:
$$(1 + y)^T = (1 + y^*)^T \{(1 - \pi)^T + f[1 - (1 - \pi)^T]\}$$

Approximation of physical default probability: $y^* \approx y + \pi'(1 - f) + rp$

Merton model for stock price: $S_T = \text{Max}(V_T - K, 0)$

Merton model for bond price: $B_T = V_T - S_T = \text{Min}(V_T, K)$

Stock valuation: $S = \text{Call} = VN(d_1) - Ke^{-r\tau}N(d_2)$

Firm value and stock volatility: $\sigma_V = (1/\Delta)\sigma_S(S/V)$

Bond valuation:

$B = \text{Risk-free bond} - \text{Put}, \quad B/Ke^{-r\tau} = [N(d_2) + (V/Ke^{-r\tau})N(-d_1)]$

Risk-neutral PD: $1 - N(d_2) = N(-d_2)$

Credit default swap, or put option:

$\text{Put} = Ke^{-r\tau} - \{Ke^{-r\tau}N(d_2) + V[1 - N(d_1)]\} = -V[N(-d_1)] + Ke^{-r\tau}[N(-d_2)]$

20.4 ANSWERS TO CHAPTER EXAMPLES

Example 20.1: FRM Exam 2007—Question 77

b. The spread is $7 - 6 = 1\%$. Dividing by the loss given default of $(1 - f) = 0.25$, we get $\pi = (y^* - y)/(1 - f) = 4\%$.

Example 20.2: FRM Exam 2007—Question 48

d. The part of the spread due to expected credit losses is $2.00 - 0.80 = 1.20\%$. Dividing by the LGD of $(1 - f) = 0.65$, we get 2%.

Example 20.3: FRM Exam 2002—Question 96

a. The credit spread should be $y^* - y = \pi(1 - f)$. Thus, $\pi(1 - f) = 2\%(1 - 40\%) = 1.2\%$. The spread over LIBOR should be 120 bp.

Example 20.4: FRM Exam 2002—Question 81

a. Changes in market prices, including bond spreads, tend to lead to changes in credit ratings. This is because market prices reflect all publicly available information about a company.

Example 20.5: Term Structure of Credit Spreads

a. First, we compute the current yield on the six-month bond, which is selling at a discount. We solve for y^* such that $99 = 104/(1 + y^*/200)$ and find $y^* = 10.10\%$. Thus the yield spread for the first bond is $10.1 - 5.5 = 4.6\%$. The second bond is at par, so the yield is $y^* = 9\%$. The spread for the second bond is $9 - 6 = 3\%$. The default rate for the first period must be greater. The recovery rate is the same for the two periods, so it does not matter for this problem.

Example 20.6: FRM Exam 2001—Question 14

b. The lender is short a put option, since exposure exists only if the value of assets falls below the amount lent.

Example 20.7: FRM Exam 2002—Question 97

a. Stock prices are the main driver of KMV's estimated default frequency (EDF), because they drive the value of equity. These models also use the volatility of asset values and the value of liabilities.

Example 20.8: FRM Exam 2003—Question 41

b. Answer a. is a limitation because there is no asset value for sovereign debt. Answer c. is also a limitation because this is a simple model. Answer d. is also a limitation as this model generates spreads that are too low. Finally, b. is not a limitation because stock prices pick up time variation in default rates, unlike credit ratings, which are "through the cycle" and thus less responsive to temporary changes in default rates.

Example 20.9: FRM Exam 2005—Question 108

a. The EDF, similarly to the risk-neutral PD, decreases when the stock prices goes up, when the leverage goes down, or when the volatility goes down. It is a transformation of the PD from a Merton-type model. The KMV framework can be extended to finding correlations, but the EDF is not sufficient.

Example 20.10: FRM Exam 2007—Question 82

b. The value of credit-sensitive debt is $B = Ke^{-(r+s)\tau}$. This increases (1) if the risk-free interest rate decreases, or (2) if the credit spread decreases, or (3) if the maturity decreases. The credit spread decreases if the value of the firm goes up, if the leverage goes down, or if the volatility goes down. Hence, the value of debt increases if the riskless rate decreases or if the volatility decreases.

Example 20.11: FRM Exam 2005—Question 134

b. We need to identify shocks that decrease the value of debt but not that of equity. An increase in the risk-free rate will decrease the value of the debt but not the equity (because this decreases leverage). An increase in volatility will have the opposite effect on debt and equity. Finally, a liquidity crisis cannot explain the divergent behavior, because, as we have seen during 2008, it would affect both corporate bonds and equity adversely.

Credit Exposure

Credit exposure is the amount at risk during the life of the financial instrument. Upon default, it is called exposure at default (EAD). When banking simply consisted of making loans, exposure was essentially the face value of the loan. In this case, the exposure is the notional amount and is fixed.

Since the development of the swap markets, the measurement of credit exposure has become more complicated. This is because swaps, like most derivatives, have an up-front value that is much smaller than the notional amount. Indeed, the initial value of a swap is typically zero, which means that at the outset, there is no credit risk because there is nothing to lose.

As the swap contract matures, however, it can turn into a positive or negative value. The asymmetry of bankruptcy treatment is such that a credit loss can only occur if the instrument has positive value, or is a claim against the defaulted counterparty. Thus, the credit exposure is the value of the asset if it is positive, like an option.

This chapter turns to the quantitative measurement of credit exposure. Section 21.1 describes the general features of credit exposure for various types of financial instruments, including loans or bonds, guarantees, credit commitments, repos, and derivatives. Section 21.2 shows how to compute the distribution of credit exposure and gives detailed examples of exposures of interest rate and currency swaps. Section 21.3 discusses exposure modifiers, or techniques that have been developed to reduce credit exposure. It shows how credit risk can be controlled by marking to market, margins, position limits, recouponing, and netting agreements. For completeness, Section 21.4 includes credit risk modifiers such as credit triggers and time puts, which also control default risk.

21.1 CREDIT EXPOSURE BY INSTRUMENT

Credit exposure is the positive part of the value of the asset at various points during its life. In particular, the current exposure is the value of the asset at the current time V_t if positive:

$$CE_t = Max(V_t, 0) \qquad (21.1)$$

The potential exposure represents the exposure on some future date, or sets of dates. Based on this definition, we can characterize the exposure of a variety of financial instruments. The measurement of current and potential exposure also motivates regulatory capital charges for credit risk, which are explained in Chapter 29.

21.1.1 Loans or Bonds

Loans or bonds are balance sheet assets whose current and potential exposure basically is the notional, or amount loaned or invested. To be more precise, this should be the market value of the asset given current interest rates, but, as we will show, this is not very far from the notional. The exposure is also the notional for receivables and trade credits, as the potential loss is the amount due.

21.1.2 Guarantees

Guarantees are off-balance sheet contracts whereby the bank has underwritten, or agrees to assume, the obligations of a third party. The exposure is the notional amount, because this will be fully drawn when default occurs. By nature, guarantees are irrevocable, that is, unconditional and binding, whatever happens.

An example of a guarantee is a contract whereby bank A makes a loan to client C only if it is guaranteed by bank B. Should C default, B is exposed to the full amount of the loan. Another example is an acceptance, whereby a bank agrees to pay the face value of a bill at maturity. Alternatively, standby facilities, or financial letters of credit, provide a guarantee to a third party of the making of a payment should the obligor default.

21.1.3 Commitments

Commitments are off-balance sheet contracts whereby the bank commits to a future transaction that may result in creating a credit exposure at a *future* date. For instance, a bank may provide a note issuance facility whereby it promises a minimum price for notes regularly issued by a borrower. If the notes cannot be placed at the market at the minimum price, the bank commits to buy them at a fixed price. Such commitments have less risk than guarantees because it is not certain that the bank will have to provide backup support.

It is also useful to distinguish between irrevocable commitments, which are unconditional and binding on the bank, and revocable commitments, where the bank has the option to revoke the contract should the counterparty's credit quality deteriorate. This option substantially decreases the credit exposure.

21.1.4 Swaps or Forwards

Swaps or forwards contracts are off-balance sheet items that can be viewed as irrevocable commitments to purchase or sell some asset on prearranged terms. The current and potential exposure will vary from zero to a large amount depending on

movements in the driving risk factors. Similar arrangements are sale–repurchase agreements (repos), whereby an institution sells an asset to another in exchange for a promise to buy it back later.

21.1.5 Long Options

Options are off–balance sheet items that may create credit exposure. The current and potential exposure also depends on movements in the driving risk factors. Here there is no possibility of negative values because options always have positive value, or zero value at worst: $V_t \geq 0$.

21.1.6 Short Options

Unlike long options, the current and potential exposure for short options is zero because the bank writing the option can incur only a negative cash flow, assuming the option premium has been fully paid.

Exposure also depends on the features of any embedded option. With an American option, for instance, the holder of an in-the-money swap may want to exercise early if the credit rating of its counterparty starts to deteriorate. This decreases the exposure relative to an equivalent European option.

EXAMPLE 21.1: FRM EXAM 2006—QUESTION 95

A credit loss on market-driven instruments such as swaps and forwards arises if:

a. Market rates move in your favor.
b. Market rates move against you.
c. Market rates move against you and the counterparty defaults.
d. Market rates move in your favor and the counterparty defaults.

EXAMPLE 21.2: FRM EXAM 2002—QUESTION 93

Which transaction does *not* result in a long-term credit risk for party A?

a. Party A makes an unsecured loan to party B.
b. Party A is a fixed-price receiver in an interest rate swap from party B.
c. Party A buys a call option on September wheat from party B.
d. Party A sells a put option on the S&P 500 index to party B.

EXAMPLE 21.3: FRM EXAM 2006—QUESTION 117

Which of the following will have the greatest potential credit exposure?

a. Long 3,000 ounces of gold for delivery in one year
b. Long 3,000 ounces of gold for delivery in two years
c. Short 3,000 ounces of gold for delivery in two years
d. Selling an at-the-money call option on 10,000 ounces of gold for delivery in two years

EXAMPLE 21.4: FRM EXAM 2004—QUESTION 8

Your company has reached its credit limit to Ford but Ford is insisting that your firm provide them some increased protection in the event a major project they are undertaking results in some unforeseen liability. Ignoring settlement risk and assuming option premiums are paid immediately at the time of the transaction, which of these strategies will *not* give rise to increased credit exposure to Ford?

a. Selling a costless collar to Ford
b. Buying an option from Ford
c. Selling an option to Ford
d. None of the above

EXAMPLE 21.5: FRM EXAM 2001—QUESTION 84

If a counterparty defaults before maturity, which of the following situations will cause a credit loss?

a. You are short euros in a one-year euro/USD forward FX contract, and the euro has appreciated.
b. You are short euros in a one-year euro/USD forward FX contract, and the euro has depreciated.
c. You sold a one-year OTC euro call option, and the euro has appreciated.
d. You sold a one-year OTC euro call option, and the euro has depreciated.

21.2 DISTRIBUTION OF CREDIT EXPOSURE

Credit exposure consists of the current exposure, which is readily observable, and the potential exposure, or future exposure, which is random. Define x as the potential value of the asset on the target date. We describe this variable by its probability density function $f(x)$. This is where market risk mingles with credit risk.

21.2.1 Expected and Worst Exposure

The expected credit exposure (ECE) is the expected value of the asset replacement value x, if positive, on a target date:

$$\text{Expected credit exposure} = \text{ECE} = \int_{-\infty}^{+\infty} \text{Max}(x, 0) f(x) dx \qquad (21.2)$$

The worst credit exposure (WCE) is the largest (worst) credit exposure at some level of confidence. It is implicitly defined as the value that is not exceeded at the given confidence level p:

$$1 - p = \int_{\text{WCE}}^{\infty} f(x) dx \qquad (21.3)$$

To model the potential credit exposure, we need to (1) model the distribution of risk factors and (2) evaluate the instrument given these risk factors. This process is identical to a market value-at-risk (VAR) computation except that the aggregation takes place at the counterparty level if contracts can be netted.

To simplify to the extreme, suppose that the payoff x, or net claim against a particular counterparty, is normally distributed with mean zero and volatility σ. The expected credit exposure is then

$$\text{ECE} = \frac{1}{2} E\left(x \mid x > 0 \right) = \frac{1}{2} \sigma \sqrt{\frac{2}{\pi}} = \frac{\sigma}{\sqrt{2\pi}} \qquad (21.4)$$

Note that we divided by 2 because there is a 50% probability that the value will be positive. The worst credit exposure at the 95% level is given by

$$\text{WCE} = 1.645\sigma \qquad (21.5)$$

Figure 21.1 illustrates the measurement of ECE and WCE for a normal distribution. Note that negative values of x are not considered.
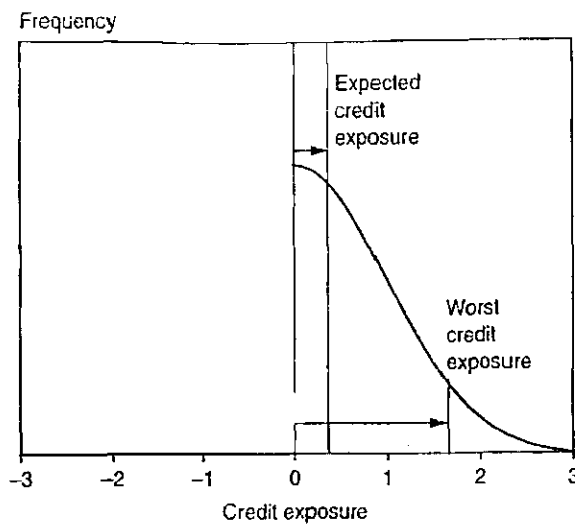
FIGURE 21.1 Expected and Worst Credit Exposures—Normal Distribution

21.2.2 Time Profile

The distribution can be summarized by the expected and worst credit exposures at each point in time. To summarize even further, we can express the average credit exposure by taking a simple arithmetic average over the life of the instrument.

The average expected credit exposure (AECE) is the average of the expected credit exposure over time, from now to maturity T:

$$AECE = (1/T) \int_{t=0}^{T} ECE_t \, dt \tag{21.6}$$

The average worst credit exposure (AWCE) is defined similarly:

$$AWCE = (1/T) \int_{t=0}^{T} WCE_t \, dt \tag{21.7}$$

21.2.3 Exposure Profile for Interest Rate Swaps

We now consider the computation of the exposure profile for an interest rate swap. In general, we need to define (1) the market risk factors, (2) the function and parameters for the joint stochastic processes, and (3) the pricing model for the swap. This is a good illustrative example for an instrument that is widely employed.

We start with a one-factor stochastic process for the interest rate, defining the movement in the rate r_t at time t as

$$dr_t = \kappa(\theta - r_t)dt + \sigma r_t^\gamma dz_t \tag{21.8}$$

as given in Chapter 4. The first term imposes mean reversion. When the current value of r_t is higher than the long-run value, the term in parentheses is negative, which creates a downward trend. More generally, the mean term could reflect the path implied in forward interest rates.

The second term defines the innovation, which can be given a normal distribution. An important issue is whether the volatility of the innovation should be constant or proportional to some power γ of the current value of the interest rate r_t. If the horizon is short, this issue is not so important because the current rate will be close to the initial rate.

When $\gamma = 0$, changes in yields are normally distributed, which is the Vasicek model (1977). As seen in a previous chapter, a typical volatility for *absolute* changes in yields is 1% per annum. A potential problem with this is that the volatility is the same whether the yield starts at 20% or 1%. As a result, the yield could turn negative, depending on the initial starting point and the strength of the mean reversion.

Another class of models is the lognormal model, which takes $\gamma = 1$. The model can then be rewritten in terms of $dr_t/r_t = d\ln(r_t)$. This specification ensures that the volatility shrinks as r gets close to zero, avoiding negative values. A typical volatility of *relative* changes in yields is 15% per annum, which is also the 1% value for changes in the level of rates divided by an initial rate of 6.7%.

For illustration purposes, we choose the normal process $\gamma = 0$ with mean reversion $\kappa = 0.02$ and volatility $\sigma = 0.25\%$ per month, which are realistic parameters based on recent U.S. data. The initial and long-run values of r are 6%. Typical simulation values are shown in Figure 21.2. Note how rates can deviate from their initial value but are pulled back to the long-term value of 6%.

This model is convenient because it leads to closed-form solutions. The distribution of future values for r is summarized in Figure 21.3 by its mean and two-tailed 90% confidence bands (called maximum and minimum values). The graph shows that the mean is 6%, which is also the long-run value. The confidence
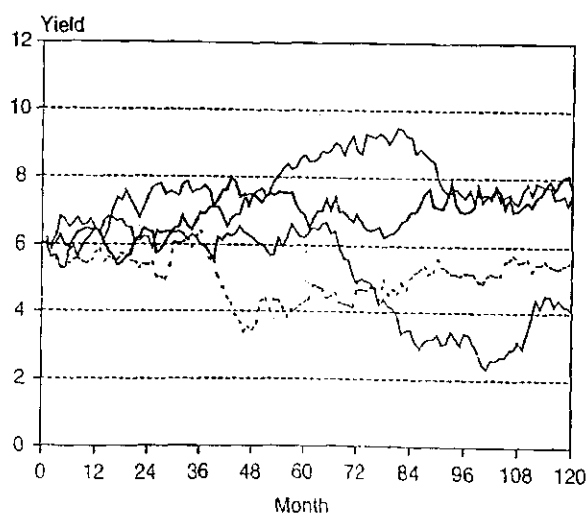
FIGURE 21.2 Simulation Paths for the Interest Rate
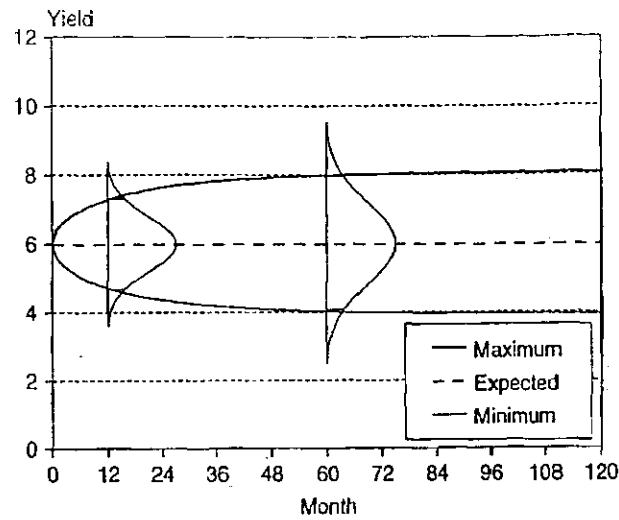
FIGURE 21.3   Distribution Profile for the Interest Rate

bands initially widen due to the increasing horizon, and then converge to a fixed value due to the mean reversion effect.

The next step is to value the swap. At each point in time, the current market value of the receive-fixed swap is the difference between the value of a fixed-coupon bond and a floating-rate note, as seen in Chapter 8:

$$V_t = B(F, t, T, c, r_t) - B(F, FRN) \tag{21.9}$$

Here, F is the notional amount or face value, c is the annualized *fixed* coupon rate, and T is the maturity date. The risk to the swap comes from the fact that the fixed leg has a coupon c that could differ from prevailing market rates. The principals are not exchanged.

Figure 21.4 illustrates the changes in cash flows that could arise from a drop in rates from 6% to 4% after five years. Assume the swap has notional of $N = \$100$ million and pays semiannually. Every six months, the receive-fixed party is owed $\$100 \times (6 - 4)\% \times 0.5 = \$1$ million until the maturity of the swap. With 10 payments remaining, this adds up to a positive credit exposure of approximately $10 million. More precisely, discounting over the life of the remaining payments gives $8.1 million as of the valuation date.

In what follows, we assume that the swap receives fixed payments that are paid at a continuous rate instead of semiannually, which simplifies the example. Otherwise, there would be discontinuities in cash-flow patterns, and we would have to consider the risk of the floating leg as well. We also use continuous compounding. Defining N as the number of remaining years, the coupon bond value is

$$B(\$100, N, c, r) = \$100\frac{c}{r}\left[1 - e^{-rN}\right] + \$100e^{-rN} \tag{21.10}$$

Cash flow ($ million)
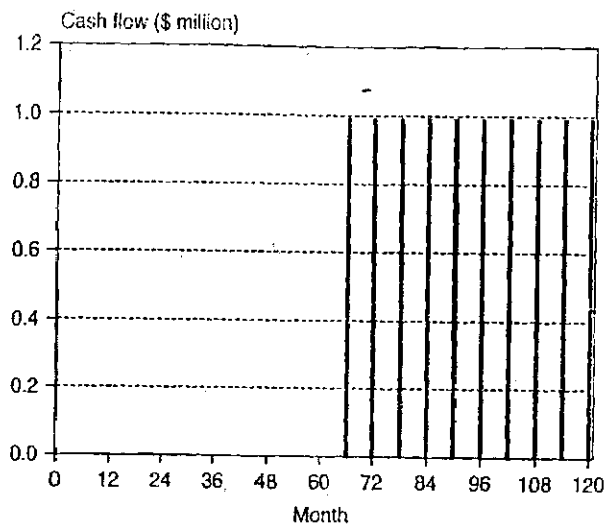
FIGURE 21.4 Net Cash Flows When Rates Fall to 4% After Five Years

as seen in the appendix to Chapter 1. The first term is the present value of the fixed-coupon cash flows discounted at the current rate r. The second term is the repayment of principal. For the special case where the coupon rate is equal to the current market rate ($c = r$), the market value is indeed $100 for this par bond. The floating-rate note can be priced in the same way, but with a coupon rate that is always equal to the current rate. Hence, its value is always at par.

To understand the exposure profile of the coupon bond, we need to consider two opposing effects as time goes by:

- The diffusion effect, which increases the uncertainty in the interest rate
- The amortization effect, which decreases the bond's duration toward zero

The latter effect is described in Figure 21.5, which shows the bond's duration converging to zero. This explains why the bond's market value converges to the face value upon maturity, whatever happens to the current interest rate.

Years
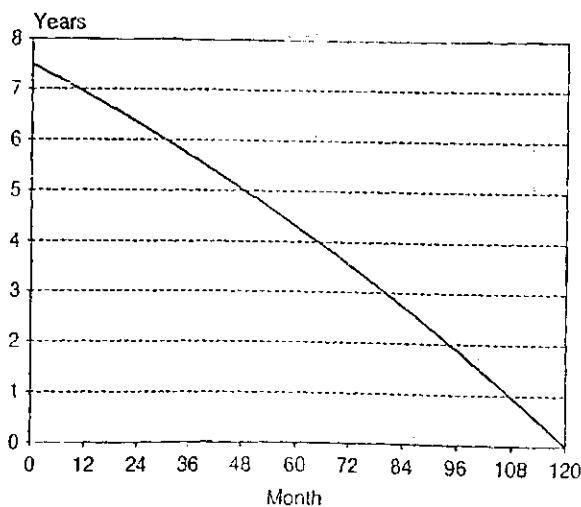
FIGURE 21.5 Duration Profile for a 10-Year Bond

Because the bond is a strictly monotonic function of the current yield, we can compute the 90% confidence bands by valuing the bond using the extreme interest rates range at each point in time. We use Equation (21.10) at each point in time in Figure 21.3. This exposure profile is shown in Figure 21.6.

Initially, the market value of the bond is $100. After two or three years, the range of values is the greatest, from $87 to $115. Thereafter, the range converges to the face value of $100. But overall, the fluctuations as a *proportion* of the face value are relatively small. Considering other approximations in the measurement of credit risk, such as the imprecision in default probability and recovery rate, assuming a constant exposure for the bond is not a bad approximation.

This is not the case, however, for the interest rate swap. Its value can be found by subtracting $100 (the value of the floating-rate note) from the value of the coupon bond. Initially, its value is zero. Thereafter, it can take on positive or negative values. Credit exposure is the positive value only. Figure 21.7 presents the profile of the expected exposure and of the maximum (worst) exposure at the one-sided 95% level. It also shows the average maximum exposure over the whole life of the swap.

Intuitively, the value of the swap is derived from the difference between the fixed and floating cash flows. Consider a swap with two remaining payments and a notional amount of $100. Its value is

$$
\begin{aligned}
V_t &= \$100\left[\frac{c}{(1+r)} + \frac{c}{(1+r)^2} + \frac{1}{(1+r)^2}\right] - \$100\left[\frac{r}{(1+r)} + \frac{r}{(1+r)^2} + \frac{1}{(1+r)^2}\right] \\
&= \$100\left[\frac{(c-r)}{(1+r)} + \frac{(c-r)}{(1+r)^2}\right]
\end{aligned}
\tag{21.11}
$$

Note how the principal payments cancel out and we are left with the discounted *net* difference between the fixed coupon and the prevailing rate $(c - r)$.
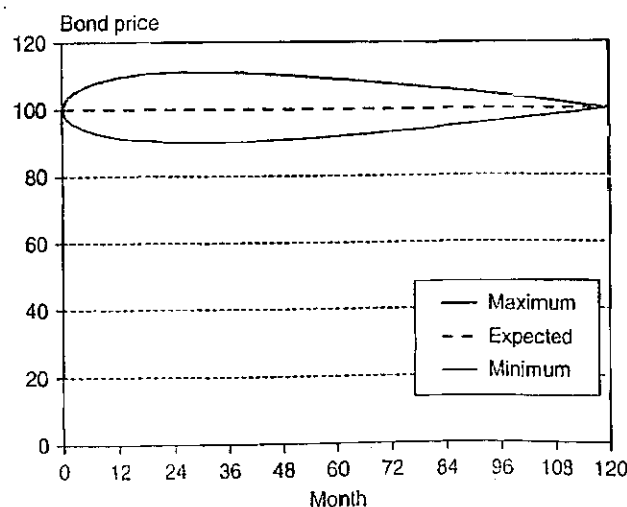
FIGURE 21.6   Exposure Profile for a 10-Year Bond
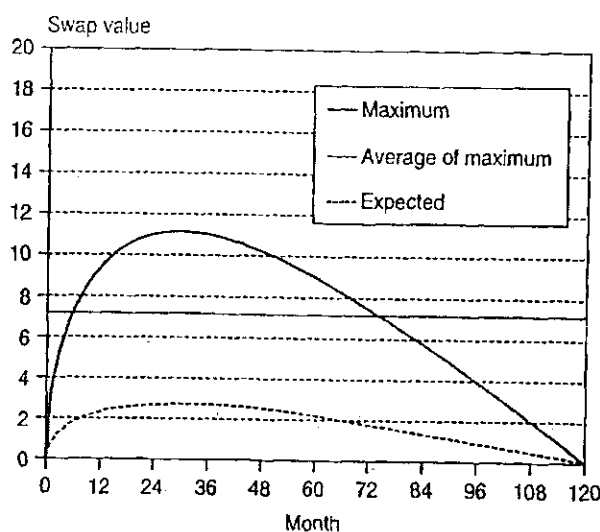
FIGURE 21.7 Exposure Profile for a 10-Year Interest Rate Swap

This information can be used to assess the expected exposure and worst exposure on a target date. The peak exposure occurs around the second year into the swap, or at about one-fourth of the swap's life. At that point, the expected exposure is about 3% to 4% of the notional, which is much less than that of the bond. The worst exposure peaks at about 10% to 15% of the notional. In practice, these values depend on the particular stochastic process used, but the exposure profiles will be qualitatively similar.

To assess the potential variation in swap values, we can make some approximations based on duration. Consider first the very short-term exposure, for which mean reversion and changes in durations are not important. The volatility of changes in rates then simply increases with the square root of time. Given a 0.25% per month volatility and 7.5-year initial duration, we can approximate the volatility of the swap value over the next year as

$$\sigma(V) = \$100 \times 7.5 \times 0.25\% \sqrt{12} = \$6.5 \text{ million}$$

Multiplying by 1.645, we get \$10.7 million, which is close to the actual \$9.4 million 95% worst exposure in a year reported in Figure 21.7.

The trade-off between declining duration and increasing risk can be formalized with a simple example. Assume that the bond's (modified) duration is proportional to the remaining life, $D = k(T - t)$ at any date t. The volatility from 0 to time t can be written as $\sigma(r_t - r_0) = \sigma \sqrt{t}$. Hence, the swap volatility is

$$\sigma(V) = [k(T - t)] \times \sigma \sqrt{t} \qquad (21.12)$$

To see where it reaches a maximum, we differentiate with respect to t:

$$\frac{d\sigma(V)}{dt} = [k(-1)]\sigma \sqrt{t} + [k(T - t)]\sigma \frac{1}{2\sqrt{t}}$$

Setting this to zero, we have

$$\sqrt{t} = (T-t)\frac{1}{2\sqrt{t}}, \quad 2t = (T-t)$$

which gives

$$t_{\text{MAX}} = (1/3)T \tag{21.13}$$

The maximum exposure occurs at one-third of the life of the swap. This occurs later than the one-fourth point reported previously because we assumed no mean reversion.

Further, we can check how this evolves with the maturity of the contract. At that point, the worst credit exposure will be

$$1.645\ \sigma(V_{\text{MAX}}) = 1.645\left[k(2/3)T\sigma\sqrt{T/3}\right] = \left[1.645k(2/3)\sigma\sqrt{1/3}\right]T^{3/2} \tag{21.14}$$

which shows that the WCE increases as $T^{3/2}$, which is faster than the maturity.

Figure 21.8 shows the exposure profile of a five-year swap. Here again, the peak exposure occurs at one-third of the swap's life. As expected, the magnitude is lower, with the peak expected exposure only about 1% of the notional.

Finally, Figure 21.9 displays the exposure profile when the initial interest rate is at 5% with a coupon of 6%. The swap starts in-the-money, with a current value of $7.9 million. With a long-run rate of 6%, the total exposure profile starts from a positive value, reaches a maximum after about two years, then converges to zero.
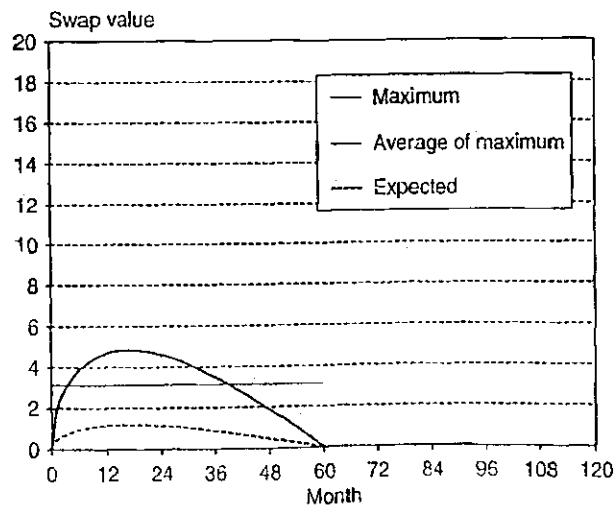
FIGURE 21.8 Exposure Profile for a Five-Year Interest Rate Swap
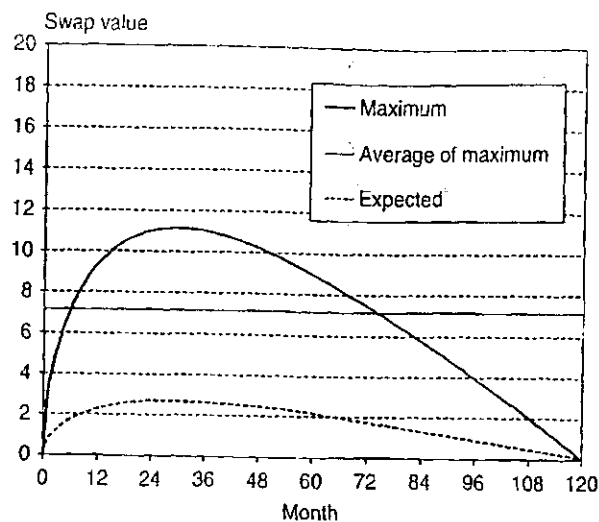
FIGURE 21.9   Exposure Profile for a 10-Year In-the-Money Swap

EXAMPLE 21.6: FRM EXAM 2004—QUESTION 43

In determining the amount of credit risk in a derivatives transaction, which of the following factors are used?

I. Notional principal amount of the underlying transaction
II. Current exposure
III. Potential exposure
IV. Peak exposure—the replacement cost in a worst case scenario
   a. I and II
   b. I, III, and IV
   c. III and IV
   d. II, III, and IV

EXAMPLE 21.7: FRM EXAM 2005—QUESTION 61

Assume that a bank enters into a USD 100 million, four-year annual pay interest rate swap, where the bank receives 6% fixed against 12-month LIBOR. Which of the following numbers best approximates the current exposure at the end of year 1 if the swap rate declines 125 basis points over the year?

   a. USD 3,420,069
   b. USD 4,458,300
   c. USD 3,341,265
   d. USD 4,331,382

EXAMPLE 21.8: PEAK EXPOSURE

Assume that the DV01 of an interest rate swap is proportional to its time to maturity (which at the initiation is equal to T). Assume that interest rate curve moves are parallel, stochastic with constant volatility, normally distributed, and independent. At what time will the maximum potential exposure be reached?

- a. $T/4$
- b. $T/3$
- c. $T/2$
- d. $3T/4$

EXAMPLE 21.9: FRM EXAM 2000—QUESTION 29

Determine at what point in the future a derivatives portfolio will reach its maximum potential exposure. All the derivatives are on one underlying, which is assumed to move in a stochastic fashion (variance in the underlying's value increases linearly with time passage). The derivatives portfolio's sensitivity to the underlying is expected to drop off as $(T - t)^2$, where T is the time from today until the last contract in the portfolio rolls off, and t is the time from today.

- a. $T/5$
- b. $T/3$
- c. $T/2$
- d. None of the above

EXAMPLE 21.10: FRM EXAM 2002—QUESTION 83

Assume that you have entered into a fixed-for-floating interest rate swap that starts today and ends in six years. Assume that the duration of your position is proportional to the time to maturity. Also assume that all changes in the yield curve are parallel shifts, and that the volatility of interest rates is proportional to the square root of time. When would the maximum potential exposure be reached?

- a. In two months
- b. In two years
- c. In six years
- d. In four years and five months

21.2.4 Exposure Profile for Currency Swaps

Exposure profiles are substantially different for other swaps. Consider, for instance, a currency swap where the notionals are 100 million U.S. dollars against 50 million British pounds (BP), set at an initial exchange rate of $S(\$/BP) = 2$.

The market value of a currency swap that receives foreign currency is

$$V_t = S_t(\$/BP)B^*(BP50, t, T, c^*, r^*) - B(\$100, t, T, c, r) \qquad (21.15)$$

Following the usual conventions, asterisks refer to foreign-currency values.

In general, this swap is exposed to domestic as well as foreign interest rate risk. When we just have two remaining coupons, the value of the swap evolves according to

$$
\begin{aligned}
V = S \times 50 &\left[\frac{c^*}{(1+r^*)} + \frac{c^*}{(1+r^*)^2} + \frac{1}{(1+r^*)^2} \right] \\
&- \$100 \left[\frac{c}{(1+r)} + \frac{c}{(1+r)^2} + \frac{1}{(1+r)^2} \right] \qquad (21.16)
\end{aligned}
$$

Note that the principals do not cancel each other, unlike an interest rate swap. Instead, they are paid at maturity in different currencies, which is a major source of credit exposure.

In what follows, we will assume for simplicity that there is no interest rate risk, or that the value of the swap is dominated by currency risk. Further, we assume that the coupons are the same in the two currencies; otherwise there would be an asymmetrical accumulation of payments. As before, we have to choose a stochastic process for the spot rate. Say this is a lognormal process with constant variance and no trend:

$$dS_t = \sigma S_t \, dz_t \qquad (21.17)$$

We choose $\sigma = 12\%$ annually, which is realistic, as seen in the chapter on market risk factors. This process ensures that the rate never becomes negative.

Figure 21.10 presents the exposure profile of a 10-year currency swap. Here there is no amortization effect, and exposure increases continuously over time. The peak exposure occurs at the end of the life of the swap. At that point, the expected exposure is about 10% of the notional, which is much higher than for the interest rate swap. The worst exposure is commensurately high, at about 45% of notional.

Although these values depend on the particular stochastic process and parameters used, this example demonstrates that credit exposures for currency swaps are far greater than for interest rate swaps, even with identical maturities.
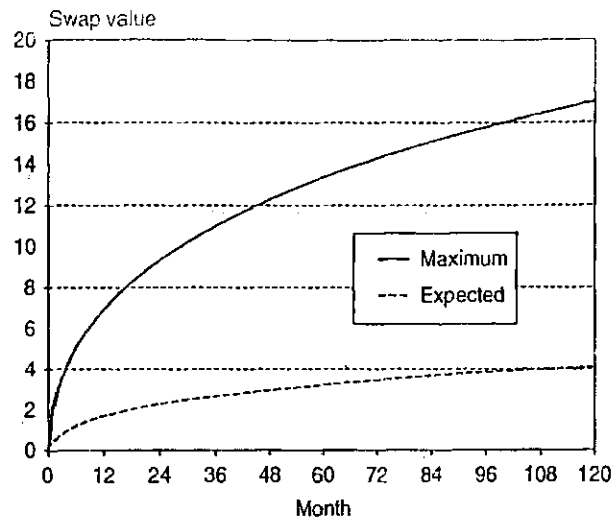
FIGURE 21.10   Exposure Profile for a 10-Year Currency Swap

21.2.5   Exposure Profile for Different Coupons

So far, we have assumed a flat term structure and equal coupon payments in different currencies, which creates a symmetric situation for the exposure for the long and short parties. In reality, these conditions will not hold, and the exposure patterns will be asymmetric.

Consider, for instance, the interest rate swap in Equation (21.11). If the term structure slopes upward, the coupon rate is greater than the floating rate, $c > r$, in which case the net payment to the party receiving fixed is initially positive. The value of the two-period swap can be analyzed by projecting floating payments at the forward rate:

$$V_t = \frac{(c - s_1)}{(1 + s_1)} + \frac{(c - f_{12})}{(1 + s_2)^2}$$

where s_1, s_2 are the one- and two-year spot rates, and f_{12} is the one- to two-year forward rate.

Example

Consider a \$100 million interest rate swap with two remaining payments. We have $s_1 = 5\%$, $s_2 = 6.03\%$ and hence, using $(1 + s_2)^2 = (1 + s_1)(1 + f_{12})$, we have $f_{12} = 7.07\%$. The coupon yield of $c = 6\%$ is such that the swap has zero initial value. The following table shows that the present value of the first payment (to the party receiving fixed) is positive and equal to \$0.9524. The second payment then must be negative, and is equal to $-\$0.9524$. The two payments exactly offset each other because the swap has zero value.

Time	Expected Spot	Expected Payment	Discounted
1	5%	$6.00 - 5.00 = +1.00$	$+0.9524$
2	7.07%	$6.00 - 7.07 = -1.07$	-0.9524
Total			-0.0000

This pattern of payments, however, creates more credit exposure to the fixed payer because it involves a payment in the first period offset by a receipt in the second. If the counterparty defaults shortly after the first payment is made, there could be a credit loss even if interest rates have not changed.

KEY CONCEPT

With a positively sloped term structure, the receiver of the floating rate (payer of the fixed rate) has a greater credit exposure than the counterparty.

A similar issue arises with currency swaps when the two coupon rates differ. Low nominal interest rates imply a higher forward exchange rate. The party that receives payments in a low-coupon currency is expected to receive greater payments later during the exchange of principal. If the counterparty defaults, there could be a credit loss even if rates have not changed.

KEY CONCEPT

The receiver of a low-coupon currency has greater credit exposure than the counterparty.

EXAMPLE 21.11: FRM EXAM 2000—QUESTION 47

Which one of the following deals would have the greatest credit exposure for a $1,000,000 deal size (assume the counterparty in each deal is an AAA-rated bank and has no settlement risk)?

a. Pay fixed in an Australian dollar (AUD) interest rate swap for one year.
b. Sell USD against AUD in a one-year forward foreign exchange contract.
c. Sell a one-year AUD cap.
d. Purchase a one-year certificate of deposit.

EXAMPLE 21.12: FRM EXAM 2001—QUESTION 8

Which of the following 10-year swaps has the highest potential credit exposure?

a. A cross-currency swap *after* two years
b. A cross-currency swap *after* nine years
c. An interest rate swap *after* two years
d. An interest rate swap *after* nine years

EXAMPLE 21.13: FRM EXAM 2004—QUESTION 14

BNP Paribas has just entered into a plain-vanilla interest-rate swap as a pay-fixed counterparty. Credit Agricole is the receive-fixed counterparty in the same swap. The forward spot curve is upward-sloping. If LIBOR starts trending down and the forward spot curve flattens, the credit risk from the swap will:

a. Increase only for BNP Paribas
b. Increase only for Credit Agricole
c. Decrease for both BNP Paribas and Credit Agricole
d. Increase for both BNP Paribas and Credit Agricole

21.3 EXPOSURE MODIFIERS

In a continuing attempt to decrease credit exposures, the industry has developed a number of methods to limit exposures. This section analyzes marking to market, margins and collateral, exposure limits, recouponing, and netting arrangements.

Other modifiers include third-party guarantees and purchasing credit derivatives. The former involve receiving, typically from a bank, a guarantee of payment should the counterparty fail. Credit derivatives will be covered in the next chapter.

21.3.1 Marking to Market

The ultimate form of reducing credit exposure is marking to market (MTM). Marking to market involves settling the variation in the contract value on a regular basis (e.g., daily). For OTC contracts, counterparties can agree to longer periods (e.g., monthly or quarterly). If the MTM treatment is symmetrical across the two counterparties, it is called two-way marking to market. Otherwise, if one party settles losses only, it is called one-way marking to market.

Marking to market has long been used by organized exchanges to deal with credit risk. The reason is that exchanges are accessible to a wide variety of investors, including retail speculators, who are more likely to default than others. On OTC markets, in contrast, institutions interacting with each other typically have an ongoing relationship. As one observer put it,

Futures markets are designed to permit trading among strangers, as against other markets which permit only trading among friends.

With daily marking to market, the *current* exposure is reduced to zero. There is still, however, *potential* exposure because the value of the contract could change before the next settlement. Potential exposure arises from (1) the time interval between MTM periods and (2) the time required for liquidating the contract when the counterparty defaults.

In the case of a retail client, the broker can generally liquidate the position fairly quickly, within a day. When positions are very large, as in the case of brokers dealing with Long-Term Capital Management (LTCM), however, the liquidation period could be much longer. Indeed, LTCM's bailout was motivated by the potential disruption to financial markets had brokers attempted to liquidate their contracts with LTCM at the same time.

Marking to market introduces other types of risks, however:

- Operational risk, which is due to the need to keep track of contract values and to make or receive payments daily
- Liquidity risk, because the institution now needs to keep enough cash to absorb variations in contract values

Margins   Potential exposure is covered by margin requirements. Margins represent cash or securities that must be advanced in order to open a position. The purpose of these funds is to provide a buffer against potential exposure.

Exchanges, for instance, require a customer to post an initial margin when establishing a new position. This margin serves as a performance bond to offset possible future losses should the customer default. Contract gains and losses are then added to the posted margin in the equity account. Whenever the value of this equity account falls below a threshold, set at a maintenance margin, new funds must be provided.

Margins are set in relation to price volatility and to the type of position, speculative or hedging. Margins increase for more volatile contracts. Margins are typically lower for hedgers because a loss on the futures position can be offset by a gain on the physical, assuming no basis risk. Some exchanges set margins at a level that covers the 99th percentile of worst daily price changes, which is a daily VAR system for credit risk.

Collateral   Over-the-counter markets may allow posting securities as collateral instead of cash. This collateral protects against current and potential exposure. Typically, the amount of the collateral will exceed the funds owed by an amount

known as the haircut. Collateral is typically managed within the ISDA credit support annex (CSA).

The haircut reflects both default risk and market risk. Safe counterparties will in general have lower haircuts. This also depends, however, on the downside risk of the asset. For instance, cash can have a haircut of zero, which means that there is full protection against current exposure. Government securities can require a haircut of 1%, 3%, and 8% for short-term, medium-term, and longer-term maturities, respectively. With greater price volatility, there is an increasing chance of losses if the counterparty defaults and the collateral loses value, which explains the increasing haircuts.

As an example, assume that hedge fund A enters swap with bank B. To mitigate A's credit risk, the two parties enter a collateral agreement that specifies the conditions under which B can ask for collateral. Now suppose the contract moves in-the-money for B, which requests $1 million in collateral from A. The funds are legally still the property of A but under the administration of B. If A defaults, B is entitled to sell the collateral and terminate the contract. Any positive excess value is returned to A. Conversely, if the collateral was not sufficient, B will have a claim against A.

21.3.2 Exposure Limits

Credit exposure can also be managed by setting position limits on the exposure to a counterparty. Ideally, these should be evaluated in a portfolio context, taking into account all the contracts between an institution and a counterparty.

To enforce limits, information on transactions must be centralized in middle-office systems. This generates an *exposure profile* for each counterparty, which can be used to manage credit line usage for several maturity buckets. Proposed new trades with the same counterparty should then be examined for their incremental effect.

These limits can be also set at the instrument level. In the case of a swap, for instance, an exposure cap requires a payment to be made whenever the value of the contract exceeds some amount. Figure 21.11 shows the effect of a $5 million cap on our 10-year swap. If, after two years, say, the contract suddenly moves into a positive value of $11 million, the counterparty would be required to make a payment of $6 million to bring the swap's outstanding value back to $5 million. This limits the worst exposure to $5 million and also lowers the average exposure.

21.3.3 Recouponing

Another method for controlling exposure at the instrument level is recouponing. Recouponing refers to a clause in the contract requiring the contract to be marked-to-market at some fixed date. This involves (1) exchanging cash to bring the MTM value to zero and (2) resetting the coupon or the exchange rate to prevailing market values.
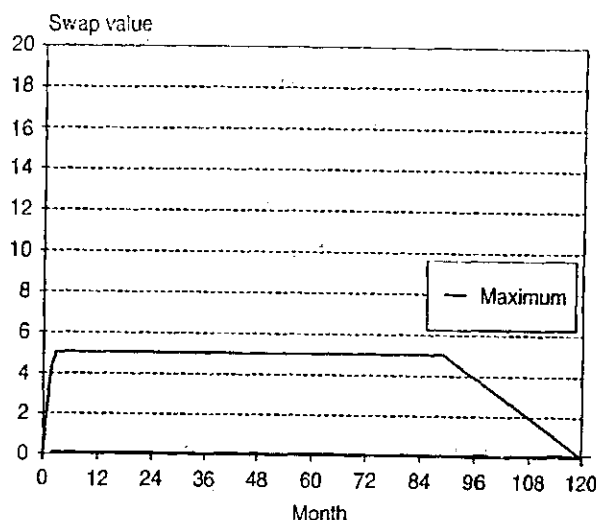
FIGURE 21.11 Effect of Exposure Cap

Figure 21.12 shows the effect of five-year recouponing on our 10-year swap. The exposure is truncated to zero after five years. Thereafter, the exposure profile is that of a swap with a remaining five-year maturity.

21.3.4 Netting Arrangements

Perhaps the most powerful mechanism for controlling exposures are netting agreements. These are now a common feature of standardized master swap agreements such as the one established in 1992 by the International Swaps and Derivatives Association (ISDA). This will be explained in more detail in Chapter 27.

The purpose of these agreements is to provide for the netting of payments across a set of contracts. In case of default, a counterparty cannot stop payments on contracts that have negative value while demanding payment on positive-value
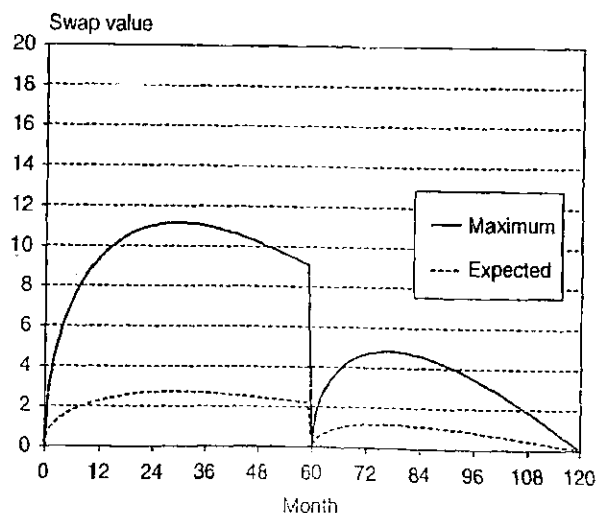
FIGURE 21.12 Effect of Recouponing After Five Years

contracts. As a result, this system reduces the exposure to the net value for all the contracts covered by the netting agreement. This prevents *cherry-picking* by the administrator of the defaulted counterparty.

Netting can be classified into three types:

1. Payment netting involves the daily off-setting of several claims in the same currency. An example is an interest rate swap, where only the net payment, floating against fixed, are exchanged.
2. Novation netting involves the cancellation of several contracts between the two parties, resulting in a replacement contract with new, net payments. As an example consider, a forward trade where A must pay 10 million euros in exchange for receiving 15 million dollars from B. In another trade, A must receive 5 million euros from B and pay 7 million dollars in exchange. Under novation, the two contracts are reduced to a payment of 5 million euros from A in exchange for a receipt of 8 million dollars from B.
3. Close-out netting involves the cancellation of all transactions under the master agreement in the event of bankruptcy or any other specified default event. The trades are then netted at market value.

Table 21.1 gives an example with four contracts. Without a netting agreement, the exposure of the first two contracts is the sum of the positive part of each, or $100 million. In contrast, if the first two fall under a netting agreement, their value would offset each other, resulting in a net exposure of $100 − $60 = $40 million. If contracts 3 and 4 do not fall under the netting agreement, the exposure is then increased to $40 + $25 = $65 million.

To summarize, the net exposure with netting is

$$\text{Net exposure} = \text{Max}(V, 0) = \text{Max}\left(\sum_{i=1}^{N} V_i, 0\right) \tag{21.18}$$

TABLE 21.1 Comparison of Exposure with and without Netting

Contract	Contract Value	Exposure	
		No Netting	With Netting for 1 and 2
Under netting agreement			
1	+$100	+$100	
2	−$60	+$0	
Total, 1 and 2	+$40	+$100	+$40
No netting agreement			
3	+$25	+$25	
4	−$15	+$0	
Grand total, 1 to 4	+$50	+$125	+$65

Without netting agreement, the gross exposure is the sum of all positive-value contracts:

$$\text{Gross exposure} = \sum_{i=1}^{N} \text{Max}(V_i, 0) \qquad (21.19)$$

This is always greater than, or equal to, the exposure under the netting agreement.

The benefit from netting depends on the number of contracts N and the extent to which contract values covary. The larger the value of N and the lower the correlation, the greater the benefit from netting. It is easy to verify from Table 21.1 that if all contracts move into positive value at the same time, or have high correlation, there will be no benefit from netting.

Figures 21.13 and 21.14 illustrate the effect of netting on a portfolio of two swaps with the same counterparty. In each case, interest rates could increase or decrease with the same probability.

In Figure 21.13, the bank is long both a receive-fixed 10-year and five-year swap. The top panel describes the worst exposure when rates fall. In this case there is positive exposure for both contracts, which we add to get the total portfolio exposure. Whether there is netting or not does not matter, because the two positions are positive at the same time. The bottom panel describes the worst exposure when rates increase. Both positions, as well as the portfolio, have zero exposure.

In Figure 21.14, the bank is long the 10-year and short the five-year swap. When rates fall, the first swap has positive value and the second has negative
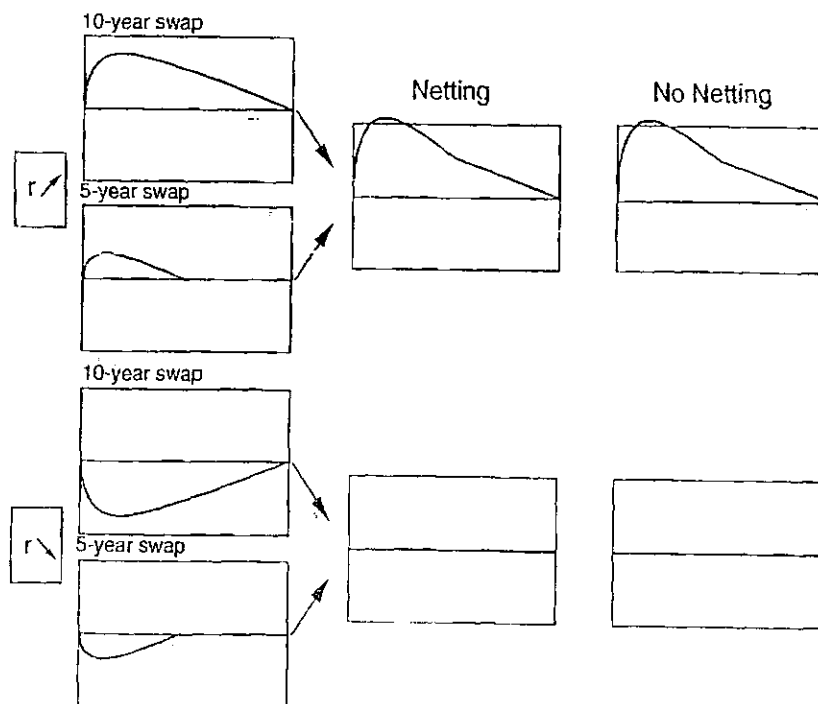
FIGURE 21.13 Netting with Two Long Positions
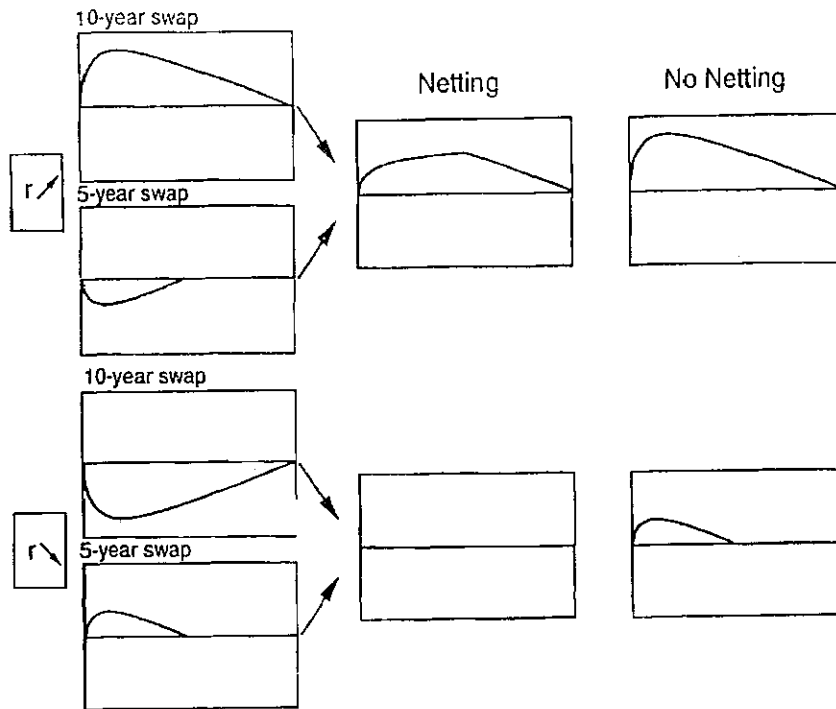
FIGURE 21.14   Netting with a Long and a Short Position

value. With netting, the worst exposure profile is reduced. In contrast, with no netting the exposure is that of the 10-year swap. Conversely, when rates increase, the swap value is negative for the first and positive for the second. With netting, the exposure profile is zero, whereas without netting it is the same as that of the five-year swap. This shows that netting is more effective with diversified positions.

Banks provide some information in their annual report about the benefit of netting for their current exposure. Without netting agreements or collateral, the gross replacement value (GRV) is reported as the sum of the worst-case exposures if all counterparties K default at the same time:

$$\text{GRV} = \sum_{k=1}^{K} (\text{Gross exposure})_k = \sum_{k=1}^{K} \left[\sum_{i=1}^{N_k} \text{Max}(V_i, 0) \right] \qquad (21.20)$$

With netting agreements and collateral, the resulting exposure is defined as the net replacement value (NRV). This is the sum, over all counterparties, of the net positive exposure:

$$\text{NRV} = \sum_{k=1}^{K} (\text{Net exposure})_k = \sum_{k=1}^{K} \left[\text{Max} \left(\sum_{i=1}^{N_k} V_i, 0 \right) \right] \qquad (21.21)$$

If collateral is held, this should be subtracted from the net exposure.

The effectiveness of netting can be assessed from BIS statistics for the OTC derivatives markets. As described in Chapter 5, the total notional amounts added

up to $596 trillion as of December 2007. The gross market value, defined as the summation of the positive part of all contracts, was estimated at $15.8 trillion. The net credit exposure was reduced to $3.3 trillion. Thus, netting reduces the exposure by 80%.

EXAMPLE 21.14: FRM EXAM 2002—QUESTION 89

If we assume that the VAR for the portfolio of trades with a given counterparty can be viewed as a measure of potential credit exposure, which of the following could *not* be used to decrease this credit exposure?

a. A netting agreement

b. Collateral

c. A credit derivative that pays out if the counterparty defaults

d. An offsetting trade with a different counterparty

EXAMPLE 21.15: FRM EXAM 2005—QUESTION 96

Which of the following statements correctly describes the impact of signing a netting agreement with a counterparty?

a. It will increase or have no effect on the total credit exposure.

b. It will decrease or have no effect on the total credit exposure.

c. It will increase exposure if exposure is net long and decrease exposure if it is net short.

d. Its impact is impossible to determine based on the available information.

EXAMPLE 21.16: FRM EXAM 2006—QUESTION 39

What are the benefits of novation?

a. Both parties are allowed to walk away from the contract in the event of default.

b. In a bilateral contract, it is specified that on default, the non-defaulting party nets gains and losses with the defaulting counterparty to a single payment for all covered transactions.

c. Financial market contracts can be terminated upon an event of default prior to the bankruptcy process.

d. Obligations are amalgamated with others.

EXAMPLE 21.17: FRM EXAM 2006—QUESTION 140

Company EFG is a large derivative market-maker that has many contracts with counterparty JKL, some transacted in the same legal jurisdiction and others across different legal jurisdictions. As a result, EFG has some contracts with JKL covered under legally enforceable netting agreement A, some contracts with JKL covered under legally enforceable netting agreement B, and some contracts with JKL with no netting agreement. Ignoring the effect of margin, if the current value (i.e., market value of the contract minus collateral and recovery value) and the netting agreement status of each contract with JKL are as shown below, what is EFG's current counterparty credit exposure to JKL?

- Contracts 1, 2, and 3 are covered under Netting Agreement A, with respective current values of USD 2,105; (-USD 3,319); USD 1,977.
- Contracts 4 to 7 are covered under Netting Agreement B, with respective current values of USD 5,876; (-USD 633); (-USD 2,335); USD 4,006.
- Contracts 8 and 9 are not covered by any Netting Agreement, with respective current values of USD 2,439; (-USD 1,504).

    a. USD 8,612
    b. USD 6,914
    c. USD 14,899
    d. USD 10,116

EXAMPLE 21.18: FRM EXAM 2003—QUESTION 24

Bank A, which is AAA-rated, trades a 10-year interest rate swap (semiannual payments) with bank B, rated A—. Because of bank B's poor credit rating, Bank A is concerned about its 10-year exposure. Which of the following measures help mitigate bank A's credit exposure to bank B?

  I. Negotiate a CSA with bank B and efficiently manage the collateral management system
  II. Execute the swap deal as a reset swap wherein the swap will be marked-to-market every six months
  III. Execute the swap deal with a break clause in the fifth year
  IV. Decrease the frequency of coupon payments from semiannual to annual
    a. I only
    b. IV only
    c. I, II, III, and IV
    d. I, II, and III

EXAMPLE 21.19: FRM EXAM 2002—QUESTION 73

Consider the following information. You have purchased 10,000 barrels of oil for delivery in one year at a price of $25/barrel. The rate of change of the price of oil is assumed to be normally distributed with zero mean and annual volatility of 30%. Margin is to be paid within two days if the credit exposure becomes greater than $50,000. There are 252 business days in the year. Assuming enforceability of the margin agreement, which of the following is the closest number to the 95% one-year credit risk of this deal governed under the margining agreement?

a. $50,000
b. $58,000
c. $61,000
d. $123,000

21.4 CREDIT RISK MODIFIERS

Credit risk modifiers operate on credit exposure, default risk, or a combination of the two. For completeness, this section discusses modifiers that affect default risk.

21.4.1 Credit Triggers

Credit triggers specify that if either counterparty's credit rating falls below a specified level, the other party has the right to have the swap cash settled. These are not exposure modifiers, but rather attempt to reduce the probability of default on that contract. For instance, if all outstanding swaps can be terminated when the counterparty rating falls below A, the probability of default is lowered to the probability that a counterparty will default when rated A or higher.

These triggers are useful when the credit rating of a firm deteriorates slowly, because few firms directly jump from investment-grade into bankruptcy. The increased protection can be estimated by analyzing transition probabilities, as discussed in a previous chapter. For example, say a transaction with an AA-rated borrower has a cumulative probability of default of 0.81% over 10 years. If the contract can be terminated whenever the rating falls to BB or below, this probability falls to 0.23%.

21.4.2 Time Puts

Time puts, or mutual termination options, permit either counterparty to terminate unconditionally the transaction on one or more dates in the contract. This feature decreases both the default risk and the exposure. It allows one counterparty to

terminate the contract if the exposure is large and the other party's rating starts to slip.

Triggers and puts, which are types of contingent requirements, can cause serious trouble, however. They create calls on liquidity precisely in states of the world where the company is faring badly, putting additional pressure on the company's liquidity. Indeed, triggers in some of Enron's securities forced the company to make large cash payments and propelled it into bankruptcy. Rather than offering protection, these clauses can trigger bankruptcy, affecting all creditors adversely.

21.5 IMPORTANT FORMULAS

Credit exposure: $CE_t = Max(V_t, 0)$

Long options: $CE_t = V_t$; short options, $CE_t = 0$

Expected credit exposure (ECE): $ECE = \int_{-\infty}^{+\infty} Max(x, 0) f(x) dx$

Worst credit exposure (WCE): $1 - p = \int_{WCE}^{\infty} f(x) dx$

Credit exposure for an interest rate swap, from:
$$V_t = B(F, t, T, c, r_t) - B(F, FRN)$$

Volatility of credit exposure for an interest rate swap: $\sigma(V_t) = [k(T - t)] \times \sigma \sqrt{t}$

Credit exposure for a currency swap, from:
$$V_t = S_t(\$/FC)B^*(F^*, t, T, c^*, r^*) - B(F, t, T, c, r)$$

Gross credit exposure: $\sum_{i=1}^{N} Max(V_i, 0)$

Net credit exposure with netting: $Max(V, 0) = Max\left(\sum_{i=1}^{N} V_i, 0\right)$

Gross replacement value (GRV):
$$GRV = \sum_{k=1}^{K} (Gross\ exposure)_k = \sum_{k=1}^{K} \left[\sum_{i=1}^{N_k} Max(V_i, 0)\right]$$

Net replacement value (NRV):
$$NRV = \sum_{k=1}^{K} (Net\ exposure)_k = \sum_{k=1}^{K} \left[Max\left(\sum_{i=1}^{N_k} V_i, 0\right)\right]$$
minus collateral held

21.6 ANSWERS TO CHAPTER EXAMPLES

Example 21.1: FRM Exam 2006—Question 95

d. For a loss to occur, the exposure must be positive, (meaning rates move in your favor, and the counterparty must default.

Example 21.2: FRM Exam 2002—Question 93

d. Selling an option does not create credit exposures, because the premium has been received up front and the option can only create a future liability.

Example 21.3: FRM Exam 2006—Question 117

b. Selling an option does not create exposure, so answer d. is wrong. Longer horizons create a potential for larger price movements, so answer a. is wrong. The potential gain from being long is greater than being short. Prices can go up several times from the initial price for a long position. For a short position, the maximum gain is if the price goes to zero.

Example 21.4: FRM Exam 2004—Question 8

c. This is the only answer that involves truly selling an option, which has no credit exposure. A collar involves the sale and purchase of an option.

Example 21.5: FRM Exam 2001—Question 84

b. Being short an option creates no credit exposure, so answers c. and d. are false. With the short forward contract, a gain will be realized if the euro has depreciated.

Example 21.6: FRM Exam 2004—Question 43

d. All measures of exposure are important, current, potential, and peak. The notional amount, however, is not at risk.

Example 21.7: FRM Exam 2005—Question 61

a. The value of the fixed-rate bond is $6/(1 + 4.75\%)^1 + 6/(1 + 4.75\%)^2 + 106/(1 + 4.75\%)^3 = 103.420$, Subtracting $100 for the floating leg gives an exposure of $3.4 million. More intuitively, the sum of the coupon difference is 3 times $(6\% - 4.75\%)\$100 = \$1.25$, or around $3.75 million without discounting.

Example 21.8: Peak Exposure

b. See Equation (21.14).

Example 21.9: FRM Exam 2000—Question 29

a. This question alters the variance profile in Equation (21.12). Taking now the variance instead of the volatility, we have $\sigma^2 = k(T - t)^4 \times t$, where k is a constant. Differentiating with respect to t,

$$\frac{d\sigma^2}{dt} = k[(-1)4(T - t)^3]t + k[(T - t)^4] = k(T - t)^3[-4t + T - t]$$

Setting this to zero, we have $t = T/5$. Intuitively, because the exposure profile drops off faster than in Equation (21.12), we must have earlier peak exposure than $T/3$.

Example 21.10: FRM Exam 2002—Question 83

b. Exposure is a function of duration, which decreases with time, and interest rate volatility, which increases with the square root of time. Define T as the original maturity and k as a constant. This give $\sigma(V_t) = k(T-t)\sqrt{t}$. Taking the derivative with respect to t gives a maximum at $t = (T/3)$. This gives $t = (6/3) =$ two years.

Example 21.11: FRM Exam 2000—Question 47

d. The CD has the whole notional at risk. Otherwise, the next greater exposure is for the forward currency contract and the interest rate swap. The short cap position has no exposure if the premium has been collected. Note that the question eliminates settlement risk for the forward contract.

Example 21.12: FRM Exam 2001—Question 8

a. The question asks about potential exposure for various swaps during their life. Interest rate swaps generally have lower exposure than currency swaps because there is no market risk on the principals. Currency swaps with longer remaining maturities have greater potential exposure. This is the case for the 10-year currency swap, which after two years has eight years remaining to maturity.

Example 21.13: FRM Exam 2004—Question 14

b. With an upward-sloping term structure, the fixed payer has greater credit exposure. He receives less initially, but receives more later. This backloading of payments increases credit exposure. Conversely, if the forward curve flattens, the fixed payer, i.e., BNP Paribas has less credit exposure. Credit Agricole must have greater credit exposure. Alternatively, if LIBOR drifts down, BNP will have to pay more, and its counterparty will have greater credit exposure.

Example 21.14: FRM Exam 2002—Question 89

d. An offsetting trade with a different party will provide no credit protection. If the first party defaults while the contract is in-the-money, there will be a credit loss.

Example 21.15: FRM Exam 2005—Question 90

b. Netting should decrease the credit exposure if contracts with the same counterparty have positive and negative values. In the worst case, that of one contract with positive value, there is no effect.

Example 21.16: FRM Exam 2006—Question 39

d. Answer a. is incorrect because this is a walk-away clause. Answer b. is incorrect because this is close-out netting. Answer c. is incorrect because this is a termination clause.

Example 21.17: FRM Exam 2006—Question 140

d. The sum of contracts 1, 2, and 3 is 763. The sum of contracts 4 to 7 is 6,914. We keep these two values because they are positive, and add the positive remaining contract value of 2,439, which gives a total of USD 10,116.

Example 21.18: FRM Exam 2003—Question 24

d. Collateral management will lower credit exposure, so Answer I. is correct. Resetting, or recouponing the swap, will also lower exposure. A break clause in five years will allow the marked-to-market, which also lowers exposure. On the other hand, decreasing the frequency of coupons will not change much the exposure. In fact, extending the period will increase exposure because there is a longer time to wait for the next payment, increasing the market will move in the favor or one counterparty.

Example 21.19: FRM Exam 2002—Question 73

c. The worst credit exposure is the $50,000 plus the worst move over two days at the 95% level. The worst potential move is $\alpha\sigma\sqrt{T} = 1.645 \times 30\% \times \sqrt{(2/252)} = 4.40\%$. Applied to the position worth $250,000, this gives a worst move of $10,991. Adding this to $50,000 gives $60,991.

Credit Derivatives and Structured Products

Credit derivatives are the latest tool in the management of portfolio credit risk. Credit derivatives are contracts whose value derives from the credit risk of an underlying obligor, corporate, sovereign, or multiname. They allow the exchange of credit risk from one counterparty to another. Credit derivatives initially grew from the need of banks to modify their credit exposure but since then have become essential portfolio management tools.

Like other derivatives, they can be traded on a standalone basis or embedded in some other instrument, such as a credit-linked note. This market has led to the expansion of structured credit products, through which portfolios of credit exposures are repackaged to better suit the needs of investors. This transformation relies on the securitization process, which was first applied to mortgage pools and was described in Chapter 7.

Section 22.1 presents an introduction to the size and rationale of these markets. Section 22.2 describes credit default swaps and their pricing. Other contracts such as total return swaps, credit spread forward and option contracts are covered in Section 22.3. Section 22.4 then presents credit structured products, including credit-linked notes and collateralized debt obligations (CDOs). Because of its importance, Section 22.5 describes the CDO market in more detail. Finally, Section 22.6 discusses the pros and cons of credit derivatives and structured products.

22.1 INTRODUCTION

22.1.1 Market Size

From 1996 to 2007, the market for credit derivatives is estimated to have grown from about $40 billion in gross notional to more than $62,000 billion, all of which is currently traded in over-the-counter (OTC) markets. As a reference, Chapter 7 has shown that the size of the global domestic corporate bond markets is $6,300 billion. Including governments, financials and international bonds add up to a total of $80,000 billion.

As is usual with OTC markets, however, only a fraction of this growth represents a net economic exchange of credit risk. A recent Fitch Ratings survey

estimates that the ratio of gross-to-net exposures is 50, implying net exposures around $1,200 billion.

Gross exposures are very high because dealers had a practice of not cancelling existing trades but instead simply added new ones with offsetting characteristics. This practice, however, creates a backlog of paperwork, increases operational risk as well as counterparty risk.

As a result, the industry is starting to implement portfolio compression, which is a process that reduces the overall size and number of items in credit portfolios, without changing the risk parameters of the portfolio. In 2008 alone, $30,000 billion worth of contracts were cancelled, leading for the first time to a decline in notional amount during the year.

22.1.2 Markets for Exchanges of Risks

Credit derivatives have grown so quickly because they provide an efficient mechanism to exchange credit risk. While modern banking is built on the sensible notion that a portfolio of loans is less risky than single ones, banks still tend to be too concentrated in geographic or industrial sectors. This is because their comparative advantage is "relationship banking," which is usually limited to a clientele that banks know best. So far, it has been difficult to lay off this credit exposure, as there is only a limited market for secondary loans. In addition, borrowers may not like to see their bank selling their loans to another party, even for diversification reasons. Credit derivatives solve this dilemma by allowing banks to keep the loans on their books and to buy protection with credit derivatives.

In fact, credit derivatives are not totally new. Bond insurance is a contract between a bond issuer and a guarantor (a bank or insurer) to provide additional payment should the issuer fail to make full and timely payment. A letter of credit is a guarantee by a bank to provide a payment to a third party should the primary credit fail on its obligations. The call feature in corporate bonds involves an option on the risk-free interest rate as well as the credit spread. Indeed the borrower can also call back the bond should its credit rating improve. At an even more basic level, a long position in a corporate bond, is equivalent to a long position in a risk-free (meaning default-free) bond plus a short position in a credit default swap (CDS).

Thus, many existing instruments embed some form of credit derivative. What is new is the transparency and trading made possible by credit derivatives. Corporate bonds, notably, are difficult to short. This position can be replicated easily, however, by the purchase of a CDS contract. Thus, credit derivatives open new possibilities for investors, hedgers, and speculators.

22.1.3 Types of Credit Derivatives

Credit derivatives are over-the-counter contracts that allow credit risk to be exchanged across counterparties. They can be classified in terms of the following:

- *The underlying credit*, which can be either a single entity (single name) or a group of entities (multiname)

* *The exercise conditions*, which can be a credit event (such as default or a rating downgrade, or an increase in credit spreads)
* *The payoff function*, which can be a fixed amount or a variable amount with a linear or nonlinear payoff

The credit derivatives market includes plain-vanilla credit default swaps, total return swaps, credit spread forwards, and options. These instruments are bilateral OTC contracts. They also appear in credit structured products, which will be discussed later in this chapter. A recent survey breaks down the market into 33% for single-name CDSs, 30% for index CDSs, 16% for synthetic CDOs, 8% for tranche index trades, 3% for credit-linked notes, and 1% for credit spread options.[1] Thus the most common instruments are CDS contracts.

22.2   CREDIT DEFAULT SWAPS

22.2.1 Definition

In a credit default swap contract, a protection buyer (say A) pays a premium to the protection seller (say B), in exchange for payment if a credit event occurs. The premium payment can be a lump sum or periodic. The contingent payment is triggered by a credit event (CE) on the underlying credit, say a bond issued by company Y. The structure of this swap is described in Figure 22.1. Thus, these contracts represent the purest form of credit derivatives, as settlement payment only occurs in "default mode" (DM).

Note that these contracts are really options, not swaps. The main difference from a regular option is that the cost of the option is paid in installments instead of up front. When the premium is paid up front, these contracts are called *default put options*.[2] The annual payment is referred to as the CDS spread.[3]

> ### Example
>
> The protection buyer, call it A, enters a one-year credit default swap on a notional of $100 million worth of 10-year bonds issued by XYZ. The swap entails an annual payment of 50bp. The bond is called the *reference credit asset*.
>
> At the beginning of the year, A pays $500,000 to the protection seller. Say that at the end of the year, company XYZ defaults on this bond, which now trades at 40 cents on the dollar. The counterparty then has to pay $60 million to A. If A holds this bond in its portfolio, the credit default swap provides protection against credit loss due to default.

---

[1] British Bankerss Association (2006), BBA Credit Derivatives Report 2006, London: BBA.

[2] Default swaps and default options are not totally identical instruments, however, because a default swap requires premium payments only until a triggering credit event occurs.

[3] This should not be confused with the bid-ask spread, which is the difference between the buying and selling rate. For instance, the bid rate may be 45bp, and the ask rate 55bp. So, the buyer would pay 0.55% annually to acquire protection. A protection seller would receive 0.45% only.
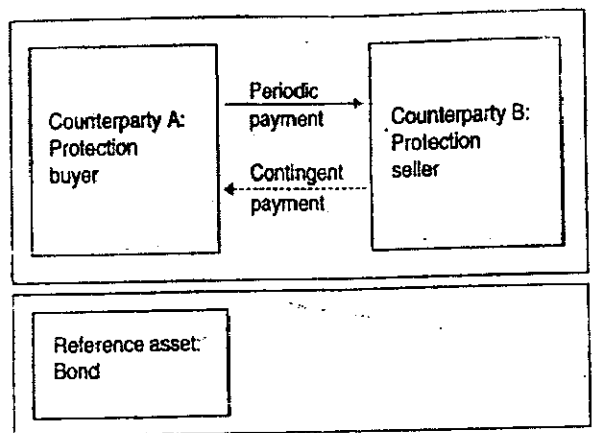
FIGURE 22.1   Credit Default Swap

Most CDS contracts are quoted in terms of an annual spread, with the payment made on a quarterly basis. Distressed names, however, can trade *up front*. For instance, on September 17, 2008, Washington Mutual was quoted at 44 points up front. This means that the buyer of protection on $100 million would have to pay $44 million up front plus the usual spread of 500bp per year. Wamu incurred a credit event on September 27, triggering payments on CDS contracts.

Default swaps are embedded in many financial products: Investing in a risky (credit-sensitive) bond is equivalent to investing in a risk-free bond plus selling a credit default swap. Say, for instance, that the risky bond sells at $90 and promises to pay $100 in one year. The risk-free bond sells at $95. Buying the risky bond is equivalent to buying the risk-free bond at $95 and selling a credit default swap worth $5 now. The up-front cost is the same, $90. If the company defaults, the final payoff will be the same.

KEY CONCEPT

A long position in a defaultable bond is economically equivalent to a long position in a default-free bond plus a short position in a CDS on the same underlying credit.

22.2.2 Settlement

Credit events must be subject to precise definitions. Chapter 19 provided such a list, drawn from the ISDA's master netting agreement. Ideally, there should be no uncertainty about the interpretation of a credit event. Otherwise, credit derivative transactions can create legal risks.

The payment on default reflects the loss to the holders of the reference asset when the credit event occurs. Define Q as the value of this payment per unit of

notional. This can take a number of forms.

- Cash settlement, or a payment equal to the strike minus the prevailing market value of the underlying bond.
- Physical delivery of the defaulted obligation in exchange for a fixed payment.
- A lump sum, or a fixed amount based on some pre-agreed recovery rate. For instance, if the CE occurs, the recovery rate is set at 40%, leading to a payment of 60% of the notional.

Thus, the payoff on a credit default swap is

$$\text{Payment} = \text{Notional} \times Q \times I(\text{CE}) \qquad (22.1)$$

where the indicator function $I(\text{CE})$ is 1 if the credit event has occurred and 0 otherwise.

The swap spread reflects both the probability of default and the loss given default, both of which are unknown. A slight variant on the usual CDS contract is the binary credit default swap, which pays a fixed amount $Q = 1$ if the credit event occurs. The two contracts can be combined to extract a market-implied estimate of the recovery rate.

With physical settlement, the contract usually defines a list of bonds that can be delivered. The bonds can trade at different prices but must all be exchanged for their face value. Naturally, the buyer of protection will select the cheapest bond, which creates a delivery option.

The growth of the CDS markets, however, has led to cases where the outstanding CDS notional far exceeds the available supply of bonds issued by a particular obligor. Cash settlement can be conducted through an auction, which defines the recovery rate. For instance, the final price for CDS contracts tied to Lehman Brothers Holdings was fixed on October 10, 2008 at a recovery rate of 8.625 cents on the dollar.[4] This meant that sellers of credit protection had to pay 91.375% of face value, an unexpectedly high fraction. Given an estimated $400 billion in outstanding CDS contracts, some feared that the settlement process would cause major disruptions to financial markets. In fact, the Depository Trust & Clearing Corporation (DTCC), which processes a large fraction of the trades, reported that only $5.2 billion had to change hands after the auction. This was because netting sharply reduced the gross exposures of $400 billion. Indeed, dividing by the estimated gross-to-net ratio of 50 gives a net exposure of only $8 billion. In addition, sellers of credit protection already had to post collateral. Thus, the CDS market has successfully handled a major default.

22.2.3 Pricing

CDS contracts can be priced by considering the present value of the cash flows on each side of the contract.

---

[4] Other notable auctions were Fannie Mae (91.51 cents for senior debt) and Washington Mutual (57 cents).

Define PV_t as the present value of a dollar paid at time t. For simplicity, assume that default occurs at the end of the year. As seen in Chapter 19, the marginal default rate from now to year t is $k_t = S_{t-1}d_t$, where S_t is the survival probability until the end of year t and d_t is the marginal probability of defaulting in year t. The survival probability is linked to the cumulative default probability $C_t = k_1 + k_2 + \cdots + k_t = 1 - S_t$.

Table 22.1 describes the annual default probabilities in the left panel. These represent typical market quotes for a credit initially rated BBB. The market-implied five-year cumulative default rate is 15.43%. Using $C_T = 1 - (1 - d)^T$, this gives an annual average default rate d of 3.30%. The second panel gives the discount factor assuming a risk-free interest rate of 6%.

Let us examine first the payoff payments. Upon default, the protection buyer receives the face value minus the recovery rate f, here assumed to be 40%. This occurs with probability k_t every year. The third panel in Table 22.1 illustrates the computations. On a notional of \$100, the PV of the expected payment in the first year is $k_1 (1 - f) \times \$100 \times PV_1 = 2.640\%(1 - 0.40) \times \$100 \times 0.9434 = 1.494$. Adding across the five years of the contract gives \$7.733.

In exchange, the protection buyer must make annual payments tied to a spread of s, defined in percent. In case of default, payments have to be made in arrears until the time of default then stops. For the first year, the PV of the expected payment is $s\$100 \times S_0 \times PV_1 = (s/100)\$100 \times 1.000 \times 0.9434 = s0.943$. Here, $S_0 = 1.000$ because we are sure to make this first payment given that default happens at the end of the year. For the second year, this is $s\$100 \times S_1 \times PV_2 = (s/100)\$100 \times 0.9736 \times 0.8900 = s0.867$. Summing across the five years gives $s3.986$.

The fair value of the spread is the number that sets the initial value of the CDS contract to zero. This solves

$$V = (\text{PV Payoff}) - s(\text{PV Spread}) = \left(\sum_{t=1}^{T} k_t(1 - f)PV_t\right) - s\left(\sum_{t=1}^{T} S_{t-1}PV_t\right) \tag{22.2}$$

In this case, the fair CDS spread solves $0 = 7.733 - s3.986$, which gives $s = 1.94\%$ because the spread was defined in percent. Note that this is very close to

TABLE 22.1 Payoffs on a Credit Default Swap

	Probability (%)				Discount Factor PV_t	Payoff Payments		Spread Payments	
Year t	Cumul. C_t	Annual d_t	Marg. k_t	Survival S_t		Expected $k_t(1 - f)$	PV	Expected sS_{t-1}	PV
1	2.64	2.640	2.640	0.9736	0.9434	1.584	1.494	$s1.000$	$s0.943$
2	5.48	2.917	2.840	0.9452	0.8900	1.704	1.517	$s0.974$	$s0.867$
3	8.57	3.269	3.090	0.9143	0.8396	1.854	1.557	$s0.945$	$s0.794$
4	11.89	3.631	3.320	0.8811	0.7921	1.992	1.578	$s0.914$	$s0.724$
5	15.43	4.018	3.540	0.8457	0.7473	2.124	1.587	$s0.881$	$s0.658$
Total		15.430			4.2124		7.733		$s3.986$

an approximation based on the annual average default rate times the loss given default, which is $3.30\% \times (1 - 0.40) = 1.98\%$.

Equation (22.2) can also be used to price an outstanding CDS contract. Assume for instance that the contract was entered with a spread of 1.50% and that the probabilities and interest rates in Table 22.1 represent current market conditions. The value of the CDS is then $V = 7.733 - 1.50 \times 3.986 = \$1.753$. This is a profit to the CDS buyer because the current spread is now greater than the locked-in value.

Also, Equation (22.2) can also be used to compute the spread duration of the contract. As an approximation, we can shock the market spreads $k_i(1 - f)$ upward by 1bp, which gives a gain of $\sum PV_t = 4.12\text{bp}$. This represents a (negative) spread duration of 4.12 years, slightly less than the maturity of the contract.

Note that the default probabilities used to price the CDS contract must be *risk-neutral* (RN) probabilities, not real-world probabilities. These RN probabilities π can be inferred from bond prices and CDS prices. For instance, assume that we observe a five-year CDS spread quote of 1.50%. Using the simplified approach, this gives $\pi = 1.50\%/(1 - 0.40) = 2.50\%$. For more precision, we could reverse the process in Table 22.1, using market quotes for one-, two-, three-, four-, and five-year CDS contracts to derive RN default probabilities for all the maturities.

Abstracting from counterparty risk, the CDS spread should be approximately equal to the difference between the yield on a corporate bond issued by the same obligor and the risk-free yield for the same maturity. If the CDS spread were markedly lower than this difference, an investor could make an arbitrage profit by buying the corporate bond, hedging pure interest-rate risk by shorting a Treasury bond, and buying the CDS contract.

KEY CONCEPT

The CDS swap spread should approximately equal the yield on a corporate bond issued by the same obligor minus the risk-free yield.

In general, however, the basis between the CDS spread and the cash yield spread is slightly positive. To some extent, this is influenced by demand and supply considerations, including the availability of arbitrage capital. Because investors can only short credit by buying CDS contracts, this pushes up the spread. Otherwise, all else equal, the basis should be wider because the delivery option, which makes buying credit protection more attractive because the buyer is long the delivery option. In contrast, counterparty credit risk should decrease the basis because there is a risk the payoff may not be made if the credit event is triggered.

22.2.4 Counterparty Risk

It is important to realize that entering a credit swap does not eliminate credit risk entirely. Instead, the protection buyer decreases exposure to the reference credit

TABLE 22.2 CDS Spreads for Different Counterparties
(Reference Obligation is Five-Year Bond Rated BBB)

| | Counterparty Credit Rating | | | |
Correlation	AAA	AA	A	BBB
0.0	194	194	194	194
0.2	191	190	189	186
0.4	187	185	181	175
0.6	182	178	171	159
0.8	177	171	157	134

Source: Adapted from Hull, J. and A. White (2001), Valuing
Credit Default Swaps II: Modeling Default Correlations. *Journal of Derivatives* 8, 12–21.

Y but assumes new credit exposure to the CDS seller. Protection will be effective
with a low correlation between the default risk of the underlying credit and of the
counterparty. Just to be sure, the contracts may involve the posting of collateral
from the protection seller.

Like options, these instruments are unfunded, meaning that each party is
responsible for making payments (i.e., premiums and settlement amount) without
recourse to other assets. In contrast, in a funded instrument, the protection seller
makes a payment that could be used to settle any potential credit event. In the
latter case, the protection buyer is not exposed to counterparty risk.

Table 22.2 illustrates the effect of the counterparty for the pricing of the CDS.
If the counterparty is default free, the CDS spread on this BBB credit should be
194bp. The spread depends on the default risk for the counterparty as well as the
correlation with the reference credit. In the worst case in the table, with a BBB
rating for the counterparty and correlation of 0.8, protection is less effective, and
the CDS spread is only 134bp.

EXAMPLE 22.1: FRM EXAM 2004—QUESTION 9

If an investor holds a five-year IBM bond, it will give him a return very close
to the return of the following position:

a. A five-year IBM credit default swap on which he pays fixed and receives
   a payment in the event of default
b. A five-year IBM credit default swap on which he receives fixed and
   makes a payment in the event of default
c. A five-year U.S. Treasury bond plus a five-year IBM credit default swap
   on which he pays fixed and receives a payment in the event of default
d. A five-year U.S. Treasury bond plus a five-year IBM credit default swap
   on which he receives fixed and makes a payment in the event of default

EXAMPLE 22.2: FRM EXAM 2007—QUESTION 18

Suppose the return on U.S. Treasuries is 3% and a risky bond is currently yielding 15%. A trader you supervise claims to be able to make an arbitrage trade earning 5% using U.S. Treasuries, the risky bond, and the credit default swap. Which of the following could be the trader's strategy, and what is the CDS premium?

a. Go long the Treasury, short the risky bond, and sell the credit default swap with a premium of 7%.
b. Go long the Treasury, short the risky bond, and buy the credit default swap with a premium of 6%.
c. Short the Treasury, invest in the risky bond, and buy the credit default swap with a premium of 7%.
d. Short the Treasury, invest in the risky bond, and sell the credit default swap with a premium of 6%.

EXAMPLE 22.3: FRM EXAM 2007—QUESTION 120

Bank A makes a USD 10 million five-year loan and wants to offset the credit exposure to the obligor. A five-year credit default swap (CDS) with the loan as the reference asset trades on the market at a swap premium of 50 basis points paid quarterly. In order to hedge its credit exposure, bank A

a. Sells the five-year CDS and receives a quarterly payment of USD 50,000.
b. Buys the five-year CDS and makes a quarterly payment of USD 12,500.
c. Buys the five-year CDS and receives a quarterly payment of USD 12,500.
d. Sells the five-year CDS and makes a quarterly payment of USD 50,000.

EXAMPLE 22.4: FRM EXAM 2004—QUESTION 50

The table below shows the bid-ask quotes by UBS for CDS spreads for companies A, B, and C. CSFB has excessive credit exposure to company C and wants to reduce it through the CDS market.

	1 Year	3 Year	5 Year
A	15/25	21/32	27/36
B	43/60	72/101	112/152
C	71/84	93/113	141/170

Since the furthest maturity of its exposure to C is three years, CSFB buys a USD 200 million three-year protection on C from UBS. In order to make its purchase of this protection cheaper, based on its views on companies A and B, CSFB decides to sell USD 300 million five-year protection on company A and to sell USD 100 million one-year protection on company B to UBS. What is the net annual premium payment made by CSFB to UBS in the first year?

a. USD 1.02 million
b. USD 0.18 million
c. USD 0.58 million
d. USD 0.62 million

EXAMPLE 22.5: FRM EXAM 1999—QUESTION 135

The Widget Company has outstanding debt of three different maturities as outlined in the table.

Maturity	Widget Co. Bonds Price	Coupon	U.S. T-Bonds Price	Coupon
one year	100	7.00%	100	6.00%
five years	100	8.50%	100	6.50%
10 years	100	9.50%	100	7.00%

All Widget Company debt ranks pari passu and contains cross default provisions. The recovery value for each bond is 20%. The correct price for a one-year CDS with the Widget Company, 9.5% 10-year bond as a reference asset is

a. 1.0% per annum
b. 2.0% per annum
c. 2.5% per annum
d. 3.5% per annum

EXAMPLE 22.6: FRM EXAM 2004—QUESTION 65

When an institution has sold exposure to another institution (i.e., purchased protection) in a CDS, it has exchanged the risk of default on the underlying asset for which of the following?

a. Default risk of the counterparty
b. Default risk of a credit exposure identified by the counterparty
c. Joint risk of default by the counterparty and of the credit exposure identified by the counterparty
d. Joint risk of default by the counterparty and the underlying asset

EXAMPLE 22.7: FRM EXAM 2007—QUESTION 85

Bank A has exposure to USD 100 million of debt issued by company R. Bank A enters into a credit default swap transaction with bank B to hedge its debt exposure to company R. Bank B would fully compensate bank A if company R defaults in exchange for a premium. Assume that the defaults of bank A, bank B, and company R are independent and that their default probabilities are 0.3%, 0.5%, and 3.6%, respectively. What is the probability that bank A will suffer a credit loss in its exposure to company R?

a. 4.1%
b. 3.6%
c. 0.0108%
d. 0.0180%

EXAMPLE 22.8: FRM EXAM 2005—QUESTION 111

You enter into a credit default swap with bank B that settles based upon the performance of company C. Assuming that bank B and company C have the same initial credit rating and everything else remains the same, what is the impact on the value of your credit default swap if bank B buys company C?

a. The credit default swap value increases.
b. The credit default swap value remains the same.
c. The credit default swap value decreases.
d. Impossible to determine based on the information provided.

22.3 OTHER CONTRACTS

22.3.1 CDS Variants

Credit default swaps can also be written on multiple names. For instance, the first-of-basket-to-default swap gives the protection buyer the right to deliver *one and only one* defaulted security out of a basket of selected securities. Because the protection buyer has more choices—that is, default can occur across a basket instead of just one reference credit—this type of protection will be more expensive than a single credit swap, all else kept equal. The price of protection also depends on the correlation between credit events. The lower the correlation, the more expensive the swap. Conversely, the higher the correlation, the lower the swap rate. To illustrate this point, consider the extreme case of perfect correlation. In such case, all underlying credits default at the same time, and this basket swap is equivalent to a regular single-name CDS.

With an *N*th-to-default swap, payment is triggered after N defaults in the underlying portfolio, but not before. When N is large, the cost of protection will be high when the default correlation is high, making it more likely that N names will default simultaneously.

CDS indices are widely used to track the performance of this market. The iTraxx indices cover the most liquid names in European and Asian credit markets. The North American and emerging markets are covered by the CDX indices. For example, the CDX.NA.IG index is composed of 125 investment-grade entities domiciled in North America. The CDX.NA.HY index covers 100 non-investment-grade (high yield) borrowers. The CDX.EM index covers borrowers from emerging markets. The indices are rebalanced every six months. Because these contracts are very liquid, and trade at tight bid–ask spreads, they provide an easy way to buy and sell marketwide or sectoral credit risk. These indices also have tradable tranches, using the CDO methodology described later in this chapter.

CDS indices trade at tight spreads, even more narrow than for single-name contracts. Assume for instance that a dealer quotes 201/203 for a five-year CDX.NA.IG. A trader wants $80,000 of protection on each of the 125 companies in the index, which add up to a notional of $10 million. The total cost is $0.0203 \times \$10,000,000 = \$203,000$.[5] If one company defaults, the buyer receives the usual CDS payment, and the notional is then reduced by $80,000.

22.3.2 Total Return Swaps

A total return swap (TRS) is a contract where one party, called the protection buyer, makes a series of payments linked to the total return on a reference asset. In exchange, the protection seller makes a series of payments tied to a reference rate, such as the yield on an equivalent Treasury issue (or LIBOR) plus a spread. If the price of the asset goes down, the protection buyer receives a payment from the counterparty; if the price goes up, a payment is due in the other direction. The structure of this swap is described in Figure 22.2.

---

[5] In practice, the pattern of payment is more complicated, with a fixed coupon and an initial price that depends on the quoted spread.
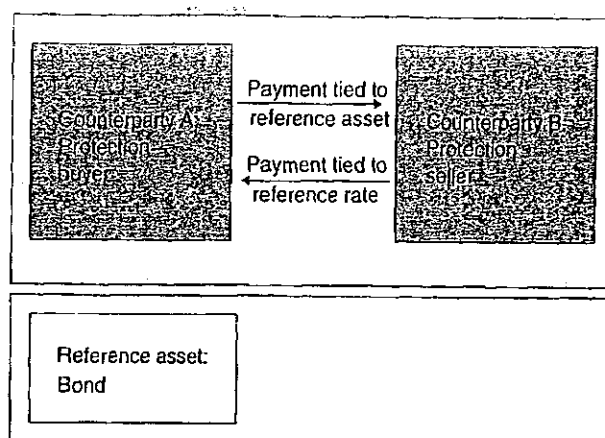
FIGURE 22.2 Total Return Swap

This type of swap is tied to changes in the market value of the underlying asset and provides protection against credit risk in a marked-to-market (MTM) framework. For the protection buyer, the TRS removes all the economic risk of the underlying asset without selling it. Unlike a CDS, the TRS involves both credit risk and market risk, the latter reflecting pure interest rate risk.

Example

Suppose that a bank, call it bank A, has made a $100 million loan to company XYZ at a fixed rate of 10%. The bank can hedge its exposure by entering a TRS with counterparty B, whereby it promises to pay the interest on the loan plus the change in the market value of the loan in exchange for LIBOR plus 50bp. If the market value of the loan decreases, the payment tied to the reference asset will become negative, providing a hedge for the bank.

Say that LIBOR is currently at 9% and that after one year, the value of the loan drops from $100 to $95 million. The *net* obligation from bank A is the sum of

- Outflow of 10% × $100 = $10 million, for the loan's interest payment
- Inflow of 9.5% × $100 = $9.5 million, for the reference payment
- Outflow of $\frac{(95-100)}{100}$% × $100 = −$5 million, for the movement in the loan's value

This sums to a net receipt of −10 + 9.5 − (−5) = $4.5 million. Bank A has been able to offset the change in the economic value of this loan by a gain on the TRS.

22.3.3 Credit Spread Forward and Options

These instruments are derivatives whose value is tied to an underlying credit spread between a risky and risk-free bond.

In a credit spread forward contract, the buyer receives the difference between the credit spread at maturity and an agreed-upon spread, if positive. Conversely,

a payment is made if the difference is negative. An example of the formula for the cash payment is

$$\text{Payment} = (S - F) \times \text{MD} \times \text{Notional} \tag{22.3}$$

where MD is the modified duration, S is the prevailing spread, and F is the agreed-upon spread. Alternatively, this could be expressed in terms of prices:

$$\text{Payment} = [P(y + F, \tau) - P(y + S, \tau)] \times \text{Notional} \tag{22.4}$$

where y is the yield to maturity of an equivalent Treasury, and $P(y + S, \tau)$ is the present value of the security with τ years to expiration, discounted at y plus a spread. Note that if $S > F$, the payment will be positive as in the previous expression.

In a credit spread option contract, the buyer pays a premium in exchange for the right to "put" any increase in the spread to the option seller at a predefined maturity:

$$\text{Payment} = \text{Max}(S - K, 0) \times \text{MD} \times \text{Notional} \tag{22.5}$$

where K is the predefined spread. The purchaser of the option buys credit protection, or the right to put the bond to the seller if it falls in value. The payout formula could also be expressed directly in terms of prices, as in Equation (22.4).

Example

A credit spread option has a notional of $100 million with a maturity of one year. The underlying security is an 8% 10-year bond issued by corporation XYZ. The current spread is 150bp against 10-year Treasuries. The option is European type with a strike of 160bp.

Assume that, at expiration, Treasury yields have moved from 6.5% to 6% and the credit spread has widened to 180bp. The price of an 8% coupon, nine-year semiannual bond discounted at $y + S = 6 + 1.8 = 7.8\%$ is $101.276. The price of the same bond discounted at $y + K = 6 + 1.6 = 7.6\%$ is $102.574. Using the notional amount, the payout is $(102.574 - 101.276)/100 \times \$100,000,000 = \$1,297,237$.

EXAMPLE 22.9: FRM EXAM 2005—QUESTION 14

Sylvia, a portfolio manager, established a Yankee bond portfolio. However, she wants to hedge the credit and interest rate risk of her portfolio. Which of the following derivatives will best fit Sylvia's need?

a. A total return swap
b. A credit default swap
c. A credit-spread option
d. A currency swap

22.4

EXAMPLE 22.10: FRM EXAM 2007—QUESTION 69

A bank holds USD 60 million of 10-year 6.5% coupon bonds that are trading at a clean price of USD 101.82. The bank is worried by the exposure due to these bonds but cannot unwind the position for fear of upsetting the client. Therefore, it purchases a total return swap (TRS) in which it receives annual LIBOR+100 bp in return for the marked-to-market return on the bond. For the first year, LIBOR sets at 6.25%, and by the end of the year the clean price of the bonds is at USD 99.35. The net receipt/payment for the bank in the total return swap will be to

a. Receive USD 1.93 million
b. Receive USD 2.23 million
c. Pay USD 2.23 million
d. Pay USD 1.93 million

EXAMPLE 22.11: FRM EXAM 2000—QUESTION 61

(Complex—use the valuation formula with prices) A credit-spread option has a notional amount of $50 million with a maturity of one year. The underlying security is a 10-year, semiannual bond with a 7% coupon and a $1,000 face value. The current spread is 120bp against 10-year Treasuries. The option is a European option with a strike of 130bp. If at expiration, Treasury yields have moved from 6% to 6.3% and the credit-spread has widened to 150bp, what will be the payout to the buyer of this credit-spread option?

a. $587,352
b. $611,893
c. $622,426
d. $639,023

22.4 STRUCTURED PRODUCTS

22.4.1 Creating Structured Products

Structured products generally can be defined as instruments created to meet specific needs of investors or borrowers that cannot be met with conventional financial instruments.

A typical example is retail demand for investments that participate in the appreciation of stock markets but also preserve capital. The payoff profile of the

product can be replicated from a combination of existing, or sometimes new instruments. In this case, for example, the payoff can be replicated by an investment in a risk-free bond with notional equal to the guaranteed capital, plus long positions in a call option, either through direct investment in a portfolio of options or indirectly replicated through dynamic trading. This instrument is a principal-protected note, and can be indexed to a variety of markets, including equities, currencies, and commodities.

In recent years, the market for credit structured products has expanded enormously. The advent of credit derivatives has made possible a flurry of innovative products where payoffs are linked to credit events.

22.4.2 Credit-Linked Notes

Credit-linked notes (CLN) are structured securities that combine a credit derivative with a regular bond. In a CLN, the buyer of protection transfers credit risk to an investor via an intermediary bond-issuing entity. This entity can be the buyer itself or a special purpose vehicle.

An example of the first case is a bank with exposure to an emerging country, say Mexico. The bank issues a note with an embedded short position in a credit default swap on Mexico. The note is a liability on the bank's balance sheet. Investors receive a high coupon but will lose some of the principal if Mexico defaults on its debt. This structure achieves its goal of reducing the bank's exposure if Mexico defaults. In this case, because the note is a liability of the bank, the investor is exposed to a default of either Mexico or of the bank.

An example of the SPV structure is provided in Figure 22.3. In this case, the investor's initial funds are placed in a top-rated investment that pays LIBOR plus a spread of Y bp. The SPV takes a short position in a credit default swap, for an additional annual receipt of X bp. The annual payment to the investor is then $LIBOR + Y + X$. In return for this higher yield, the investor must be willing to lose some of the principal should a default event occur.

Relative to a regular investment in, say, a note issued by the government of Mexico, this structure may carry a higher yield if the CDS spread is greater than the bond yield spread. This structure may also be attractive to investors who are precluded from investing directly in derivatives.
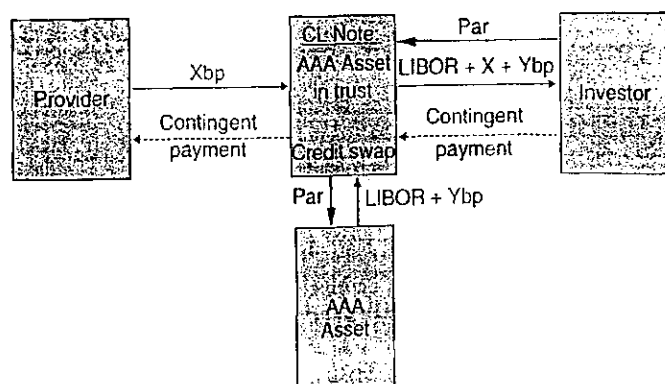
FIGURE 22.3 Credit-Linked Note

22.4.3 Collateralized Debt Obligations

Much of financial engineering is about repackaging financial instruments to make them more palatable to investors, creating value in the process. In the 1980s, collateralized mortgage obligations (CMOs) brought mortgage-backed securities to the masses by repackaging their cash flows into tranches with different characteristics.

The same magic is performed with collateralized debt obligations (CDOs), which are securities backed by a pool of debt. Collateralized bond obligations (CBOs) and collateralized loan obligations (CLOs) are backed by bonds and loans, respectively. Figure 22.4 illustrates a typical CDO structure.[6]

The first step is to place a package of corporate bonds in a special-purpose vehicle (SPV). Assume that we have a total of $1,000 million, representing exposures of $10 million to 100 entities. Multiple tranches are then issued by the SPV, with a specified waterfall structure, or priority of payments to the various tranches. Tranches are categorized as senior, mezzanine, and subordinated or equity. In the simplest structure, the SPV is ideally a passive entity. It redistributes cash flows according to well-defined rules. There is no need for other management action.

In this example, 80% of the capital structure is apportioned to tranche A, which has the highest credit rating of Aaa, using Moody's rating, or AAA. It pays LIBOR + 45bp, for example. Other tranches have lower priority and rating. These intermediate, *mezzanine*, tranches are typically rated A, Baa, Ba, or B (A, BBB, BB, B, using S&P's ratings). For instance, tranche C would absorb losses from 3% to 10%. These numbers are called, respectively, the attachment point and the detachment point.

At the bottom comes the equity tranche, which is not rated. Due to leverage, the return on the equity tranche can be very high if there is no default. In
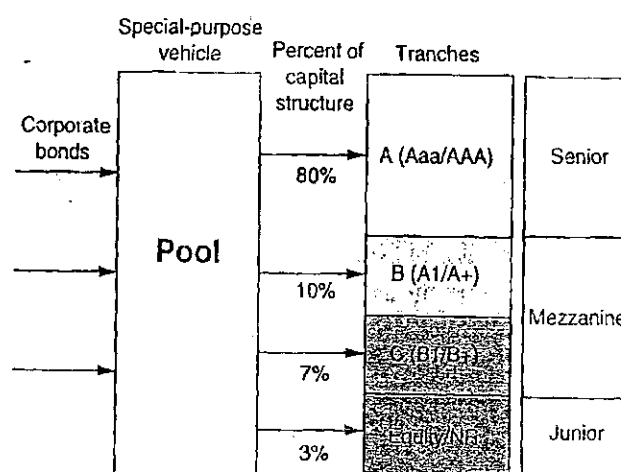
FIGURE 22.4 Collateralized Debt Obligation Structure

---

[6] This structure has similarities with the CLN structure. The differences are that CDOs are always issued by a SPV, involve a pool with a large number of underlying assets, and are usually tranched.

exchange, this is exposed to the first dollar loss in the portfolio. Special conventions apply to trading in equity tranches. The investor first pays the notional amount, which is $30 million in this case. In exchange, this protection seller receives a spread, called running spread, and an up-front fee. This fee is quoted in percent and is typically around 40% for an investment-grade, CDO. In this case, the investor would get 40% × $30 = $12 million up front, and the running spread, say 500bp.

Cumulative losses of $20 million would reduce the notional of the equity tranche to $30 − $20 = $10 million. The running spread then applies to the new notional of $10 million.

For losses amounting to $45 million, the first tranche is wiped out, and investors in tranche C receive only $70 − $15 = $55 million back. Thus, the rating enhancement for the senior classes is achieved through prioritizing the cash flows. Rating agencies have developed internal models to rate the senior tranches based on the probability of shortfalls due to defaults.

Whatever transformation is brought about, the resulting package must obey some basic laws of conservation. For the underlying and resulting securities, we must have the same cash flows at each point in time, apart from transaction costs. As a result, this implies (1) the same total market value, and (2) the same risk profile, both for interest rate and default risk. The weighted duration of the final package must equal that of the underlying securities. The expected default rate, averaged by market values, must be the same. So, if some tranches are less risky, others must bear more risk. Like CMOs, CDOs are structured so that most of the tranches have less risk than the collateral. Inevitably, the remaining residual tranche is more risky. This is sometimes called "toxic waste." If this residual is cheap enough, however, some investors should be willing to buy it. Oftentimes, the institution sponsoring the CDO will retain the most subordinate equity tranche to convince investors of the quality of the pool. Credit investors have developed sophisticated trading strategies that involve going long and short different tranches of these capital structures.

Example: Correlation Trading

Take a synthetic $1 billion CDO with 100 names worth $10 million each. Say that the equity tranche is $30 million, which represents the first 3% of losses. For simplicity, suppose that all premiums are measured in net present value terms and that there is no recovery. The investor gets paid $15 million up front for assuming the equity risk, so his worst net loss on the tranche is $15 million. He then hedges by buying CDS on the same 100 names, with notional of $3 million each. The present value of the spread is 2%, which gives a payment of $300 × 2% = $6 million. If there is no default, the principal is returned, and the net gain is $15 − $6 = $9 million. If all 100 names default, the position loses the principal of the equity and gains the CDS payments, which gives ($15 − $30) − $6 + $300 = $279 million. Of course, this is very unlikely.

We need to explore other scenarios that could generate losses. If only three names default, the equity tranche is wiped out. The investor then exercises three CDS contracts and unwinds the 97 remaining CDS hedges, which are no longer necessary. As a worst-case situation, suppose the CDS spreads have tightened and that the contracts are sold for $4.8 million. This translates into a net loss of $(\$15 - \$30) - \$6 + (3 \times \$3) + \$4.8 = -\$7.2$ million.

EXAMPLE 22.12: FRM EXAM 2004—QUESTION 63

A CDO, consisting of three tranches, has an underlying portfolio of n corporate bonds with a total principal of USD N million. Tranche 1 has 10% of N and absorbs the first 10% of the default losses. Tranche 2 has 20% of N and absorbs the next 20% of default losses. The final tranche 3 has 70% of N and absorbs the residual default loss. Which of the following statements are true?

I. Tranche 2 has the highest yield.
II. Tranche 1 is usually called "toxic waste."
III. Tranche 3 would typically be rated as AAA by S&P.
IV. Tranche 3 has the lowest yield.
   a. I only
   b. IV only
   c. II, III, and IV only
   d. II and IV only

EXAMPLE 22.13: FRM EXAM 2001—QUESTION 12

A pool of high yield bonds is placed in a SPV and three tranches (including the equity tranche) of bonds are issued collateralized by the bonds to create a Collateralized Bond Obligation (CBO). Which of the following is true?

   a. At fair value the value of the issued bonds should be less than the collateral.
   b. At fair value the total default probability, weighted by size of issue, of the issued bonds should equal the default probability of the collateral pool.
   c. The equity tranche of the CBO has the least risk of default.
   d. The yield on the low risk tranche must be greater than the yield on the collateral pool.

EXAMPLE 22.14: FRM EXAM 2002—QUESTION 32

A CBO (Collateralized Bond Obligation) consists of several tranches of notes from a repackaging of corporate bonds, ranging from equity to super senior. Which of the following is generally true of these structures?

    a. The total yield of all the CBO tranches is slightly less than the underlying repackaged bonds to allow the issuer to recover their fees/costs/profits.

    b. The super senior tranche has expected loss rate higher than the junior tranche.

    c. The super senior tranche is typically rated below AAA and sold to bond investors.

    d. The equity tranche does not absorb the first losses of the structure.

EXAMPLE 22.15: FRM EXAM 2007—QUESTION 130

A three-year, credit-linked note (CLN) with underlying company Z has a LIBOR + 60bps semi-annual coupon. The face value of the CLN is USD 100. LIBOR is 5% for all maturities. The current three-year CDS spread for company Z is 90bps. The fair value of the CLN is closest to

    a. USD 100.00
    b. USD 111.05
    c. USD 101.65
    d. USD 99.19

22.5 CDO MARKET

22.5.1 Balance Sheet and Arbitrage CDOs

Table 22.3 describes the evolution of the CDO market from 2004 to 2008. From 2004 to 2006, the market doubled every year, reaching $521 billion in new issues during 2006. As a result of the credit crisis, however, issuances have spiraled down.

CDO transactions are typically classified by purpose, as balance sheet or arbitrage. The primary goal of balance sheet CDOs is to move loans off the balance sheet of commercial banks to lower regulatory capital requirements.

In contrast, arbitrage CDOs are designed to capture the spread between the portfolio of underlying securities and that of highly rated, overlying, tranches. Because CDO senior tranches should be relatively safe due to diversification effects,

TABLE 22.8 Evolution of the CDO Market Annual Issues ($ Millions)

Explanation	2004	2006	2008
By Type:			
Cash flow	119,531	410,504	40,321
Synthetic	37,237	66,503	1,191
Market value	650	43,638	14,581
By Purpose:			
Arbitrage	146,998	454,971	45,124
Balance sheet	10,420	65,674	10,967
Total	157,418	520,645	56,093

Source: Bond Market Association

they pay a tight spread over LIBOR. The arbitrage profit then goes into the equity tranche (but also into management and investment banking fees).

Senior tranches also seem attractive for investors. Generally, AA-rated corporate borrowers pay LIBOR. Higher credit pay rates below LIBOR. A typical AAA-rated senior CDO tranche, however, pays a higher rate than LIBOR. This explains why investors were attracted to this market, putting blind faith into the credit ratings. Some commentators wondered "whether this is a true arbitrage or involves some type of model risk." By now, it is clear that the models used by the credit rating agencies were flawed (see also Chapter 19).

22.5.2 Cash Flow and Synthetic CDOs

Credit risk transfer can be achieved by cash flow or synthetic structures. The example in Figure 22.4 is typical of traditional, or *funded*, cash flow CDOs. The physical assets are sold to a SPV and the underlying cash flows used to back payments to the issued notes.

In contrast, the credit risk exposure of synthetic CDOs is achieved with credit default swaps. We know that a long position in a defaultable bond is equivalent to a long position in a default-free bond plus a short position in a CDS. Synthetic CDOs create higher yields by first, funding or placing the initial investment in default-free, or Treasury, securities, and second, selling a group of CDSs to replicate a cash flow CDO.

Synthetic CDOs offer several advantages. First, they are easier to manage than cash flow CDOs. In case of bankruptcy of one of the underlying credits, the management of a cash flow SPV has to take part in the bankruptcy process. With a short CDS position that is cash settled, there is no need for the SPV to get involved with the bankruptcy process. Second, the issue does not need to be fully funded. In a full capital structure CDO, the total notional amount of notes issued is equal to the total notional amount of the underlying portfolio. So, it is fully funded. In contrast, a single tranche CDO is a bespoke transaction where the bank and the investor agree on the terms of a deal including size, credit rating, and underlying

credits.[7] Effectively, the bank holds the rest of the capital structure and does not place it.

Take as an example a $10 million tranche rated A+ paying a coupon of six-month LIBOR plus 111bp, on a reference portfolio of $1,000 million of 100 North American investment-grade entities. The attachment point is 5%. The detachment point is 6%. The investor will receive the promised payments as long as cumulative losses in the reference portfolio remain below 5%. Above that, the investor will have to take a loss. For example, if cumulative dollar losses on the portfolio amount to $52 million, or 5.2% of total, the investor will lose $(5.2 - 5.0)/(6.0 - 5.0) = 20\%$ of his capital, which is $2 million in this case. The coupon then applies to the reduced notional amount. A bank can sell such structure to an investor without fully funding it.

22.5.3 Cash Flow and Market Value CDOs

In the case of cash flow CDOs, payments to investors solely come from collateral cash flows. In contrast, with market value CDOs, payments are made from collateral cash flows as well as sales of collateral. If the market value of the collateral falls below some level, payments to the equity tranche are suspended. This creates more flexibility for the portfolio manager.

Credit rating agencies analyze the quality of credit structures using overcollateralization ratios (OC). This ratio measures how many times the collateral can cover the SPV liabilities. For a market value CDO, define V as the market value of assets and D as the par value of liabilities. The OC ratio is then defined as:

$$OC = \frac{V}{D} \tag{22.6}$$

This must be high enough to ensure sufficient coverage of liabilities.

Alternatively, the par value of cumulative tranches, starting from the top, must be kept below the market-value of assets times an advance rate. In our CDO example, the notional for the first tranche is $800 million. This must be kept below the value of assets, say $1,000 million, times an advance rate of 85%. Because $800 < $850, the first tranche in the structure passes the test. For the first two tranches, the advance rate is 95%, and the test gives $800 + $100 < $950, so this passes the test as well. A structure that fails the overcolleralization test risks downgrading. Such failure can be cured by selling some of the assets and repaying some of the tranches or issuing more equity.

For a cash flow CDO, the ratio uses the par value of total assets in the numerator. Another ratio, the interest rate coverage ratio (IC), is also used to assess the quality of a credit structure. This is computed as the total interest payment to be received by the collateral divided by the interest liability of each tranche and above.

---

[7] The term "bespoke" was originally used to describe clothing made to customer's specification. The term comes from the word *bespeak*, meaning to ask for or order something.

22.5.4 Static and Managed CDOs

Finally, CDOs differ in the management of the asset pool. In static CDOs, the asset pool is basically fixed. In contrast, with managed CDOs, a portfolio manager is allowed to trade actively the underlying assets.

This has all the usual benefits and disadvantages of active management. One benefit is the ability to unwind assets with decreasing credit quality, to buy undervalued securities, and to sell overvalued securities. With managed CDOs, investors face credit risk as well as poor management risk, however. In addition, they pay management fees.

22.5.5 Other Products

As the market for CDOs has expanded, new products have appeared. For instance, a CDO can invest in CDO tranches instead of individual credits. This is a CDO-squared structure. The main benefit of this structure is the greater degree of diversity. A typical single-layer CDO references 50 to 100 corporate credits. A CDO-squared has five to 10 one-layer CDOs, and is thus exposed to 250 to 1,000 names. There was even talk of a further structure, called CDO-cubed.

The market now also trades credit default swaps on asset-backed securities (ABS) tranches, called ABCDS. Most commonly, the assets are backed by home equity and commercial property loans. In the past, it was difficult to short such ABSs. Buying an ABCDS is equivalent to acquiring protection, or shorting the security. This opens up new possibilities to implement relative value trades or to hedge this type of risk. The market has developed rapidly thanks to standardized ISDA documentation as well as the establishment of a benchmark index, the ABX index, which contains 20 home equity securities.

ABCDS are complex instruments. A corporate CDS makes a payment if the underlying company suffers a credit event. In contrast, with an ABS, the issuing SPV cannot go bankrupt but defaults can occur for individual loans in the pool. Also, the notional amount is not fixed but amortizes over time as principal is paid back on the loans.

These ABCDSs are similar to CDSs on CDOs, which are credit default swaps on CDO tranches, usually the senior ones. These instruments provide an efficient way to short sell the market. Dealers who are arranging cash CDOs can buy the CDSs to hedge their exposure, for example.

Another recent innovation is the constant proportional debt obligation (CPDO). CPDOs offer protection on a portfolio of corporate credits such as the iTraxx European CDS index. The transaction is highly levered and dynamically adjusted, getting rid of the credits that deteriorate over time and changing the leverage as spreads vary. This creates a structure that is rated AAA yet pays LIBOR plus 200bp. These new instruments looked very attractive in a benign environment of stable or falling credit spreads. For risk managers, however, their risk profile is difficult to assess due to their dynamic nature. The credit crisis caused considerable losses to many CPDOs, some of which have already defaulted.

EXAMPLE 22.16: FRM EXAM 2003—QUESTION 7

A standard synthetic CDO references a portfolio of 10 corporate names. Assume the following. The total reference notional is X, and the term is Y years. The reference notional per individual reference credit name is $X/10$. The default correlations between the individual credit names are all equal to one. The single-name CDS spread for each individual name is 100 bp, for a term of Y years. The assumed recovery rate on default for all individual reference credits is zero in all cases. The synthetic CDO comprises two tranches, a 50% junior tranche priced at a spread J, and a 50% senior tranche priced at spread S. All else constant, if the default correlations between the individual reference credit names are reduced from 1.0 to 0.7, what is the effect on the relationship between the junior tranche spread J and the senior tranche spread S?

a. The relationship remains the same.
b. S increases relative to J.
c. J increases relative to S.
d. The effect cannot be determined given the data supplied.

EXAMPLE 22.17: FRM EXAM 2007—QUESTION 81

A bank is considering buying (i.e., selling protection on) an AAA-rated super senior tranche [10% − 11%] of a synthetic collateralized debt obligation (CDO) referencing an investment-grade portfolio. The pricing of the tranche assumes a fixed recovery of 40% for all names. All else being equal, which one of the following four changes will make the principal invested more risky?

a. An increase in subordination of 1%, i.e., investing in the [11% − 12%] tranche
b. An increase in the tranche thickness from 1% to 3%, i.e., investing in the [10% − 13%] tranche
c. Using a recovery rate assumption of 50%
d. An increase in default correlation between names in the portfolio

---

EXAMPLE 22.18: FRM EXAM 2007—QUESTION 10

Consider the following homogeneous reference portfolio in a synthetic CDO: Number of reference entities, 100; CDS spread, $s = 150$bps; recovery rate $f = 50\%$. Assume that defaults are independent. On a single name the annual default probability is constant over five years and obeys the relation: $s = (1 - f)$PD. What is the expected number of defaulting entities over the next five years, and which of the following tranches would be entirely wiped out (loses 100% of the principal invested) by the expected number of defaulting entities?

a. 14 defaults and a [3% – 14%] tranche would be wiped out.
b. Three defaults and a [0% – 1%] tranche would be wiped out.
c. Seven defaults and a [2% – 3%] tranche would be wiped out.
d. 14 defaults and a [6% – 7%] tranche would be wiped out.

22.6 CONCLUSIONS

Credit products are by far the fastest growing segment of financial derivatives. Credit default swaps have become mainstream products, and now are actively traded for a large variety of names.

The rapid growth of the credit derivatives market is the best testimony of their usefulness. These instruments are superior risk management tools, allowing the *transfer of risks*. Table 22.4 breaks down the market by participants, taken from the BBA survey. a survey by the British Bankerss Association (BBA).

Banks are net buyers of credit protection, which is a hedge against their lending business. This helps explain why banks weathered the 2001 recession rather well, in spite of large corporate (WorldCom and Enron) and sovereign (Argentina) defaults. Most of these bank exposures had been sold. On the other side are insurance companies, which are net sellers of credit protection. This is akin to selling insurance.

TABLE 22.4 Buyers and Sellers of Credit Protection

Type of Institution	Buyer	Seller	Net
Banks	59	44	+15
Insurers	6	17	−11
Hedge funds	28	32	−4
Others	7	7	0
Total	100	100	0

Source: BBA Credit Derivatives Report 2006

On the other hand, structured products have spread risk all over the world, which has created contagion effects when subprime-backed assets started to go bad. In addition, insurance companies such as AIG have ended up selling too much protection. Because of its size, a failure by AIG would have probably caused systemic risk (the case of AIG is discussed in Chapter 25).

Standard credit derivatives such as credit default swaps, however, have many benefits. During 2007 and 2008, this market has remained fairly liquid unlike the cash bond markets. This CDS market creates transaction prices that provide useful information about the cost of credit to outside observers. In other words, they provide *price discovery*. CDS contracts also allow *transactional efficiency*, because they have lower transaction costs than the cash markets.

On the downside, the growth of this market has created *operational risk* because of backlogs in the processing of trades. Regulators have pushed the industry to improve the operational infrastructures, including more automated trade processing.

In addition, *counterparty risk* has become an issue in the wake of Lehman's failure. Lehman was a major player in the CDS market. This explains the push for a *centralized clearinghouse*. As in the case of the CLS Bank discussed in Chapter 18, this would allow multilateral netting of contracts and generally decrease counterparty risk. This would also make trading more transparent and, for some contracts, more liquid.

Credit derivatives also introduce a new element of risk, which is *legal risk*. Parties may not agree on the terms of the trade in case of default. Even with full confirmation of the trade, parties sometimes squabble over the definition of a credit event. Such disagreement occurred during the Russian default as well as notable debt restructurings and demergers. The widespread use of ISDA confirmation agreements helps resolve some of this uncertainty.

Overall, credit default swaps are likely to continue to thrive because they provide many benefits to financial market participants. In contrast, the future of complex credit securitizations is more clouded. This market has evolved from *regulatory arbitrage*—that is, attempts to defeat capital requirements by laying off loan credit risk through securitizations. As discussed in Chapter 7, the recent credit crisis has revealed serious flaws in the securitization process, which created complex instruments. Evaluating these complex structures requires sophisticated portfolio credit risk models, which are covered in the next chapter. Indeed the losses suffered on many of these structures were the trigger for the credit crisis that started in 2007.

22.7 IMPORTANT FORMULAS

Payoff on a credit default swap:
   Payment $=$ Notional $\times Q \times I(\mathrm{CE})$

Payoff on a credit spread forward contract:
   Payment $= (S - F) \times \mathrm{MD} \times$ Notional
   Payment $= [P(y + F, \tau) - P(y + S, \tau)] \times$ Notional

Valuation of a CDS contract:

$$V = (\text{PV Payoff}) - s(\text{PV Spread}) = \left(\sum_{t=1}^{T} k_t(1-f)\text{PV}_t\right)$$
$$- s\left(\sum_{t=1}^{T} S_{t-1}\text{PV}_t\right)$$

22.8 ANSWERS TO CHAPTER EXAMPLES

Example 22.1: FRM Exam 2004—Question 9

d. A long corporate bond position is equivalent to a long Treasury bond position plus a short CDS.

Example 22.2: FRM Exam 2007—Question 18

c. The bond yield spread is $15 - 3 = 12\%$. So, the trader could buy the corporate bond and hedge the interest rate risk by shorting the Treasury. To protect against default risk, he should buy a CDS on the same obligor at a spread of y. The total profit must be $12\% - y = 5\%$. Hence, the CDS spread must be 7%.

Example 22.3: FRM Exam 2007—Question 120

b. The bank should buy the swap to protect against default. The quarterly payment will be $10M \times 0.50\%/4 = \$12,500$.

Example 22.4: FRM Exam 1999—Question 135

a. Because all bonds rank equally, all defaults occur at the same time and have the same loss given default. Therefore the cash flow on the one-year credit swap can be replicated (including any risk premium) by going long the one-year Widget bond and short the 1-year T-bond.

Example 22.5: FRM Exam 2004—Question 50

a. The payment is $200 \times 113 - 300 \times 27 - 100 \times 43$, which translates into $1.02 million.

Example 22.6: FRM Exam 2004—Question 65

d. The protection buyer is exposed to the joint risk of default by the counterparty and underlying credit. If only one defaults, there is no credit risk.

Example 22.7: FRM Exam 2007—Question 85

d. For a loss to occur, both bank B and company R must default. The joint probability of default by B and R is 0.5% times 3.6%, which gives 0.018%.

Example 22.8: FRM Exam 2005—Question 111

c. If bank B buys company C, the two entities B and C will default at the same time. This increase in the default correlation makes the CDS contract less valuable. In Table 22.2, the fair CDS spread decreases when the correlation increases. Given that the existing CDS contract has a fixed spread, this event should decrease the value of the outstanding contract.

Example 22.9: FRM Exam 2005—Question 14

a. A TRS will provide protection against both interest rate and credit risk, as it is indexed to the bond portfolio value. A CDS or CS option only provide protection against credit risk. There is no currency risk in Yankee bonds, which are denominated in dollars, anyway.

Example 22.10: FRM Exam 2007—Question 69

a. On the LIBOR leg, the bank receives 7.25. In exchange, it pays the return on the bond, which is the coupon of 6.5% plus the relative return of $(99.35 - 101.82)/100 = -2.47\%$. This gives a receipt of $\$60 \times (7.25 - 6.5 + 2.47)/100 = 1.932$.

Example 22.11: FRM Exam 2000—Question 61

c. We need to value the bond with remaining semiannual payments for nine years using two yields, $y + S = 6.30 + 1.50 = 7.80\%$ and $y + K = 6.30 + 1.30 = 7.60\%$. This gives $\$948.95$ and $\$961.40$, respectively. The total payout is then $\$50,000,000 \times [\$961.40 - \$948.95]/\$1,000 = \$622,424$.

Example 22.12: FRM Exam 2004—Question 63

c. The equity tranche, tranche 1, must have the highest yield, and is sometimes called "toxic waste" because it has the highest risk. Conversely, tranche 3 would have the highest credit rating and the lowest yield.

Example 22.13: FRM Exam 2001—Question 12

b. The market values and weighted probability of default should be equal for the collateral and various tranches. So, a. is wrong. The equity tranche has the highest risk of default, so c. is wrong. The yield on the low-risk tranche must be the lowest, so d. is wrong.

Example 22.14: FRM Exam 2002—Question 32

a. In the absence of transaction costs or fees, the yield on the underlying portfolio should be equal to the weighted average of the yields on the different tranches.

With costs, however, the CBO yield will be slightly less. Otherwise, the senior tranche is typically rated AAA, has the lowest loss rate of all tranches, and absorbs the last loss on the structure.

Example 22.15: FRM Exam 2007—Question 130

d. Because the current CDS spread is greater than the coupon, the CLN must be selling at a discount. The only solution is d. More precisely, we can use the spread duration from Equation (22.2), which is the sum of the present value factor over three years. Assuming a flat term structure, this is $\sum PV_t = 0.952 + 0.907 + 0.864 = 2.72$ years. Multiplying by $(90 - 60) = 30$bp gives a fall of 0.81%, which gives $99.19.

Example 22.16: FRM Exam 2003—Question 7

c. If the correlation is one, all names will default at the same time, and the junior and senior tranche will be equally affected. Hence, their spread should be 100bp, which is the same as for the collateral. With lower correlations, the losses will be absorbed first by the junior tranche. Therefore, the spread on the junior tranche should be higher, which is offset by a lower spread for the senior tranches.

Example 22.17: FRM Exam 2007—Question 81

d. Increasing the subordination will make the senior tranche less risky because there is a thicker layer beneath to absorb losses. Increasing the thickness of the tranche will make it less likely to be wiped out, so is less risky. An increase in the default correlation will increase the risk. In the limit, if all assets default at the same time, all tranches will suffer a loss.

Example 22.18: FRM Exam 2007—Question 10

d. The annual marginal PD is $d = 1.5\%/(1 - 0.50) = 3.00\%$. Hence, the cumulative PD for the five years is $d + S_1 d + S_2 d + S_3 d + S_4 d = 3\%(1 + 0.970 + 0.941 + 0.913 + 0.885) = 14.1\%$, where the survival rates are $S_1 = (1 - 3\%) = 0.970$, $S_2 = S_1(1 - 3\% = 0.941$, and so on. The expected number of defaults is therefore $100 \times 14.1\%$, or 14. With a recovery rate of 50%, the expected loss is 7% of the notional. So, all the tranches up to the 7% point are wiped out.

Managing Credit Risk

Previous chapters have explained how to estimate the various input to portfolio credit risk models, including default probabilities, credit exposures, and recovery rates for individual credits. We now turn to the measurement of credit risk for the overall portfolio.

In the past, credit risk was measured on a standalone basis, in terms of a "yes" or "no" decision by a credit officer. Some consideration was given to portfolio effects through very crude credit limits at the overall level. Portfolio theory, however, teaches us that risk should be viewed in the context of the contribution to the total risk of a portfolio, not in isolation. The new credit risk models measure risk on a portfolio basis.

While this fundamental focus on diversification also exists in market risk, credit risk is markedly more complex. In particular, it is difficult to estimate probabilities and correlations of default events. These correlations, however, are essential drivers of diversification benefits.

Section 23.1 introduces the distribution of credit losses. This has two major features. The first is the expected credit loss, which is essential information for pricing and reserving purposes, as explained in Section 23.2. The second component is the unexpected credit loss, or worst deviation from the expected loss at some confidence level. Section 23.3 shows how this credit value at risk (CVAR), like market VAR, can be used to determine the amount of capital necessary to support a position. Section 23.4 then provides an overview of recently developed credit risk models, including CreditMetrics, CreditRisk+, the KMV model, and Credit Portfolio View. Finally, Section 23.5 gives some concluding comments. Given the complexity of these models, it is essential for risk managers to understand their weak spots and limitations.

23.1 MEASURING THE DISTRIBUTION OF CREDIT LOSSES

23.1.1 Steps

The previous chapters provided a detailed analysis of the various components of credit models, which include default probabilities, credit exposures, and recovery rates. We can now pool this information to measure the distribution of losses

due to credit risk. For simplicity, we initially consider only losses in default mode (DM), that is, losses due to defaults instead of changes in market values.

For one instrument, the potential credit loss is

$$\text{Credit loss} = b \times \text{CE} \times \text{LGD} \tag{23.1}$$

which involves the random variable b that takes on the value of 1 when the discrete state of default occurs, with probability of default (PD) p; the credit exposure, also called exposure at default (EAD); and the loss given default (LGD). With this definition, the credit loss is positive.

For a portfolio of N counterparties, the credit loss (CL) is

$$\text{CL} = \sum_{i=1}^{N} b_i \times \text{CE}_i \times \text{LGD}_i \tag{23.2}$$

where CE_i is now the total credit exposure to counterparty i, across all contracts and taking into account netting agreements.

The distribution of credit loss is quite complex. Typically, information about credit risk is described by the net replacement value (NRV), which is

$$\text{NRV} = \sum_{i=i}^{N} \text{CE}_i \tag{23.3}$$

evaluated at the current time. This is the most that could be lost if all parties defaulted at the same time ($b_i = 1$) and if there was no recovery ($\text{LGD}_i = 1$). This is not very informative, however. The NRV, which is often disclosed in annual reports, is equivalent to using notionals to describe the risks of derivatives portfolios. It does not take into account the probability of default or correlations across defaults and exposures.

Chapter 18 gave an example of a loss distribution for a simple portfolio with three counterparties. This example was tractable, as we could enumerate all possible states. In general, we need to consider many more credit events. We also need to account for movements and co-movements in risk factors, which drive exposures, uncertain recovery rates, and correlations among defaults. This can be done with the help of *Monte Carlo simulations*. Once this is performed for the entire portfolio, we obtain a distribution of credit losses on a target date. Figure 23.1 describes a typical distribution of credit profits and losses. A later section will illustrate the construction of this distribution as provided by commercial models.

We note that the distribution of credit P&L is *highly skewed to the left*, in contrast to that of market risk factors, which is in general roughly symmetrical. This credit distribution is similar to a short position in an option. This is one of the essential insights of the Merton model, which equates a risky bond to a risk-free bond plus a short position in an option.

Frequency distribution
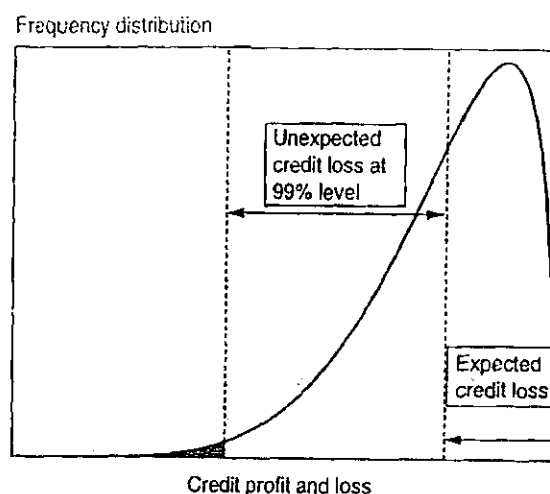
FIGURE 23.1   Distribution of Credit Losses

23.1.2 Major Features

This distribution can be described by:

- Expected credit loss (ECL). The expected credit loss represents the average credit loss. The *pricing* of the portfolio should be such that it covers the expected loss. In other words, the price should be advantageous enough to offset average credit losses. In the case of a bond, the price should be low enough, or the yield high enough, to compensate for expected losses. In the case of a derivative, the bank that takes on the credit risk should factor the expected loss into the pricing of its product. Loan loss reserves should be accumulated as a credit provision against expected losses. Focusing only on the default variables, the ECL depends solely on default probabilities.
- Unexpected credit loss (UCL). The worst credit loss represents the loss that will not be exceeded at some level of confidence, typically 99.9%. This is basically the quantile of the distribution. Taking the deviation from the expected loss gives the unexpected credit loss. The institution should have enough equity capital to cover the unexpected loss. Focusing only on the default variables, the UCL depends both on default probabilities and default correlations.

23.1.3 Effect of Correlations

The key to this approach is measuring risk at the top, portfolio level. This approach can also reveal the effect of correlations between and across risk types.

At the top of the list are the correlations across default events b_i. With low correlations and many obligors, the distribution will be narrow, as illustrated in Chapter 18. In this case, a bank could leverage up its equity several times. In its simplest version, the Basel Accord requires a minimum ratio of equity to assets of 8%, implying a maximum leverage ratio of 12.5. In contrast, high correlations will lead to simultaneous defaults, which extend the tail of the distribution and

increase the UCL. In the limit, with perfect correlations, the worst loss at a fixed confidence level is the entire notional amount of the portfolio. In this case, the bank cannot have any leverage. Its assets must be covered by the same amount of equity.

Correlation can also occur across the default event and exposure, i.e., between b_i and CE. For example, wrong-way trades are positions where the exposure is positively correlated with the probability of default. Before the Asian crisis, for instance, many U.S. banks had lent to Asian companies in dollars, or entered equivalent swaps. Many of these Asian companies did not have dollar revenues but instead were speculating, reinvesting the funds in the local currency. When currencies devalued, the positions were in-the-money for the U.S. banks, but could not be collected because the counterparties had defaulted. Conversely, right-way trades occur when the transaction is a *hedge* for the counterparty—for instance, when a loss on its side of the trade offsets an operating gain.

KEY CONCEPT

Credit risk is lowered for right–way trades, where the counterparty is using the trade as a hedge. Conversely, wrong–way trades create a positive correlation between the credit exposure and the probability of default.

Another pernicious correlation is between the default event and the loss given default, i.e., between b_i and LGD. Figure 23.2 plots the recovery rate from Moody's over the period 1982 to 2007 against the speculative-grade default rate during that year. During recession years, such as 1990, 1991, and 2001, the recovery rate on unsecured senior bonds was markedly lower than during other years. The default rate was also very high during these years. This effect will extend the tail of the credit loss distribution. If ignored, the credit VAR measure will underestimate risk.

EXAMPLE 23.1: CREDIT PROVISIONS

Credit provisions should be taken to cover all of the following *except*

a. Nonperforming loans
b. The expected loss of a loan portfolio
c. An amount equal to the VAR of the credit portfolio
d. Excess credit profits earned during below-average-loss years
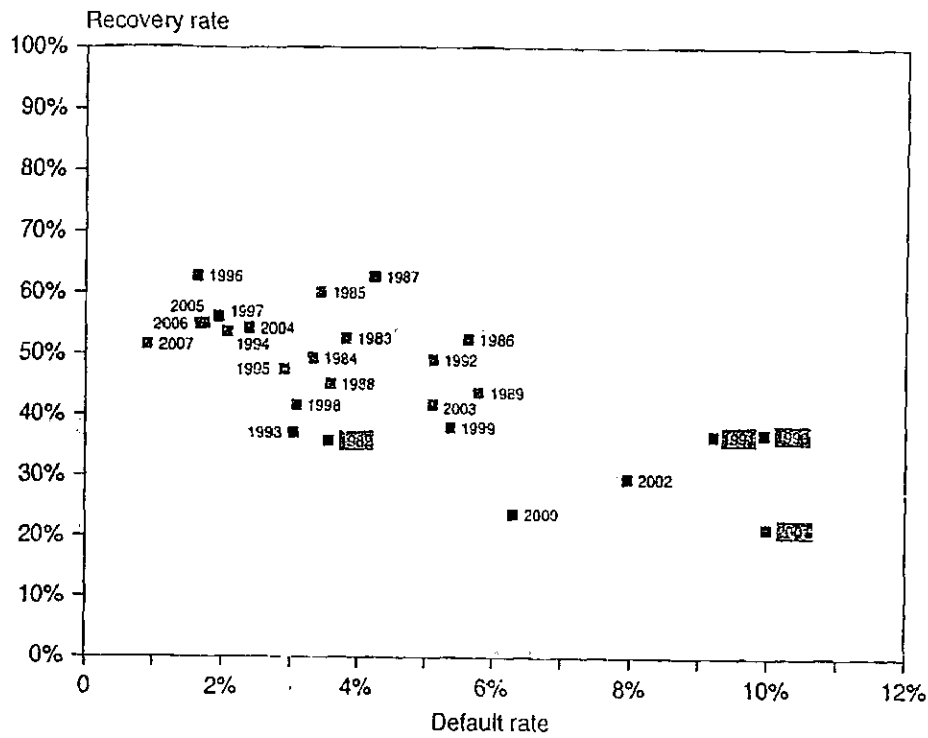
FIGURE 23.2   Recovery Rates and Default Rates

EXAMPLE 23.2: FRM EXAM 2002—QUESTION 74

Following is a set of identical transactions. Assuming all counterparties have the same credit rating, which transaction should preferably be executed?

a. Buying gas from a trading firm
b. Buying gas from a gas producer
c. Buying gas from a distributor
d. Indifferent between a., b., and c.

23.2   MEASURING EXPECTED CREDIT LOSS

23.2.1   Expected Loss over a Target Horizon

For pricing purposes, we need to measure the expected credit loss, which is

$$E[\text{CL}] = \int f(b, \text{CE}, \text{LGD})(b \times \text{CE} \times \text{LGD}) \, db \, d\text{CE} \, d\text{LGD} \qquad (23.4)$$

If the random variables are independent, the joint density reduces to the product of densities. We have

$$E[CL] = \left[\int f(b)(b)\, db\right]\left[\int f(CE)(CE)\, dCE\right]\left[\int f(LGD)(LGD)\, dLGD\right] \quad (23.5)$$

which is the product of the expected values. In other words,

$$\text{Expected credit loss} = \text{Prob[default]} \times E[\text{Credit exposure}] \times E[\text{LGD}] \quad (23.6)$$

As an example, the actuarial expected credit loss on a BBB-rated $100 million five-year bond with 47% recovery rate is $E[CL] = 2.28\% \times \$100,000,000 \times (1 - 47\%) = \$1.2$ million. Note that this expected loss is the same whether the bank has one $100 million exposure or 100 exposures worth $1 million each. The distributions, however, will be very different with more credits.

23.2.2 The Time Profile of Expected Loss

So far, we have focused on a fixed horizon, say a year. For pricing purposes, however, we need to consider the total credit loss over the life of the asset. This should involve the time profile of the exposure, the probability of default, and the discounting factor. Define PV_t as the present value of a dollar paid at time t.

The present value of expected credit losses (PVECL) is obtained as the sum of the discounted expected credit losses at each time step:

$$PVECL = \sum_t E[CL_t] \times PV_t = \sum_t [k_t \times ECE_t \times (1 - f)] \times PV_t \quad (23.7)$$

where the probability of default is $k_t = S_{t-1}\, d_t$, or the probability of defaulting at time t, conditional on not having defaulted before.

Alternatively, we could simplify by using the average default probability and average exposure over the life of the asset:

$$PVECL_A = \text{Ave}[k_t] \times \text{Ave}[ECE_t] \times (1 - f) \times \left[\sum_t PV_t\right] \quad (23.8)$$

This approach, however, may not be as good an approximation when default risk and exposure profile are correlated over time. For example, currency swaps with highly rated counterparties have an exposure and a default probability that increase with time. Due to this correlation, taking the product of the averages understates credit risk. In other cases, it could overstate credit risk.

An even simpler approach, when ECE is constant, considers the final maturity T only, using the cumulative default rate c_T and the discount factor PV_T:

$$PVECL_F = c_T \times ECE \times (1 - f) \times PV_T \quad (23.9)$$

TABLE 23.1 Computation of Expected Credit Loss for a Swap

| Year t | P(default) (%) | | | Exposure ECE_t | LGD $(1 - f)$ | Discount PV_t | Total $PVECL_t$ |
	c_t	d_t	k_t				
1	0.22	0.220	0.220	$1,660,000	0.55	0.9434	$1,895
2	0.54	0.321	0.320	$1,497,000	0.55	0.8900	$2,345
3	0.88	0.342	0.340	$1,069,000	0.55	0.8396	$1,678
4	1.55	0.676	0.670	$554,000	0.55	0.7921	$1,617
5	2.28	0.741	0.730	$0	0.55	0.7473	$0
Total			2.280			4.2124	$7,535
Average			0.456	$956,000			
$PVECL_A$			0.456	$956,000	0.55	4.2124	= $10,100

23.2.3 Examples

Table 23.1 shows how to compute the PVECL. We consider a five-year interest rate swap with a counterparty initially rated BBB and a notional of $100 million. The discount factor is 6% and the recovery rate 45%. We also assume that default can occur only at the end of each year. This analysis is similar to that for a credit default swap in Chapter 22. For simplicity, we use here real-world default probabilities.

In the first column, we have the cumulative default probability, c_t, for a BBB-rated credit from years 1 to 5, expressed as a percentage. The second column shows the marginal probability of defaulting during that year, d_t, and the third column shows the probability of defaulting in each year, conditional on not having defaulted before, $k_t = S_{t-1}d_t$. The fourth column reports the end-of-year expected credit exposure, ECE_t. The fifth column shows the constant LGD. The sixth column displays the present value factor, PV_t.

The final column gives the product $[k_t ECE_t(1 - f)PV_t]$. The first entry, for example is $0.220\% \times \$1,660,000 \times 0.55 \times 0.9434 = \$1,895$. Summing across years gives $7,535 on a swap with a notional of $100 million, or 0.007% of principal. This is very small, less than 1 basis point. So the expected credit loss on an interest rate swap is minuscule. Basically, the expected credit loss is very low due to the small exposure profile. For a regular bond or currency swap, the expected loss is much greater.

The last line shows a shortcut to the measurement of expected credit losses based on averages, from Equation (23.8). The average annual default probability is 0.456. Multiplying by the average exposure, $956,000, the LGD, and the sum of the discount rates gives $10,100. This is on the same order of magnitude as the exact calculation.

Table 23.2 details the computation for a bond assuming a constant exposure of $100 million. The expected credit loss is $1.020 million, about 100 times larger than for the swap. This is because the exposure is also about 100 times larger.

As in the previous table, the next line shows results based on averages. Here the expected credit loss is $1.056 million, very close to the exact number, as there is no variation in credit exposures over time.

TABLE 23.2   Computation of Expected Credit Loss for a Bond

	P(default) (%)			Exposure	LGD	Discount	Total
Year t	c_t	d_t	k_t	ECE_t	$(1-f)$	PV_t	$PVECL_t$
1	0.22	0.220	0.220	$100,000,000	0.55	0.9434	$114,151
2	0.54	0.321	0.320	$100,000,000	0.55	0.8900	$156,639
3	0.88	0.342	0.340	$100,000,000	0.55	0.8396	$157,009
4	1.55	0.676	0.670	$100,000,000	0.55	0.7921	$291,887
5	2.28	0.741	0.730	$100,000,000	0.55	0.7473	$300,024
Total			2.280			4.2124	$1,019,710
Average			0.456	$100,000,000			
$PVECL_A$			0.456	×$100,000,000	×0.55	×4.2124	= $1,056,461
$PVECL_F$	2.280			×$100,000,000	×0.55	×0.7473	= $937,062

We could also take the usual shortcut and simply compute an expected credit loss given by the cumulative five-year default rate times $100 million times the loss given default, which is $1.254 million. Discounting to the present, we get $0.937 million, close to the previous result.

EXAMPLE 23.3: FRM EXAM 2003—QUESTION 26

Which of the following loans has the lowest credit risk?

Loan	1 Year Probability of Default	Loss Given Default	Remaining Term (Months)
a.	1.99%	60%	3
b.	0.90%	70%	9
c.	1.00%	75%	6
d.	0.75%	50%	12

EXAMPLE 23.4: FRM EXAM 2007—QUESTION 38

Mr. Rosenquist, asset manager, holds a portfolio of SEK 200 million, which consists of BBB-rated bonds. Assume that the one-year probability of default is 4%, the recovery rate is 60%, and defaults are uncorrelated over the years. What is the two-year cumulative expected credit loss on Mr. Rosenquist's portfolio?

a. SEK 6.40 million

b. SEK 6.27 million

c. SEK 9.60 million

d. SEK 9.48 million

23.3 MEASURING CREDIT VAR

23.3.1 Credit VAR over a Target Horizon

Credit VAR is defined as the unexpected credit loss at some confidence level. Using the measure of credit loss in Equation (23.1), we construct a distribution of the credit loss $f(CL)$ over a target horizon. At a given confidence c, the worst credit loss (WCL) is defined such that

$$1 - c = \int_{WCL}^{\infty} f(x)dx \qquad (23.10)$$

The credit VAR is then measured as the deviation from ECL

$$CVAR = UCL = WCL - ECL \qquad (23.11)$$

where all losses are defined as positive numbers.

This CVAR number should be viewed as the economic capital to be held as a buffer against *unexpected* losses. Its application is fundamentally different from the *expected* credit loss, which is additive across obligors and can be aggregated over time.

Instead, the CVAR is measured over a target horizon, say one year, which is deemed sufficient for the bank to take corrective actions should credit problems start to develop. Corrective action can take the form of exposure, reduction, or adjustment of economic capital, all of which take considerably longer than the typical horizon for market risk.

23.3.2 Using CVAR to Manage the Portfolio

Once credit VAR is measured, it can be managed. The portfolio manager can examine the trades that contribute most to CVAR. If these trades are not particularly profitable, they should be eliminated.

The marginal contribution to risk can also be used to analyze the incremental effect of a proposed trade on the total portfolio risk. As in the case of market risk, individual credits should be evaluated on the basis not only of their standalone risk, but also of their contribution to the portfolio risk. For the same expected return, a trade that lowers risk should be preferable over one that adds to the portfolio risk. Such trade-offs can be made only with a formal measurement of portfolio credit risk, however.

This marginal analysis can also help to establish the remuneration of capital required to support the position. Say the distribution has an ECL of $1 billion and UCL of $5 billion. The bank then needs to set aside $5 billion just to cover random deviations from expected credit losses. This equity capital, however, will require remuneration. So, the pricing of loans should cover not only expected losses, but also the remuneration of this economic capital. This is what we call a *risk premium* and explains why observed credit spreads are larger than necessary simply to cover actuarial losses.

EXAMPLE 23.5: CREDIT VAR FOR ONE BOND

A risk analyst is trying to estimate the credit VAR (CVAR) for a risky bond. CVAR is defined as the maximum unexpected loss at a confidence level of 99.9% over a one-month horizon. Assume that the bond is valued at $1,000,000 one month forward, and the one-year cumulative default probability is 2% for this bond. What is the best estimate of the CVAR for the bond, assuming no recovery?

a. $20,000
b. $1,682
c. $998,318
d. $0

EXAMPLE 23.6: CREDIT VAR FOR TWO BONDS

A risk analyst is trying to estimate the credit VAR for a portfolio of two risky bonds. The credit VAR is defined as the maximum unexpected loss at a confidence level of 99.9% over a one-month horizon. Assume that each bond is valued at $500,000 one month forward, and the one-year cumulative default probability is 2% for each of these bonds. What is the best estimate of the credit VAR for this portfolio, assuming no default correlation and no recovery?

a. $841
b. $1,682
c. $998,318
d. $498,318

EXAMPLE 23.7: FRM EXAM 2005—QUESTION 122

You are the credit risk manager for Bank Happy. Bank Happy holds Treasuries for USD 500 million, one large loan that has a positive probability of default for USD 400 million and another loan that has a positive probability of default for USD 100 million. The defaults are uncorrelated. The bank computes a credit VAR at 1% using CreditRisk+. Which of the following statements made about the VAR by the analyst who works for you is necessarily *wrong*?

a. The VAR or WCL can be equal to zero.
b. The expected loss on the portfolio exceeds the VAR.
c. The expected loss on the portfolio is necessarily smaller than the VAR.
d. None of the above statements is wrong.

23.4   PORTFOLIO CREDIT RISK MODELS

Portfolio credit risk models can be classified according to their approaches. This section also describes the four main portfolio credit models.

23.4.1   Approaches to Portfolio Credit Risk Models

Table 23.3 summarizes the essential features of portfolio credit risk models in the industry.

Model Type   Top-down models group credit risks using single statistics. They aggregate many sources of risk viewed as *homogeneous* into an overall portfolio risk, without going into the details of individual transactions. This approach is appropriate for retail portfolios with large numbers of credits, but less so for corporate or sovereign loans. Even within retail portfolios, top-down models may hide specific risks, by industry or geographic location.

Bottom-Up models account for features of each instrument. This approach is most similar to the structural decomposition of positions that characterizes market VAR systems. It is appropriate for corporate and capital market portfolios. Bottom-up models are also most useful for taking corrective action, because the risk structure can be reverse-engineered to modify the risk profile.

Risk Definitions   Default-mode models consider only outright default as a credit event. Hence any movement in the market value of the bond or in the credit rating is irrelevant.

Mark-to-market models consider changes in market values and ratings changes, including defaults. These fair market value models provide a better assessment of

TABLE 23.3   Comparison of Credit Risk Models

	CreditMetrics	CreditRisk+	KMV	CreditPf.View
Originator	J.P. Morgan	Credit Suisse	KMV	McKinsey
Model type	Bottom-up	Bottom-up	Bottom-up	Top-down
Risk definition	Market value (MTM)	Default losses (DM)	Default losses (MTM/DM)	Market value (MTM)
Risk drivers	Asset values	Default rates	Asset values	Macro factors
Credit events	Rating change/ default	Default	Continuous default prob.	Rating change/ default
Probability	Unconditional	Unconditional	Conditional	Conditional
Volatility	Constant	Variable	Variable	Variable
Correlation	From equities (structural)	Default process (reduced-form)	From equities (structural)	From macro. factors
Recovery rates	Random	Constant within band	Random	Random
Solution	Simulation/ analytic	Analytic	Analytic	Simulation

risk, which is consistent with the holding period defined in terms of the liquidation period.

Models of Default Probability   Conditional models incorporate changing macroeconomic factors into the default probability through a functional relationship. Notably, we observe that the rate of default increases in a recession.

Unconditional models have fixed default probabilities and tend to focus on borrower- or factor-specific information. Some changes in the environment, however, can be allowed by manually changing the input parameters.

Models of Default Correlations   Because default correlations are not directly observed for the obligors in the portfolio, they must be inferred from a model.

Structural models explain correlations by the joint movements of assets, for example, stock prices. For each obligor, this price is the random variable that represents movements in default probabilities.

Reduced-form models explain correlations by assuming a particular functional relationship between the default probability and "background factors." For example, the correlation between defaults across obligors can be modeled by the loadings on common risk factors, say, industrial and country.

23.4.2 CreditMetrics

CreditMetrics, published in April 1997 by J.P. Morgan, was an early portfolio credit risk model. The system is a "bottom-up" approach where credit risk is driven by movements in bond ratings taken from a transition matrix.

In this model, credit quality is measured by a *latent variable*, unobserved, which can be interpreted as the value of assets of the obligor. This value is related to the value of the equity, which is the source of correlations across obligors because equity prices are observable. When the value of the asset falls below some floor, the obligor is assumed to be in a state of default. Thus, this class of models includes three types of random variables, (1) equity value, (2) asset value, and (3) default indicator.

The components of the system are described in Figure 23.3.

Measurement of Exposure by Instrument   This starts from the user's portfolio, decomposing all instruments by their exposure and assessing the effect of market volatility on expected exposures on the target date. The range of covered instruments includes bonds and loans, swaps, receivables, commitments, and letters of credit.

Distribution of Individual Default Risk   This step starts with assigning each instrument to a particular credit rating. Credit events are then defined by rating migrations, which include default, through a matrix of migration probabilities.
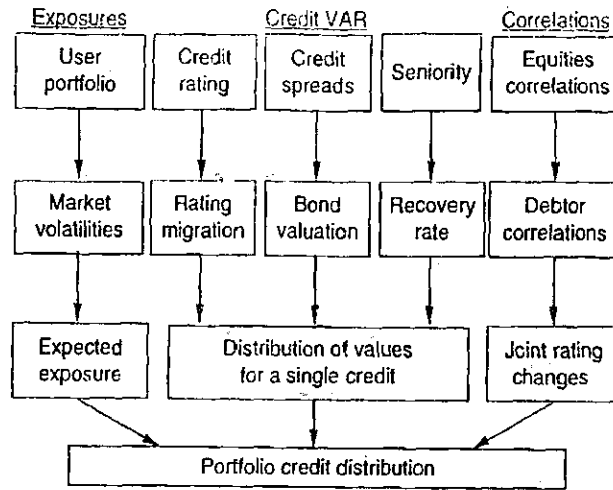
FIGURE 23.3 Structure of CreditMetrics

Thus movements in default probabilities are discrete. After the credit event, the instrument is valued using credit spreads for each rating class. In the case of default, the distributions of recovery rates are used from historical data for various seniority classes.

This is illustrated in Figure 23.4. We start from a bond or credit instrument with an initial rating of BBB. Over the horizon, the rating can jump to eight new values, including default. For each rating, the value of the instrument is recomputed, for example, to $109.37 if the rating goes to AAA, or to the recovery value of $51.13 in case of default. Given the state probabilities and associated values, we can compute an expected bond value, which is $107.09, and a standard deviation of $2.99.

Changes in the credit rating are driven by the latent factor, which is the asset value. Each asset value has a standard normal distribution with cutoff points

	Probability (p_i)	Value (V_i)	Exp. $\Sigma(p_iV_i)$	Var. $\Sigma p_i(V_i\text{-}m)^2$
AAA	0.02%	$109.37	0.02	0.00
AA	0.33%	$109.19	0.36	0.01
A	5.95%	$108.66	6.47	0.15
BBB	86.93%	$107.55	93.49	0.19
BB	5.30%	$102.02	5.41	1.36
B	1.17%	$98.10	1.15	0.95
CCC	0.12%	$83.64	0.10	0.66
Default	0.18%	$51.13	0.09	5.64

Sum = 100.00% $m = \$107.09$ $\sigma^2 = 8.95$

$SD = \$2.99$

FIGURE 23.4 Building the Distribution of Bond Values

TABLE 23.4 Cutoff Values for Simulations

Rating i	Prob. p_i	Cum. Prob. $N(z_i)$	Cutoff z_i
AAA	0.02%	100.00%	
AA	0.33%	99.98%	3.54
A	5.95%	99.65%	2.70
BBB	86.93%	93.70%	1.53
BB	5.30%	6.77%	−1.49
B	1.17%	1.47%	−2.18
CCC	0.12%	0.30%	−2.75
Default	0.18%	0.18%	−2.91

selected to represent the probabilities of changes in credit ratings. Table 23.4 illustrates the computations for our BBB credit. From Figure 23.4, there is a 0.18% probability of going from BBB into the state of default. We choose z_1 such that the area to its left is $N(z_1) = 0.18\%$. This gives $z_1 = -2.91$. Next, we need to choose z_2 so that the probability of falling between z_1 and z_2 is 0.12%, or that the total left-tail probability is $N(z_2) = 0.18\% + 0.12\% = 0.30\%$. This gives $z_2 = -2.75$, and so on.

Correlations among Defaults Correlations among defaults are inferred from correlations between asset values. These in turn are taken from correlations across *equity indices*. Each obligor is mapped to an industry and a geographical sector, using pre-assigned weights. Correlations are inferred from the co-movements of the common risk factors, using a database with some 152 country–industry indices, 28 country indices, and 19 worldwide–industry indices.

As an example, company 1 may be such that 90% of its volatility comes from the U.S. chemical industry. Using standardized returns, we can write

$$r_1 = 0.90\, r_{US,Ch} + k_1 \epsilon_1$$

where the residual ϵ is uncorrelated with other variables. Because the total volatility is normalized to 1, we must have $k_1 = \sqrt{1 - 0.9^2} = 0.44$.

Next, suppose that company 2 has a 74% weight on the German insurance index and 15% on the German banking index:

$$r_2 = 0.74\, r_{GE,In} + 0.15\, r_{GE,Ba} + k_2 \epsilon_2$$

The correlation between asset values for the two companies is

$$\rho(r_1, r_2) = (0.90 \times 0.74)\rho(r_{US,Ch}, r_{GE,In}) + (0.90 \times 0.15)\rho(r_{US,Ch}, r_{GE,Ba})$$
$$\rho(r_1, r_2) = (0.90 \times 0.74)0.15 + (0.90 \times 0.15)0.08 = 0.11$$

CreditMetrics then uses simulations of the joint asset values, assuming a multivariate normal distribution with the prespecified correlations. Thus, the approach

relies on a normal copula. This gives a total value for the portfolio and a distribution of credit losses over an annual horizon.

These simulations can also be used to compute correlations among default events. Because defaults are much less common than rating changes, the correlation is typically much less than the correlation between asset values. CreditMetrics reports that asset correlations in the range of 40% to 60% will typically translate into default correlations of 2% to 4%.[1]

Another drawback of this approach is that it does not integrate credit and market risk. Losses are generated only by changes in credit states, not by market movements. There is no uncertainty over market exposures. For swaps, for instance, the exposure on the target date is taken from the expected exposure. Bonds are revalued using today's forward rate and current credit spreads, applied to the credit rating on the horizon. So there is no interest rate risk.

23.4.3 CreditRisk+

CreditRisk+ was made public by Credit Suisse in October 1997. The approach is drastically different from CreditMetrics. It is based on a purely actuarial approach derived from the property insurance literature.

CreditRisk+ is a default mode (DM) model rather than a mark-to-market (MTM) model. Only two states of the world are considered—default and no-default.

The model starts with an assumption of a large number of identical loans n with independent default probability p. The total loss, $x = \sum_{i=1}^{n} b_i$, then follows a binomial distribution, which can be approximated by a Poisson distribution with intensity $\lambda = np$:

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!} \qquad (23.12)$$

Default correlations are introduced by assuming that the intensity itself is random. A high draw in λ then increases the probability of default for each obligor. This default intensity can also be time-varying, in which case it is modeled as a function of factors that change over time. CreditRisk+ accounts for variability in default rates by dividing the portfolio into homogeneous sectors within which obligors share the same systematic risk factor.

The other component of the approach is the severity of losses. This is roughly modeled by sorting assets by severity bands, say loans around $20,000 for the first band, $40,000 for the second band, and so on. A distribution of losses is then obtained for each band. These distributions are combined across bands to generate an overall distribution of default losses.

---

[1] This result, however, is driven by the joint normality assumption, which is not totally realistic. Other distributions can generate a greater likelihood of simultaneous defaults. This is obviously of major importance for credit portfolios.

The method provides a quick analytical solution to the distribution of credit losses with minimal data inputs. As with CreditMetrics, however, there is no uncertainty over market exposures.

23.4.4 Moody's KMV

Moody's KMV provides forecasts of estimated default frequencies (EDFs) for approximately 30,000 public firms globally.[2] Much of its technology is considered proprietary and is unpublished.

The model is an application of the Merton approach, which views the firm's equity E as a call option on the firm's assets

$$E = c(A, K, r, \sigma_A, \tau) \qquad (23.13)$$

In practice, KMV defines the floor K as the value of all short-term liabilities (one year and under) plus half the book value of all long-term debt. The value of assets is taken as the market value of equity plus the book value of all debt $A = nS + D$.

As seen in Section 20.2.2, this equation has to be iteratively estimated from observable variables, in particular the stock price S and its volatility σ_S to get the asset volatility. KMV computes a normalized distance to default (DD), which is essentially the distance between the current value of assets and the boundary point. Suppose, for instance, that $A = \$100$ million, $K = \$80$ million, and $\sigma_A = \$10$ million. We have

$$DD = z = \frac{A - K}{\sigma_A} = \frac{\$100 - \$80}{\$10} = 2 \qquad (23.14)$$

The main drivers of DD are (1) the level of the stock price, (2) the amount of leverage, and (3) the volatility of asset value. Lower stock prices, higher leverage, and higher asset volatility will decrease the DD measure.

In the final step, KMV uses this information to report an estimated default frequency (EDF), or default probability. If we assume normally distributed returns, for example, the probability of a standard normal variate z falling below -2 is about PD $= 2.3\%$. In practice, the EDFs are calibrated to actual default data, which gives objective (as opposed to risk-neutral) probabilities of default.

KMV generates default correlations between obligors directly from their equity prices, unlike CreditMetrics. First, returns on asset values are computed from changes in equity and debt values for the obligor. Second, these returns are regressed against a set of macroeconomic factors, country, and industry indices. Finally, this factor model is used to generate joint random variables representing obligor asset values, again using the industry standard normal copula.

The strength of this approach is that it relies on what is perhaps the best market data for a company—its stock price. Thus, it works best for *public firms*.

---

[2] KMV was founded by S. Kealhofer, J. McQuown, and O. Vasicek (hence the abbreviation KMV) to provide credit risk services. KMV started as a private firm based in San Francisco in 1989 and was acquired by Moody's in April 2002.

KMV also provides a model for *private companies*, which is based on accounting data for the firm and the industry as well as equity information, but only for public firms in the same industry. As expected, EDFs for private companies are considerably less accurate.

23.4.5 Credit Portfolio View

The last model we consider is Credit Portfolio View (CPV), published by the consulting firm McKinsey in 1997. The focus of this top-down model is on the effect of macroeconomic factors on portfolio credit risk.

This approach models loss distributions from the number and size of credits in subportfolios, typically consisting of customer segments. Instead of considering fixed transition probabilities, this model conditions the default probability on the state of the economy, allowing increases in defaults during recessions. The default probability p_t at time t is driven by a set of macroeconomic variables x^k for various countries and industries through a linear combination called y_t. The functional relationship to y_t, called the *logit model*, ensures that the probability is always between 0 and 1:

$$p_t = 1/[1 + \exp(y_t)], \qquad y_t = \alpha + \sum \beta^k x_t^k \qquad (23.15)$$

Using a multifactor model, each debtor is assigned to a country, industry, and rating segment. Uncertainty in recovery rates is also factored in. The model uses numerical simulations to construct the distribution of default losses for the portfolio. While useful for modeling default probabilities conditioned on the state of the economy, this approach is mainly top-down and does not generate sufficient detail of credit risk for corporate portfolios.

23.4.6 Comparisons

The International Swaps and Derivatives Association (ISDA) conducted a comparative survey of credit risk models.[3] The empirical study consisted of three portfolios of one-year loans with a total notional of $66.3 billion each.

1. High-credit-quality, diversified portfolio (500 names)
2. High-credit-quality, concentrated portfolio (100 names)
3. Low-credit-quality, diversified portfolio (500 names)

The models are listed in Table 23.5 and include CreditMetrics, CreditRisk+, and two internal models, all with a one-year horizon and 99% confidence level. Also reported are the charges from the Basel I "standard" rules, which will be explained in a later chapter. Suffice it to say that these rules make no allowance for variation in credit quality or diversification effects. Instead, the capital charge is based on 8% of the loan notional.

---

[3] ISDA (1998), *Credit Risk and Regulatory Capital*, New York: ISDA.

TABLE 28.5 Capital Charges from Various Credit Risk
Models

	Assuming Zero Correlation		
	Portfolio A	Portfolio B	Portfolio C
CreditMetrics	777	2,093	1,989
CreditRisk+	789	2,020	2,074
Internal model 1	767	1,967	1,907
Internal model 2	724	1,906	1,756
Basel I rules	5,304	5,304	5,304
	Assessing Correlations		
	Portfolio A	Portfolio B	Portfolio C
CreditMetrics	2,264	2,941	11,436
CreditRisk+	1,638	2,574	10,000
Internal model 1	1,373	2,366	9,654
Basel I rules	5,304	5,304	5,304

The top of the table examines the case of zero correlations. The Basel rules
yield the same capital charge, irrespective of quality or diversification effects. The
charge is also uniformly higher than most others, at $5,304 million, which is 8%
of the notional.

Generally, the four credit portfolio models show remarkable consistency in
capital charges. Portfolios A and B have the same credit quality, but B is more
concentrated. Portfolio A has indeed lower CVAR, approximately $800 million
against $2,000 million for B. This reflects the benefit from greater diversification.
Portfolios A and C have the same number of names, but C has lower credit quality.
This increases CVAR from around $800 million to $2,000 million.

The bottom panel assesses empirical correlations, which are typically positive.
The Basel charges are unchanged, as expected because they do not account for
correlations. Internal models show capital charges to be systematically higher
than in the previous case. There is also more dispersion in results across models,
however. It is interesting to see, in particular, that the economic capital charge for
portfolio C, with low credit quality, is typically twice the Basel charge. Such results
demonstrate that the Basel rules can lead to inappropriate credit risk charges. As
a result, banks subject to these capital requirements may shift the risk profile to
lower-rated credits until their economic capital is in line with regulatory capital.
This shift to lower credit quality was certainly not an objective of the original
Basel rules. This lack of sensitivity is what led to the new Basel Accord, which will
be discussed in Chapter 29.

More recently, another study compared the capital charges from the new Basel
II rules with three commercial models, taking great care to align parameters.[4] The
base portfolio consists of $100 billion in loans to 3,000 obligors spread across
different industries and countries, with an average credit rating of BBB.

---

[4] IACPM and ISDA (2006), *Convergence of Economic Capital Models*, New York: ISDA.

TABLE 23.6   Comparison of Credit Risk Models

	Expected Loss	Capital at 99.9%
KMV (PM)	563	3,791
CreditMetrics (CM)	562	3,533
CreditRisk+	564	3,662
Basel II	607	3,345

Table 23.6 shows the results in default mode. The Basel II rules, under the advanced approach, require economic capital of about \$3.3 billion, close to 4% of the notional. The three commercial models give remarkably close results. The report concludes that "if assumptions are aligned, there is not much difference between the valuation methods from PM and CM." Of course, this is also because these models are based on the same joint density function, using a normal copula, and calibrated to the same historical data.

23.5 CONCLUSIONS

Portfolio credit risk models take market risk models one step further. Here again, the fundamental intuition is that diversification across obligors, regions, and industries should lower portfolio risk.

The problem with internal portfolio credit risk models, however, is their complexity. Unlike market risk, where the risk manager can observe a history of movements in risk factors, there is generally no history of default for a particular obligor. Hence, default probabilities have to be modeled indirectly. The problem is even more difficult for estimating default correlations, as well as the shape of the joint density of defaults. Prior to 2007, these models had not been tested over a full market cycle, which should include a recession. This is particularly important because downturns lead to higher default probabilities, higher default correlations, and lower recovery rates.[5]

Finally, it will be difficult to verify such models given that they generate measures of economic capital over a long horizon, typically one year, and at a very high confidence level, typically 99.9%. In contrast, market risk models produce a daily VAR at the 99% confidence level, which should produce on average two or three exceptions a year. Thus, backtests of credit risk VAR estimates are not feasible, unlike for market risk.

This explains why regulators had considerable doubts about the precision of these models and, as a result, did not allow commercial banks to use their internal portfolio models as the basis for their credit risk charge. Indeed the losses suffered during the recession that started in 2007 have been much greater than

---

[5] The academic literature has long emphasized the sensitivity of credit risk models to correlations between defaults. See Das, Duffie, Kapadia, and Saita (2007), Common Failings: How Corporate Defaults are Correlated, *Journal of Finance*; Jorion and Zhang (2009), Credit Contagion from Counterparty Risk, *Journal of Finance*.

worst-case scenarios. Even though considerable progress has been made in our understanding of credit risk, it is fair to conclude that much more work is needed to develop robust models.

EXAMPLE 23.8: FRM EXAM 2004—QUESTION 11

When determining the standard deviation of value due to credit quality changes for a single exposure, the CreditMetrics model uses three primary factors. Which of the following is *not* one of the factors used in this model?

a. Credit ratings
b. Seniority
c. Equity prices
d. Credit spreads

EXAMPLE 23.9: FRM EXAM 2002—QUESTION 129

A bank computes the distribution of its loan portfolio marked-to-market value one year from now using the CreditMetrics approach of computing values for rating transition outcomes using (1) a rating agency transition matrix, (2) current forward curves, and (3) correlations among rating transition outcomes derived from stock returns of the obligors. In computing firm-wide risk using this distribution of its loan portfolio, the bank is most likely to understate its risk because it ignores

a. The term structure of interest rates
b. Rating drift
c. Spread risk
d. The negative correlation between the Treasury rates and credit spreads

---

EXAMPLE 23.10: FRM EXAM 2003—QUESTION 92

KMV measures the normalized distance from default. How is this defined?

a. (Expected assets − Weighted debt) / (Volatility of assets)
b. Equity/ (Volatility of equity)
c. Probability of stock price falling below a threshold
d. Leverage times stock price volatility

EXAMPLE 23.11: FRM EXAM 2004—QUESTION 20

A firm's assets are currently valued at $500 million and its current liabilities are $300 million. The standard deviation of asset values is $80 million. The firm has no other debt. What will be the approximate distance to default using the KMV calculation?

a. 2 standard deviations
b. 2.5 standard deviations
c. 6.25 standard deviations
d. Cannot be determined

EXAMPLE 23.12: FRM EXAM 2007—QUESTION 59

You are given the following information about a firm. The market value of assets at time 0 is 1,000; at time 1 is 1,200. Short-term debt is 500; long-term debt is 300. The annualized asset volatility is 10%. According to the KMV model, what are the default point and the distance to default at time 1?

a. 800 and 3.33
b. 650 and 7.50
c. 650 and 4.58
d. 500 and 5.83

EXAMPLE 23.13: FRM EXAM 2006—QUESTION 69

Which of the following model(s) calculates the change in portfolio value due to rating migration of the underlying instruments?

a. CreditRisk+
b. CreditMetrics
c. KMV
d. Both a. and c.

EXAMPLE 23.14: FRM EXAM 2005—QUESTION 36

Which of the following credit risk models uses the option theoretic approach for modeling correlation between the credit risky assets?

a. CreditRisk+
b. CreditMetrics
c. KMV for public firms
d. Both CreditMetrics and KMV for public firms

23.6 IMPORTANT FORMULAS

Credit loss: $\sum_{i=1}^{N} b_i \times CE_i \times LGD_i$

Expected credit loss: $ECL = \text{Prob[default]} \times E[\text{Credit exposure}] \times E[LGD]$

Present value of expected credit losses (PVECL):
$$PVECL = \sum_t E[CL_t] \times PV_t = \sum_t [k_t \times ECE_t \times (1 - f)] \times PV_t$$

Approximation to PVECL: $PVECL_F = c_T \times ECE \times (1 - f) \times PV_T$

Credit VAR: $CVAR = WCL - ECL$

KMV's normalized distance from default: $z = (A - K)/\sigma_A$

Credit portfolio view's default probability: $p_t = 1/[1 + \exp(y_t)]$, $y_t = \alpha + \sum \beta^k x_t^k$

23.7 ANSWERS TO CHAPTER EXAMPLES

Example 23.1: Credit Provisions

c. Credit provisions should be made for actual and expected losses. Capital, however, is supposed to provide a cushion against unexpected losses based on VAR.

Example 23.2: FRM Exam 2002—Question 74

b. This is an example of right-way trade. To have lower credit risk, it would be preferable to engage in a trade where there is a lower probability of a default by the counterparty when the contract is in-the-money. This will happen if the counterparty enters a transaction to hedge an operating exposure. For instance, a gas producer has a natural operating exposure to gas. If the producer sells gas at a fixed price, the swap will lose money if the market price of gas goes up. In this situation, however, there is little risk of default because the producer is sitting on an inventory of gas. A trading firm or distributor could go bankrupt if the transaction loses money.

Example 23.3: FRM Exam 2003—Question 26

a. The one-year PD needs to be adjusted to the maturity of the loan, using $(1 - d^m)^T$, where d^m is computed from $(1 - d^m)^{12} = (1 - d)$.

Loan	PD to Maturity	Loss Given Default	EL
a.	0.50%	60%	0.301%
b.	0.68%	70%	0.473%
c.	0.50%	75%	0.376%
d.	0.75%	50%	0.375%

Example 23.4: FRM Exam 2007—Question 38

b. The survival rate over two years is $S_2 = (1 - 4\%)^2 = 92.16\%$, which implies a cumulative two-year default rate of 7.84%. Put differently, the first-year PD is 4%, then $(1 - 4\%)4\% = 3.84\%$. Multiplying by 200 and 40% gives 6.27.

Example 23.5: Credit VAR for One Bond

c. First, we have to transform the annual default probability into a monthly probability. Using $(1 - 2\%) = (1 - d)^{12}$, we find $d = 0.00168$, which assumes a constant probability of default during the year. Next, we compute the expected credit loss, which is $d \times \$1,000,000 = \$1,682$. Finally, we calculate the WCL at the 99.9% confidence level, which is the lowest number CL_i such that $P(CL \leq CL_i) \geq 99.9\%$. We have $P(CL = 0) = 99.83\%$; $P(CL \leq 1,000,000) = 100.00\%$. Therefore, the WCL is \$1,000,000, and the CVAR is $\$1,000,000 - \$1,682 = \$998,318$.

Example 23.6: Credit VAR for Two Bonds

d. As in the previous question, the monthly default probability is 0.0168. The following table shows the distribution of credit losses.

Default	Probability (p_i)	Loss L_i	$p_i L_i$	$1 - \sum p_i$
2 bonds	$d^2 = 0.00000282$	\$1,000,000	\$2.8	100.00000%
1 bond	$2d(1 - d) = 0.00335862$	\$500,000	\$1,679.3	99.99972%
0 bonds	$(1 - d)^2 = 0.99663854$	\$0	\$0.0	99.66385%
Total	1.00000000		\$1,682.1	

This gives an expected loss of \$1,682, the same as before. Next, \$500,000 is the WCL at a minimum 99.9% confidence level because the total probability of observing a number equal to or lower than this is greater than 99.9%. The CVAR is then $\$500,000 - \$1,682 = \$498,318$.

Example 23.7: FRM Exam 2005—Question 122

c. The credit VAR could be zero. For instance, assume that the PD is 0.003. The joint probability of no default is then $(1 - 0.003)(1 - 0.003) = 99.4\%$. Because this is greater than the 99% confidence level, the worst loss is zero. The expected loss, however, would be 0.3% assuming zero recovery, which is greater than VAR.

Example 23.8: FRM Exam 2004—Question 11

c. CreditMetrics uses credit ratings, the transition matrix, recovery rates, and LGD for various seniority, but not equity prices for the obligor.

Example 23.9: FRM Exam 2002—Question 129

c. CreditMetrics ignores spread risk. It does account for ratings drift and the term structure of interest rates, albeit not their volatility.

Example 23.10: FRM Exam 2003—Question 92

a. The distance-to-default measure is a standardized variable that measures how much the value of firm assets exceeds the liabilities.

Example 23.11: FRM Exam 2004—Question 20

b. Using Equation (23.14), the DD is $(500 - 300)/80 = 2.5$ standard deviations.

Example 23.12: FRM Exam 2007—Question 59

c. The default point is given by short-term liabilities plus half of long-term liabilities, which is $500 + 300/2 = 650$. The distance to default at point 1 is $(V - K)/\sigma_V = (1{,}200 - 650)/(1{,}200 \times 0.10) = 4.58$.

Example 23.13: FRM Exam 2006—Question 69

b. Only CreditMetrics uses the rating migration. KMV uses the distance to default. CreditRisk+ uses random variables drawn from a Poisson distribution.

Example 23.14: FRM Exam 2005—Question 36

c. KMV estimates default probabilities using the Merton approach based on the company's stock price.

Legal, Operational, and Integrated Risk Management
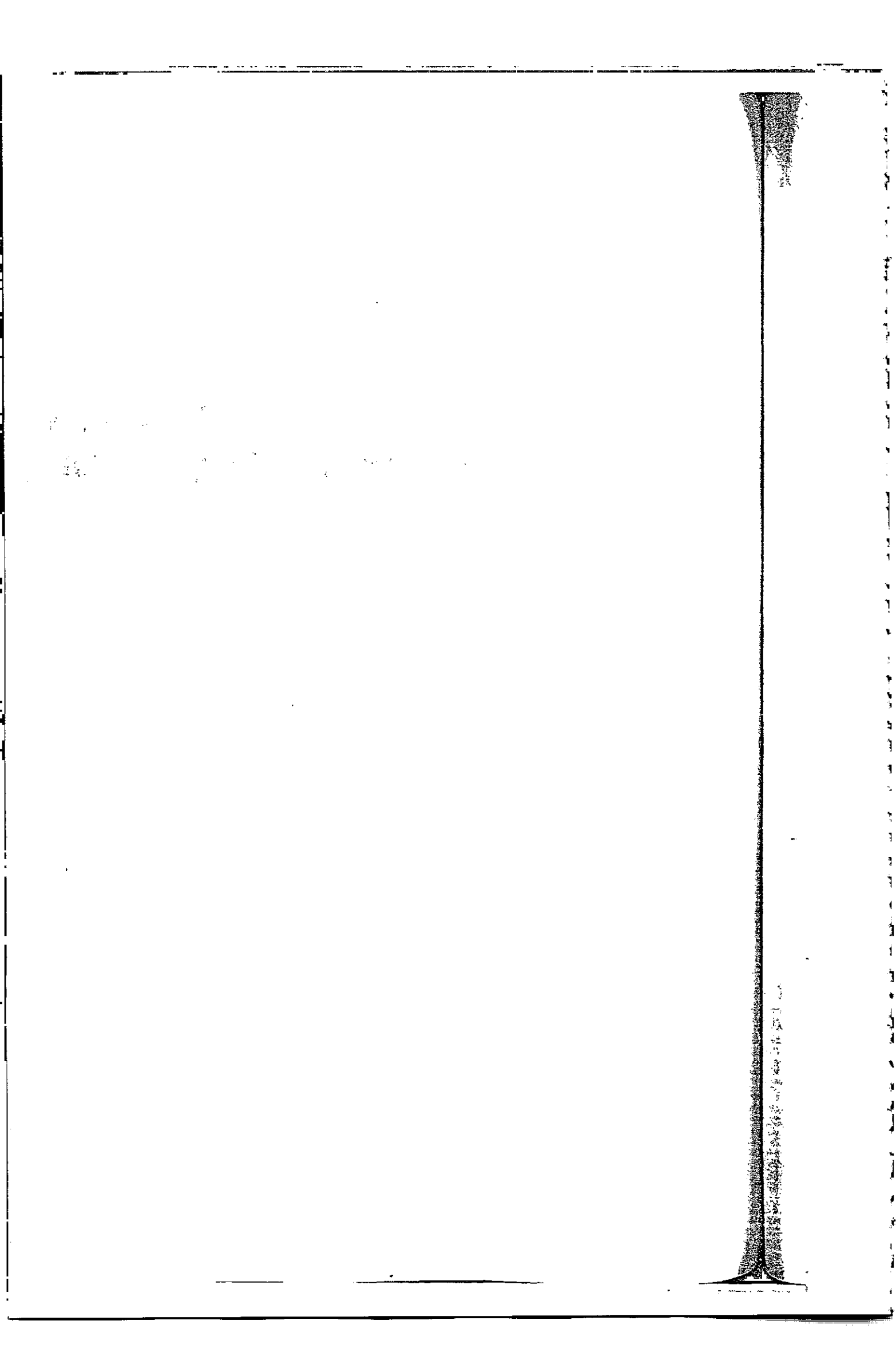
Operational Risk

The financial industry has developed standard methods to measure and manage market and credit risks. The industry is turning next to operational risk, which has proved to be an important cause of financial losses. Indeed, most company-specific financial disasters can be attributed to a combination of market and credit risk along with some failure of controls, which is a form of operational risk.

As in the case of market and credit risk, the financial industry is being pushed in the direction of better control of operational risk by bank regulators. For the first time, the Basel Committee has established capital charges for operational risk, in exchange for lowering them on market and credit risk. This new charge, which is explained in Chapter 29, would constitute approximately 12% of the total capital requirement.[1] As a result, this is forcing the banking industry to pay close attention to operational risk.

As with market and credit risk, the management of operational risk follows a sequence of logical steps: (1) identification, (2) assessment, (3) monitoring, and (4) control or mitigation.

Historically, operational risk has been managed by internal control mechanisms within business lines, supplemented by the audit function. The industry is now starting to use specific structures and control processes specifically tailored to operational risk.

To introduce operational risk, Section 24.1 summarizes lessons from well-known financial disasters. Given this information, Section 24.2 turns to definitions of operational risk. Various measurement approaches are discussed in Section 24.3. Finally, Section 24.4 shows how to use the distribution of operational losses to manage this risk better and offers some concluding comments.

24.1 THE IMPORTANCE OF OPERATIONAL RISK

The Basel Committee recently reported that "[a]n informal survey ... highlights the growing realization of the significance of risks other than credit and market risks, such as operational risk, which have been at the heart of some important banking problems in recent years." These problems are described in case histories next.

---

[1] See Basel Committee on Banking Supervision (2003), *Sound Practices for the Management and Supervision of Operational Risk*, Basel: BIS.

24.1.1 Case Histories

- *January 2008—SocGen (4.9 billion euros loss).* A rogue trader, Jerome Kerviel, systematically deceives systems, taking unauthorized positions worth up to 49 billion euros in stock index futures. The bank has enough capital to absorb the loss but its reputation is damaged.
- *February 2002—Allied Irish Bank ($691 million loss).* A rogue trader, John Rusnack, hides three years of losing trades on the yen/dollar exchange rate at the U.S. subsidiary. The bank's reputation is damaged.
- *March 1997—NatWest ($127 million loss).* A swaption trader, Kyriacos Papouis, deliberately covers up losses by mispricing and overvaluing option contracts. The bank's reputation is damaged. NatWest is eventually taken over by the Royal Bank of Scotland.
- *September 1996—Morgan Grenfell Asset Management ($720 million loss).* A fund manager, Peter Young, exceeds his guidelines, leading to a large loss. Deutsche Bank, the German owner of MGAM, agrees to compensate the investors in the fund.
- *June 1996—Sumitomo ($2.6 billion loss).* A copper trader amasses unreported losses over three years. Yasuo Hamanaka, known as "Mr. Five Percent," after the proportion of the copper market he controlled, is sentenced to prison for forgery and fraud. The bank's reputation is severely damaged.
- *September 1995—Daiwa ($1.1 billion loss).* A bond trader, Toshihide Igushi, amasses unreported losses over 11 years at the U.S. subsidiary. The bank is declared insolvent.
- *February 1995—Barings ($1.3 billion loss).* Nick Leeson, a derivatives trader, amasses unreported losses over two years. Barings goes bankrupt.
- *October 1994—Bankers Trust ($150 million loss).* The bank becomes embroiled in a high-profile lawsuit with a customer that accuses it of improper selling practices. Bankers settles, but its reputation is badly damaged. It is later bought out by Deutsche Bank.

Many of these spectacular losses can be traced to a rogue trader, or a case of internal fraud. These failures involve a mix of market risk and operational risk, i.e., failure to supervise properly. It should be noted that the cost of these events has been quite high. They led to large direct monetary losses, sometimes even to bankruptcy. In addition to these direct costs, banks often suffered large indirect losses due to reputational damage.

24.1.2 Business Lines

These failures have occurred across a variety of business lines. Some are more exposed than others to market risk or credit risk. All have some exposure to operational risk, however.

Commercial banking is exposed mainly to credit risk, less so to operational risk, and least to market risk. Investment banking, trading, and treasury management have greater exposure to market risk. On the other hand, business lines such

TABLE 24.1 Examples of Operational Risks

Type of Risk	Definition	Market Bank	Credit Bank
Operations risk	Losses due to complex systems and processes	High risk	Medium risk
Ops. settlement risk	Lost interest/fines due to failed settlements	High risk	Low risk
Model risk	Losses due to imperfect model or data	High risk	Low risk
Fraud risk	Reputational/financial damage due to fraud	High risk	Low risk
Misselling risk	Losses due to unsuitable sales	Medium risk	Medium risk
Legal risk	Reputational/financial damage due to fraud	High risk	Medium risk

Source: Financial Services Authority (1999), "Allocating Regulatory Capital for Operational Risk," London: FSA.

as retail brokerage and asset management are exposed primarily to operational risk. Asset managers assume no direct market risk since they act as agents for the investors. If they act in breach of guidelines, however, they may be liable to clients for their losses, which represents operational risk.

Table 24.1 presents a partial list of risks for market banks that are primarily involved in trading, and credit banks that specialize in lending activities. The table shows that different lines of business are characterized by very different exposures to the listed risks. Credit banks deal with relatively standard products, such as mortgages, with little trading. Hence, they have medium operations risk and low operations settlement risk. This is in contrast with trading banks, with constantly changing products and large trading volume, for which both risks are high. Trading banks also have high model risk, because of the complexity of products, and high fraud risk, because of the autonomy given to traders. In contrast, these two risks are low for credit banks.

For trading banks that deal with so-called sophisticated investors, misselling risk has low probability but high value; hence, it is a medium risk. A good example is Merrill Lynch settling with Orange County for about $400 million following allegations that the broker had sold the county unsuitable investments. For credit banks that deal with retail investors, this risk has higher probability but lower value: hence, it is a medium risk. Legal risks are high for market banks and medium for credit banks due to the more litigious environment of corporations relative to retail investors.

24.2 IDENTIFYING OPERATIONAL RISK

One could argue that operational risk has no clear-cut definition, unlike market risk and credit risk. There was a long debate as to the proper definition of operational risk, or even whether it makes sense to attempt to measure it.

After much industry consultation, the Basel Committee has settled on a definition that is becoming an industry standard. Operational risk is defined as

the risk of loss resulting from inadequate or failed internal processes, people and systems, or from external events

This includes the usual internal business events but also external events such as external fraud, security breaches, regulatory effects, or natural disasters. It includes legal risk, which arises when a transaction proves unenforceable in law. This definition, however, excludes strategic and reputational risk, which would be very difficult to measure, anyway.

The British Bankers' Association provides further detail for this definition. Table 24.2 breaks down operational risk into categories of people risk, process risk, system risk, and external risk. Among these, a notable risk for complex products is model risk, which is due to the use of wrong models for valuing and hedging assets. This is an internal risk that combines lack of knowledge (people) with product complexity/valuation errors (process) and perhaps programming errors (systems).

TABLE 24.2   Operational Risk Classification

Internal Risks		
People	Processes	Systems
Employee collusion/fraud	Accounting error	Data quality
Employee error	Capacity risk	Programming errors
Employee misdeed	Contract risk	Security breach
Employer liability	Misselling/suitability	Strategic risks
Employment law	Product complexity	(platform/suppliers)
Health and safety	Project risk	System capacity
Industrial action	Reporting error	System compatibility
Lack of knowledge/skills	Settlement/payment error	System delivery
Loss of key personnel	Transaction error	System failure
	Valuation error	System suitability

External Risks	
External	Physical
Legal	Fire
Money laundering	Natural disaster
Outsourcing	Physical security
Political	Terrorism
Regulatory	Theft
Supplier risk	
Tax	

Source: British Bankers' Association survey.

EXAMPLE 24.1: FRM EXAM 2004—QUESTION 39

Which of the following is *not* a type of operational risk as defined by Basel II?

a. Human error and internal fraud
b. Destruction by fire or other external catastrophes
c. Damaged reputation due to a failed merger
d. Failure or breakdown in internal control processes

EXAMPLE 24.2: FRM EXAM 2002—QUESTION 133

Which one of the following cases or events can be considered as resulting from operational risk?

a. A bank reports losses on a diversified portfolio of stocks during the stock market decline.
b. The bank becomes embroiled in a high-profile lawsuit with a customer that accuses it of improper selling practices.
c. The bank reports the loss of $1.5 billion due to rises in interest rates.
d. A U.S. investor makes a loss as the yen depreciates relative to the dollar.

EXAMPLE 24.3: FRM EXAM 2003—QUESTION 65

Which of these outcomes is *not* associated with an operational risk process?

a. The sale of call options is being booked as a purchase.
b. A monthly volatility is inputted in a model that requires a daily volatility.
c. A loss is incurred on an option portfolio because ex post volatility exceeded expected volatility.
d. A volatility estimate is based on a time-series that includes a price that exceeds the other prices by a factor of 100.

EXAMPLE 24.4: FRM EXAM 2007—QUESTION 56

All the following are operational risk loss events, *except*:

a. An individual shows up at a branch presenting a check written by a customer for an amount substantially exceeding the customer's low checking account balance. When the bank calls the customer to ask him for the funds, the phone is disconnected and the bank cannot recover the funds.

b. A bank, acting as a trustee for a loan pool, receives less than the projected funds due to delayed repayment of certain loans.

c. During an adverse market movement, the computer network system becomes overwhelmed, and only intermittent pricing information is available to the bank's trading desk, leading to large losses as traders become unable to alter their hedges in response to falling prices.

d. A loan officer inaccurately enters client financial information into the bank's proprietary credit risk model.

EXAMPLE 24.5: FRM EXAM 2007—QUESTION 139

The risk of the occurrence of a significant difference between the mark-to-model value of a complex and/or illiquid instrument and the price at which the same instrument is revealed to have traded in the market is referred to as:

a. Liquidity risk
b. Dynamic risk
c. Model risk
d. Mark-to-market risk

24.3 ASSESSING OPERATIONAL RISK

Once identified, operational risk should be measured, or rather "assessed" if it is less amenable to precise quantification than market or credit risk. Various approaches can be broadly classified into top-down models and bottom-up models.

24.3.1 Comparison of Approaches

Top-down models attempt to measure operational risk at the broadest level, that is, using firm-wide or industry-wide data. Results are then used to determine the amount of capital that needs to be set aside as a buffer against this risk. This capital is allocated to business units.

Bottom-up models start at the individual business unit or process level. The results are then aggregated to determine the risk profile of the institution. The

main benefit of bottom-up models is that they lead to a better understanding of the causes of operational losses, as in the case of VAR-based market risk systems.

Tools used to manage operational risk can be classified into six categories:

1. Audit oversight. This consists of reviews of business processes by an external audit department.
2. Critical self-assessment. Each business unit identifies the nature and degree of operational risk. These *subjective* evaluations include expected frequency and severity of losses, as well as a description of how risk is controlled. The tools used for this type of process include checklists, questionnaires, and facilitated workshops. The results are then aggregated, in a bottom-up approach.
3. Key risk indicators. These consist of simple measures that provide an indication of whether risks are changing over time.

    These *early warning signs* can include audit scores, staff turnover, trade volumes, and so on. The assumption is that operational risk events are more likely to occur when these indicators increase. These *objective* measures allow the risk manager to forecast losses through the application of regression techniques, for example.
4. Earnings volatility. This approach consists of taking a time series of earnings, after stripping the effect of market and credit risk, and computing its volatility. This measure is simple to use but has numerous problems. This risk measure also includes fluctuations due to business and macroeconomic risks, which fall outside operational risk. Also, such a measure is backward-looking and does not account for improvement or degradation in the quality of controls.
5. Causal networks. These describe how losses can occur from a cascade of different causes. Causes and effects are linked through conditional probabilities. Simulations are then run on the network, generating a distribution of losses. Such bottom-up models improve the understanding of losses since they focus on drivers of risk. The process is explained in the appendix.
6. Actuarial models. These combine the distribution of frequency of losses with their severity distribution to produce an *objective* distribution of losses due to operational risk. These can be either bottom-up or top-down models.

EXAMPLE 24.6: FRM EXAM 2003—QUESTION 50

Which of the following is a weakness of the top-down approach to measuring operational risk?

a. It fails to consider historical information.

b. You cannot use earnings volatility as an indicator of risk potential in this approach.

c. Information on specific sources of risk is not provided.

d. It is based on the specific mapping of business units and not the overall organization.

24.3.2 Actuarial Models

Actuarial models estimate the objective distribution of losses from historical data and are widely used in the insurance industry. Such models combine two distributions: loss frequencies and loss severities. The loss frequency distribution describes the number of loss events over a fixed interval of time. The loss severity distribution describes the size of the loss once it occurs.

Loss severities can be tabulated from historical data, for instance, measures of the loss severity y_k, at time k. These measures can be adjusted for inflation and some measure of current business activity. Define P_k as the consumer price index at time k and V_k as a business activity measure such as the number of trades. We could assume that the severity is proportional to the volume of business V and to the price level. The *scaled* loss is measured as of time t as

$$x_t = y_k \times \frac{P_t}{P_k} \times \frac{V_t}{V_k} \qquad (24.1)$$

Loss severity distributions have very long tails, representing the possibility of very large losses. Ideally, they should include internal and external data.

Next, define the loss frequency distribution by the variable n, which represents the number of occurrences of losses over the period. The density function is

$$\text{p.d.f. of loss frequency} = f(n), \quad n = 0, 1, 2, \ldots \qquad (24.2)$$

If x (or X) is the loss severity when a loss occurs, its density is

$$\text{p.d.f. of loss severity} = g(x \mid n = 1), \quad x \geq 0 \qquad (24.3)$$

Finally, the total loss over the period is given by the sum of individual losses over a random number of occurrences:

$$S_n = \sum_{i=1}^{n} X_i \qquad (24.4)$$

Table 24.3 provides a simple example of two such distributions. Our task is now to combine these two distributions into one—that of total losses over the period.

TABLE 24.3 Sample Loss Frequency and Severity Distributions

Frequency Distribution		Severity Distribution	
Probability	Frequency	Probability	Severity
0.6	0	0.5	$1,000
0.3	1	0.3	$10,000
0.1	2	0.2	$100,000
Expectation	0.5	Expectation	$23,500

Assuming that the frequency and severity of losses are independent, the two distributions can be combined into a distribution of aggregate loss through a process known as convolution. Convolution can be implemented, for instance, through tabulation. Tabulation consists of systematically recording all possible combinations with their associated probabilities and is illustrated in Table 24.4. Generally, convolution must be implemented by numerical methods, as there are too many combinations of variables for a systematic tabulation.

We start with the obvious case, no loss, which has probability 0.6. Next, we go through all possible realizations of one loss only. From Table 24.3, we see that a loss of $1,000 can occur with total probability of $P(n = 1) \times P(x = \$1,000) = 0.3 \times 0.5 = 0.15$. Similarly, for one-time losses of $10,000 and $100,000, the probabilities are 0.09 and 0.06, respectively. We then go through all occurrences of two losses, which can result from many different combinations. For instance, a loss of $1,000 can occur twice, for a total of $2,000, with a probability of $0.1 \times 0.5 \times 0.5 = 0.025$. We can have a loss of $1,000 and $10,000, for a total of $11,000, with probability $0.1 \times 0.5 \times 0.3 = 0.015$. We repeat these steps until we exhaust all combinations.

TABLE 24.4   Tabulation of Loss Distribution

Number of Losses	First Loss	Second Loss	Total Loss	Probability
0	0	0	0	0.600
1	1,000	0	1,000	0.150
1	10,000	0	10,000	0.090
1	100,000	0	100,000	0.060
2	1,000	1,000	2,000	0.025
2	1,000	10,000	11,000	0.015
2	1,000	100,000	101,000	0.010
2	10,000	1,000	11,000	0.015
2	10,000	10,000	20,000	0.009
2	10,000	100,000	110,000	0.006
2	100,000	1,000	101,000	0.010
2	100,000	10,000	110,000	0.006
2	100,000	100,000	200,000	0.004

Sorted Losses	Cumulative Probability
0	60.0%
1,000	75.0%
2,000	77.5%
10,000	86.5%
11,000	89.5%
20,000	90.4%
100,000	96.4%
101,000	98.4%
110,000	99.6%
200,000	100.0%
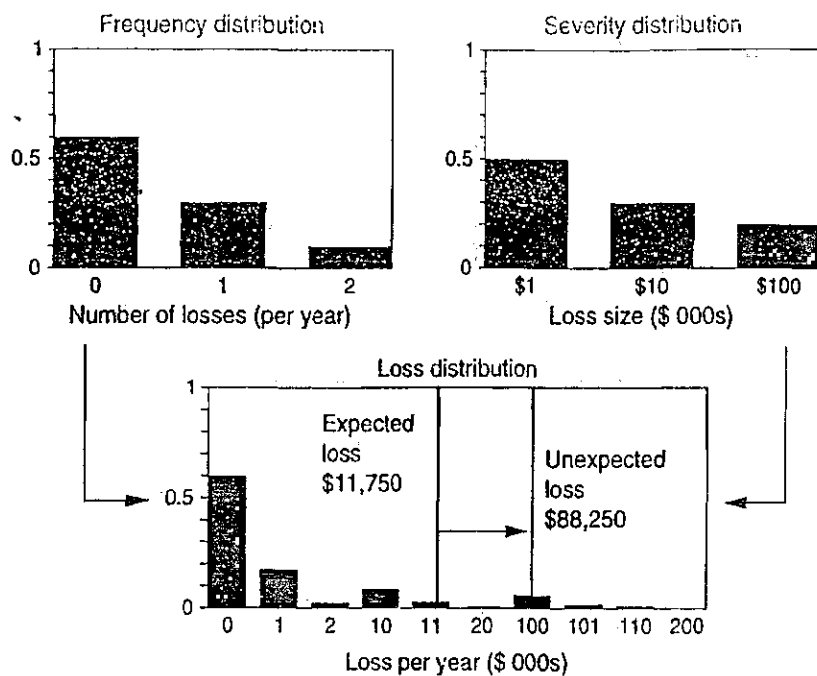
FIGURE 24.1 Construction of the Loss Distribution

The resulting distribution is displayed in Figure 24.1, and in the lower panel of Table 24.4. As usual for operational risk, losses are recorded as positive values. It is interesting to note that the very simple distributions in Table 24.3, with only three realizations, create a complex loss distribution. We can compute the expected loss, which is simply the product of expected values for the two distributions, or $E[S] = E[N] \times E[X] = 0.5 \times \$23,500 = \$11,750$. Risk management, however, is about the unexpected. So, the risk manager should also report the lowest number such that the probability is greater than 95%. This is $100,000, with a probability of 96.4%. Hence, the unexpected loss is $100,000 − $11,750 = $88,250. If operational VAR must include the expected loss, this is simply $100,000. This is the default measure under Basel II, which measures VAR at the 99.9% level of confidence over one year.

EXAMPLE 24.7: FRM EXAM 2007—QUESTION 138

The severity distribution of operational losses usually has the following shape:

a. Symmetrical with short tails
b. Long-tailed to the right
c. Uniform
d. Symmetrical with long tails

24.4 MANAGING OPERATIONAL RISK

24.4.1 Capital Allocation and Insurance

Like market VAR, the distribution of operational losses can be used to estimate expected losses as well as the amount of capital required to support this financial risk. Figure 24.2 highlights important attributes of a distribution of losses due to operational risk.

The expected loss represents the size of operational losses that should be expected to occur. Typically, this is dominated by high-frequency, low-severity events. This type of loss is generally absorbed as an ongoing cost and managed through internal controls. Such losses are rarely disclosed.

The unexpected loss represents the deviation between the quantile loss at some confidence level and the expected loss. Typically, this represents lower-frequency, higher-severity events. This type of loss is generally offset against capital reserves or transferred to an outside insurance company, when available. Such losses are sometimes disclosed publicly but often with little detail.

The stress loss represents a loss in excess of the unexpected loss. By definition, such losses are very infrequent but extremely damaging to the institution. The Barings bankruptcy can be attributed, for instance, in large part to operational risk. This type of loss cannot be easily offset through capital allocation, as it would require too much capital. Ideally, it should be transferred to an insurance company. Due to their severity, such losses are disclosed publicly.

However, purchasing insurance is no panacea. The insurance payment would have to be made very quickly and in full. The bank could fail while waiting for payment or arguing over the size of compensation. In addition, the premium may be very high. This is because once the insurance is acquired, the purchaser has less
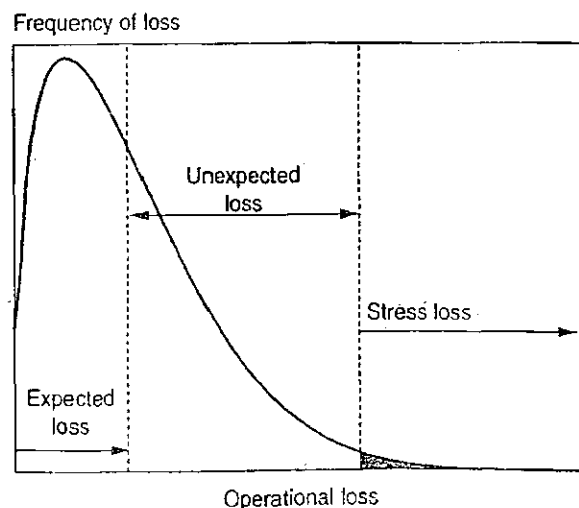
FIGURE 24.2 Distribution of Operational Losses

incentive to control losses. This problem is called moral hazard. The insurer will be aware of this and will increase the premium accordingly. The premium may also be high because of the adverse selection problem. This describes a situation where banks vary in the quality of their controls. Banks with poor controls are more likely to purchase insurance than banks with good controls. Because the insurance company does not know what type of bank it is dealing with, it will increase the average premium. Insurance with a deductible amount, i.e., where the bank would have to share the first layer of losses, only provides a partial solution to these problems.

24.4.2 Mitigating Operational Risk

The approach so far has been to take operational risk as given. Such measures are extremely useful because they highlight the size of losses due to operational risk. Armed with this information, the institution can then decide whether it is worth spending resources on decreasing operational risk.

Say that a bank is wondering whether to install a straight-through processing system, which automatically captures trades in the front office and transmits them to the back office. Such a system eliminates manual intervention and the potential for human errors, thereby decreasing losses due to operational risk. The bank should purchase the system if its cost is less than its operational risk benefit.

More generally, reduction of operational risk can occur in terms of the frequency of losses and/or the size of losses when they occur. Operational risk is also contained by a firm-wide risk management framework. In a later chapter, we will discuss *best practices* in risk management, which are designed to provide some protection against operational risk.

Consider, for instance, a transaction in a plain-vanilla, five-year interest rate swap. This simple instrument generates a large number of cash flows, each of which has the potential for errors. At initiation, the trade needs to be booked and confirmed with the counterparty. It must be valued so that a P&L can be attributed to the trading unit. With biannual payments, the swap will generate 10 cash flows along with 10 rate resets and net payment computations. These payments need to be computed with absolute accuracy—that is, to the last cent. Errors can range from minor issues, such as paying a day late, to major problems, such as failure to hedge or fraudulent valuation by the trader.

The swap will also create some market risk, which may need to be hedged. The position needs to be transmitted to the market risk management system, which will monitor the total position and risk of the trader and of the institution as a whole. In addition, the current and potential credit exposure must be regularly measured and added to all other trades with the same counterparty. Errors in this risk measurement process can lead to excessive exposure to market and/or credit risk.

Operational risk can be minimized in a number of ways.[2] Internal control methods consist of

- *Separation of functions.* Individuals responsible for committing transactions should not perform clearance and accounting functions.
- *Dual entries.* Entries (inputs) should be matched from two different sources— that is, the trade ticket and the confirmation by the back office.
- *Reconciliations.* Results (outputs) should be matched from different sources— for instance, the trader's profit estimate and the computation by the middle office.
- *Tickler systems.* Important dates for a transaction (e.g., settlement and exercise dates) should be entered into a calendar system that automatically generates a message before the due date.
- *Controls over amendments.* Any amendment to original deal tickets should be subject to the same strict controls as original trade tickets.

External control methods consist of

- *Confirmations.* Trade tickets need to be confirmed with the counterparty, which provides an independent check on the transaction.
- *Verification of prices.* To value positions, prices should be obtained from external sources. This implies that an institution should have the capability of valuing a transaction in-house before entering it.
- *Authorization.* The counterparty should be provided with a list of personnel authorized to trade, as well as a list of allowed transactions.
- *Settlement.* The payment process itself can indicate if some of the terms of the transaction have been incorrectly recorded—for instance, if the first cash payments on a swap are not matched across counterparties.
- *Internal and external audits.* These examinations provide useful information on potential weakness areas in the organizational structure or business process.

24.4.3 Conceptual Issues

The management of operational risk is beset by conceptual problems. First, unlike market and credit risk, operational risk is largely internal to financial institutions. Because institutions are understandably reluctant to advertise their mistakes, it is more difficult to collect data on operational losses. Another problem is that losses may not be directly applicable to another institution, as they were incurred under possibly different business profiles and internal controls. Internal data also has survival bias, because they will not contain cases of losses that could bankrupt the institution.

Second, market and credit risk can be conceptually separated into exposures and risk factors. Exposures can be easily measured and controlled. In contrast, the

---

[2] See Brewer (1997), Minimizing Operations Risk, in R. Schwartz & C. Smith (eds.), *Derivatives Handbook*, New York: Wiley.

link between risk factors and the likelihood and size of operational losses is not so easy to establish. Here, the line of causation runs through internal controls.

Third, very large operational losses, which can threaten the stability of an institution, are relatively rare (thankfully so). This leads to a very small number of observations in the tails. This "thin tails" problem makes it very difficult to come up with a robust "value for operational risk" at a high confidence level. As a result, there is still some skepticism as to whether operational risk can be subject to the same quantification as market and credit risks.

EXAMPLE 24.11: FRM EXAM 2002—QUESTION 102

Capital is used to protect the bank from which of the following risks?

a. Risks with an extreme financial impact
b. High-frequency low-loss events
c. Low-frequency risks with significant financial impact
d. High-frequency uncorrelated events

EXAMPLE 24.12: FRM EXAM 2001—QUESTION 49

Which of the terms below is used in the insurance industry to refer to the effect of a reduction in the control of losses by an individual who is insured because of the protection provided by insurance?

a. Control trap
b. Moral hazard
c. Adverse selection
d. Control hazard

EXAMPLE 24.13: FRM EXAM 2003—QUESTION 48

Which of the options below does *not* describe a problem faced by banks when purchasing insurance as a hedge against operational risk?

a. The fact that the loss reimbursement period can take several years
b. The credit rating of insurers
c. The different perspective of operational risk between banks and insurers
d. Not having an operational VAR

EXAMPLE 24.14: FRM EXAM 2005—QUESTION 48

Insurance is an effective tool to transfer which of types of operational risks?

a. High frequency, low severity
b. Low frequency, high severity
c. Operational losses whose magnitude is affected by the actions of the company
d. Operational losses for which insurance companies only sell policies with low limits

EXAMPLE 24.15: FRM EXAM 2005—QUESTION 52

Which of the following statements are valid about hedging operational risk?

I. A primary disadvantage of insurance as an operational risk management tool is the limitation of policy coverage.
II. If an operational risk hedge works properly, a firm will avoid damage to its reputation from a high-severity operational risk event.
III. While all insurance contracts suffer from the problem of moral hazard, deductibles help reduce this problem.
IV. Catastrophe (cat) bonds allow a firm to hedge operational risks associated with natural disasters.
    a. I, III, and IV only
    b. I, II, and IV only
    c. II and III only
    d. III and IV only

24.5 ANSWERS TO CHAPTER EXAMPLES

Example 24.1: FRM Exam 2004—Question 39

c. Damaged reputation due to a failed merger is a business risk. Also, reputational risk is not a type of operational loss.

Example 24.2: FRM Exam 2002—Question 133

b. Answers a., c., and d. correspond to the market risk of stocks, fixed-income securities, and currencies, respectively. Lawsuits, on the other hand, are part of operational risk.

Example 24.3: FRM Exam 2003—Question 65

c. Choices a., b., and d. are operational losses. Answer c. is the result of a bet on volatility, which is market risk.

Example 24.4: FRM Exam 2007—Question 56

b. Statement a. represents external fraud, which is included in operational risk. Statement c. represents a systems failure. Statement d. is a failure in internal processes.

Example 24.5: FRM Exam 2007—Question 139

c. This is a situation where the model price is significantly different from the market price, which is model risk. Liquidity risk could also explain part of the difference, but this is less likely to be the case given the emphasis on the complexity of the instrument.

Example 24.6: FRM Exam 2003—Question 50

c. The top-down approach is based on historical information from earnings volatility, so answers a. and b. are incorrect. Answer d. is also incorrect because the top-down approach is not based on specific mappings.

Example 24.7: FRM Exam 2007—Question 138

b. Loss severity distributions are bounded by zero but should include very large losses. So, they are asymmetrical with long right tails.

Example 24.8: FRM Exam 2000—Question 64

a. Constructing the operational loss requires the probability, or frequency, of the event as well as estimates of potential loss sizes. Answer b. is wrong as measurement of operational risk is still developing. Answer c. is wrong as the business unit is also responsible for controlling operational risk. Answer d. is wrong as losses can occur due to a combination of operational and market or credit risks.

Example 24.9: FRM Exam 2003—Question 33

a. Because VAR should include EL, there is no need to compute EL separately. The table shows that the lowest loss, such that the cumulative probability is 95% or more, is $100,000.

Loss	Probability		Cumulative
0	0.5	= 0.500	50.0%
1,000	0.3 × 0.6	= 0.180	68.0%
2,000	0.2 × 0.6 × 0.6	= 0.072	75.2%
10,000	0.3 × 0.3	= 0.090	84.2%
11,000	0.2 × 0.6 × 0.3 × 2	= 0.072	91.4%
20,000	0.2 × 0.3 × 0.3	= 0.018	93.2%
100,000	0.3 × 0.1	= 0.030	96.2%
101,000	0.2 × 0.1 × 0.6 × 2	= 0.024	98.6%
110,000	0.2 × 0.1 × 0.3 × 2	= 0.012	99.8%
200,000	0.2 × 0.1 × 0.1	= 0.002	100.0%

Example 24.10: FRM Exam 2006—Question 118

a. Operational loss distributions do have heavier tails than the lognormal, so statement I. is correct. Statements II. and III. are correct as well. Market risk distributions are typically short-term and can involve continuous distribution, as opposed to operational losses, which are discrete. Statement IV. is incorrect because the confidence level is set by the user.

Example 24.11: FRM Exam 2002—Question 102

c. Capital is supposed to absorb risks that have significant financial impact on the firm. Risks with extreme financial impact, such as systemic risk, cannot be absorbed by capital alone, so answer a. is wrong. Low-loss events are unimportant, so b. is wrong. Uncorrelated events tend to diversify, so d. is wrong.

Example 24.12: FRM Exam 2001—Question 49

b. Moral hazard arises when insured individuals have no incentive to control their losses because they are insured.

Example 24.13: FRM Exam 2003—Question 48

d. Answers a., b., and c. describe problems arising from the purchase of insurance against operational risk. This is irrespective of whether the bank has an operational VAR model.

Example 24.14: FRM Exam 2005—Question 48

b. The purpose of insurance is to reimburse large losses, or operational risk events with high severity. Answer c. is incorrect because this type of moral hazard should result in much higher premiums.

Example 24.15: FRM Exam 2005—Question 52

a. All the statements are valid, except for II. Even if a firm implements a hedge or purchases insurance, the news of a large operational loss will still damage its reputation.

APPENDIX: CAUSAL NETWORKS

Causal networks explain losses in terms of a sequence of random variables. Each variable itself can be due to a combination of other variables. For instance, settlement losses can be viewed as caused by a combination of (1) exposure and (2) time delay. In turn, exposure depends on (1) the value of the transaction and (2) whether it is a buy or a sell. Next, the causal factor for time delay can be chosen as (1) the exchange, (2) the domicile, (3) the counterparty, (4) the product, and (5) daily volume.

These links are displayed through graphical models based on process work flows. One approach is the Bayesian network. Here each node represents a random variable and each arrow represents a causal link.

Causes and effects are related through conditional probabilities, an application of Bayes' theorem. For instance, suppose we want to predict the probability of a settlement failure, or *fail*. Set $y = 1$ if there is a failure and $y = 0$ otherwise. The causal factor is, say, the quality of the back-office team, which can be either good or bad. Set $x = 1$ if the team is bad. Assume there is a 20% probability that the team is bad. If the team is good, the conditional probability of a fail is $P(y = 1 \mid x = 0) = 0.1$. If the team is bad, this probability is higher, $P(y = 1 \mid x = 1) = 0.7$. We can now construct the unconditional probability of a fail, which is

$$P(y = 1) = P(y = 1 \mid x = 0)P(x = 0) + P(y = 1 \mid x = 1)P(x = 1) \qquad (24.5)$$

which is here $P(y = 1) = 0.1 \times (1 - 0.20) + 0.7 \times 0.20 = 0.22$. Armed with this information, we can now evaluate the benefit of changing the team from bad to good through training, for example, or through new hires. Or we could assess the probability that the team is bad given that a fail has occurred. Using Bayes' rule, this is

$$P(x = 1 \mid y = 1) = \frac{P(y = 1, x = 1)}{P(y = 1)} = \frac{P(y = 1 \mid x = 1)P(x = 1)}{P(y = 1)} \qquad (24.6)$$

which is here $P(x = 1 \mid y = 1) = \frac{0.7 \times 0.20}{0.22} = 0.64$. In other words, the probability that the team is bad has increased from 20% to 64% based on the observed fail. Such an observation is useful for process diagnostics.

Once all initial nodes have been assigned probabilities, the Bayesian network is complete. The bank can now perform Monte Carlo simulations over the network, starting from the initial variables and continuing to the operational loss, to derive a distribution of losses.
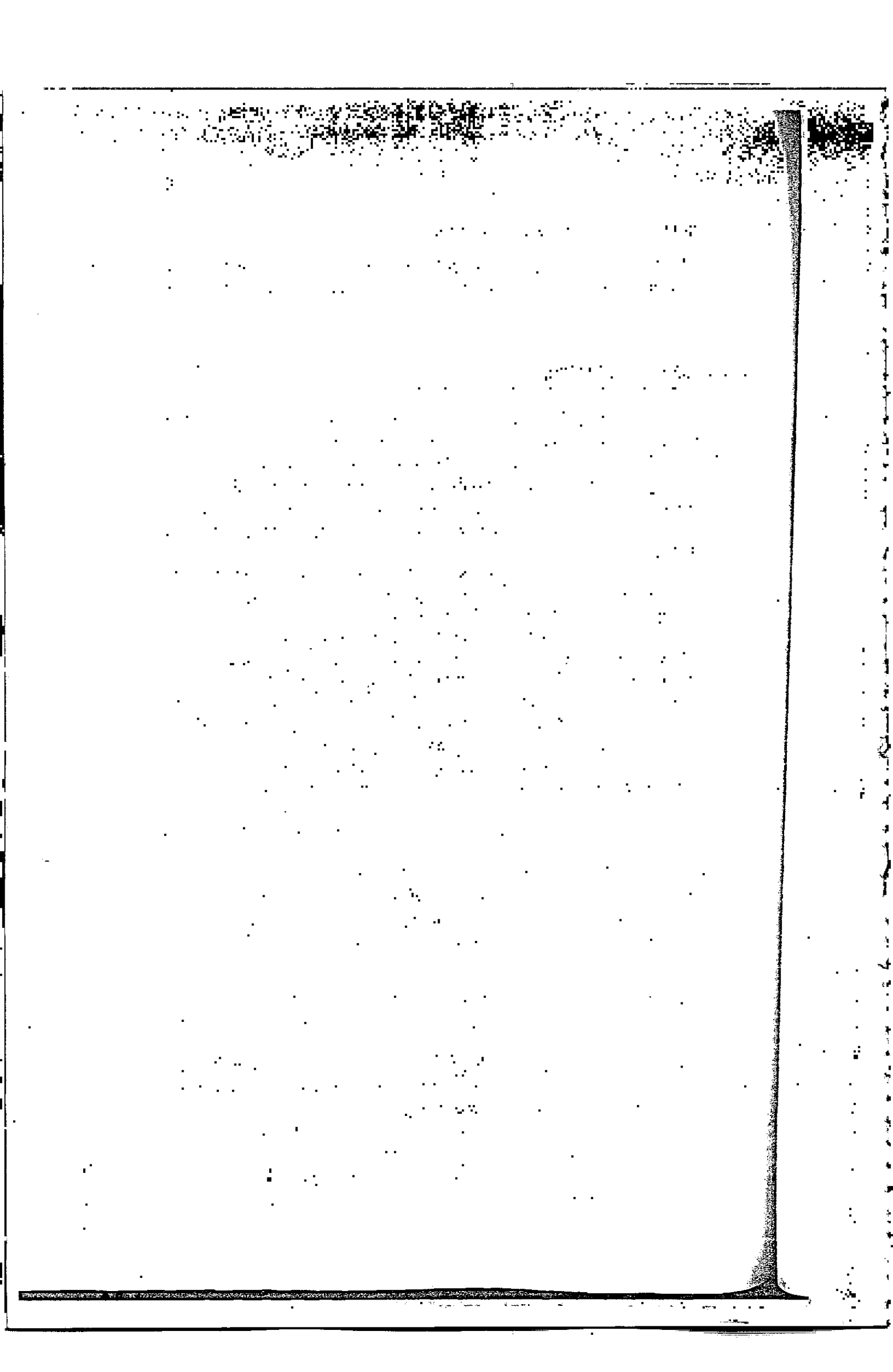
Liquidity Risk

Liquidity risk is an important source of financial risk, as we have witnessed in the latest credit crisis. The crisis of confidence that started with subprime losses suddenly accelerated after the Lehman bankruptcy. Many debt holders refused to roll over their investments, creating massive funding problems for financial institutions. These problems were compounded by their difficulties in selling assets to meeting funding needs.

Liquidity risk, unfortunately, is less amenable to formal risk measurement, unlike market risk, credit risk, and operational risk. This is why the Basel Committee did not institute formal capital charges against liquidity risk. Yet, it stated that "Liquidity is crucial to the ongoing viability of any banking organization. Banks' capital positions can have an effect on their ability to obtain liquidity, especially in a crisis."[1] Thus, it is crucial for financial institutions to assess, monitor, and manage their liquidity risk.

Section 25.1 describes sources of liquidity risk, which involve both asset liquidity risk and funding risk. Section 25.2 analyzes asset liquidity risk. The ability to liquidate assets to generate cash depends on market conditions, including bid–ask spreads and market impact, as well as the liquidation time horizon. To some extent, VAR can be expanded into a liquidity-adjusted VAR. Section 25.3 then analyzes funding liquidity risk. This is illustrated using the example of Northern Rock, the failed British bank. Funding claims can be evaluated by examining the structure of liabilities as well as off-balance sheet items. Finally, Section 25.4 discusses how banks can assess and control liquidity risk, primarily using gap analysis. It highlights the importance of contingency funding plans and illustrates disclosures about liquidity risk management.

25.1 SOURCES OF LIQUIDITY RISK

Lack of liquidity can cause the failure of an institution, even when it is technically solvent, i.e., when the value of its assets exceeds that of liabilities. We will see in Chapter 28 that commercial banks have an inherent liquidity imbalance between their assets (long-term loans) and their liabilities (retail deposits and capital market

---

[1] Paragraph 741 in BCBS (2006), *International Convergence of Capital Measurement and Capital Standards*, Basel: BIS.

debt). As a result, a crisis of confidence might lead to depositors demanding their money right away. Even if the bank has sufficient assets to cover deposits, it might not be able to liquidate its assets fast enough, or at reasonable prices, to meet the redemption requests. Similarly, hedge funds need to manage carefully the liquidity risk inherent in some balance sheets.

Liquidity risk consists of both asset liquidity risk and funding liquidity risk. The Committee of European Banking Supervisors (CEBS) provides these definitions.[2]

- Asset liquidity risk, also called market/product liquidity risk. This is the risk that a position cannot easily be unwound or offset at short notice without significantly influencing the market price, because of inadequate market depth or market disruption.
- Funding liquidity risk. This is the current or prospective risk arising from an institution's inability to meet its liabilities and obligations as they come due without incurring unacceptable losses.

These two types of risk interact with each other if the portfolio contains illiquid assets that must be sold at distressed prices to meet funding requirements.

25.2 ASSET LIQUIDITY RISK

25.2.1 Assessing Asset Liquidity Risk

To evaluate asset liquidity risk, we start with a characterization of market conditions for the asset to be traded. The bid–ask spread measures the round-trip transaction cost of buying and selling an amount within normal market size (NMS). If $P(\text{ask})$ is the ask price, $P(\text{bid})$ the bid price, and $P(\text{mid}) = [P(\text{ask}) + P(\text{bid})]/2$ the mid price, the spread is defined in relative terms as

$$S = \frac{[P(\text{ask}) - P(\text{bid})]}{P(\text{mid})} \qquad (25.1)$$

Assets with good liquidity will have tight bid–ask spreads. Tightness is a measure of the divergence between actual transaction prices and quoted mid-market prices. Liquid assets are also characterized by good depth, which is a measure of the volume of trades possible without affecting prices too much (e.g., at the bid/offer prices). This is in contrast to thinness.

For larger transactions, asset liquidity can be assessed by a price-quantity function, called market impact, which describes how the price is affected by the quantity transacted. Sometimes, this is called endogenous liquidity, meaning that the price drop depends on the size of the position. In contrast, positions within normal market sizes are characterized by exogenous liquidity.

---

[2] CEBS (2008), *Second Part of CEBS's Technical Advice to the European Commission on Liquidity Risk Management*, London. Available at http://www.c-ebs.org.

When selling a large block of an asset in a liquid market, prices may drop temporarily but should recover quickly. Resiliency is a measure of the speed at which price fluctuations from trades are dissipated.

For liquid assets, such as the Treasury market, this function is rather flat, meaning that large volumes of transactions do not affect prices much. For instance, one can generally transact $10 million of a Treasury bond at a cost of one-half the bid–ask spread of 0.10%, which translates into a cost of $10,000,000 × 0.10%/2 = $5,000, which is very low.

In contrast, illiquid assets are those where spreads are wide and where transactions can quickly affect prices. For example, bank loans are traded over-the-counter and can have spreads as wide as 10%. A sale of $10 million would then push prices down by 5%, which is a cost of $10,000,000 × 10%/2 = $500,000, which is much higher than in the previous example. A sale of twice this amount might incur an even larger price drop, such as 8%, to clear the market. Thus, the prices of illiquid assets are more affected by current demand and supply conditions. As a result, they are usually more volatile than liquid assets, provided trading occurs.

The latter example also shows that liquidity is a function of the time horizon. If the price-quantity function is steep, an immediate sale will force price down by a large amount. A patient investor, on the other hand, would fetch a better price by splitting the sale order over several days, thereby incurring a lower market impact.

Figure 25.1 compares the price-quantity functions for a liquid and an illiquid asset. For the liquid asset, the bid–ask spread is tight, the market has more depth, implying larger normal market sizes, and the lines representing the market impact have lower slopes.

Generally, assets that have greater trading volumes are more liquid. Trading volume reflects differences of opinion across investors but also depends on the
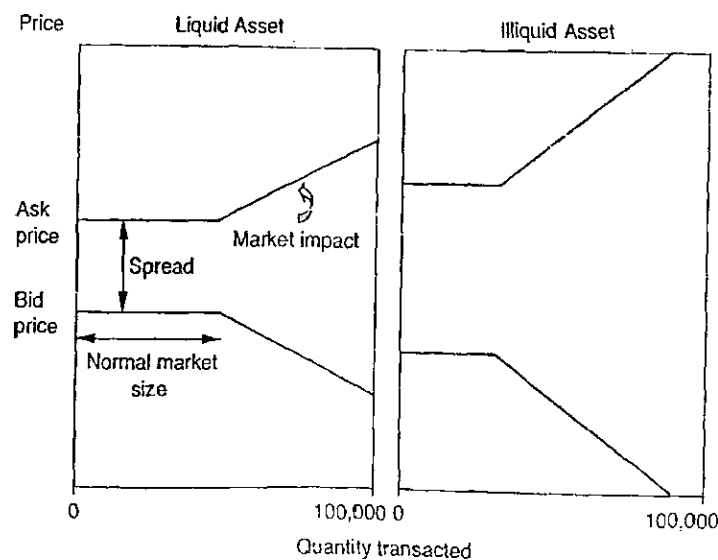
FIGURE 25.1 Comparison of Liquid and Illiquid Assets

presence of active speculators. Hedge funds, in particular, actively trade in many markets, which increases market liquidity.

Assets that are simple to price are also more liquid. At one extreme are Treasury bonds with fixed coupons, which are simple instruments and therefore easy to evaluate. At the other would be structured notes with complicated payoffs, which are harder for participants to evaluate and hedge. As a result, spreads on such notes will be much wider than on T-bonds.

Liquidity varies across asset classes and can be security-specific. Securities that have greater outstanding amounts or are issued more recently are generally more liquid. On-the-run securities are those that are issued most recently and hence are more active and liquid. Other securities are called off-the-run. Consider, for instance, the latest issued 30-year U.S. Treasury bond. This is on-the-run until another 30-year bond is issued, at which time it becomes off-the-run. Compared to the newest on-the-run, however, these two securities have the same credit risk (i.e., that of default by the U.S. government) and market risk (i.e., both have maturities close to 30 years). Because they are so similar, their yield spread must be a liquidity premium.

Asset liquidation costs also depend on asset fungibility. Contracts traded on centralized exchanges, such as futures or common stocks, easily can be resold to the highest bidder and therefore are fungible. On the other hand, privately negotiated derivatives require the agreement of the original counterparty to unwind the trade. In this situation, the counterparty may demand a discount to cancel the position.

In summary, asset liquidity risk depends on several factors: (1) market conditions (bid-ask spreads and market impact); (2) liquidation time horizon; (3) asset and security type; and (4) asset fungibility.

Illiquidity can also be market-wide and time-varying. Large-scale changes in market liquidity seem to occur on a regular basis, including the bond market rout of 1994, the Russian/LTCM crisis of 1998, and the credit crisis that started in 2007. Such crises are characterized by a flight to quality, which occurs when there is a shift in demand away from low-grade securities toward high-grade securities, in particular government bonds. The low-grade market then becomes illiquid with depressed prices. This is reflected in an increase in the yield spread between corporate and government issues.

25.2.2 Liquidity-Adjusted VAR

Asset liquidity risk is less amenable to formal measurement than traditional market risk. Illiquidity can be loosely factored into VAR measures, by increasing the horizon or by selectively increasing volatilities. These adjustments, however, are mainly adhoc.

We can attempt to incorporate the effect of bid/ask spreads in risk measures. When the spread S is fixed, liquidity-adjusted VAR (LVAR) can be defined as:

$$LVAR = VAR + L_1 = W\left[\alpha\sigma + \frac{1}{2}(S)\right] \qquad (25.2)$$

where W is the initial wealth, or portfolio value. If VAR is to be measured from zero (relative to the initial portfolio value), instead of away from the mean, we need to subtract μ from $\alpha\sigma$.

For instance, assume we have $10 million invested in a 30-year Treasury bond, with daily volatility of $\sigma = 1\%$ and spread of $S = 0.10\%$. The one-day LVAR at the 95% confidence level is

$$\$10,000,000 \left[\left(1.645 \times 0.01 \right) + \frac{1}{2} \left(0.0010 \right) \right]$$

$$= \$164,500 + \$5,000 = \$169,500$$

Here, this correction term is small. In contrast, the correction term is $500,000 for the bank loan in the previous example.

If bid–ask spreads vary substantially, Equation (25.2) can be adjusted to account for the worst increase in spread at some confidence level. The distribution of the spread can be described by its mean $\bar{S}$ and standard deviation σ_S. The worst-case LVAR is then:

$$\text{LVAR} = \text{VAR} + L_2 = W \left[\alpha\sigma + \frac{1}{2} \left(\bar{S} + \alpha'\sigma_S \right) \right] \tag{25.3}$$

By adding up the worst-case losses for market and liquidity risk, this effectively assumes a high correlation between these two risk drivers.

In practice, estimating the distribution of spreads is a challenge. Spreads tend to be stable for long periods and then explode in periods of crisis. Therefore, the distribution of spreads is highly non-normal. To measure risk at the top level, the risk manager also needs estimates of correlations across spreads. In addition, this analysis assumes that the quantity transacted is within normal market sizes. Otherwise, an immediate forced sale of a large quantity would also incur market impact.

25.2.3 Illiquidity and Risk Measures

Asset illiquidity poses special problems for risk measurement. In illiquid markets, fewer trades imply that prices do not move much. As a result, prices at the end of a reporting period generally will not represent market-clearing transactions and will tend to be sluggish in the absence of regular trading. This creates downward biases in measures of volatility and correlations with other asset classes.

In addition, news slowly impacts prices, creating positive autocorrelation in returns, which invalidate the square root of time rule when extrapolating VAR to longer horizons. These effects are further discussed in Section 17.4.2, in the context of hedge funds.

EXAMPLE 25.1: FRM EXAM 2003—QUESTION 15

Which of the following statements regarding liquidity risk is *correct*?

a. Asset liquidity risk arises when a financial institution cannot meet payment obligations.

b. Flight to quality is usually reflected in a decrease in the yield spread between corporate and government issues.

c. Yield spread between on-the-run and off-the-run securities mainly captures the liquidity premium, and not the market and credit risk premium.

d. Funding liquidity risk can be managed by setting limits on certain asset markets or products and by means of diversification.

EXAMPLE 25.2: FRM EXAM 2002—QUESTION 36

The following statements compare a highly liquid asset against an (otherwise similar) illiquid asset. Which statement is most likely to be *false*?

a. It is possible to trade a larger quantity of the liquid asset without affecting the price.

b. The liquid asset has a smaller bid–ask spread.

c. The liquid asset has higher price volatility since it trades more often.

d. The liquid asset has higher trading volume.

EXAMPLE 25.3: FRM EXAM 2007—QUESTION 78

A mutual fund investing in common stocks has adopted a liquidity risk measure limiting each of its holdings to a maximum of 30% of its 30-day average value traded. If the fund size is USD 3 billion, what is the maximum weight that the fund can hold in a stock with a 30-day average value traded of USD 2.4 million?

a. 24.00%

b. 0.08%

c. 0.024%

d. 80.0%

EXAMPLE 25.4: FRM EXAM 2000—QUESTION 74

In a market crash the following are usually *true?*

I. Fixed-income portfolios hedged with short Treasury bonds and futures lose less than those hedged with interest rate swaps given equivalent durations.
II. Bid-offer spreads widen because of lower liquidity.
III. The spreads between off-the-run bonds and benchmark issues widen.
    a. I, II, and III
    b. II and III
    c. I and III
    d. None of the above

EXAMPLE 25.5: FRM EXAM 2007—QUESTION 116

You are holding 100 Wheelbarrow Company shares with a current price of $50. The daily mean and volatility of the stock return is 1% and 2%, respectively. VAR should be measured relative to the initial wealth. The bid–ask spread of the stock varies over time. The daily mean and volatility of the spread is 0.5% and 1%, respectively. Both the return and spread are normally distributed. Calculate the daily liquidity-adjusted VAR (LVAR) at a 99% confidence level.

    a. USD 254
    b. USD 229
    c. USD 325
    d. USD 275

25.3 FUNDING LIQUIDITY RISK

25.3.1 Indicators of Liquidity Risk

Liquidity risk has been a major risk factor in the credit crisis that started in 2007. As commercial and investment banks started to accumulate losses, initially due to subprime asset-backed securities, banks were reluctant to lend in fear of counterparty default.

Conditions in money markets can be gauged, for example, by comparing the three-month Treasury-bill rate, the three-month London interbank offer rate

(LIBOR), and the overnight federal funds rate.[3] To ensure comparability, all rates are in U.S. dollars. The T-bill rate has no credit risk—other than that of the U.S. government. In contrast, LIBOR and fed funds are for unsecured loans.

The difference between LIBOR and fed funds is a term spread. It can be viewed as the price of an option to call a loan. A bank that has lent overnight can choose not to renew the loan if bad news were to strike; conversely, a bank committed to a three-month loan has no such option. As usual, the value of an option increases in uncertain times, which explains why this term spread has sharply increased.

The credit spread between Eurodollar LIBOR and Treasuries is known as the TED spread. This reflects expected credit losses as well as a liquidity risk premium.

Figure 25.2 describes the behavior of these interest rates during 2007 and 2008. The sharp reduction in the fed funds rate shows that the Federal Reserve has aggressively eased monetary policy. Treasury yields have correspondingly gone down as well. LIBOR rates, however, have stayed stubbornly high, reflecting tight conditions in credit markets. In particular, the TED spread, which is usually around 25bp, has widened sharply, exceeding 500bp after the Lehman bankruptcy on September 15, 2008. Firms with a more shaky credit rating have had to face even higher rates—that is, when they were able to obtain any funding at all.
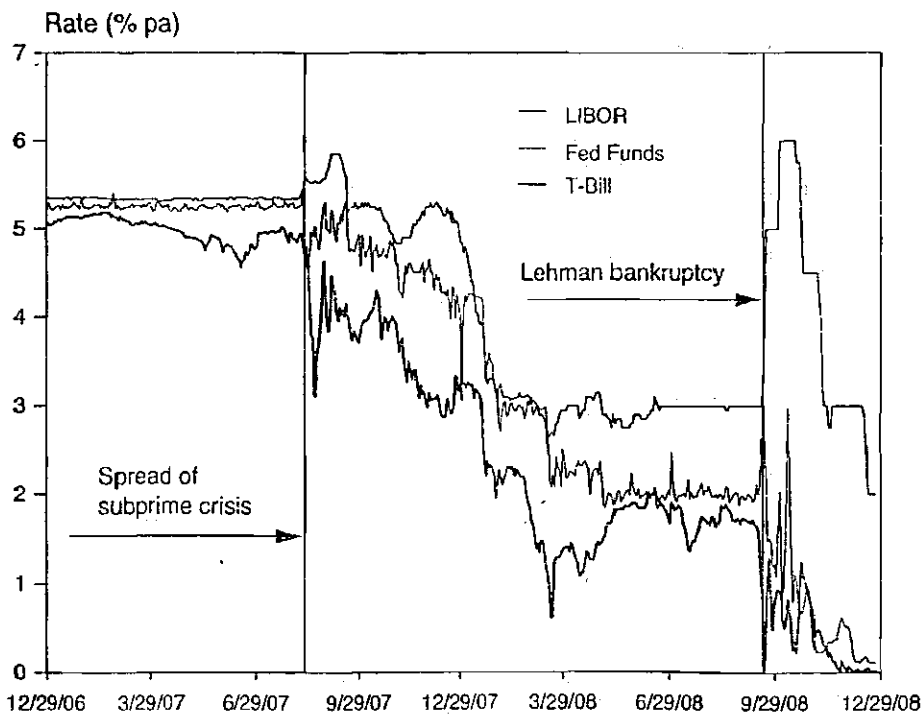
FIGURE 25.2 Comparison of Short-Term Dollar Interest Rates

---

[3] Fed funds are unsecured loans of reserve balances at the Federal Reserve Bank that banking institutions make to one another, usually overnight. The rate at which these transactions occur is called the fed funds rate. The central bank sets a target level for the fed funds rate, which is its primary tool for monetary policy.

25.3.2 Assessing Funding Liquidity Risk

The Basel Committee provides a comprehensive definition of this risk.[4]

> Funding liquidity risk *is the risk that the firm will not be able to meet efficiently both expected and unexpected current and future cash flow and collateral needs without affecting either daily operations or the financial condition of the firm.*

The consequences of this risk can be illustrated by the Northern Rock example.

Example: Northern Rock's Liquidity Risk

Northern Rock (NR) is a bank that was counted among the top five mortgage lenders in Britain. As shown in the table, total assets added up to 113.5 billion British pounds (BP) as of June 2007, about BP 81 billion of which was funded through capital markets, and only BP 30 billion through customer deposits. NR's business model was unusually reliant on funding from capital markets instead of retail deposits. Capital market funding, however, is more volatile than retail deposits. The bank had used this unusual structure to fuel its fast growth.

Northern Rock's Balance Sheet			
Assets		Debt	
Loans	96.7	Retail deposits	30.1
Cash	0.8	Debt securities	71.0
Securities	8.0	Other	10.1
Total	113.5	Total	111.2

Northern Rock stated that it was meeting FSA liquidity rules, which require enough liquidity for at least five business days. The bank had not expected such a long and widespread credit crisis, however.

During August 2007, NR started to run into difficulties rolling over its short-term debt and issuing securitized loans. Higher rates on new capital, when available, started to squeeze margins, leading to a free-fall in the bank's share price. Regulators believed that the bank was still solvent, however. On September 13, it was announced that the Bank of England had granted emergency financial support to Northern Rock. This news started a bank run. Because deposits are only partially insured in Britain, depositors panicked and withdrew billions of deposits in the following days. On Monday, September 17, the Chancellor of the Exchequer announced that the government would fully guarantee all deposits.

By the end of the year, the bank had been unable to roll over BP 8 billion in short-term debt and had lost BP 15 billion in customer accounts. The loan from the Bank of England had grown to BP 27 billion. After two unsuccessful bids to sell Northern Rock, it was nationalized on February 22, 2008.

Northern Rock was a victim of funding liquidity risk, as it had funded long-term loans by short-term debt that it could not roll over.

---

[4] BCBS (2008), *Principles for Sound Liquidity Risk Management and Supervision*, Basel: BIS.

Table 25.1 provides a general framework for assessing liquidity risk. Funding liquidity risk arises from the liability side, either for on–balance sheet or off–balance sheet items. Liabilities can be classified into *stable* or *volatile*, where these terms refer to the predictability of cash flows.

For a public corporation, equity is stable.[5] Financial institutions can manage their equity liquidity profile by means of dividend policies, share repurchases, and new issues.

Next, we turn to the debt, which can be divided into unsecured and secured funding. For instance, Northern Rock had issued BP 45.7 billion in securitized notes and 8.1 billion in covered bonds. Investors should be more willing to provide funds secured by assets. In contrast, unsecured funding is subject to default risk by the issuer.

Within the unsecured funding category, retail deposits are more stable than capital market instruments.[6] For example, under conditions of stress, investors in money market instruments may demand higher compensation for risk, or require to roll over their investment for shorter maturities, or even refuse to extend financing at all.

Among off–balance sheet liabilities, loan commitments, letters of credit, and financial guarantees provided by a bank will create a contingent claim on liquidity if drawn. Derivatives may also create cash flow needs if counterparties demand more collateral as the position moves out-of-the-money, or if contracts contain credit trigger clauses that require additional collateral in the case of a downgrade. Indeed special purpose vehicles (SPVs) can also create contingent liquidity exposures. Some structures, such as bank-sponsored conduits, have explicit backing

TABLE 25.1   Managing Bank Liquidity Risk

Balance Sheet	
Assets	**Debt**
• Highly liquid (cash...)	• Unsecured funding
• Other, unencumbered	• Retail deposits
(pledgeable as collateral)	• Capital markets
• Other, encumbered	• Secured funding
	Equity
	• New stock issues

Off–Balance Sheet	
Assets	**Liabilities**
• Derivatives	• Derivatives
• Credit lines purchased	• Guarantees provided
	• Commitments
	• SPVs

[5] On the other hand, hedge funds need to worry about redemptions from equity investors, as discussed in Chapter 17.
[6] With the advent of Internet-based banking, however, deposits can move faster across banks.

from the issuing bank and can draw upon liquidity lines if the SPV is not able to roll over its debt. Other structures, such as structured investment vehicles (SIVs), may not have explicit backing, but the bank nevertheless may choose to provide liquidity support for business or reputational reasons.

Focusing now on the asset side of the balance sheet, funding gaps can be met with asset sales. Cash or liquid assets provide a cushion that can be used immediately. Unencumbered securities, defined as those that do not have claim against them, can be sold, perhaps at a discount that reflects asset liquidity risk. Alternatively, they can be sold through a repurchase agreement as collateral against cash either with a private counterparty or with the central bank if permitted.

In addition, cash can be forthcoming from derivatives positions that move in-the-money. Institutions can also establish bank credit lines, which can be drawn upon in case of liquidity needs.

Finally, this on– and off–balance sheet information should be integrated with cash flows. In particular, the banking system has expanded securitization as a means to reduce assets on the balance sheet. During the recent credit crisis, however, banks were forced to postpone some securitization, leading to a buildup of a large loan inventory that had to be financed.

Example: AIG's Liquidity Risk

American International Group (AIG) is a global insurance conglomerate that once was among the largest public companies in the world. Due to its strong revenues and capital base, AIG long enjoyed a top credit rating of AAA. As a result, it had allowed its Financial Products division to take on increasingly large positions. It sold credit default swaps on the senior tranches of CDOs for which, because of its high credit rating, it did not have to post collateral.

On March 14, 2005, AIG's CEO was forced to step down amid allegations of questionable business practices. The next day, its credit rating was downgraded to AA+. The downgrade triggered provisions requiring posting $1.2 billion in additional collateral for its swaps. At the time, this was still manageable, given that AIG had equity of about $80 billion.

Its CDS portfolio, however, continued to grow, reaching $500 billion. As the subprime crisis started to unfold, the CDOs tranches lost value quickly. AIG announced that it had lost $13 billion in the first half of 2008.

On September 15, 2008 S&P lowered AIG's credit rating from AA– to A–. As a result, AIG was required to post an additional $20 billion in collateral, which it did not have. Because a collapse of AIG would have had systemic consequences, the U.S. government stepped in and provided a $85 billion loan. In October, this amount had to be increased by $38 billion. In November, the U.S. Treasury invested an additional $40 billion in newly issued AIG senior preferred stock under the Troubled Asset Relief Program (TARP). This was the largest public bailout of a private company. Apparently, leaders at the Financial Products division had failed to prepare for downgrades in AIG's credit rating.

25.4 MANAGING LIQUIDITY RISK

25.4.1 Steps in Liquidity Risk Management

Liquidity risk management requires robust internal governance, implemented by adequate tools to identify, measure, monitor, and manage liquidity risk. The Board of Directors is ultimately responsible for the institution's liquidity strategy.

While there is no single measure of liquidity risk, a range of metrics can be used to assess liquidity risk. Liquidity risk management starts with operational liquidity, which lays out the *daily* payment queue, forecasting all cash inflows and outflows. This is no simple affair, however. In recent years, improvements in the design of payment and settlement systems, e.g., real-time gross settlement systems and the CLS bank for netting foreign currency payments, have compressed the time between payments. While such systems reduce credit risk and operational risk, they do place more strain on liquidity management.

The next step is *tactical* management, which assesses access to unsecured funding sources and the liquidity characteristics of the asset inventory. This involves an evaluation of asset liquidity risk.

Finally, this information is integrated in a *strategic* perspective, which starts from current assets and liabilities as well as off–balance sheet items. This information is used to build a funding matrix, which details the funding requirements for various maturities. Any funding gap should be covered by plans to raise additional funds, either through borrowing or by asset sales.

25.4.2 Funding Gaps

Table 25.2 gives a hypothetical example in a pure run-off mode, i.e., with no new business nor rollover of funding. Here, the funding matrix starts from balance sheet items, loans, retail deposits, other short-term debt, and long-term debt.

These items have cash flows that are either fixed or stochastic and maturities that are either fixed or stochastic. For example, coupons and amortization

TABLE 25.2   Example of Funding Gap Analysis

				Time Profile				
	Balance	O/N	7D	14D	1M	3M	1Y	Cumulative
Funding matrix								
Loans	100	5	5	3	15	5	5	38
Retail deposits	−50	−5	−5	−5	−8	−5	−5	−33
Short-term debt	−30	−10	−5	−5	−5	−5	0	−30
Long-term debt	−30	0	0	0	−5	0	0	−5
Total: Funding gap		−10	−5	−7	−3	−5	0	−30
Gap closure								
Cash	5	5	0	0	0	0	0	5
Unencumbered secs.	20	10	8	2	0	0	0	20
Total		15	8	2	0	0	0	25
Net funding gap		5	3	−5	−3	−5	0	−5
Cumulative		5	8	3	0	−5	−5	

payments of fixed-rate debt have both fixed cash flows and maturities. A second category are items such as floating rate loans and bonds where cash flows are stochastic but maturities are deterministic. A third category are items such as callable bonds or loans with flexible redemption schedules where maturities are stochastic but cash flows are deterministic. In the last category are items with stochastic cash flows and maturities, such as retail deposits, drawdowns on committed credit lines and revolving loans. Stochastic cash flows or maturities require modeling based on market experience and product knowledge.

In Table 25.2, the initial balance of loans is 100. The table lays out the cash inflows from this amount for various maturity buckets. A portion of these loans normally will be paid back over the next year. Next are retail deposits, short-term and long-term debt, for which cash outflows are forecast over the horizon. Note that all of the short-term debt is expected to be repaid within one year. The total creates a time profile of the funding gap. In this case, the cumulative funding gap to one year is −30.

The next part of the table lays out gap closure items. For instance, cash can be used immediately to cover funding outflows. Unencumbered securities can be sold over time as a function of their asset liquidity risk. The sum of the funding gap and the gap closure items creates the expected net funding gap. The longer the period with a positive flow, the safer the bank. In this case, the survival period, which is the time until which the cumulative net funding gap becomes negative, is one month.

25.4.3 Stress Tests

Risk management is about dealing with the unexpected, however. Hence, institutions should also evaluate stress scenarios where cash flows deviate from their expected path and sources of funding are unexpectedly cut off.[7] Institutions should consider a broad range of scenarios, including institution-specific, country-specific, and market-wide scenarios. An example of country-specific scenario would be sudden restrictions on currency convertibility.

In addition, in situations of extreme stress, funding liquidity risk is likely to interact negatively with asset liquidity risk as it may become more difficult to sell assets under these conditions. Due to the self-fulfilling nature of reputation risk, an institution's perceived liquidity problems can undermine its ability to sell its assets at a reasonable cost.

25.4.4 Controlling Liquidity Risk

Liquidity risk can be controlled by various means, including reliance on more stable sources of funding and diversification across sources of funds, geographical location, and debt maturities. Similarly, asset liquidity risk can be controlled by setting limits on certain markets or products and by means of diversification. Funding gaps should also be subject to limits over various horizons. Some

---

[7] In theory, the entire distribution of cash flows could be estimated instead of a stress test, which is just a particular realization. This leads to liquidity-at-risk (LAR) models. The problem, however, is that liquidity behavior is difficult to model and that recent historical data may not be relevant.

regulators do require minimum levels of liquid assets, limits on maturity mismatches, or limits on the reliance on a particular funding source.

Another tool to control liquidity risk is to penalize business units or instruments that can generate claims on liquidity. According to the SSG (2008) report, institutions that performed better than others during the crisis had adopted a firm-wide perspective that explicitly accounted for liquidity risk.[8] These firms has charged business lines appropriately for building contingent liquidity exposures to reflect the cost of obtaining liquidity in a more difficult market environment. Institutions that did well also had effective management of funding liquidity, capital, and their balance sheet.

25.4.5 Contingency Funding Plans

The goal of contingency funding plans (CFP) is to establish a plan of action should one of the liquidity stress scenarios develop. In a crisis situation, management usually does not have much time to react, which is why a pre-established plan is useful. The CFP should define the trigger events, clear lines of responsibilities for implementation, and a plan for alternative sources of funding. It should also take into account the reputational effects of announcement of the execution of the funding plan.

Generally, public disclosures about liquidity risk management procedures should help reassure investors that the institution has developed a process to deal with liquidity risk. A bank that is perceived as having liquidity is less likely to lose the confidence of funds providers. The following is an example of liquidity risk disclosures by Deutsche Bank.

Example: Deutsche Bank's Liquidity Management

Deutsche Bank (DB) is a leading German commercial bank. As of December 2007, it had 2,020 billion euros in assets and 329 billion euros in risk-weighted assets. With 37 billion euros in shareholder equity, its tier 1 capital ratio was 8.6%.

DB's liquidity risk management approach starts at the *intraday level* forecasting daily cash flows and factoring in access to central banks. It then covers *tactical* liquidity risk management dealing with the access to unsecured funding sources and the liquidity characteristics of its assets. For example, the bank generates 25% of its unsecured funding from retail deposits and 20% from capital markets. In terms of asset liquidity, the bank assigns liquidity values to different assets; it also holds a portfolio of 25 billion euros in highly liquid securities to protect against short-term liquidity squeezes. Finally, the *strategic* perspective comprises the maturity profile of all assets and liabilities on its balance sheet as well as its issuance strategy.

The bank employs stress testing and scenario analysis to evaluate the impact of sudden stress events on its liquidity position. The hypothetical events encompass

[8] See Senior Supervisor Group (2008), *Observations on Risk Management Practices during the Recent Market Turbulence*, Basel: BIS.

external shocks, such as market risk events, emerging market crises, and systemic shocks, as well as internal shocks, such as operational risk events and ratings downgrades (e.g., from AA to AA– for a 1-notch downgrade). Under each of these scenarios, the bank assumes that all maturing loans to customers will need to be rolled over and require funding, whereas rollover of liabilities will be partially impaired resulting in a funding gap. The bank then models the steps it would take to counterbalance the resulting net shortfall in funding, which includes selling assets and switching from unsecured to secured funding. For each scenario, the table shows the cumulative funding gap over an eight-week horizon, in billions of euros, and how much counterbalancing liquidity could be generated.

Scenario	Funding Gap	Gap Closure
Market risk	5.5	98.9
Emerging markets	27.7	117.1
Systemic shock	20.4	70.9
Operational risk	13.9	106.7
1-notch downgrade	28.1	129.3
3-notch downgrade	108.6	129.3

EXAMPLE 25.6: FRM EXAM 2007—QUESTION 57

You have been asked to review a memo on how market liquidity is affected by shocks to the financial system. Which of the following observations made in the memo is *incorrect*?

a. In periods of acute market stress, market liquidity typically increases in the most liquid markets, creating a self-correcting loop that will ultimately remove downward pressure on asset prices.

b. Evaporation of market liquidity is an important factor in determining whether and at what speed financial disturbances become financial shocks with potentially systemic threats.

c. Market shocks may not be reflected in marked-to-market portfolio values immediately for portfolios with illiquid assets. As a result, it is possible for market shocks to have delayed effects on financial institutions.

d. The impact of a market shock on the liquidity of a specific asset depends on the characteristics of the investors who own the asset.

25.5 IMPORTANT FORMULAS

Relative bid–ask spread: $S = \frac{[P(\text{ask}) - P(\text{bid})]}{P(\text{mid})}$

Liquidity-adjusted VAR (LVAR):

$$LVAR = VAR + L_1 = W[\alpha\sigma + \tfrac{1}{2}(S)]$$

Worst-case liquidity-adjusted VAR (LVAR):
$$LVAR = VAR + L_2 = W[\alpha\sigma + \tfrac{1}{2}(\bar{S} + \alpha'\sigma_S)]$$

25.6 ANSWERS TO CHAPTER EXAMPLES

Example 25.1: FRM Exam 2003—Question 15

c. The yield spread between on-the-run and off-the-run reflects a liquidity premium because the bonds are otherwise nearly identical. In answers a. and d., asset and funding risk should be interchanged. Finally, for b., a flight to quality increases the yield spread.

Example 25.2: FRM Exam 2002—Question 36

c. Compare two stocks. The liquid stock typically has higher trading volumes and lower bid-ask spreads, so b. and d. are true. It also has greater depth, meaning that large quantities can be traded without affecting prices too much, so a. is true. As a result, the remaining answer c. must be wrong. There is no necessary relationship between trading activity and volatility.

Example 25.3 FRM Exam 2007—Question 78

c. The maximum weight w is given by $3,000 \times w = 30\% \times 2.4$, or $w = 0.024\%$.

Example 25.4: FRM Exam 2000—Question 74

b. In a crash, bid offer spreads widen, as do liquidity spreads. Answer I. is incorrect because Treasuries usually rally more than swaps, which leads to *greater* losses for a portfolio short Treasuries than swaps.

Example 25.5: FRM Exam 2007—Question 116

a. The regular VAR relative to the initial portfolio value is $VAR = W(\alpha\sigma - \mu) = \$5,000(2.33 \times 2\% - 1\%) = \$183$. (Note that this estimate of the mean is abnormally high.) To this must be added $L_2 = \tfrac{1}{2}W(\bar{S} + \alpha'\sigma_S) = \tfrac{1}{2}\$5,000(0.5\% + 2.33 \times 1\%) = \$70.75$, for a total of $254.

Example 25.6: FRM Exam 2007—Question 57

a. Answer b. is correct, as proved by the events of 2007. Answer c. correctly states that the prices of illiquid assets reflect a delayed reaction to events. Answer d. explains that asset liquidity depends on investor positions, which is correct. An asset that is mainly owned by leveraged investors can experience a sharp swing in prices if the investors are forced to sell.

Firm-Wide Risk Management

This chapter turns to best practices for firm-wide management of financial risks. The financial industry has come to realize that risk management should be implemented on a firm-wide basis, across business lines and types of risk. This is due to a number of factors, including (1) increased exposures to more global sources of risk as institutions expand their operations, (2) interactions between risk factors, and (3) linkages in products across types of market risks as well as types of financial risks. These linkages make it important to consider correlations among risks and products.

Interactions between types of risk bear emphasis, as they are too often ignored. The industry has made great strides in recent years in the measurement of market and credit risk. Once measured, risk can be penalized, as with risk-adjusted return on capital (RAROC) measures. The danger is that this creates an incentive to move risk to areas where it is not well measured or controlled.

This explains the trend toward integrated, or firm-wide, risk management. Integrated risk management provides a consistent and global picture of risk across the whole institution. This requires measuring risk across all business units and all risk factors, using consistent methodologies, systems, and data.

Section 26.1 reviews different types of financial risks. Section 26.2 discusses the three pillars of global risk management: best-practices policies, methods, and infrastructure. Section 26.3 then turns to a description of organizational structures that are consistent with these best practices. Section 26.4 shows how traders can be controlled through compensation adjustment and limits. Finally, Section 26.5 shows how risk measures can be integrated in the measurement of performance of traders and business units through RAROC-type measures.

26.1 INTEGRATED RISK MANAGEMENT

26.1.1 Types of Risk

We first briefly review various types of financial risks.

- Market risk is the risk of loss due to movements in the level or volatility of market prices. This is covered in Chapters 10 to 15.
- Liquidity risk takes two forms, asset liquidity risk and funding liquidity risk. Asset liquidity risk, also known as market/product liquidity risk, arises when a

623

transaction cannot be conducted at prevailing market prices due to the size of the position relative to normal trading lots. Funding liquidity risk, also known as cash-flow risk, refers to the inability to meet payment obligations. Asset liquidity risk generally falls under the market risk management function. This is covered in Chapter 25.

- Credit risk is the risk of loss due to the fact that counterparties may be unwilling or unable to fulfill their contractual obligations. This is covered in Chapters 18 to 23.

- Operational risk is generally defined as the risk of loss resulting from failed or inadequate internal processes, systems, and people, or from external events. This is covered in Chapter 24.

Institutions should attempt to measure all of these risks in a consistent fashion and across the entire firm. Otherwise, risk will tend to flow to areas where it is least penalized, or with the weakest measures.

The previous chapter illustrated this point with liquidity risk during the recent credit crisis. Banks that did not properly assess this risk nor penalize business units for the claims on liquidity they generated ended up with much worse liquidity problems than others. Another example are the banks that had weak systems to measure the risk of senior tranches of subprime-backed debt. Because their systems showed very little risk, these institutions did not monitor large build-ups in these securities. During 2007, for example, UBS ended up with losses of $19 billion on these securities, which were found in the CDO warehousing book, in the trading book, in the liquid Treasury book, and in a hedge fund subsidiary.[1] Apparently, there was no monitoring of net nor gross concentrations of positions in this asset class at the firm-wide level.

26.1.2 Risk Interactions

Risk categories do not always fit into neat, separate silos. Operational risk can create market and credit risk, and vice versa. For instance, collateral payments in swaps decrease credit risk by marking to market on a regular basis but create a greater need for cash flow management, which increases operational and liquidity risk. The reverse can also occur as an operational failure, such as incorrect confirmation of a trade, can lead to inappropriate hedging or greater market risk. Incorrect data entry of swap terms can create incorrect market risk measurement as well as incorrect credit exposures.

Another important example is the interaction between market risk and credit risk. Wrong-way trades are those where market risk amplifies credit risk. Consider, for example, a swap between a bank and a speculator. If the bank loses money on the swap, credit risk is not an issue. On the other hand, if the bank makes a large profit on the swap, this must be at the expense of the speculator. If the loss to the other party is sufficiently large, the speculator could default precisely because of the swap. Therefore, such trades are inherently more dangerous than those where

---

[1] UBS (2008), *Shareholder Report on UBS's Write-Downs*, Zurich: UBS.

the counterparty is a hedger. This is because the loss on the swap to the hedger should be offset by a gain on the hedged position. As a result, such trades are safer for the bank.

This is easier said than done, however. In theory, interactions between different types of risk should be taken into account. In practice, banks that now report VAR estimates for market, credit, and operational risk most often simply add up the three risk measures to get an estimate of the bank's total risk. This consolidation, however, overstates the risk because it assumes that the worst loss will occur simultaneously across the three risk categories.

26.1.3 Illustration

Table 26.1 illustrates a firm-wide economic capital analysis reported by Deutsche Bank. The bank estimates separately the worst loss at a 99.98% level of confidence over one year for the three risk categories, market, credit, and operational risk. Market risk includes both trading risk as well as that of the banking book, e.g., the interest rate risk of loans and deposits. Nontrading market risk is sometimes referred to as asset/liability management, or mismatch, (ALM) risk. For trading risk, the traditional one-day, 99% VAR measures have been extrapolated to the parameters for economic capital.

Credit risk accounts for the largest fraction of the economic capital, as is typical of a commercial bank or universal bank with substantial lending activities. In latter years, operational risk is second in importance, before market risk, especially when considering trading risk only. Recent surveys have confirmed this ranking for a broad sample of global banks. In terms of importance, first comes credit risk, then operational risk, then market risk.[2] In contrast, market risk is relatively more important for investment banks and life insurance companies.

The table also shows that, during the first three years, the bank made no allowance for diversification effects across risk categories. This effectively assumes perfect correlations. Over time, the bank has refined its risk models. For 2004 and 2005, the bank estimated a diversification effect across credit and market risk.

TABLE 26.1 Firm-Wide Economic Capital—Deutsche Bank (Millions of Euros)

By source	2007	2006	2005	2004	2003	2002	2001	2000
(1) Credit risk	8,506	7,351	7,125	5,971	7,363	8,942	9,064	8,200
(2) Market risk	3,481	2,994	3,042	5,476	5,912	9,057	7,257	3,700
Trading	1,763	1,605	1,595	1,581	972	765		
Nontrading	1,718	1,389	1,447	3,895	4,940	8,292		
Diversification (1,2)			(563)	(870)				
(3) Operational risk	3,974	3,323	2,270	2,243	2,282	2,449	2,538	2,800
Diversification (1,2,3)	(2,651)	(2,158)						
Total	13,310	11,509	11,874	12,820	15,557	20,448	18,859	14,700
Actual core capital	28,320	23,539	21,898	18,727	21,618	22,742	24,803	23,504

[2] Kuritzkes, A., T. Schuermann, and S. Weiner (2003), Risk Measurement, Risk Management, and Capital Adequacy in Financial Institutions, *Working Paper*, Wharton.

For 2006 and 2007, the diversification effect was extended to all three categories. This is achieved by copula methods, which join the marginal distributions, and use Monte Carlo analysis to construct the firm-wide risk distribution. As of 2007, total economic capital was estimated at 13 billion euros. The bank had actual core capital of 28.3 billion euros, which is well above its estimate of economic capital.[3]

For all their apparent elegance, however, these models have limitations. These are complex models that require many assumptions and simplifications, as well as fitting distributions and parameters. This creates model risk. In addition, because the confidence level is so high, the estimates of economic capital must be rather imprecisely measured. Yet financial reports never provide any information on this. Finally, these measures inherit all of the drawbacks of internal credit risk models, which have not been tested on a full credit cycle.

EXAMPLE 26.1: FRM EXAM 2005—QUESTION 110

Which of the following problems are *not* serious obstacles to estimating a firm-wide risk measure?

a. Operational, market, and credit risks follow different statistical distributions.
b. Firms typically have different time horizons for their analysis of operational, market, and credit risks.
c. Operational risks are not financial risks, so that it is not possible to compute a return on operational risk.
d. All of the above.

EXAMPLE 26.2: FRM EXAM 2002—QUESTION 103

Consider a bank that wants to have an amount of capital so that it can absorb unexpected losses corresponding to a firm-wide VAR at the 1% level. It measures firm-wide VAR by adding up the VARs for market risk, operational risk, and credit risk. There is a risk that the bank has too little capital because

a. It does not take into account the correlations among risks.
b. It ignores risks that are not market, operational, or credit risks.
c. It mistakenly uses VAR to measure operational risk because operational risks that matter are rare events.
d. It is meaningless to add VARs.

---

[3] Core capital is mainly book equity, as explained in Chapter 29.

EXAMPLE 26.3: FRM EXAM 2006—QUESTION 109

Large banks typically allocate risk capital for credit, operational, and market/ALM risks. Which of the following statements ranks the typical amount of risk capital allocated to these different risks correctly?

a. Market/ALM risk requires more risk capital than credit risk.
b. Credit risk requires more risk capital than market/ALM risk, which requires more risk capital than operational risk.
c. Market/ALM risk requires more risk capital than operational risk but less than credit risk.
d. Credit risk requires more risk capital than operational risk, which requires more risk capital than market/ALM.

EXAMPLE 26.4: FRM EXAM 2005—QUESTION 33

Counterparty A is an American company with manufacturing operations in Indonesia and its main customers in the United States, while counterparty B is an American company that manufactures its goods domestically and exports solely to Indonesia. Which one of the following transactions with either counterparty will be a wrong-way exposure for a bank?

a. A five-year plain-vanilla IDR/USD cross-currency swap between the bank and counterparty A where the bank is USD interest rate receiver.
b. A five-year plain-vanilla IDR/USD currency option sold by the bank to counterparty A for it to buy IDR at a certain rate.
c. A five-year plain-vanilla IDR/USD cross-currency swap between the bank and counterparty B where the bank is USD interest rate receiver.
d. A five-year plain-vanilla IDR/USD currency option bought by the bank from counterparty B for the bank to buy IDR at a certain rate.

26.2 BEST PRACTICES REPORTS

Best practices in the industry have evolved from the lessons of financial disasters. Some well-publicized losses in the early 1990s led to the threat of regulatory action against derivatives.

Financial institutions then realized that it was in their best interests to promote a set of best practices to forestall regulatory action. This led to the Group of Thirty (G-30) report, which was issued in July 1993. The 1995 Barings failure

was followed by an in-depth report from the Bank of England in July. Similarly, the 1998 near-failure of Long-Term Capital Management (LTCM) was analyzed in a report produced by the Counterparty Risk Management Policy Group (CRMPG) in June 1999, followed by updates in 2005 and in 2008. These reports added to the collective wisdom about best practices.

26.2.1 The G-30 Report

The Group of Thirty (G-30) is a private, nonprofit association, consisting of senior representatives of the private and public sectors and academia. In the wake of the derivatives disasters of the early 1990s, the G-30 issued a report in 1993 that has become a milestone document for risk management.[4] The report provides a set of 24 sound management practices. The most important ones are summarized here.

- **Role of senior management.** *Dealers and end-users should use derivatives in a manner consistent with the overall risk management and capital policies approved by their boards of directors.... Policies governing derivatives use should be clearly defined, including the purposes for which these transactions are to be undertaken. Senior management should approve procedures and controls to implement these policies, and management at all levels should enforce them.*

   In other words, derivatives policies should be set by top management.
- **Marking-to-market.** *Dealers should mark their derivatives positions to market, on at least a daily basis, for risk management purposes.*

   In other words, marking to market is the most appropriate valuation technique. Countless mistakes have resulted when institutions valued instruments using a historical, accrual method.
- **Measuring market risk.** *Dealers should use a consistent measure to calculate daily the market risk of their derivatives positions and compare it with market risk limits. Market risk is best measured as "value at risk" using probability analysis based on a common confidence interval and time horizon.*

   This recommendation endorsed VAR as the most appropriate quantitative measure of market risk.
- **Stress simulations.** *Dealers should regularly perform simulations to determine how their portfolios would perform under stress conditions.*
- **Investing and funding forecasts.** *Dealers should periodically forecast the cash investing and funding requirements arising from their derivatives portfolios.*
- **Independent market risk management.** *Dealers should have a market risk management function, with clear independence and authority, to ensure that the following responsibilities are carried out: risk limits; stress tests; revenue reports; back-testing VAR; review of pricing models and reconciliation procedures.*
- **Independent credit risk management.** *Dealers and end-users should have a credit risk management function with clear independence and authority,....*

---

[4] Group of Thirty (1993), *Derivatives: Practices and Principles*, New York: Group of Thirty.

responsible for: approving credit exposure measurement standards; setting credit limits and monitoring their use; reviewing credits and concentrations of credit risk; reviewing and monitoring risk reduction arrangements.

These recommendations stress the need for risk management functions with "clear independence and authority."

26.2.2 The Bank of England Report on Barings

Violation of the fundamental principle of separation of functions was the primary cause of the Barings failure. Nick Leeson had control over both the front office and the back office. This organizational structure allowed him to falsify trading entries, hiding losses in a special account.

But new lessons were also described in the main report on Barings, produced by the Bank of England (BoE).[5] The report mentioned for the first time reputational risk. This is the risk of indirect losses to earnings arising from negative public opinion. These losses are distinct from the direct monetary loss ascribed to an event.

The BoE report listed several lessons from this disaster.

- **Duty to understand.** Management teams have a duty to understand fully the businesses they manage. Senior Barings management later claimed they did not fully understand the nature of their business (which is equivalent to claiming financial insanity, or that one is not responsible for financial losses due to a lack of understanding).
- **Clear responsibility.** Responsibility for each business activity must be clearly established. Barings had a *matrix* structure, with responsibilities assigned by product and region, which made it harder to assign responsibility to one person.
- **Relevant internal controls.** Internal controls, including clear segregation of duties, is fundamental to any effective risk control system.
- **Quick resolution of weaknesses.** Any weakness identified by an internal or external audit must be addressed quickly. In the Barings case, an internal audit report in the summer of 1994 had identified the lack of segregation of duties as a significant weakness. Yet this was not addressed by Barings top management.

26.2.3 The CRMPG Report on LTCM

The near-failure of the hedge fund Long-Term Capital Management (LTCM) also led to useful lessons for the industry. The Counterparty Risk Management Policy Group (CRMPG) was established in the wake of the LTCM setback to strengthen practices related to the management of financial risks.

The CRMPG consists of senior-level practitioners from the financial industry, including many banks that provided funding to LTCM. The industry came under criticism for allowing LTCM to build up so much leverage. Apparently, loans to

---

[5] Bank of England (1995), *Report of the Board of Banking Supervision Inquiry into the Circumstances of the Collapse of Barings*, London: HMSO Publications.

LTCM were fully collateralized as to their current, but not potential, exposure. In fact, it was fear of the disruption of markets and the potential for large losses that led the New York Federal Reserve Bank to orchestrate a bailout of LTCM.

In response, the CRMPG report provides a set of recommendations, which are summarized here.[6]

- **Information sharing.** Financial institutions should obtain more information from their counterparties, especially when significant credit exposures are involved. This includes the capital condition and market risk of the counterparty. This information should be kept confidential.
- **Leverage, market risk, and liquidity.** Financial risk managers should monitor the risks of large counterparties better, focusing on the interactions between leverage, liquidity, and market risk.
- **Liquidation-based estimates of exposure.** When exposures are large, information on exposures based on marked-to-market values should be supplemented by liquidation-based values. This should include current and potential exposures.
- **Stress testing.** Institutions should stress-test their market and credit exposure, taking into account the concentration risk to groups of counterparties and the risk that liquidating positions could move the markets.
- **Collateralization.** Loans to highly leveraged institutions should require appropriate collateral, taking into account liquidation costs.
- **Management responsibilities.** Senior management should convey clearly its tolerance for risk, expressed in terms of potential losses. The function of risk managers is then to design a reporting system that enables senior management to monitor the risk profile.

The report makes a number of other recommendations related to market practices and conventions, as well as regulatory reporting. In particular, the report identifies areas for improvements in standard industry documents, which should help to ensure that netting arrangements are carried out in a timely fashion.

Perhaps the most important lesson from LTCM for lenders is the relationship between market risk and credit risk. The G-30 report recommends the establishment of market and credit risk functions but does not discuss integration of these functions. When LTCM was about to fail, lenders realized that they had no protection for potential exposure and that many of their positions were similar to those of LTCM. Had LTCM defaulted (a credit event), lenders could have lost billions of dollars from market risk.

The second lesson from LTCM is the need for risk managers to make adjustments for large or illiquid positions. The third lesson from LTCM is that institutions should perform systematic stress tests, because VAR models based on recent history can fail to capture the extent of losses in a disrupted market. This seems obvious, as VAR purports to give only a first-order magnitude of the size of losses in a normal market environment.

---

[6] Counterparty Risk Management Policy Group (1999), *Improving Counterparty Risk Management Practices*, New York: CRMPG.

26.2.4 The CRMPG II and III Reports

The CRMPG provided an update in a next report, called CRMPG II.[7] CRMPG II notes that many of the earlier recommendations had been put into practice. In particular, there is now a greater focus on liquidity-based adjustments to close-out values. However, the report notes that market developments have introduced new risks, including the dispersion of credit products with embedded leverage among industry participants, which lessens transparency.

In response to the credit crisis that started in 2007, the CRMPG has provided a new set of recommendations, in a report called CRMPG III.[8]

- **Improved corporate governance.** Comparing how various institutions fared during the crisis, it has been apparent that the culture of corporate governance is important. Institutions that rely heavily on judgment, communication, and coordination across the entire firm did better than others. Also, the system of incentives has produced behavior that focused on short-term profits at the expense of financial stability. Incentives need to be better aligned with the long-term success of institutions and their risk tolerance.
- **Risk monitoring.** Financial institutions should be able to monitor risk concentrations to asset classes on a net and gross basis and to provide coherent reports to top management.
- **Estimating risk appetite.** Financial institutions should conduct regularly comprehensive exercises aimed at estimating risk appetite, using stress tests and a combination of qualitative and quantitative factors.
- **Focusing on contagion.** Financial institutions should regularly assess the effect of contagion, or the channels and linkages through which local financial disturbances can take on systemic characteristics.
- **Enhancing oversight.** Large financial institutions are subject to oversight by their boards of directors and by official supervisory bodies. Given the complexities of such institutions, the CRMPG recommends annual meetings of these two groups to share views on the condition of the institution.

EXAMPLE 26.5: FRM EXAM 2004—QUESTION 47

The failure of Barings Bank is a typical example of a lack in control pertaining to which one of the following risks:

a. Liquidity risk

b. Credit risk

c. Operational risk

d. Foreign exchange risk

---

[7] Counterparty Risk Management Policy Group (2005), *Toward Greater Financial Stability: A Private Sector Perspective*, New York: CRMPG.

[8] Counterparty Risk Management Policy Group (2008), *Containing Systemic Risk: The Road to Reform*, New York: CRMPG.

26.3 ORGANIZATIONAL STRUCTURE

To be effective, the organizational structure must be designed to reflect the policy of effective firm-wide risk management. Figure 26.1 reflects a typical organizational structure of an old-style commercial bank.

Here risk is monitored mainly by the business lines. Within the credit function, the risk manager approves transactions, sets exposure limits, and monitors the exposure limits as well as the counterparty's financial health. Treasury and trading implement proprietary trading and hedging. Within this unit, the risk manager measures and monitors positions. Line management deals with business and product strategy. It also controls operations. Finally, the audit function, external or internal, provides an independent review of business processes.

There are numerous problems with such a structure. Perhaps the main one is that market risk management reports to trading, which violates the principle of independence of risk management. In addition, the decentralization of risk management among separate lines leads to a lack of coordination and failure to capture correlations between different types of risk. The credit risk manager, for instance, will prefer an instrument that transforms credit risk into operational risk, which is under another manager's watch. Situations where credit risk and market risk exacerbate each other (as in the case of LTCM) will also be missed. Finally, models and databases may be inconsistent across lines.

To maintain independence, risk managers should report not to traders but directly to top management. Ideally, the risk management function should be a firm-wide function, covering market, credit, and operational risks. Such a structure will avoid situations where risks are pushed from one area, where they are well measured, toward other areas. Firm-wide risk management should also be able to capture interactions between different types of risks.
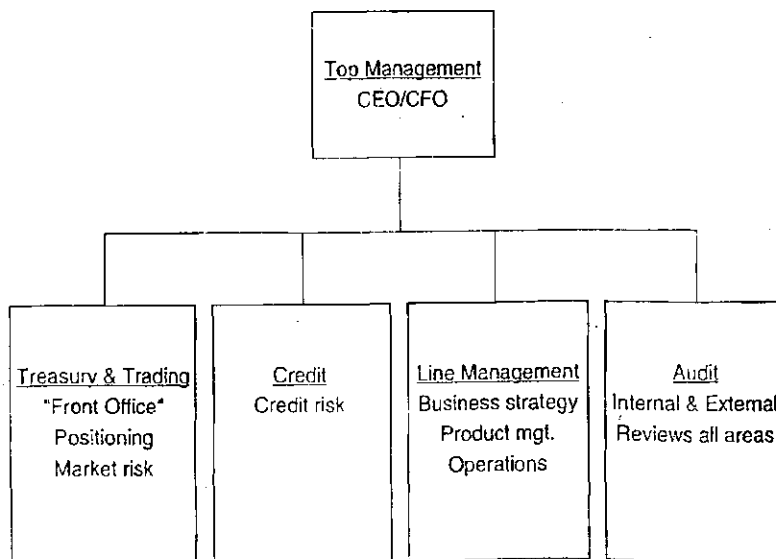
FIGURE 26.1   Old-Style Organizational Structure

EXAMPLE 26.6: BEST PRACTICES

When would it be prudent for a trader to direct accounting entries?

a. Never

b. When senior management of the firm and the board of directors are aware and have approved the practice on an exception basis

c. When audit controls are such that the entries are reviewed on a regular basis to ensure detection of irregularities

d. Solely during such times as staffing turnover requires the trader to back-fill until additional personnel can be hired and trained

EXAMPLE 26.7: FRM EXAM 2005—QUESTION 17

Which of the following is *not* a proper practice of risk management and control for a financial institution with assets in excess of $100 million?

a. A firm's sole mechanism to monitor the implementation of the control policies defined by the board is an external audit firm.

b. A subcommittee of the board is responsible for the approval of risk limits, risk management policies, and delegation of exceptional approval authorities.

c. Senior management is responsible for the day-to-day oversight of the firm's activities, implementing appropriate risk management and control policies, and monitoring the risks and exposures of the firm.

d. Senior management is responsible for establishing written documentation about control procedures at each level of the control hierarchy.
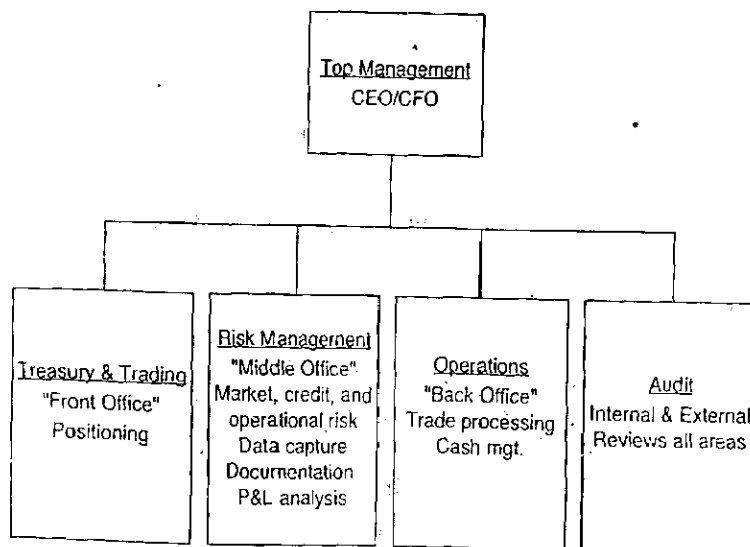
FIGURE 26.2 Modern Organizational Structure

```
                           ┌─────────────────┐
                           │ Risk Management │
                           │       EVP       │
                           └────────┬────────┘
                                    │
          ┌────────────┬────────────┼────────────┐
          │            │            │            │
    ┌──────────┐ ┌──────────┐ ┌──────────┐ ┌──────────┐
    │Market Risk│ │Credit Risk│ │Operational│ │ Risk MIS │
    │Trading room│ │Trading and│ │   Risk   │ │Analytics,│
    │           │ │banking books│ │          │ │  RAROC   │
    └──────────┘ └──────────┘ └──────────┘ └──────────┘
```
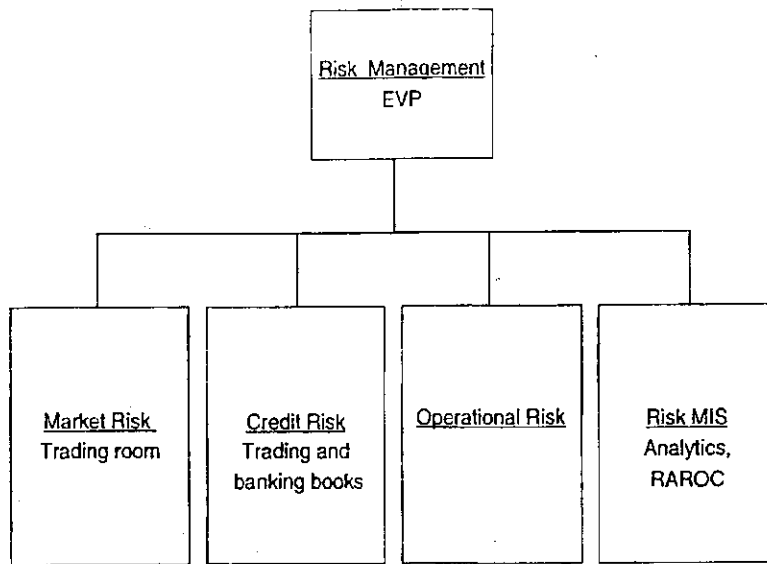
FIGURE 26.3   Risk Management Organizational Structure

The philosophy of separation of functions and independence of risk man-. agement must be embodied in the organizational structure of the institution. Figure 26.2 describes one such implementation. The most important aspect of this flowchart is that the risk management unit is independent of the trading unit.

The front office is concerned with positioning and perhaps some local hedging, subject to position and VAR limits established by risk management. The back office deals with trade processing and reconciliation as well as cash management. The middle office has expanded functions, which include risk measurement and control.

The chief risk officer is responsible for

- Establishing risk management policies, methodologies, and procedures consistent with firm-wide policies
- Reviewing and approving models used for pricing and risk measurement
- Measuring risk on a global basis as well as monitoring exposures and movements in risk factors
- Enforcing risk limits with traders
- Communicating risk management results to senior management

Figure 26.3 describes the centralization of the risk management function under an executive vice president or chief risk officer. The figure shows the units reporting to this new function. To this officer report *market risk management*, which monitors risk in the trading book; *credit risk management*, which monitors risk in the banking and trading books; *operational risk management*, which monitors operational risks; and systems. The latter unit deals with *risk management information systems (MIS)*, which include hardware, software, and data capture; *analytics*, which develops and tests risk management methodologies; and RAROC, which ensures that economic capital is allocated according to risk.

26.4 CONTROLLING TRADERS

26.4.1 Trader Compensation

The compensation structure for traders should also be given due thought. Usually, traders are paid a bonus that is directly related to their performance—for instance, 20% of profits—when positive. Note that the design of this compensation contract is asymmetrical, like that of an option. If the trader is successful, he can become a millionaire at a very young age. If the trader loses money, he is simply fired. In many cases, the trader will find another employer since he now has "experience."

Such a compensation scheme is designed to attract the very best talents into trading. The downside is that the trader, who is now long an option, has an incentive to increase the value of this option by increasing the risk of the positions. This, however, may not be in the best interests of the company.

Such a tendency for risk taking can be controlled by various means:

- By modifying the structure of the compensation contract to better align the interests of the trader and the company (e.g., by paying with company stock or tying compensation to longer-term performance)
- By subtracting a risk-based capital charge from trading profits, as in a RAROC-type system
- By appointing an independent risk manager

To be effective, the compensation structure for *risk managers* must be independent of how well traders perform. The compensation for risk managers needs to be attractive enough to draw talented individuals, however.

26.4.2 Trader Limits

To some extent, trading risk can be managed by appropriately altering the incentives of traders. Alternatively, this risk can be controlled by imposing limits. These can be separated into backward-looking and forward-looking limits. The former consist of stop-loss limits. The latter consist of exposure or VAR limits.

Stop-loss limits are restrictions on traders' positions that are imposed after a trader has accumulated losses. Because their design is backward-looking, they cannot prevent losses from occurring. What they do prevent, however, are attempts by traders who lose money to recover their losses by "doubling their bets," that is, taking bigger bets in the hope that a future gain will be sufficient to wipe out a string of previous losses. These limits may also be useful if markets are trending, which would amplify the losses.

Exposure limits are systematically imposed on traders as a means to control losses before they occur. These are defined in terms of notional principal, duration, or other exposure measures. For example, the maximum position for a yen trader could be set at the equivalent of $10 million. These limits are typically set by considering the worst loss a unit could absorb, combined with an extreme move in the risk factor.

The problem with such limits is that they do not account for diversification or movements in market risks. Also, complex products for which the notional does not represent the worst loss lend themselves to a form of limit "arbitrage," where the trader abides by the letter of the guideline but not its spirit. For instance, a trader may have a $10 million limit on notes with maturities up to five years. Typically, such notes will have duration of, say, four years. The spirit of the limit is to cap the interest rate exposure. The trader, however, may circumvent the spirit of the limit by investing in inverse floaters with a duration of 12 years.

VAR limits are becoming a more common addition to conventional limits. These account for diversification and time variation in risk. For example, the VAR limit for a business unit may be less than the sum of the VAR limits for individual desks due to diversification. In practice, VAR limits are also susceptible to arbitrage, so they are used together with exposure limits.

EXAMPLE 26.8: FRM EXAM 2002—QUESTION 132

The following is *not* a problem of having one employee perform trading functions and back-office functions.

a. The employee gets paid more because she performs two functions.
b. The employee can hide trading mistakes when processing the trades.
c. The employee can hide the size of her book.
d. The employee's firm may not know its true exposure.

EXAMPLE 26.9: FRM EXAM 2000—QUESTION 69

Which of the following strategies can contribute to minimizing operational risk?

I. Individuals responsible for committing to transactions should perform clearance and accounting functions.
II. To value current positions, price information should be obtained from external sources.
III. Compensation schemes for traders should be directly linked to calendar revenues.
IV. Trade tickets need to be confirmed with the counterparty.
   a. I and II
   b. II and IV
   c. III and IV
   d. I, II, and III

EXAMPLE 26.10: FRM EXAM 2007—QUESTION 36

To control risk-taking by traders, your bank links trader compensation with their compliance with imposed VAR limits on their trading book. Why should your bank be careful in tying compensation to the VAR of each trader?

a. It encourages traders to select positions with high estimated risks, which leads to an underestimation of the VAR limits.

b. It encourages traders to select positions with high estimated risks, which leads to an overestimation of the VAR limits.

c. It encourages traders to select positions with low estimated risks, which leads to an underestimation of the VAR limits.

d. It encourages traders to select positions with low estimated risks, which leads to an overestimation of the VAR limits.

26.5 RISK-ADJUSTED PERFORMANCE AND RAROC

The ability to measure risk has profound implications for performance measurement. In the past, performance was measured by yardsticks such as return on assets (ROA), which adjusts profits for the associated book value of assets, or return on equity (ROE), which adjusts profits for the associated book value of equity. Such measures are simple to compute but are fundamentally flawed because they ignore risks. As a result, these measures could lead to dangerous behavior, such as expanding operations in markets or lines of business where expected returns are high but where risks are much higher.

Risk managers now have tools to control for this behavior. They can assess the total risk of an operation in terms of economic capital required to support all categories of risk, including market, credit, and operational risk. This capital, also called risk capital, is basically a value-at-risk (VAR) measure at a high confidence level.

Armed with this information, institutions can make better-informed decisions about business lines. Each activity should provide sufficient profit to compensate for the risks involved. Thus, product pricing should account not only for expected losses but also for the remuneration of risk capital.

Some activities may require large amounts of risk capital, which in turn requires higher returns. This is the essence of risk-adjusted return on capital (RAROC) measures. The central objective is to establish benchmarks to evaluate the economic return of business activities. This includes transactions, products, customer trades, and business lines, as well as the entire business.

26.5.1 Risk Capital

RAROC was developed by Bankers Trust in the late 1970s. The bank was faced with the problem of evaluating traders involved in activities with different risk profiles.

RAROC is part of the family of risk-adjusted performance measures (RAPM). Consider, for instance, two traders that each returned a profit of $10 million over the last year. The first is a foreign currency trader, the second a bond trader. The question is, how do we compare their performance? This is important in providing appropriate compensation as well as deciding which line of activity to expand.

Assume the FX and bond traders have notional amount and volatility as described in Table 26.2. The notional amount is also the market value of their book. The bond trader deals in larger amounts, $200 million, but in a market with lower volatility, at 4% per annum, against $100 million and 12% for the FX trader. The risk capital (RC) can be computed as a VAR measure, say at the 99% level over a year, as Bankers Trust did. Assuming normal distributions, this translates into a risk capital of

$$RC = VAR = \$100,000,000 \times 0.12 \times 2.33 = \$28 \text{ million}$$

for the FX trader and $19 million for the bond trader. More precisely, Bankers Trust computes risk capital from a weekly standard deviation σ_w as

$$RC = 2.33 \times \sigma_w \times \sqrt{52} \times (1 - \text{Tax rate}) \times \text{Notional} \qquad (26.1)$$

which includes a tax factor that determines the amount required on an after-tax basis.

The risk-adjusted performance is then measured as the dollar profit divided by the risk capital,

$$RAPM = \frac{\text{Profit}}{\text{RC}} \qquad (26.2)$$

and is shown in the last column. Thus the bond trader is actually performing better than the FX trader, as the activity requires less risk capital. More generally, risk capital should account for credit risk, operational risk, and any interaction.

It should be noted that this approach views risk on a standalone basis, that is, using each product's volatility. In theory, for capital allocation purposes, risk

TABLE 26.2   Computing RAPM

	Profit	Notional	Volatility	VAR	RAPM
FX trader	$10	$100	12%	$28	36%
Bond trader	$10	$200	4%	$19	54%

should be viewed in the context of the bank's whole portfolio and measured in terms of its marginal contribution to the bank's overall risk. In practice, however, it is best to charge traders for risks under their control, which means the volatility of their portfolios.

26.5.2 RAROC

This RAROC methodology can be applied at the level of a transaction, of a desk, or of a business unit. In each case, the first step is to compute the economic capital (EC) required to support the operation. This includes market, credit, and operational risk.

RAROC is formally defined as

$$RAROC = Net\ Profit/EC = [Expected\ Profit - Costs + k(EC)]/EC \qquad (26.3)$$

where the net profit includes: (1) revenues, net of expected losses; (2) minus any direct operating cost; (3) minus financing costs; (4) plus the return on economic capital.

EXAMPLE 26.11: FRM EXAM 2006—QUESTION 3

A risk manager for ABC Bank has compiled the following data regarding a bond trader and an equity trader. Assume that the returns are normally distributed and that there are 52 trading weeks per year. ABC Bank computes its capital using a 99% VAR. Dollar amounts are in millions.

	After-Tax Profit	Net Book Market Value	Weekly Volatility	Tax Rate
Bond Trader	USD 8	USD 120	1.10%	40%
Equity Trader	USD 18	USD 180	1.94%	40%

Calculate the risk-adjusted performance measure (RAPM) for the bond trader.

a. 25.24%

b. 36.08%

c. 60.15%

d. 84.92%

EXAMPLE 26.12: FRM EXAM 2006—QUESTION 4

Continuing with the same ABC Bank data, which of the following statements are correct in relation to the equity trader?

I. The equity trader has an annual, after-tax VAR at a 99% confidence level of USD 33.2 million.
II. In comparing the RAROC for both traders, the equity trader is performing better than the bond trader.
   a. I only
   b. II only
   c. Both
   d. Neither

EXAMPLE 26.13: FRM EXAM 2007—QUESTION 124

The bank you work for has a RAROC model. The RAROC model, computed for each specific activity, measures the ratio of the expected yearly net income to the yearly VAR risk estimate. You are asked to estimate the RAROC of its $500 million loan business. The average interest rate is 10%. All loans have the same probability of default of 2% with a loss given default of 50%. Operating costs are $10 million. The funding cost of the business is $30 million. RAROC is estimated using a credit-VAR for loan businesses, in this case, 7.5%.
   The economic capital is invested and earns 6%. The RAROC is:

   a. 19.33%
   b. 46.00%
   c. 32.67%
   d. 13.33%

26.6 IMPORTANT FORMULAS

Economic risk capital (RC): $RC = VAR$

Risk-adjusted performance measure (RAPM): $RAPM = \frac{Profit}{RC}$

Risk-adjusted return on capital (RAROC):
$$RAROC = [Expected\ Profit - Costs + k(EC)]/EC$$

26.7 ANSWERS TO CHAPTER EXAMPLES

Example 26.1: FRM Exam 2005—Question 110

d. Statement a. is not a problem, as risk managers use simulations to bring together distributions for different types of risk. Statement b. is not a problem. Indeed different horizons are used, but these can be translated to a common horizon, typically annual. Statement c. is not a problem, as demonstrated in Chapter 24.

Example 26.2: FRM Exam 2002—Question 103

b. VAR can be added across different types of risk, but this will provide a conservative estimate of capital as diversification effects are ignored. So answer a. would be for *too much* capital. Answer c. is not correct because rare events can be factored into operational VAR. Most likely, the bank may have too little capital for other types of risk than those measured by these three categories.

Example 26.3: FRM Exam 2006—Question 109

d. For most global banks, the order of importance is, first, credit risk, then operational risk, then market/ALM risk. Also, answers b. and c. are the same.

Example 26.4: FRM Exam 2005—Question 33

c. This is an example of a wrong-way exposure, where a gain on the instrument for the bank is associated with a higher PD for its counterparty. If the IDR depreciates, company A will make a profit because its costs will go down in dollars. Conversely for company B, because its dollar revenues will decrease. Under c., the company pays USD and receives IDR. This transaction will create a loss if the IDR depreciates. In this situation, company B will lose money as well on its exports. Hence, this is a wrong-way trade.

Example 26.5: FRM Exam 2004—Question 47

c. The Barings failure falls in the category of operational risk because of a breakdown in procedures. The trader, Nick Leeson, had control of the back office.

Example 26.6: Best Practices

a. As one risk manager has said, this is one of the few instances where *never* means *absolutely never*. Allowing traders to tabulate their own profits and losses is a recipe for disaster.

Example 26.7: FRM Exam 2005—Question 17

a. Control policies also need to be verified by an internal audit function.

Example 26.8: FRM Exam 2002—Question 132

a. Answers b., c., and d. all can lead to a situation where the trader loses money and hides the losses. Answer a. is not a problem per se.

Example 26.9: FRM Exam 2000—Question 69

b. Answer I violates the principle of separation of functions. Answer III. may create problems of traders taking too much risk. Answer II. advises the use of external sources for valuing positions, as traders may affect internal price data.

Example 26.10: FRM Exam 2007—Question 36

c. Traders may engage in VAR arbitrage, trying to exploit weaknesses in VAR measures. With a VAR limit, they may seek positions that have low measured VAR, in which case the VAR limits will be less effective.

Example 26.11: FRM Exam 2006—Question 3

c. The 99% VAR is $2.33 \times 1.10\% \times \sqrt{52} \times (1 - 40\%) \times \$120 = \$13.3$m. Hence, RAPM $= 8/13.3 = 60.1\%$.

Example 26.12: FRM Exam 2006—Question 4

d. The equity trader's VAR is $2.33 \times 1.94\% \times \sqrt{52} \times (1 - 40\%) \times \$180 = \$35.2$ million, so statement I. is incorrect. The RAPM is $18/35.2$, or 51.1%, which is worse than that of the bond trader, so statement II. is incorrect as well.

Example 26.13: FRM Exam 2007—Question 124

a. First, we compute the numerator. The net interest is, after expected losses, $\$500 \times (10\% - 2\%(1 - 50\%)) = \$45$. Next, we compute economic capital, or $\$500 \times 7.5\% = \$37.5$. To revenues, we then add the return on economic capital, or $\$37.5 \times 6\% = \$2.25$. From this, we deduct operating and funding costs, which gives $\$47.25 - 10 - 30 = \$7.25$. Finally, we divide by $\$37.5$ and get 19.33%.

Legal Issues

We now turn to legal issues in risk management. Legal risk can be defined as the risk that a contract is not legally enforceable or documented correctly. More generally, this is "the risk that a transaction cannot be consummated because of some legal barrier, such as inadequate documentation, a regulatory prohibition on a specific counterparty, and non-enforceability of bilateral and multilateral close-out netting and collateral arrangements in bankruptcy."[1] This includes changes in law, mistakes, liabilities of agents, and political risks.

Legal risk invariably arises when the counterparty lost money on a transaction and is more likely to sue as a result. Legal risk is also intimately related to credit risk, as situations of default require enforcement of contracts, which creates legal uncertainty.

This chapter will focus on legal risk for derivatives, although many of the concepts developed here also apply to legal risks for other financial instruments, such as loans or bonds. Section 27.1 briefly reviews the history of legal risks in the derivatives markets. Section 27.2 discusses netting, an important feature of swaps that has been developed to control market, credit, and legal risk. Section 27.3 summarizes the master netting agreement established by the International Swaps and Derivatives Association (ISDA) in 1992. Readers, however, should also be familiar with the full text of the agreement. Finally, Section 27.4 contains a glossary of useful legal terms.

27.1 LEGAL RISKS WITH DERIVATIVES

While legal risks have always existed in derivatives contracts, they became more significant with the inception of the swap markets. Unlike exchange-traded futures, which are standardized, the essence of the over-the-counter market is to *tailor* contracts to the counterparty. This, however, requires not only customizing financial terms (prices, quantities, maturities) but also the legal documentation to the counterparty, which creates additional risk.

Legal risks are also intermingled with market and credit risks. When a counterparty loses a large amount of money on a transaction, reflecting market risk,

---

[1] See the Federal Reserve Board's in-depth guide, *Trading and Capital Markets Activities Manual* (1998), Section 1000.1.

there may be a tendency to resort to legal action as a means to recover some of the losses. For example, when Procter & Gamble lost $157 million on swaps arranged by Bankers Trust, the company sued the bank and recovered its losses.

Another famous example of legal risk is the case of Hammersmith & Fulham. City councils in the United Kingdom had entered into a series of interest rate swaps, which turned out to produce major losses as U.K. interest rates almost doubled from 1988 to 1989. The swaps were later ruled invalid by the U.K. high court. The court decreed that the city councils did not have the authority to enter these transactions, which were found to be ultra vires (or "beyond the power" of the cities to enter). All the contracts were deemed void and hence the cities were not responsible for the losses. As a result, losses of $178 million had to be absorbed by their counterparty banks.

After this experience, banks have tried to control their legal risks by verifying that the counterparty indeed has the right to enter into a transaction. Even so, this is not always easy to assess. Before the Hammersmith verdict, for instance, many lawyers were convinced that the swaps in question would withstand legal scrutiny.

Until recently, the Hammersmith loss was the greatest single credit loss in the swap markets. For instance, a study by the ISDA noted that total losses amounted to only $358 million by the end of 1991. About 50% of this sum was due to the Hammersmith case.

Even so, these losses are relatively small compared with the size of the market. The total of $358 million represents only 0.012% of the notional amount of $4.3 trillion at that time. As we have learned, however, notionals provide an exaggerated measure of the size of derivatives markets. A more relevant measure is the credit exposure, which is the maximum amount that can be lost. Current exposure is measured by marked-to-market values, which amounted to $77.5 billion in 1991. Compared with this number, the loss percentage is still very small, only 0.46%.

For more recent data, we can turn to information provided by the Office of the Comptroller of the Currency (OCC) for U.S. commercial banks.[2] The OCC provides quarterly reports on the charge-offs from derivatives (or credit losses). Figure 27.1 presents quarterly charge-offs since 1996. By the end of the sample period, these losses had accumulated to approximately $2,600 million.

The peak quarterly losses occurred in the third quarter of 1998, as a result of the Asian financial crisis and the Russian default. Even this number, $445 million, represents only 0.0014% of the total notional of $33 trillion, or 0.11% of the total credit exposure at that time. Another perspective would be to compare this peak number with the charge-offs on loans, which was 0.49% in the same quarter. Overall, derivatives credit losses are very small relative to the size of these markets. The low incidence of losses results from two main factors. First, the credit quality of the derivatives counterparty is typically high. Second, most of the exposures are collateralized on a daily basis.

---

[2] The OCC is an agency overseeing U.S. commercial banks. Chapter 28 will present an overview of bank regulators.
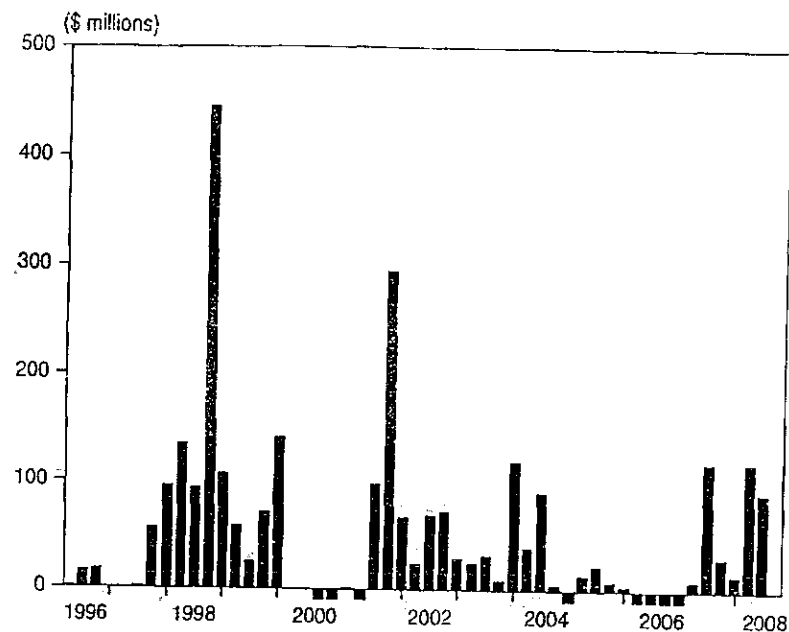
FIGURE 27.1 Charge-Offs on Derivatives: U.S. Commercial Banks
Source: Office of the Comptroller of the Currency Bank Derivatives Reports

Legal risks can arise from a number of sources:

- *A failure in contracting.* This can happen if the contract is not properly authorized or executed, as in the Hammersmith case. Even in the United States, there was some uncertainty as to the legal status of swaps until recently. The Commodity Exchange Act did not make it clear that swaps are legally distinct from futures contracts. If swaps had been ruled to be futures contract, they could have been found illegal and thus void. This changed only with the passage of the Commodity Futures Modernization Act of 2000, which secured legal certainty for OTC derivatives transactions.
- *A failure in contract documentation.* Mistakes can arise in contract documentation, such as incorrect number of entries.
- *Bankruptcy risks.* By nature, the bankruptcy process is fraught with uncertainties. For instance, the bankruptcy court could "cherry-pick" the contracts, or choose to honor the contracts having the greatest value for the defaulting party only, to the detriment of counterparties.

Special protection is accorded, however, for the set-off of margin payments and liquidation of collateral under securities contracts and commodities contracts. In the United States, close-out netting agreements (to be defined in the next section) are specifically exempted from the automatic stay provision that applies upon the filing of a bankruptcy petition. This protection was adopted by the Financial Institutions Reform Recovery and Enforcement Act (FIRREA) of 1989, which also confirmed the right to access the collateral posted by the defaulting counterparty.

Even so, there is often uncertainty in the application of these laws. The case of Long-Term Capital Management (LTCM) is a good example, because LTCM was chartered in the Cayman Islands. Had LTCM declared bankruptcy

in the Cayman Islands, there would have been legal uncertainty as to whether counterparty banks would have the right to liquidate their collateral under the U.S. Bankruptcy Code. This uncertainty is reportedly one reason why the same banks wanted to avoid a messy bankruptcy scenario and agreed to bail out LTCM.

■ *Changes in laws and regulations.* Contracts may contain clauses protecting one party against changes in tax or regulatory treatments. As an example, coupons on Eurobonds are exempt from withholding taxes. If the country of the bond issuer suddenly imposes new taxes, the issuer may be subject to a so-called gross-up clause that requires it to pay the investor additional money to make up for the new tax.[3] Changes in the regulatory environment may also lead to changes in the value of contracts.

EXAMPLE 27.1: FRM EXAM 2002—QUESTION 60

Lawsuits involving derivatives to major corporations are most likely to involve which of the following issues?

a. The type of derivative
b. Broker size
c. Breach of fiduciary duty
d. Enforceability of contract

27.2 NETTING

As we have seen in analyzing credit risk, netting has developed over time as a powerful mechanism to reduce credit exposure. The purpose of netting is to offset transactions between two parties, with settlement of the *net* difference in cash flows across all contracts covered by a netting agreement. In the case of bankruptcy, however, netting is fully beneficial only when enforced by the courts.

ISDA keeps track of countries that have adopted or are considering changes in legislation to allow netting. It has obtained legal opinions that netting would be upheld in most leading jurisdictions. Similarly, the Bank for International Settlements has issued a report concluding that bilateral netting is likely to be effective in G-10 countries.[4]

27.2.1 Netting under the Basel Accord

In 1995 the Basel Committee on Banking Supervision (BCBS) lowered capital charges to recognize, and encourage, netting agreements.[5] The BCBS recognizes

---

[3] Additional complications may arise if the issuer has the right to redeem the bond at par. If the bond is trading at a premium, this provides a windfall profit for the issuer.

[4] Bank for International Settlements (1990), *Report of the Committee on Interbank Netting Schemes of the Central Banks of the Group of Ten Countries (Lamfalussy Report)*, available at http://www.bis.org/publ/cpss04.pdf.

[5] See BCBS (1995), *Basel Capital Accord: Treatment of Potential Exposure for Off-Balance Sheet Items.* Available at http://www.bis.org/publ/bcbs18.pdf .

netting under novation, which substitutes outstanding debt payments for new ones that provide for *net* payment obligations. Under novation, any obligation between a bank and its counterparty to deliver a given currency on a given value date is automatically amalgamated with all other obligations for the same currency and value date, legally substituting one single amount for the previous gross obligations.

Another form is the close-out netting agreement, which is a bilateral contract specifying that upon default, the non-defaulting party nets gains and losses with the defaulting counterparty to a single payment for all covered transactions.

The ability to terminate financial market contracts upon an event of default is central to the effective management of financial risk. Without a close-out or termination clause, counterparties would helplessly watch their contracts fluctuate in value during the bankruptcy process, which could take years.

The Basel Accord recognizes netting, as long as the bank can assure its national supervisor that it has:

(1) *A netting contract or agreement with the counterparty which creates a single legal obligation, covering all included transactions, such that the bank would have either a claim to receive or obligation to pay only the net sum of the positive and negative mark-to-market values of included individual transactions in the event a counterparty fails to perform due to any of the following: default, bankruptcy, liquidation or similar circumstances*

(2) *Written and reasoned legal opinions that, in the event of a legal challenge, the relevant courts and administrative authorities would find the bank's exposure to be such a net amount under:*
   –the law of the jurisdiction in which the counterparty is chartered and, if the foreign branch of a counterparty is involved, then also under the law of the jurisdiction in which the branch is located;
   –the law that governs the individual transactions; and
   –the law that governs any contract or agreement necessary to effect the netting.
   The national supervisor, after consultation when necessary with other relevant supervisors, must be satisfied that the netting is enforceable under the laws of each of the relevant jurisdictions.

(3) *Procedures in place to ensure that the legal characteristics of netting arrangements are kept under review in the light of possible changes in relevant law*

27.2.2 Walk-Away Clauses

Netting, however, attracts a favorable capital treatment only for contracts without walk-away clauses. These clauses, also known as limited two-way payment provisions, allow both parties to walk away from the contract in case of default.

Consider, for example, the collapse in 1990 of the Drexel Burnham Lambert Group (DBL Group), which placed its swap subsidiary, DBL Products, in default. Some swaps were out-of-the-money for DBL Products, in which case counterparties had a claim against DBL Products. This placed them in the same position as other unsecured senior creditors, which seems normal.

Other swaps, however, were in-the-money for DBL Products, which means that counterparties owed money. In theory, the walk-away clause would have permitted them to reap a windfall profit, randomly benefiting from the misfortune of others, which seems questionable.

Even so, nearly all in-the-money contracts were fully paid. Counterparties settled to avoid expensive litigation over the enforceability of these contracts. Financial institutions also recognized that walk-away clauses create uncertainty for financial markets. Contracts have now evolved to contain a full two-way payment provision, which provides for full payment to the counterparty, subject to a bankruptcy distribution rule.

The final nail in the coffin for the walk-away clause was the ruling by the Basel Committee that such contracts are not provided any regulatory relief in terms of lower capital requirement.

27.2.3 Netting and Exchange Margins

Netting also applies to the credit risk that futures traders face from their brokers. Clients deposit margins with their brokers. Assuming the broker is a clearing member, the broker in turn deposits margins with the clearinghouse.

If a broker goes bankrupt, clients could lose the part of their margins held by the broker. In the United States, two clearinghouses (CME and NYMEX) collect *gross margins*, that is, a separate margin for all client positions. Others collect *net margins*, allowing the broker to offset long and short positions by different customers. This netting decreases the margin held by the clearinghouse. Normally, a gross margin system is safer for the client because a greater fraction of the margin is held by the clearinghouse. The risk of a net margin system is lessened, however, if the broker properly *segregates* client accounts by holding them separately from its own accounts.

EXAMPLE 27.2: FRM EXAM 2002—QUESTION 117

You are an investment manager trying to decide whether the Chicago Mercantile Exchange, the Chicago Board of Trade, or the OTC marketplace is where you will place part of your portfolio hedge. You will have to make an OTC transaction with your broker in any case. You also are considering a direct OTC deal with your broker for the whole hedge. You want to carry out the transaction that will result in the lowest possible exposure to your broker. Assuming that the size of the OTC hedge, if you use an exchange, is the same regardless of the exchange and that the effectiveness of the hedge is the same absent counterparty risks, how would you hedge?

    a. Hedge on the Chicago Mercantile Exchange and with your broker.

    b. Hedge on the Chicago Board of Trade and with your broker.

    c. It doesn't matter, as the broker exposure is the same for each exchange.

    d. Hedge your portfolio with a series of over-the-counter transactions, with your broker as counterparty.

27.3 ISDA MASTER NETTING AGREEMENT

At the beginning of the 1980s, swaps were tailor-made financial contracts that required documentation to be drafted on a case-by-case basis. This was very time-consuming and costly, and it introduced a time lag between the commercial agreement and the signing of the legally binding contract.

In response, the industry developed standardized terms for swaps. As with futures, this made it easier to offset the contracts, increasing liquidity and decreasing legal uncertainty. Out of this effort came the master netting agreement established by the ISDA in 1987 and revised in 1992 and in 2002. This form establishes (1) a template for a standardized contract, which is supplemented by (2) a schedule to the master agreement and (3) the actual confirmation of contract. Parties have the flexibility to select parts of the agreement or to amend the base document through the schedule. The more specific clauses (e.g., confirmation) override more general clauses. Thus, the order of precedence in the case of conflict is, first the confirmation, then the schedule, and finally the master agreement. In addition, the credit support annex (CSA) manages the exchange of collateral between parties.

The ISDA master agreement contains the following provisions:

- A list of *obligations*, detailing the mechanics of payment conditions (section 2 in the ISDA agreement), including the netting of obligations.
- A list of *credit provisions*, which describe events of default and termination (section 5), early termination (section 6), and credit support provisions (e.g., the system of collateral payments). The event of default includes:
  - Failure to pay
  - Breach of agreement
  - Credit support default (e.g., failure to provide collateral when due)
  - Misrepresentation
  - Default under a specified transaction
  - Cross-default, which is optional
  - Acts pertaining to bankruptcy or liquidation
  - Mergers without the successor assuming the obligation to perform under the swap

  Termination includes:
  - An illegality in which a party is unable to perform due to a change in law or regulation
  - A tax event such as a change in tax law that causes a party to make an additional payment (called gross-up)
  - A tax event upon merger
  - A credit event upon merger where the creditworthiness of the successor is materially weaker than the original entity
- A list of contractual *boilerplate statements*, including representations (section 3), agreements (section 4), transfer provisions (section 7), governing law (section 13), and so on.

Although the ISDA forms attempt to provide comprehensive and standardized coverage of swap events, they cannot anticipate every eventuality. When Russia defaulted on its domestic-currency debt on August 17, 1998, it imposed a moratorium on foreign-currency debt payments as well as a 90-day freeze on forward foreign exchange contracts. It has maintained payment on its foreign debt, however. Whether this constitutes a credit event on the foreign debt was not clearly defined by the swap agreements in place. This has created considerable disagreement over the interpretation of standard contracts. By 1999, the ISDA had published a revised set of definitions for credit derivatives that considers both sovereign and non-sovereign entities. This list is provided in the credit derivatives chapter.

EXAMPLE 27.3: FRM EXAM 2001—QUESTION 124

Most credit derivatives contracts

 a. Are based on English law.
 b. Are written on a one-off basis.
 c. Have a clause about restructuring.
 d. Are based on the ISDA agreement.

EXAMPLE 27.4: FRM EXAM 2000—QUESTION 22

A typical master netting agreement as established by the ISDA will contain all of the following *except* a list of

 a. Obligations.
 b. Historical market prices.
 c. Credit provisions.
 d. Contractual boilerplate statements.

EXAMPLE 27.5: FRM EXAM 2004—QUESTION 62

Which of the following are *not* considered events of termination under the ISDA Master Agreement?

 I. Misrepresentation
 II. Tax event upon corporate take-over
 III. Change in tax law that results in gross-up
 IV. Bankruptcy
    a. I and IV only
    b. I and III only
    c. II and III only
    d. II and IV only

27.4 GLOSSARY

27.4.1 General Legal Terms

Civil law: (1) Legal system whose law is centered around a comprehensive legislative code (e.g., such as that established by Napoléon in France). (2) In the United States, law under which a person (the plaintiff) may sue another person (the defendant) to obtain redress for a wrong committed by the defendant, for example, a breach of contract. This is in contrast with criminal law.

Common law: System of law derived from the English system of laws "common to the population," produced primarily by a group of judges to harmonize their decisions with those in other parts of the country. It was introduced after the Norman conquest of England as a means of unifying the country. Common law builds on precedents. This is in contrast to the French-type system of civil law.

Criminal law: Law that defines public offenses against the state or government and prescribes their punishment. This is a part of public law, which also includes constitutional and administrative law.

27.4.2 Bankruptcy Terms

Absolute priority rule (APR): Hierarchical rule for the distribution of a firm's assets: Payments go first to secured creditors, then to priority creditors (e.g., to cover taxes and bankruptcy costs), then to unsecured creditors (such as bondholders and bank depositors), then to subordinated-debt holders, and finally to stockholders. (See also the credit derivatives chapter.)

Automatic stay: In bankruptcy, the suspension of legal actions (other than the bankruptcy proceeding itself) until the bankruptcy case is over.

Bankruptcy: A legal process under which (1) a financially troubled debtor is declared to be insolvent, or incapable of meeting debt payments; (2) the assets of the debtor are distributed to creditors according to bankruptcy law; and (3) the debtor, if honest, is discharged from liability for remaining unpaid debt.

The word *bankruptcy* comes from the Italian *banca rotta*, or "broken bench." The tradition was that when a medieval trader failed to pay his creditors, his trading bench was broken.

Liquidating proceeding: A bankruptcy proceeding in which the debtor's assets are converted to cash and distributed to creditors. In the United States, liquidation is covered under Chapter 7 of the U.S. Bankruptcy Code.

Reorganization proceeding: A bankruptcy proceeding in which the troubled firm may stay in business as it reorganizes in a process of financial rehabilitation. In the United States, reorganization is covered under Chapter 11 of the U.S. Bankruptcy Code. A majority of creditors and equity holders must approve the plan; otherwise, liquidation proceeds under Chapter 7.

27.4.3 Contract Terms

Acceleration clause: A provision in a promissory note permitting the debtor to make, or the creditor to receive, payment before the due date.

Close-out or termination clause: A provision that gives the right to terminate a contract upon certain specified events and to calculate a termination amount due to, or due from, the defaulting party.

Covenant: A contractual provision whereby one party promises to take certain specific actions (positive covenant) or to refrain from taking certain actions (negative covenant). Bond covenants contain clauses prohibiting, for instance, the creditor from selling major assets or paying too large a dividend to stockholders.

Cross-default clause: A contractual provision whereby default on a contract occurs whenever the counterparty defaults on *any* other obligation.

Negative pledge clause: A provision that prevents the subordination of a contract to secured creditors, by pledging assets for new debt, for instance.

Netting: A provision that gives the right to *set off*, or net, claims or payment obligations between two or more parties, with the goal of arriving at a single net payment.

Novation: The extinguishment of a party's obligation (e.g., the debt of the obligee) through an agreement between the old obligor, a new obligor, and the obligee to substitute the old obligor for a new one.

Pari passu: Equal ranking (from Latin), meaning that all creditors within the same class will be treated equally. This term often used in bankruptcy proceedings where creditors are paid pro rata in accordance with the amount of their claims.

Secured transaction: An arrangement such that the creditor is provided with a backup source of payment if the debtor defaults.

Security agreement: An agreement between a debtor and a creditor whereby the creditor receives security interest, or property, to secure debt payments.

Ultra vires: Outside the power of a person or corporation (from Latin). This is in contrast to *intra vires*.

27.5 ANSWERS TO CHAPTER EXAMPLES

Example 27.1: FRM Exam 2002—Question 60

d. Most derivatives lawsuits arise from interpretation of the provisions of the contract. There is generally no fiduciary duty issue, as most contracts are with major corporations, which are supposed to be more informed than individual investors.

Example 27.2: FRM Exam 2002—Question 117

a. The CME clearinghouse (along with the NYMEX) collects gross margins, that is, separate margins for all clients. Other exchanges collect net margins, that is, allow the broker to mingle client positions. The gross margin system is safer for the client.

Example 27.3: FRM Exam 2001—Question 124

d. Most derivatives contracts are based on the standard form provided by the ISDA, which ensures uniformity in contracts and reduces legal uncertainty.

Example 27.4: FRM Exam 2000—Question 22

b. A master agreement will contain a list of obligations, credit provisions, and boilerplate statements. There is no reason to have historical market prices.

Example 27.5: FRM Exam 2004—Question 62

a. Default includes misrepresentation and bankruptcy. Termination includes tax events.

Regulation and Compliance
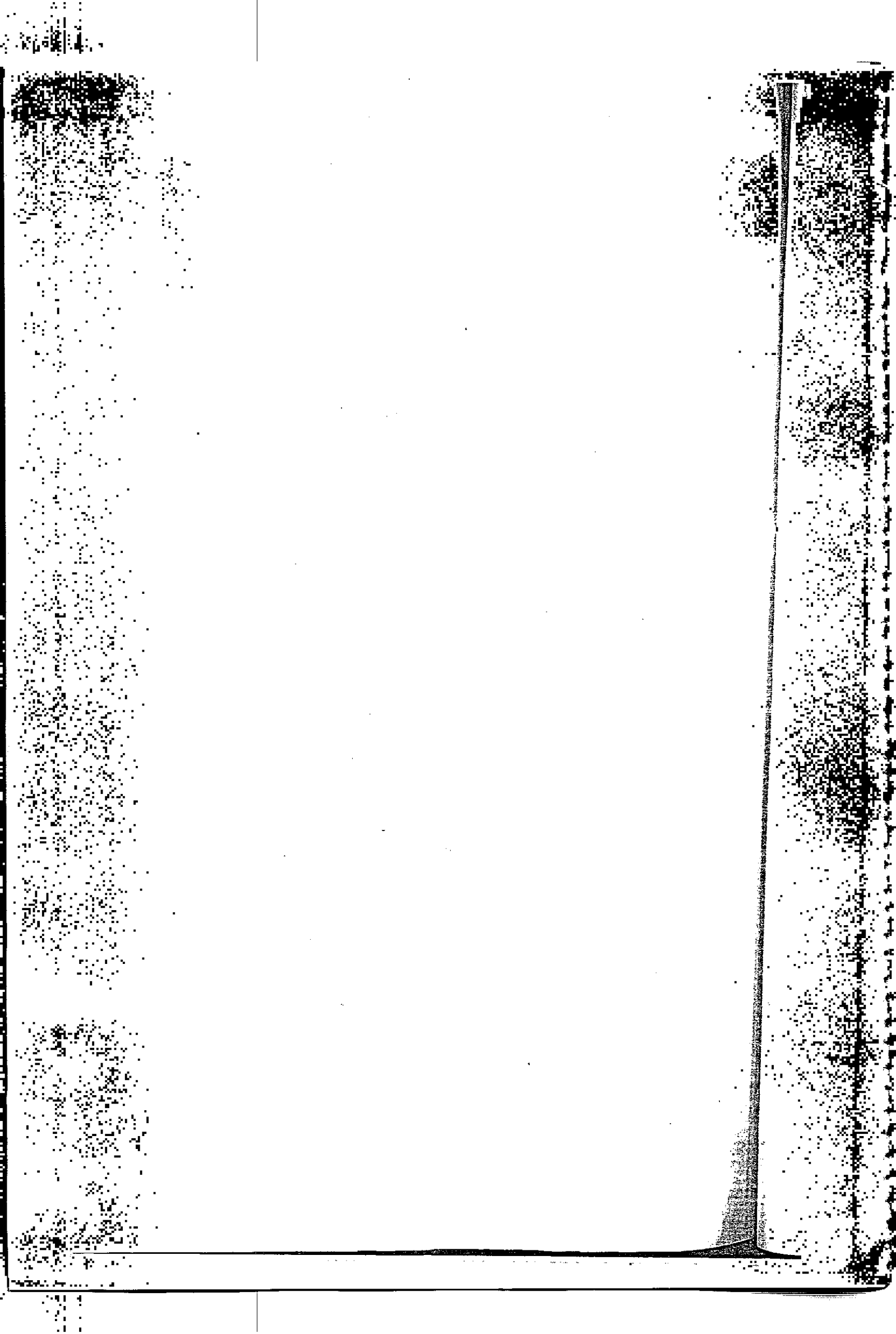
Regulation of Financial Institutions

We now tackle the last part of this book, which deals with regulatory capital. Banks and securities houses must now comply with risk-based capital requirements. These regulatory capital requirements have been the catalyst for advances in risk management of the last decade. They have spurred the industry into better understanding and management of their risks. In turn, regulators are now forced to upgrade their requirements to keep up with modern developments in risk management. Analyzing the rationale behind these regulations yields interesting insights into broader issues that we have not addressed yet, such as systemic risk.

Section 28.1 provides a broad classification of financial institutions subject to regulation. Section 28.2 then discusses systemic risk, which is a major rationale for the regulation of financial institutions. Next, Sections 28.3 and 28.4 describe the regulation of commercial banks and securities houses, respectively. Finally, Section 28.5 concludes with a summary of the tools and objectives of financial regulation.

28.1 DEFINITION OF FINANCIAL INSTITUTIONS

Financial institutions are fundamentally different from other firms. When an industrial corporation goes bankrupt, shareholders, bondholders, and other creditors suffer financial losses. The overall effects of the failure, however, are limited to direct stakeholders. In contrast, the failure of a financial institution can be potentially much more harmful.

But first, we need to define financial institutions. These include:

- Commercial banks, whose primary function is to hold customer deposits and to extend credit to businesses, households, or governments.[1]
- Securities houses, whose primary function is to intermediate in securities markets. These include investment banks, which specialize in the initial sale of

---

[1] Similar intermediaries are savings institutions, which specialize in residential mortgages, and credit unions, which extend mortgage and consumer credit. These are generally local and relatively small institutions whose failure is unlikely to destabilize financial markets.

securities in the primary markets,[2] and broker-dealers, whose primary function is to assist in the trading of securities in the secondary markets.[3]

- Insurance companies, which provide property and casualty (P&C) or life insurance coverage.

In some countries, the first two types are separated and subject to different regulators. This was the case in the United States until the recent repeal of the Glass–Steagall Act, which separated banking and securities functions. This is an example of asset restrictions on financial institutions. In other countries with a so-called universal bank model, a bank can engage in traditional banking and securities activities.

Financial institutions also include other intermediaries that constitute the "buy side" of Wall Street, as opposed to banks and brokers, the "sell side" that intermediates in financial markets. The buy side consists of professional (as opposed to private) investors, called institutional investors, which include insurance companies, pension and endowment funds, investment companies (e.g., mutual and closed-end funds), and hedge funds. These are subject to different regulatory requirements from banks and securities houses.

At the outset, we should ask the question of whether the regulation of financial institutions is at all necessary. After all, other industries are not regulated (except for antitrust reasons, i.e., to avoid monopolies such as in the recent Microsoft case). Private corporations already have their own governance mechanism, which is shareholder supervision. Shouldn't shareholders decide on the appropriate risk-return profile for the company in which they have invested their own funds? Why should governments intervene in free markets? Why do we need regulators?

28.2 SYSTEMIC RISK

Unlike other entities, banks and securities houses play a special role of intermediation. They facilitate payment flows across customers and maintain markets for financial instruments. This very role, however, can also make bank failures much more disruptive for the economy than the failure of other entities. The primary threat is systemic risk.

Systemic risk is defined as the risk of a sudden shock that would damage the financial system and create ripple effects throughout the economy. Systemic risk involves contagious transmission of the shock due to actual or suspected exposure to a failing bank. This is usually accompanied by a flight to quality, which reflects an increased demand for government securities, pushing up the relative cost of capital for the corporate sector. If prolonged, this can lead to a fall in investment

---

[2] The term *bank* in *investment bank* is a misnomer, since these institutions do not extend credit like commercial banks.

[3] Brokers act as pure intermediaries and simply match buyers with sellers. As a result, they take no market risk. In contrast, dealers stand ready to buy and sell securities at given prices. Therefore, they must maintain an inventory of securities and are exposed to market risk.

spending. Higher rates will also dampen consumption spending, slowing down the entire economy.

Indeed, failures in the domestic banking system have been particularly damaging. Among emerging markets, domestic financial collapses have often cost more than 10% of a country's gross domestic product (GDP). In each case, the government (rather, the taxpayer) has paid for the failure, hoping that this would be less costly than allowing a domestic banking failure to spread to the rest of the economy.

This threat largely explains why governments have come to the support of the banking system after September 2008. In October, the British government launched a 500 billion pound bailout plan, of which 50 billion was injected as new capital in banks. Many European governments followed suit, as well as the U.S. government, where the Emergency Economic Stabilization Act of 2008 allowed spending up to $700 billion on the financial system. The goal was not to bail out Wall Street bankers, but rather to avoid collateral damage to the rest of the economy.

Systemic risk can come from two sources:

- *Panicky behavior of depositors or investors.* This can arise from the failure of an institution or a political shock. In a bank run, depositors become worried about the stability of their bank (when there is no deposit insurance) and demand an immediate return of their funds, which may lead to a failure of the bank. Similarly, a sudden drop in securities prices may lead to margin calls, forcing leveraged investors to liquidate their positions, which puts further pressure on prices. Some institutions may fail, resulting in a loss of liquidity and a credit crunch.
- *Interruptions in the payment system.* This can arise from the failure of an institution or from a technological breakdown in the payment system. Banks and securities houses are central to the payment system by which transactions for goods, services, and assets are cleared and settled. When an institution cannot pay, it may expose the payment system to a breakdown.

28.3 REGULATION OF COMMERCIAL BANKS

28.3.1 Bank Runs and Deposit Insurance

Our experience with systemic risk is profoundly marked by the banking crisis of the 1930s in the United States. The banking system was subject to bank runs, when depositors lost faith in the ability of their deposit bank to make full payment and consequently "ran to the bank" to withdraw their funds.

This can happen even though the bank is perfectly solvent, that is, have assets (e.g., loans, real estate) whose value exceeds its liabilities (e.g., demand deposits). Because such assets are illiquid, however, the bank may not be able to meet redemptions immediately, leading to default. Indeed during the U.S. banking crisis of the 1930s, one bank in three failed, causing a severe contraction of credit.

In response, the United States established federal deposit insurance in 1933. The insurance fund protects investors if their bank fails, thereby eliminating the need for a bank run. This scheme was widely credited for stopping bank runs. By now, most countries have a compulsory deposit insurance program. As of 1980, all U.S. bank accounts were insured up to a limit of $100,000.

The fear of bank failure was a major factor during the credit crisis that started in 2007. Northern Rock, a British bank, ran into difficulties rolling over its short-term debt during 2007. Because deposit insurance in Britain was limited to 90% of the first 35,000 pounds, depositors queued to get their money back. Northern received a liquidity support facility from the Bank of England and was eventually nationalized in February 2008.

During the sudden worsening of the crisis that started with Lehman's failure in September 2008, depositors became increasingly worried about the safety of the banking system. In response, many governments increased their deposit insurance limit. In October, European governments increased the limit to 50,000 euros or above. The U.S. temporarily increased the limit to $250,000.

The problem with deposit insurance, however, is that some of the financial risk is now passed on to the deposit insurance fund (i.e., ultimately the government or taxpayer). This creates a moral hazard problem.[4] The banks can alter their behavior, increasing their risks at the expense of the deposit insurance fund. This possibility justifies regulation of insured institutions.

28.3.2 Payment Failures

Turning next to the other source of systemic risk, the prime example of a break-down in the payment system was the June 1974 failure of Bankhaus Herstatt, a small German bank active in the foreign exchange market. The bank was shut down by noon, U.S. time, after having received payments in German marks. In exchange, the counterparty banks were due to receive payment on the same afternoon in U.S. dollars. These payments never came, however, creating a serious liquidity squeeze for counterparties. This event caused severe disruption in the payment system and was perhaps the most extreme shock experienced in the foreign exchange market. What has become known as Herstatt risk has led to a concerted effort by bank regulators to try to avoid such situations, which ultimately gave birth to the Basel Committee on Banking Supervision (BCBS).

28.3.3 Bank Regulators

The BCBS consists of central bankers from the Group of Ten (G-10) countries.[5] Its primary objective is to promote the safety and soundness of the global financial

---

[4] Moral hazard is the name given to problematical (immoral) behavior, which increases the possibility of negative outcomes (hazards).

[5] The Basel Committee's members are senior officials from the G-10 (Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Sweden, United Kingdom, and the United States plus Luxembourg and Switzerland), who meet four times a year, usually in Basel, under the aegis of the Bank for International Settlements. Its web site is http://www.bis.org.

system, that is, to try to control systemic risk. Another objective is to create a system that ensures a level playing field for global financial institutions.[6]

The Basel Committee has established minimum risk-based capital standards that apply to so-called core institutions. These represent internationally active *commercial* banks, which are major players in large-value payment systems. The capital adequacy rules are described in a series of documents known as the Basel Accord, which will be analyzed in the following chapters.

It should be emphasized that core institutions are ultimately regulated by their domestic banking regulators. Although pronouncements of the Basel Committee are not legally binding, member countries have implemented them. Even countries that are not part of the Basel Committee often feel obligated to abide by the same regulations. By now, over 100 countries have adopted the framework of the Basel Accord. In fact, the Accord applies to all internationally active commercial banks.

In the United States, for instance, commercial banks are regulated by the Board of Governors of the Federal Reserve System (the "Fed"),[7] the Office of the Comptroller of the Currency (OCC),[8] and the Federal Deposit Insurance Corporation (FDIC).[9] This fragmentation of supervision is somewhat puzzling but is common among U.S. agencies.

In the United Kingdom, the regulatory framework is more logical, with only one regulator for banks, securities markets, and insurance firms—the Financial Services Authority (FSA).[10] This all-powerful regulator was created in October 1997, taking over banking supervision from the Bank of England.

Banks in the European Union (EU) are subject to minimum standards, which are binding over all member countries.[11] A new Capital Adequacy Directive (CAD) implementing Basel II was adopted in June 2006 and applies to all EU banks.

In Japan, supervision of financial markets, including banking, securities business, and insurance, rests with the Financial Services Agency (FSA), established in July 2000. This responsibility is shared with the central bank, or Bank of Japan, which conducts monetary policy and ensures the stability of the financial system by monitoring financial institutions.[12]

---

[6] At that time, one concern was that Japanese banks were expanding into global markets and were able to undercut their competitors due to more lenient Japanese regulations.

[7] The Federal Reserve supervises all bank holding companies and state-chartered banks that are members of the Federal Reserve System. Its web site is http://www.federalreserve.gov.

[8] The principal function of the OCC is to supervise U.S. national banks and branches and agencies of foreign banks in the United States. National banks are defined as those chartered by the federal government, as opposed to state banks. The OCC is a bureau of the Treasury Department. Its web site is http://www.occ.treas.gov.

[9] The FDIC is a U.S. government agency whose mission is to maintain the stability and public confidence in the nation's financial system. It has provided deposit insurance since 1933. Its web site is http://www.fdic.gov.

[10] Its web site is http://www.fsa.gov.uk.

[11] The EU includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom, plus former Eastern bloc countries. This covers all countries in Western Europe except for Switzerland and Norway.

[12] The web sites for the FSA and the Bank of Japan are http://www.fsa.go.jp and http://www.boj.or.jp.

28.3.4 Regulatory Capital

The Basel Accord sets *minimum* risk-based levels of capital for core institutions. The measurement of regulatory capital will be examined in detail in the next chapter. National authorities, however, are free to adopt arrangements that set higher levels or other criteria. The Federal Reserve board, for example, has an additional requirement based on the bank's leverage ratio.[13] This places a constraint on the degree to which a banking organization can leverage its equity capital base. Failure to meet the capital-adequacy requirements triggers regulatory action, affecting the types of activities in which institutions can engage and requiring prompt corrective action (PCA), including the possible appointment of a receiver.

Note that regulatory capital can differ from economic capital, which is the amount of capital that a financial institution estimates is needed to secure survival in a worst-case scenario. Banks should have enough actual, equity capital to cover both regulatory capital (otherwise, the regulator would step in) and economic capital (otherwise, the bank would taking too much risk by its own measures). Ideally, the two measures should be in line with each other and reflect the true risks that the bank is taking.

To summarize, the regulation of commercial banks is motivated by two objectives:

1. *Minimizing systemic risk*
2. *Protecting the deposit insurance fund*

28.4 REGULATION OF SECURITIES HOUSES

The regulation of securities houses substantially differs from that of commercial banks. Broker–dealers hold securities on the asset and liability side (usually called long and short) of their balance sheet. Because securities are much more liquid than bank loans, there is no rationale for bank runs.

The objectives of regulation for securities houses are

- *Protecting the customer.* One goal is to protect the firm's customers against a default of their broker–dealer. The rationale here is that small investors (e.g., the traditional "widows and orphans") are less capable of informed investment decisions. Another goal is to protect consumers against excessive prices or opportunistic behavior by financial intermediaries.
- *Ensuring the integrity of markets.* The goal is to ensure that failure by one institution does not destabilize financial markets, causing systemic risk.

Let us first examine the consumer protection argument. First, it must be emphasized that investors are risk takers by definition. As Philip McBride

---

[13] The ratio of (tier 1) capital to total consolidated assets must be greater than 3% plus an additional cushion of 100 to 200 basis points. Tier 1 capital will be defined in the following chapter.

Johnson, former chairman of the Commodities Futures Trading Commission, has put it,

> *Regulation is not meant to insulate investors from the consequences of free economic forces, or from their own poor judgment, but rather from abuses perpetrated by other persons.*

Regulation, however, is generally considered necessary when the market fails in two respects, either through excessive prices or opportunistic behavior.

In a free market with informed customers, prices can be excessive only if sellers collude to maintain high prices. This is why there is a need for antitrust legislation to prevent collusion among financial intermediaries.

Opportunistic behavior can arise if sellers have more information than buyers, which can result from access to inside information. This justifies laws against trading on inside information. Or brokers may have conflicts of interest that push them to give bad advice to their clients for the brokers' personal profit. Likewise, accounting standards and disclosure rules help to reduce asymmetries of information in financial markets, which is ultimately socially beneficial as it increases participation in financial markets.

Finally, brokers are subject to suitability standards when making recommendations to clients. Broker–dealers are obligated to recommend only transactions that are suitable to the client's financial situation, investment objectives, and sophistication. Unsuitable recommendations may constitute fraud, which is punishable by law.

Securities regulators require a prudent capital reserve to achieve the goals of protecting consumers and markets. The purpose of this capital is to ensure an orderly *liquidation* of the institution, in contrast to banks, for which capital is measured on an ongoing basis. These minimum reserves are calculated using different methods that use the total amount of debt, the total amount of money owed customers, and, more recently, measures of market risk based on VAR.

As with commercial banks, securities regulators meet in a global forum, the International Organization of Securities Commissions (IOSCO), based in Montréal.[14] Its Technical Committee addresses regulatory problems related to international securities transactions. The IOSCO and the Basel Committee collaborate on common regulatory issues. Likewise, regulatory authority rests with a domestic supervisor, for example the Securities and Exchange Commission (SEC) in the United States.[15]

---

[14] Its web site is http://www.iosco.org.

[15] The SEC is a U.S. federal agency that has wide authority to oversee the nation's security markets. Among other responsibilities, it regulates the financial reporting practices of public corporations. To make information reporting more transparent, the SEC now requires registrants to disclose quantitative information on market risks using one of three possible alternatives: (i) a tabular presentation of expected cash flows and contract terms summarized by risk category, (ii) a sensitivity analysis expressing possible losses for hypothetical changes in market prices, and (iii) a VAR measure. Its web site is http://www.sec.gov.

Securities regulation is based on either the "comprehensive approach" or the "simplified approach." The comprehensive approach is a system of capital charges detailed by the regulator. In contrast, the simplified approach uses a VAR model.

In the United States, the SEC uses the comprehensive approach with its net capital rule, Rule 15c3-1 under the Securities Exchange Act of 1934. A broker–dealer must satisfy a minimum capital ratio based on the calculated ratio of capital to debt or receivables. This ratio is 6.67% of aggregate debt, or 2% of the total amount of money owed by customers. To compute net capital, only liquid assets are considered, minus haircuts, which provide a further margin of safety in case of default and reflect market risk, liquidity risk, and counterparty risk.

The SEC's net capital rule, however, is widely viewed as conservative. As a result, it has become too expensive to operate derivatives activities under these rules. In January 1999, the SEC issued a ruling that created a class of OTC derivatives dealers, which are dealers active in OTC derivative markets. To bring their regulatory requirements in line with foreign firms and U.S. banks, the SEC created risk-based capital rules based on internal VAR models, which parallel the Basel rules.

28.5 TOOLS AND OBJECTIVES OF REGULATION

Table 28.1 provides a summary of the tools and objectives of financial regulation. Systemic risk is controlled through capital adequacy rules, asset restrictions, and disclosure standards. Consumer protection is achieved through capital standards, disclosure rules, and conflict of interest rules.

Capital adequacy and disclosure rules can help to achieve both objectives. Disclosure reduces asymmetries in capital markets, which, in turn, protects consumers. In addition, more disclosure can also stabilize capital markets. Firms that fail to reveal much information about their activities may be susceptible to market rumors, possibly resulting in loss of business or funding difficulties. Indeed, the turmoil that surrounded the near-failure of Long-Term Capital Management

TABLE 28.1   Tools and Objectives of Financial Regulation

Tools	Objectives	
	Systemic Risk	Consumer Protection
Capital standards	√	√
Disclosure standards	√	√
Asset restrictions	√	
Antitrust enforcement		√
Conflict-of-interest rules		√

Source: Herring and Litan (1995), *Financial regulation in the global economy.* Washington, DC: Brookings Institution.

illustrates the panic behavior of banks that suspect that a financial institution with large positions similar to theirs may fail.

EXAMPLE 28.1: FRM EXAM 2004—QUESTION 64

Government insurance of bank deposits provides banks with an incentive to engage in higher risk business activities. This incentive creates:

a. An insurance arbitrage
b. An insurance fraud
c. A moral hazard
d. A moral risk

EXAMPLE 28.2: INSTITUTIONS SUBJECT TO THE BASEL ACCORD

The Basel Capital Accord applies to these entities:

a. National banks chartered in the United States
b. All internationally active commercial banks
c. All banks and securities firms in the G-10 countries plus Luxembourg
d. Banks regulated by the Swiss banking regulatory authorities

---

EXAMPLE 28.3: INSTITUTIONS SUBJECT TO THE CAD

Which of the following financial institutions needs to comply with the provisions of CAD, the Capital Adequacy Directive? This question concerns the main home-country operations of these banks, not certain overseas subsidiaries or branches.

a. J. P. Morgan (a U.S. bank)
b. Credit Suisse First Boston (a Swiss bank)
c. Deutsche Bank (a German bank)
d. Sumitomo Bank (a Japanese bank)

EXAMPLE 28.4: FRM EXAM 2007—QUESTION 72

Which of the following statements regarding economic capital are true?

I. Economic capital is designed to provide a cushion against unexpected losses at a specified confidence level over a set time horizon.

II. Since regulatory capital models and economic capital models have different objectives, economic capital models cannot help regulators in setting regulatory capital requirements.

III. Firms whose capital exceeds their required regulatory capital are firms that employ their capital inefficiently, and their shareholders would benefit if they used all of their excess capital to repurchase shares or increase dividends.

IV. Economic capital can be used to validate a firm's regulatory capital requirement against its own assessment of the risks it is running.

    a. I, II, and III only

    b. III and IV only

    c. I and IV only

    d. I, III, and IV only

28.6 ANSWERS TO CHAPTER EXAMPLES

Example 28.1: FRM Exam 2004—Question 64

c. Moral hazard occurs when one has less incentive to control risk because the cost of a loss is partially borne by another party, as in the case of deposit insurance.

Example 28.2: Institutions Subject to Basel Accord

b. The capital accord applies to commercial banks with international activities.

Example 28.3: Institutions Subject to the CAD

c. The Capital Adequacy Directive applies to banks within the European Union. Of the four countries listed, only Germany belongs to the EU.

Example 28.4: FRM Exam 2007—Question 72

c. Economic capital is typically constructed using a measure of extreme loss at a high confidence level, so Statement I. is correct. Some measures of economic capital are more sensitive to risk and therefore can be helpful to guide regulatory capital, so Statement II. is false. Firms may have economic capital that is much lower than regulatory catpial but are still required to maintain regulatory capital, so Statement III. is false. Finally, economic capital is certainly useful to assess the efficiency of regulatory capital, so Statement IV. is true.

The Basel Accord

The Basel Capital Accord, concluded on July 15, 1988, represents a landmark financial agreement for the regulation of internationally active commercial banks. It instituted for the first time minimum levels of capital to be held by international banks against financial risks.

Initially, the capital charges were based on a set of standard, rigid rules defined by the Basel Committee on Banking Supervision (BCBS). These risk-based capital adequacy requirements evolved over time, first covering credit risk and then market risks. The latest rules by the Basel Committee, called Basel II and finalized in June 2004, represent an extensive revision of the capital charges that creates more risk-sensitive capital requirements. The new rules also add a charge against operational risks. Overall, Basel II represents a substantial improvement in capital requirements.

The Basel Accord is agreed by all members of the Committee and is endorsed by the Central Bank Governors and Heads of Banking Supervision of the G-10 countries. Although strictly speaking it only applies to internationally active bank within the G-10, these minimum capital requirements have been applied to banks in more than 100 countries as well. U.S. regulators will apply Basel II to the largest U.S. banks only. In contrast, the European Union (EU) has adopted the Basel II rules into EU law, which applies to all banks within the Union.

Section 29.1 provides a broad overview of the Basel Accord. Section 29.2 details the original Basel capital requirements, with particular emphasis on credit risk. Market risk is a complex subject in itself and will be developed in the next chapter. Section 29.3 illustrates the application of capital adequacy ratios for Citibank. Section 29.4 discusses major drawbacks of the original Basel Accord and describes the main components of the New Accord. Finally, Section 29.5 concludes with a general evaluation of capital requirements.

29.1 STEPS IN THE BASEL ACCORD

29.1.1 The Basel I Accord

The original goal of the 1988 Basel Accord, which came into force in 1992, was to minimum capital requirements for commercial banks as a buffer against financial losses. Thus its primary objective was to promote the safety and soundness of the global financial system. A secondary objective was to create a level playing

field for internationally-active banks by setting uniform minimum standards. The risk-based capital charges roughly attempted to create a greater penalty for riskier assets.

Initially, the 1988 Basel Accord only covered credit risk. The Accord set a minimum level of capital expressed as a ratio of the total risk-weighted (RW) assets, which include on–balance sheet and off–balance sheet items. Banks have to hold capital that covers at least 8% of their risk-weighted assets. The purpose of this capital is to serve as a buffer against unexpected financial losses, thereby protecting depositors and financial markets.

29.1.2 The 1996 Amendment

In 1996, the Basel Committee amended the Capital Accord to incorporate market risks. This amendment, which came into force at the end of 1997, added a capital charge for market risk. Banks are allowed to use either a standardized model or an internal models approach (IMA), based on their own risk management system.

The amendment separates the bank's assets into two categories, the trading book and banking book. The trading book represents the bank portfolio with financial instruments that are intentionally held for short-term resale and typically marked-to-market. The banking book consists of other instruments, mainly loans, that are held to maturity and typically valued on a historical cost basis.

The 1996 amendment adds a capital charge for (1) the market risk of trading books, and (2) the currency and commodity risk of the banking book. In exchange, the credit risk charge excludes debt and equity securities in the trading book and positions in commodities (apart from the specific risk charge.) As before, it still includes all OTC derivatives, whether in the trading or banking books.

29.1.3 The Basel II Accord

Capital markets have witnessed enormous changes since the initial Capital Accord of 1988. Increasingly, these credit risk charges have appeared outdated and, even worse, may be promoting unsound behavior by some banks.

In June 2004, the Basel Committee finalized a comprehensive revision to the Basel Accord. In the European Union, the new Capital Adequacy Directive implementing Basel II applies to all banks in the EU, starting in 2007, with the most advanced methods being available from 2008. U.S. regulators apply Basel II to a small number of large banks, with other banks subject to a revised version of Basel I because this is a simpler system. Basel II implementation starts in 2008, with a three-year transition period during which U.S. regulators reserve the right to change the application of rules.

The new framework is based on three pillars, viewed as mutually reinforcing:

1. *Pillar 1: Minimum capital requirement.* These are meant to cover credit, market, and operational risk. Relative to the 1988 Accord, banks have now a wider choice of models for computing their risk charges. The BCBS, however, still tried to keep constant the total level of capital in the global banking system, at 8% of risk-weighted assets.

2. *Pillar 2: Supervisory review process.* Relative to the previous framework, supervisors are given an expanded role. Supervisors need to ensure that
   - Banks have a process in place for assessing their capital in relation to risks.
   - Banks indeed operate above the minimum regulatory capital ratios.
   - Corrective action is taken as soon as possible when problems develop.
3. *Pillar 3: Market discipline.* The New Accord emphasizes the importance of risk disclosures in financial statements. Such disclosures enable market participants to evaluate banks' risk profile and the adequacy of their capital positions. The new framework sets out disclosure requirements and recommendations. Banks that fail to meet disclosure requirements will not qualify for using internal models. As internal models generally lead to lower capital charges, this provides a strong incentive for complying with disclosure requirements. In essence, the trade-off for greater reliance on a bank's own models is greater transparency.

The New Accord provides for finer measurement of credit risk, which will generally lead to lower capital requirements. In order to maintain the overall level of bank capital, however, new capital charges are set against operational risk. Capital adequacy will be measured as follows:

$$\frac{\text{Total Capital}}{\text{Credit Risk} + \text{Market Risk} + \text{Operational Risk}} = \text{Bank's Capital Ratio} > 8\%$$

$$(29.1)$$

As before, credit risk in the denominator is measured by the sum of risk-weighted assets for credit risk. The other items are measured from the multiplication of the market risk charge (MRC) and operational risk charge (ORC) by $(1/8\%) = 12.5$. For instance, if a bank has \$875 in risk-weighted assets and MRC $= \$10$ and ORC $= \$20$, the denominator would be computed as $\$875 + [(\$10 + \$20) \times 12.5] = \$1,250$. The bank then has to hold at least $8\% \times \$1,250 = \$100$ in capital to satisfy the minimum requirement. This is equivalent to saying that the total charge must be at least $8\% \times \$875 + \$10 + \$20 = \$70 + \$10 + \$20 = \$100$.

Figure 29.1 summarizes the coverage of credit, market, and operational risk charges for the banking and trading books. Banks will also have access to a menu of methods to compute their risk charges. These are described in Table 29.1.

We should note that pillar 1 omits some important bank risks. In particular, interest rate risk in the banking book is not covered. Banks are exposed to repricing risk, which arises from differences in the maturity and repricing of assets and liabilities. For instance, a bank funding a long-term fixed-rate loan by short-term deposits could suffer a repricing loss if interest rates increase. Measuring this risk, however, requires modeling the complex behavior of deposits, which do not have a fixed maturity date, as well as the prepayment optionality of loans. Due to these difficulties, there is substantial heterogeneity across banks in the methods used to monitor and manage this risk. As a result, the BCBS has decided that this risk falls under pillar 2. Institutions that are perceived to have more of this risk can be subject to higher capital charges.
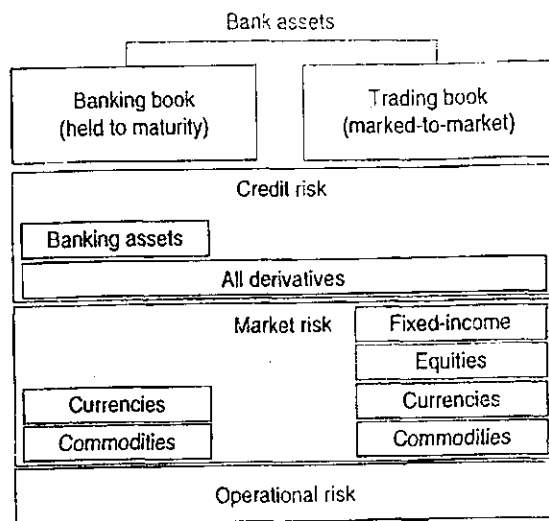
FIGURE 29.1 Summary of Basel II Risk Charges

Finally, there is no formal capital charge for liquidity risk, due to the difficulty of formal measurement of this risk. Yet, the Basel Committee recognizes that "liquidity is crucial to the ongoing viability of any banking organization."

EXAMPLE 29.1: APPLICABLE MARKET RISKS

For regulatory capital calculation purposes, what market risks must be incorporated into a bank's VAR estimate?

a. Risks in the trading account relating to interest rate risk and equity risk

b. Risks in the trading account relating to interest rate risk and equity risk and risks throughout the bank related to foreign exchange and commodity risks

c. Risk throughout the bank related to interest rate risk, equity risk, foreign exchange risk, and commodity risk

d. Interest rate risk, equity risk, foreign exchange risk, and commodity risk in the trading account only

---

TABLE 29.1 Menu of Approaches to Measure Risk

Risk Category	Allowed Approach
Credit	Standardized Approach (based on the 1988 Accord) Foundation Internal Ratings-Based Approach Advanced Internal Ratings-Based Approach
Market	Standardized Approach Internal Models Approach
Operational	Basic Indicator Approach Standardized Approach Advanced Measurement Approach

29.2 THE 1988 BASEL ACCORD

29.2.1 Risk Capital

The 1988 capital adequacy rules require any internationally active bank to carry capital of at least 8% of its total risk-weighted assets. This applies to commercial banks on a consolidated basis. So, for instance, holding companies that are parents of banking groups have to satisfy the capital adequacy requirements.

In the Basel Accord, "capital" has a broader interpretation than the book value of equity. The key purpose of capital is its ability to absorb losses, providing some protection to creditors and depositors. Hence, to be effective, capital must be permanent, must not impose mandatory fixed charges against earnings, and must allow for legal subordination to the rights of creditors and depositors.

The Basel Accord recognizes three forms of capital.

1 Tier 1 Capital, or "Core" Capital
   Tier 1 capital includes equity capital and disclosed reserves, most notably after-tax retained earnings. Such capital is regarded as a buffer of the highest quality. Goodwill is subtracted.[1]
   - Equity capital, or shareholders' funds. This consists of issued and fully paid common stock and nonredeemable, noncumulative preference shares (also called preferred stock).
   - Disclosed reserves correspond to share premiums, retained profits, and general reserves.
2 Tier 2 Capital, or "Supplementary" Capital
   Tier 2 capital includes components of the balance sheet that provides some protection against losses but ultimately must be redeemed or contain a mandatory charge against future income. These include:
   - Undisclosed reserves, or hidden reserves that are allowed by the accounting standards of some countries. These are reserves that passed through the earnings statement but remain unpublished. Due to this lack of transparency, as well as the fact that many countries refuse to recognize undisclosed reserves, undisclosed reserves are not part of core capital.
   - Asset revaluation reserves, which arise, for instance, from long-term holdings of equity securities that are valued at historical acquisition costs. Such capital could be used to absorb losses on a going-concern basis, subject to some discount to reflect market volatility and future taxes in case of sales.
   - General provisions/loan loss reserves, which are held against future unidentified losses. These are the result of loan loss allowances, which are deductions taken against interest income in anticipation of probable credit losses. These deductions reduce retained profits in tier 1 capital but may qualify as tier 2 capital to the extent that they do not reflect a known deterioration in

---

[1] This is an accounting entry which, after an acquisition, goes into book equity to represent the excess of the purchase value over book value. It is omitted because it is not a buffer against losses.

particular assets (in which case they are "specific.")[2] General provisions will play a special role under Basel II.

- Hybrid debt capital instruments, which combine some characteristics of equity and of debt. When they are unsecured, subordinated, and fully paid-up, they are allowed into supplementary capital. These include, for instance, cumulative preference shares.
- Subordinated term debt, with a minimum original maturity of five years, and subject to a discount of 20% during the last five years. Subordinated debt would be junior in right of payment to all other debt in the event of liquidation.

3 **Tier 3 Capital, for Market Risk Only**
Tier 3 capital consists of short-term subordinated debt with a maturity of at least two years. This is eligible to cover market risk only.

There are additional restrictions on the relative amount of various categories. Of the 8% capital charge for credit risk, at least 50% must be covered by tier 1 capital. Next, the amount of tier 3 capital is limited to 250% of tier 1 capital allocated to support market risks (tier 2 capital can be substituted for tier 3 capital if needed). Other restrictions apply to various elements of the three tiers.

For credit risk, the eligible capital must exceed the regulatory capital, or

$$\text{Eligible Tier 1 Capital for CR} + \text{Allowed Tier 2 Capital} \geq \text{CRC} \qquad (29.2)$$

A similar constraint applies to market risk capital:

$$\text{Eligible Tier 1 Capital for MR} + \text{Allowed Tier 3 (or 2) Capital} \geq \text{MRC} \qquad (29.3)$$

A worked-out example later will be given later. Next, we look at the construction of risk charges.

EXAMPLE 29.2: FRM EXAM 2002—QUESTION 71

What is the best definition of tier 1 regulatory capital?

a. Equity capital, retained earnings, disclosed reserves
b. Subordinated debt, undisclosed reserves
c. Equity capital, subordinated debt with a maturity greater than five years
d. Long-term debt, revaluation reserves

---

[2] As credit losses occur, they are charged against this reserve instead of profits, which helps to smooth out earnings.

EXAMPLE 29.3: FRM EXAM 2007—QUESTION 53

Consider a bank balance sheet with (1) common stock of USD 600,000,000;
(2) unrealized long-term marketable equity securities gain: USD 5,000,000;
(3) allowance in anticipation of possible credit losses: USD 5,000,000; (4)
goodwill: USD 30,000,000.

Based solely on the above information, the Tier 1 and Tier 2 capital
numbers are, respectively:

a. USD 595,000,000 USD 45,000,000
b. USD 570,000,000 USD 10,000,000
c. USD 600,000,000 USD 15,000,000
d. USD 630,000,000 USD 20,000,000

EXAMPLE 29.4: FRM EXAM 2004—QUESTION 29

Consider the following financial data for a bank, in millions of dollars:
Shareholders' funds: 627.4. Retained earnings: 65.6. Undisclosed reserves:
33.5. Goodwill: 21.3. Subordinated debt: 180.0. Specific provisions: 11.7.
The ratio of tier 2 to tier 1 capital is:

a. 30.81%
b. 31.78%
c. 33.53%
d. 34.03%

29.2.2 On–Balance Sheet Risk Charges

We first examine on–balance sheet assets, which consist primarily of loans for
most credit institutions. Ideally, the capital charges should recognize differences
in asset credit quality.

Indeed, the 1988 Basel Accord applies to the notional of each asset a risk
capital weight taken from four categories, as described in Table 29.2. Each dollar
of risk-weighted notional exposure must be covered by 8% capital.

These categories provide an extremely rough classification of credit risk. For
instance, claims on Organization for Economic Cooperation and Development
(OECD) central governments, such as holdings of U.S. Treasuries, are assigned a
weight of zero since these assets have presumably no default risk.[3] Cash held is also

---

[3] The OCED currently consists of 30 counties, including Austria, Belgium, Canada, Denmark,
France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, The Netherlands, Norway, Portugal,

TABLE 29.2   Risk Capital Weights by Asset Class

Weights	Asset Type
0%	Cash held
	Claims on OECD central governments
	Claims on central governments in national currency
20%	Cash to be received
	Claims on OECD banks and regulated securities firms
	Claims on non-OECD banks below one year
	Claims on multilateral development banks
	Claims on foreign OECD public-sector entities
50%	Residential mortgage loans
100%	Claims on the private sector (corporate debt, equity, ...)
	Claims on non-OECD banks above one year
	Real estate
	Plant and equipment

assigned a zero weight. At the other extreme, claims on corporations, including loans, bonds, and equities, receive a 100% weight, whatever the risk of default or maturity of the loan.

The credit risk charge (CRC) is then defined for balance sheet items (BS) as

$$CRC(BS) = 8\% \times (RWA) = 8\% \times \left(\sum_i RW_i \times Notional_i\right) \qquad (29.4)$$

where RWA represents risk-weighted assets, and RW_i is the risk weight attached to asset i.

EXAMPLE 29.5: FRM EXAM 2001—QUESTION 38

A bank subject to the Basel I Accord makes a loan of $100m to a firm with a risk weight of 50%. What is the basic on–balance sheet credit risk charge? -

a. $8m
b. $4m
c. $2m
d. $1m

29.2.3 Off–Balance Sheet Risk Charges

By the late 1980s, focusing on balance sheet items only missed an important component of the credit risk of the banking system, which is the exposure to swaps.

Spain, Sweden, Switzerland, Turkey, United Kingdom, United States, Japan, Finland, Australia, New Zealand, Mexico, Czech Republic, Hungary, South Korea, Poland, and Slovakia, in order of accession.

The first swaps were transacted in 1981. By 1990, the outstanding notional of open positions had grown to $3,500 billion, which seems enormous. Some allowance had to be made for the credit risk of swaps. Unlike loans, however, the notional amount does not represent the maximum loss.

To account for such off–balance sheet (OBS) items, the Basel Accord computes a "credit exposure" that is equivalent to the notional for a loan, through credit conversion factors (CCFs). The Accord identifies five broad categories.

1. Instruments that substitute for loans (e.g. guarantees, bankers' acceptances, and standby letters of credit serving as guarantees for loans and securities) carry the full 100% weight (or credit conversion factor). The rationale is that the exposure is not different from a loan. Take a financial letter of credit (LC), for instance, which provides irrevocable access to bank funds for a client. When the client approaches credit distress, it will almost assuredly draw down the letter of credit. Like a loan, the full notional is at risk. This category also includes asset sales with recourse, where the credit risk remains with the bank, and forward asset purchases.
2. Transaction-related contingencies (e.g., performance bonds or commercial letters of credit related to particular transactions) carry a 50% factor. The rationale is that a performance letter of credit is typically secured by some income stream and has lower risk than a general financial LC.
3. Short-term, self-liquidating trade-related liabilities (e.g., documentary credits collateralized by the underlying shipments) carry a 20% factor.
4. Commitments with maturity greater than a year (such as credit lines), as well as note issuance facilities (NIFs), carry a 50% credit conversion factor. Shorter-term commitments or revocable commitments have a zero weight. Note that this applies to the unfunded portion of commitments only, as the funded portion is an outstanding loan and appears on the balance sheet. Under Basel II, shorter-term commitments now receive a CCF of 20%.
5. Other derivatives, such as swaps, forwards and options on currency, interest rate, equity, and commodity products are given special treatment given the complexity of their exposures.

For the first four categories, the position is replaced by a credit equivalent, computed as

$$\text{Credit Equivalent} = \text{Credit Conversion Factor} \times \text{Notional} \tag{29.5}$$

For the last category (derivatives), the credit exposure is computed as the sum of the current, net replacement value (NRV) plus an add-on that is supposed to capture future or potential exposure:

$$\text{Credit Exposure} = \text{NRV} + \text{Add-On}$$
$$\text{Add-On} = \text{Notional} \times \text{Add-On Factor} \times (0.4 + 0.6 \times \text{NGR}) \tag{29.6}$$

Here, the add-on factor depends on the tenor (maturity) and type of contract, as listed in Table 29.3 (NGR will be defined later). It roughly accounts for the maximum credit exposure which, as we have seen before, depends on the volatility of the risk factor and the maturity. As we have seen, volatility is highest for commodities, then equity, then currencies, then fixed-income instruments. This explains why the add-on factor is greater for currency, equity, and commodity swaps than for interest rate instruments, and also increases with maturity.

More precisely, the numbers have been obtained from simulation experiments (such as those in Chapter 21) that measure the 80th percentile worst loss over the life of a matched pair of swaps. The matching of pairs reflects the hedging practice of swap dealers and effectively divides the exposure in two, since only one swap can be in-the-money. Take, for instance, a currency swap with five-year initial maturity. Assuming exchange rates are normally distributed and ignoring interest rate risk, the maximum credit exposure as a fraction of the notional should be

$$\mathrm{WCE} = \frac{1}{2} \times 0.842 \times \sigma\sqrt{5}, \qquad\qquad (29.7)$$

where the $\frac{1}{2}$ factor reflects swap matching and the 0.842 factor corresponds to a one-sided 80% confidence level. Assuming a 10% annual volatility, this gives WCE = 9.4%. This is in line with the add-on of 7.5% in Table 29.3.

Further simulations by the Bank of England and the New York Fed have shown that these numbers also roughly correspond to a 95th percentile loss over a six-month horizon. In the case of a new five-year interest-rate swap, for instance, the worst exposure over the life at the 80th% level is 1.49%; the worst exposure over six months at the 95th% level is 1.58%. This is in line with the add-on of 1.5% for this category.

Next, the NGR factor in Equation (29.6) represents the net-to-gross ratio, or ratio of current net market value to gross market value, which is always between 0 and 1. The purpose of this factor is to reduce the capital requirement for contracts that fall under a legally valid netting agreement. Without netting agreements in place, i.e., with NGR = 1, the multiplier (0.4 + 0.6 × NGR) is equal to one. There is no reduction in the add-on.

TABLE 29.3 Add-On Factors for Potential Credit Exposure (Percent of Notional)

Residual Maturity (tenor)	Contract				
	Interest Rate	Exchange Rate, Gold	Equity	Precious Metals	Other Commodities
< 1 year	0.0	1.0	6.0	7.0	10.0
1 – 5 year	0.5	5.0	8.0	7.0	12.0
> 5 year	1.5	7.5	10.0	8.0	15.0

On the other hand, take a situation where a bank has two swaps with the same counterparty currently valued at $+100$ and at -60. The gross replacement value is the sum of positive values, which is 100. The net value is 40, creating a NGR ratio of 0.4. The multiplier $(0.4 + 0.6 \times NGR)$ is equal to 0.64.

At the other extreme, if all contracts currently net out to zero, $NGR = 0$, and the multiplier $(0.4 + 0.6 \times NGR)$ is equal to 0.4. The purpose of this minimum of 0.4 is to provide protection against *potential movements* in the NGR which, even if currently zero, could change over time.

The computation of risk-weighted assets is then obtained by applying counterparty risk weights to the credit exposure in Equation (29.6). Since most counterparties for such transactions tend to be excellent credit, the risk weights from Table 29.3 are multiplied by 50%. The credit risk charge for OBS items is defined as

$$CRC(OBS) = 8\% \times \left(\sum_i RW_i \times 50\% \times Credit\ Exposure_i \right) \qquad (29.8)$$

Example: The Credit Charge for a Swap

Consider a $100 million interest-rate swap with a domestic corporation. Assume a residual maturity of four years and a current market value of $1 million. What is the credit risk charge?

Answer
Since there is no netting, the factor $(0.4 + 0.6 \times NGR) = 1$. From Table 29.3, we find an add-on factor of 0.5. The credit exposure is then $CE = \$1,000,000 + \$100,000,000 \times 0.5\% \times 1 = \$1,500,000$. This number must be multiplied by the counterparty-specific risk weight and one-half of 8% to derive the minimum level of capital needed to support the swap. This gives $60,000.

EXAMPLE 29.6: FRM EXAM 2001—QUESTION 45

The Basel Accord computes the credit exposure of derivatives using both replacement cost and an "add-on" to cover potential future exposure. Which of the following is the correct credit risk charge for a purchased seven-year OTC equity index option of $50m notional with a current mark to market of $15m with no netting and a counterparty weighting of 100%?

a. $1.6m

b. $1.2m

c. $150,000

d. $1m

EXAMPLE 29.7: FRM EXAM 2000—QUESTION 134

The capital requirement for an unfunded, short-term (under one year) credit commitment is

a. 0%

b. 4%

c. 8%

d. 100%

29.2.4 Total Risk Charge

Finally, the total risk charge is computed as the sum of the credit risk charges, both for on–balance sheet and off–balance sheet items, plus the market risk charge. Define MRC as the market risk charge, which will be detailed in the next chapter.

To translate all numbers into similar risk-adjusted assets, the MRC is transformed into a risk-adjusted asset equivalent, by dividing the MRC by 8%. For instance, if MRC is computed as $1,832 million, the risk-adjusted asset number would be $22.9 billion, which is taken as equivalent to the notional of loans.

We can then simply sum the risk-adjusted assets across all risk categories to find the total risk charge (TRC):

$$TRC = CRC + MRC = 8\% \times (Total\ Risk - Adjusted\ Assets) \qquad (29.9)$$

subject to various restrictions on the use of different tiers. The New Accord adds an operational risk charge to this.

Table 29.4 gives an example. The total risk-adjusted assets for credit risk are 7,500. The market risk charge is 350 which translates into 350/8% = 4,375 in risk assets. The credit risk charge is 8% of 7,500, or 600. Of this, no more than 50% can be accounted by tier 2 capital. So, we could have 300 in tier 1 capital plus 300

TABLE 29.4 Computation of Capital Requirements

Category	Risk Assets	Capital Charge (8%)	Minimum Capital, Required	Available Capital	Minimum Capital, Actual	Eligible Capital
Credit risk	7,500	600	Tier 1: 300		Tier 1: 500	
			Tier 2: 300		Tier 2: 100	
Market risk	4,375	350	Tier 1: 100		Tier 1: 100	
			Tier 3: 250		Tier 3: 250	o
Tier 1				700		700
Tier 2				100		100
Tier 3				600		250
Total	11,875	950		1,400	950	1,050
Capital ratio						8.8%

in tier 2 capital covering credit risk. For market risk, we know the maximum ratio of tier 3 to tier 1 capital is 250 to 100..Hence, with a 350 market risk charge, we can have a maximum allocation of 250 for tier 3 for every 100 of tier 1.

The next step is to match these numbers with the available capital. Assume the bank has capital available of 700, 100, and 600 in tiers 1, 2, and 3, respectively. For credit risk, we only have 100 in tier 2 capital, so that the remaining 500 must be in the form of tier 1 capital. For market risk, we apply the maximum of 250 in tier 3 capital, so that the remainder of 100 comes from tier 1 capital.

This leaves a buffer of excess capital. We can compute the capital ratio using all eligible capital. All of tier 1 capital is eligible, plus 100 in tier 2, plus 250 in tier 3. This sums to a total of 1,050, which translates into an "eligible" capital ratio of 1,050/11,875 = 8.8%. The bank has also 600 − 250 = 350 in unused tier 3 capital.

29.3 ILLUSTRATION

As an illustration, let us examine the capital adequacy requirements for Citibank, which was once the biggest global commercial bank.

Table 29.5 summarizes on–balance sheet and off–balance sheet items as of December 2007. The bank has total assets of $1,252 billion, consisting of cash

TABLE 29.5   Citibank's Credit Risk-Weighted Assets

On–Balance Sheet Assets ($ Billion)

Item	Notional	Not Cov'd	Risk Weight Category			
			0%	20%	50%	100%
Cash and due	86.2	0.0	38.1	43.9	0.0	4.1
Securities	166.2	(1.9)	73.3	85.8	2.3	6.7
Loans and leases	677.1	(10.7)	15.3	96.2	154.0	422.3
Trading assets	215.4	215.4	0.0	0.0	0.0	0.0
All other assets	106.7	21.6	3.9	14.4	1.1	65.7
Total on–BS	1251.7	224.6	130.6	240.2	157.3	498.9

Off–Balance Sheet Items ($ Billion)

Item	Notional	Conv. Factor	Credit Equiv.	Risk Weight Category			
				0%	20%	50%	100%
Financial standby LC	84.5	1.00	84.5	11.4	32.9	3.3	36.9
Performance standby LC	17.1	0.50	8.6	2.0	1.3	0.0	5.3
Commercial LC	9.0	0.20	1.8	0.1	0.4	0.0	1.3
Securities lent	152.8	1.00	152.8	152.5	0.2	0.0	0.0
Other credit substitutes	0.3	—	2.6	0.0	0.0	0.0	2.6
Other off–balance sheet	2.5	1.00	2.5	0.0	0.0	0.8	1.6
Unused commitmt. >1 yr	187.1	0.50	93.6	2.5	36.0	5.0	50.0
Derivative contracts	27,235		271.1	11.7	124.9	134.6	0.0
Total off–BS			617.4	180.3	195.7	143.8	97.7

TABLE 29.6   Citibank's Risk-Weighted Assets

Risk-Weighted Assets ($ Billion)					
	Risk Weight Category				
Item	0%	20%	50%	100%	Total
On–BS and off–BS items	310.9	435.9	301.1	596.6	
Credit RW assets	0.0	87.2	150.5	596.6	834.3
Market RW assets					78.3
Others					0.0
Total RW assets					912.6

equivalents, securities, loans, trading assets, and other assets. The notional for each asset is assigned to one of the four risk weighted categories, ranging from 0% to 100%. For example, out of the $166 billion in securities, $73 billion have a zero risk weight because these represent, for instance, positions in OECD government bonds. Of the remainder, $86 billion has a 20% weight, $2 billion has a 50% weight, and $7 billion has a 100% weight. Most of the loans carry a risk weight of 100%. Trading assets are excluded from this computation because they carry a market risk charge only.

The second panel of the table displays off–balance sheet information. The second column displays the notional, the third the conversion factor, and the fourth the credit equivalent, which is the product of the previous two. As described in the previous section, the conversion factors are 1.00 for financial LCs and securities lent, 0.50 for performance LCs and unused commitments greater than one year, and 0.20 for commercial LCs.[4]

Finally, note the huge size of the notional derivatives position. At $27,235 billion, it is several times the size of Citibank's total assets of $1,252 billion and dwarfs its equity of $99 billion. The notional amounts, however, give no indication of the risk. The credit equivalent amount, which consists of the net replacement value plus the add-on, is $271 billion, a much lower number.

From this information, we can compute the total risk-weighted assets and capital adequacy ratios. This is shown in Tables 29.6 and 29.7. The first line adds up on–balance sheet and off–balance sheet items for each category. Multiplication by the risk weights gives the second line. The total RW assets for credit risk are $834 billion, which consists of $625 billion for on–BS items and $209 billion for off–BS items. To this, we add the RW assets for market risk, or $78 billion. Thus, most of Citibank's regulatory risk capital covers credit risk. Market risk represents less than 10% of the total.

The total RW assets add up to $913 billion. Applying the 8% ratio, we find a minimum regulatory capital of $73 billion. In fact, the available risk capital

---

[4] The category "credit substitutes" represent residual interests, such as the equity tranche from securitizations of assets, which are subject to a dollar-for-dollar capital requirement. This implies a credit conversion factor of (1/8%) = 12.50. U.S. regulators have imposed this high capital requirement to reflect the higher risk of such residual interests, whose value can be wiped out easily in case of losses on the underlying assets.

TABLE 29.7 Citibank's Capital Requirements

Capital	Amount ($ Billion)	Ratio (Percent)
Equity	99.3	
Goodwill	−21.2	
Others	3.9	
Tier 1	82.0	9.0%
Sub.debt	28.2	
LL. Allowance	11.4	
Others	0.1	
Tier 2	39.7	4.3%
Total	121.6	13.3%
Tier 1 Leverage		6.7%

adds up to $122 billion, which represents a 13.3% ratio, comfortably above the regulatory minimum. The ratio for a well capitalized bank would be 10%. Apparently, the regulatory constraint is not binding.[5]

As long as the regulatory capital ratio is not binding, the bank could decide on a target optimal capital ratio, based on a careful consideration of the trade-off between increasing expected returns and increasing risks. In other words, this requires a measure of economic capital. If the current capital ratio is viewed as too high relative to this target, the bank could shrink its capital base through dividend payments or share repurchases. Like other major banks, Citibank has decided to hold more capital than the minimum regulatory standard of 8%. Conversely, if the bank thinks it requires more capital, it could issue new shares.

Table 29.8 summarizes the financial statements for Citigroup starting in 2003 and ending in the third quarter of 2008. Citigroup is a bank holding company that wholly owns Citibank, which is the commercial bank. From 2003 to 2006, the company has expanded aggressively, as indicated by the growth in assets, but has maintained a capital ratio in a narrow range. This was managed in part by dividend payments and share repurchases.

Starting in 2007, however, the company's situation deteriorated quickly, with a sharp downturn in net income and fall in capital ratio. As a result, the company stopped repurchasing shares and instead issued a large amount of tier 1 capital in order to maintain its capital ratio. In January 2008, Standard & Poor's lowered Citigroup's credit rating from AA to AA−.

On October 28, 2008, the company raised $25 billion through the U.S. Treasury Troubled Asset Relief Program (TARP), which was in essence a government bailout. The company also decreased its quarterly dividend from $0.54 per share in 2007 to $0.32 during the first three quarters of 2008 and to $0.16 in October. The credit rating was further lowered by two notches to A in December 2008. Over the year, Citi's stock price went from $29.4 to $6.7.

---

[5] In addition, based on total assets, tier 1 leverage is 6.7%, which is above the minimum ratio set by the Federal Reserve board, as described in the previous chapter.

TABLE 29.8   Citigroup's Summary Financials ($ Billion)

	2003	2004	2005	2006	2007	2008Q3
Assets	1,264	1,484	1,494	1,884	2,188	2,050
Risk-weighted assets	750	852	885	1,058	1,253	1,176
Equity	98	109	113	120	114	126
Capital ratio	12.0%	11.9%	12.0%	11.7%	10.7%	11.7%
Net income	17.8	17.0	24.6	21.5	3.6	–11.0
Dividends	5.8	8.4	9.2	9.8	10.8	6.0
Tier 1 issuance	2.3	0.9	1.4	1.8	12.9	37.3
Share repurchase	2.4	1.8	12.8	7.0	0.7	0.0

29.4   THE NEW BASEL ACCORD

The Basel Accord has been widely viewed as successful in raising banking capital ratios. As a result of the Accord, the aggregate tier 1 ratio increased from $840 to $1,500 billion from 1990 to 1998 for the 1,000 largest banks.

29.4.1   Issues with the 1988 Basel Accord

Over time, however, these regulations have shown their age. The system has led to regulatory arbitrage, which can be broadly defined as bank activities aimed at getting around these regulations. Lending patterns have been transformed, generally in the direction of taking on more credit risk to drive the economic capital up to the level of regulatory capital.

To illustrate, consider a situation where a bank can make a loan of $100 million to an investment-grade company rated AAA or to a speculative-grade company rated CCC. Irrespective of the credit quality, the bank is forced to hold regulatory capital of $8 million, so it has to borrow $92 million. Suppose the rate of return on the AAA loan is 6%, after expenses. The cost of borrowing is close, at 5.7%. The dollar return to shareholders is then $100,000,000 × 6% − $92,000,000 × 5.7% = $756,000. Compared to a capital base of $8 million, this represents a rate of return of 9.5% only, which may be insufficient for shareholders. The bank could support this loan with a much smaller capital base. For instance, a capital base of $2 million would require borrowing $98 million and would yield a return of $100,000,000 × 6% − $98,000,000 × 5.7% = $414,000, assuming the cost of debt remains the same. This translates into a rate of return of 20.7%, which is much more acceptable. The bank, however, is unable to lower its capital due to the binding regulatory requirement.

Suppose now the rate of return on the CCC loan is 7%, after expenses and expected credit losses. The dollar return to shareholders is now $1.756 million, which represents a 22.0% rate of return. In this situation, the bank has an incentive to increase the risk of its loan in order to bring the economic capital more in line with its regulatory capital. This simple example has shown that regulation may perversely induce banks to shift lending to lower-rated borrowers.

In addition to inadequate differentiation of credit risk, the 1988 Accord did not recognize credit mitigation techniques, nor diversification effects for credit risk. Some of these drawbacks have been corrected with Basel II.

EXAMPLE 29.8: RETURN ON BANK EQUITY

A bank that funds itself at LIBOR − 5bp purchases an A+ rated corporate floating coupon loan paying LIBOR + 15bp. Based on the Basel I minimum capital requirements, what is the annualized return on regulatory capital for this loan?

a. 2.5%

b. 5.0%

c. 11%

d. None of the above

29.4.2 Definition of Capital

The new Basel Accord, dubbed Basel II, was finalized in June 2004. It gives banks a choice between a standardized approach, which is a simple extension of the Basel I rules, and a more complex internal ratings-based (IRB) approach.

For the former, capital is still defined as before. However, general provisions or loan loss reserves can be included in tier 2 only subject to a limit of 1.25% of risk-weighted assets.

For the IRB approach, in contrast, the new Accord distinguishes between Expected Loss (EL) and Unexpected Loss (UL). Capital is supposed to absorb unexpected losses, which means that it cannot support expected losses as well. Banks typically fund accounts called general provisions, or loan loss reserves to absorb expected credit losses. Hence, Basel II withdraws general provisions from tier 2 capital.[6]

29.4.3 The Credit Risk Charge

As before, the credit risk charge is computed as the sum of individual credit charges:

$$\text{CRC} = 8\% \times \left(\sum_i \text{RW}_i \times \text{N}_i \right) \qquad (29.10)$$

---

[6] If total expected losses are less than eligible provisions. however, the difference may be recognized in tier 2 capital, up to a maximum of 0.6% of risk-weighted assets. On the other hand, if total expected losses exceed eligible provisions, the bank must deduct the difference from capital (50% from tier 1 and 50% from tier 2).

In general, the capital charges are calibrated to correspond to the amount of capital required to support a 99.9% confidence level over a one-year horizon.

A notable feature of this approach is that it is *additive*. The main reason for this specification is that it should be robust. However, it is not clear why the addition of individual capital charges should lead to a capital charge that reflects a 99.9% VAR measure for the entire portfolio. After all, summing individual VARs certainly does not add up to the portfolio VAR. It turns out, however, that when default correlations are generated by a single factor, this decomposition adds up to a good approximation of the portfolio risk. The capital charges have been chosen by the Basel Committee to represent bank portfolios of typical sizes with typical default correlations. This analysis is similar to the component VAR decomposition presented in Chapter 16.

Banks have now a choice of three approaches for the risk weights.

(1) Standardized Approach
This is an extension of the 1988 Accord, but with finer classification of categories for credit risk, based on external credit ratings, provided by external credit assessment institutions. Table 29.9 describes the new weights, which now fall into five categories for banks and sovereigns, and four categories for corporates. For sovereigns, OECD membership is no longer given preferential status. For banks, two options are available. The first assigns a risk weight one notch below that of the sovereign—the other uses an external credit assessment. The new Accord also removes the 50% risk weight cap on derivatives.

(2) Foundation Internal Ratings-Based Approach (FIRB Approach)
Under the internal ratings-based approach (IRB), banks are allowed to use their internal estimate of creditworthiness, subject to regulatory standards. Under the foundation approach, banks estimate the probability of default (PD) and supervisors supply other inputs, which carry over from the standardized approach.

TABLE 29.9   Risk Weights: Standardized Approach

Claim	Credit Rating					
	AAA/ AA–	A+/ A–	BBB+/ BBB–	BB+/ B–	Below B–	Unrated
Sovereign	0%	20%	50%	100%	150%	100%
Banks-option 1	20%	50%	100%	100%	150%	100%
Banks-option 2	20%	50%	50%	100%	150%	50%
Short-term	20%	20%	20%	50%	150%	20%
Claim	AAA/ AA–	A+/ A–	BBB+/ BB–		Below BB–	Unrated
Corporates	20%	50%	100%		150%	100%

Note: Under option 1, the bank rating is based on the sovereign country in which it is incorporated. Under option 2, the bank rating is based on an external credit assessment. Short-term claims are defined as having an original maturity less than three months.

TABLE 29.10   IRB Risk Weights

Probability of Default	Corporate	Residential Mortgage	Other Retail
0.03%	14.44%	4.15%	4.45%
0.10%	29.65%	10.69%	11.16%
0.25%	49.47%	21.30%	21.15%
0.50%	69.61%	35.08%	32.36%
0.75%	82.78%	46.46%	40.10%
1.00%	92.32%	56.40%	45.77%
2.00%	114.86%	87.94%	57.99%
3.00%	128.44%	111.99%	62.79%
4.00%	139.58%	131.63%	65.01%
5.00%	149.86%	148.22%	66.42%
0.00%	193.09%	204.41%	75.54%
20.00%	238.23%	253.12%	100.28%
50.00%	217.87%	226.62%	105.94%

Note: Illustrative weights for LGD = 45%, maturity of 2.5 years, and large corporate exposures (firms with turnover greater than 50 million euros).

Table 29.10 illustrates the link between PD and the risk weights for various asset classes. For instance, a corporate loan with a 1.00% probability of default would be assigned a risk weight of 92.32%, which is close to the standard risk weight of 100% from Basel I. Note that retail loans have much lower risk weights than the other categories, reflecting their greater diversification.[7]

(3) Advanced Internal Ratings-Based Approach (AIRB Approach)

Under the advanced approach, banks can supply other inputs as well. These include loss given default (LGD) and exposure at default (EAD). The combination of PDs and LGDs for all applicable exposures are then mapped into regulatory risk weights. The capital charge is obtained by multiplying the risk weight by EAD by 8%. The advanced IRB approach applies only to sovereign, banks and corporate exposures and not to retail portfolios.

Adoption of Approach   Banks with simple portfolios can follow the standardized approach. More advanced banks are expected to adopt an IRB approach. To be eligible for the IRB approach, a bank must demonstrate to its supervisor that it meets a set of minimum requirements. Most importantly, the internal rating system must be consistent and reliable. Also, banks cannot allocate borrowers across rating systems, or cherry-pick ratings to minimize capital requirements. In addition, the bank-developed rating system must be approved at the highest level and subject to independent oversight.

---

[7] Also, these weights cover unexpected losses; as the PD increases to very high levels, the weights start to decrease because most of the losses are expected and hence should be covered by general provisions.

Once a bank adopts an IRB approach, it is expected to extend it eventually across all asset classes and across the entire banking group. Banks adopting the IRB approach are expected to continue to employ it; a voluntary return to the standardized approach is only permitted in special cases, as approved by the supervisor.

Credit Risk Mitigation   The New Accord also recognizes credit risk mitigation (CRM) techniques, such as collateralization, third-party guarantees, credit derivatives, and netting. Collateralized credit exposures are those where the borrower has posted assets as collateral. Recognition is only given to cash, gold, listed equities, investment-grade debt, sovereign securities rated BB— or better, or mutual funds investing in the same assets.

Under the Standardized Approach, two treatments are possible. In the simple approach, the risk of the collateral is simply substituted for that of the counterparty, generally subject to a 20% floor. In contrast, the comprehensive approach is more accurate and will lead to lower capital charges.

Even if the exposure is exactly matched by the collateral, there is some credit risk due to the volatility of values during a default. In the worst case, the value of the exposure could go up and that of the collateralized assets could go down. This volatility effect is measured by a haircut parameter (H) that is instrument-specific and approximates the 99% VAR over a ten-day period. For equities, for example, $H = 25\%$. For cash, this is zero.

The exposure after risk mitigation is then

$$E^* = E \times (1 + H_e) - C \times (1 - H_c - H_{fx}) \qquad (29.11)$$

if positive, where E is the value of the uncollateralized exposure, C is the current market value of the collateral held, H_e is the haircut appropriate to the exposure, H_c is the haircut appropriate to the collateral, and H_{fx} is the haircut appropriate for a currency mismatch between the two.

Under the Foundation Approach, only the comprehensive approach is allowed. The effective loss given default (LGD*) is derived from the usual LGD and the current value of the exposure E and the exposure after risk mitigation E^*

$$LGD^* = LGD \times (E^*/E) \qquad (29.12)$$

Other forms of CRM are guarantees and credit derivatives, which are a form of protection against obligor default provided by a third party, called the guarantor. Capital relief, however, is only granted if there is no uncertainty as to the quality of the guarantee. Protection must be direct, explicit, irrevocable, and unconditional. In such a situation, one can apply the principle of substitution. In other words, if bank A buys credit protection against a default of company B from bank C, it may substitute C's credit risk for B's risk. It will do so if the credit rating of bank C is better than that of B.

TABLE 29.11 Risk Weights for Securitizations: Standardized Approach

	AAA/ AA−	A+/ A−	BBB+/ BBB−	BB+/ BB−	B+ and below or unrated
Tranche	20%	50%	100%	350%	1,250% (deduction)

An allowance can be made, however, for the low probability of double default. In order for bank A to incur a credit loss, both B and C must default. The likelihood of such a double default occurrence is generally very low. For instance, if defaults are independent, the probability of a credit loss is given by the *product* of the two default probabilities. In July 2005, the BCBS adopted new capital requirements that account for double default effects

$$RW_{DD} = RW_0(0.15 + 160 \times PD_g) \tag{29.13}$$

where RW_0 is the original capital requirement and PD_g is the probability of default of the guarantor, bank C in this case.

Securitization Finally, the New Accord also deals explicitly with securitization, which involves the economic or legal transfer of assets to a third party, typically called special purpose vehicle (SPV). Examples are asset-backed securities such as collateralized loan obligations, where the underlying asset is a pool of bank loans. Because of the high regulatory cost of keeping loans on their balance sheets, banks are now routinely transforming loans into tradeable securities. The securitization process is explained in Chapter 7.

A bank can remove these assets from its balance sheet only after a true sale, which is defined using clean break criteria. These are satisfied if a number of conditions are all met: (1) significant credit risk must be transferred to third parties, (2) the seller does not maintain effective or indirect control over the assets,[8] (3) the securities are not an obligation of the seller, and (4) the holders of the SPV have the right to pledge or exchange those interests. Two other technical conditions are also involved.

If these conditions are all met, then the bank can remove the assets from its balance sheet and becomes subject to new risk weights for securitization tranches. These are described in Table 29.11 under the Standardized Approach. For example, the risk weight for a BBB-rated tranche is 100%. For the lowest-rated tranches, the bank must hold capital equal to the notional amount, which implies a risk weight of $(1/8\%) = 1250\%$.

---

[8] In particular, the transferred assets must be legally separated from the seller so that it does not have additional obligations in case the SPV goes bankrupt. Also, the seller cannot maintain effective control either by being able to repurchase the assets at a profit or by being obligated to retain the risk of the transferred assets.

EXAMPLE 29.9: FRM EXAM 2004—QUESTION 67

Which of the following statements about the Basel II capital requirements is *false*?

a. It increases the risk sensitivity of minimum capital requirements for internationally active banks.

b. It only addresses credit risk and market risk.

c. United States insurance companies are not required to comply with Basel II capital requirements.

d. Banks are not allowed to use their internal models for credit risk in determining the capital requirements for credit risk.

EXAMPLE 29.10: FRM EXAM 2006—QUESTION 79

Which of the following is *not* allowed by the Basel II Accord for measuring credit risk?

a. Banks may measure credit risk in a standardized manner supported by external credit assessments.

b. Banks may measure credit risk using a bank's internal credit rating systems, subject to the explicit approval of the bank's supervisor.

c. Using external ratings for certain assets and internal ratings for the remaining assets.

d. Banks may measure credit risk using their own internal estimates of the probability of default, loss given default, the exposure at default, and effective maturity, subject to certain minimum conditions, disclosure requirements, and supervisory approval.

EXAMPLE 29.11: FRM EXAM 2006—QUESTION 108

Which of the following statements is *not* correct about the Foundation IRB and the Advanced IRB approaches for credit risk capital charges in Basel II?

a. Under the Advanced IRB approach, banks are allowed to use their own estimates of PD, LGD, EAD, and correlation coefficient but must use the risk-weighted functions provided by the supervisors.

b. Under the Foundation IRB approach, banks provide their own estimates of PD and rely on supervisory estimates for other risk components.

c. Banks adopting the Advanced IRB approach are expected to continue to employ this approach. A voluntary return to the standardized approach is only permitted in extraordinary circumstances.

d. Under both Foundation IRB and Advanced IRB approaches, the expected loss is not included in the credit risk capital charge.

EXAMPLE 29.12: FRM EXAM 2006—QUESTION 90

Under the comprehensive approach for the Foundation Internal Ratings Based approach under Basel II, which of the following methods is used for calculating the effective loss given default (LGD*) where:

LGD* is the effective loss given default (considering risk mitigation measures). LGD is that of the senior unsecured exposure before recognition of collateral. E is the current value of the exposure (i.e., cash lent or securities lent or posted). E^* is the exposure value after risk mitigation.

 a. $LGD^* = LGD \times (E^*/E)$
 b. $LGD^* = LGD \times (E^*) * (E)$
 c. $LGD^* = LGD \times (E^* + E)$
 d. $LGD^* = LGD \times (E^* - E)$

29.4.4 The Operational Risk Charge

One of the most significant, and controversial, additions to the New Accord is the operational risk charge (ORC). The Basel Committee expects that the ORC will represent on average 12% of the total capital charge.

The new rules give three alternatives methods. The simplest is called the basic indicator approach. This is based on an aggregate measure of business activity. The capital charge equals a fixed percentage, called alpha factor, of the exposure indicator defined as gross income (GI):[9]

$$ORC^{BIA} = \alpha \times GI \tag{29.14}$$

where α has been set at 15%. The advantage of this method is that it is simple, transparent, and uses readily available data. The problem is that it does not account for the quality of controls. As a result, this approach is expected to be mainly used by non-sophisticated banks.

The second method is the standardized approach. Here, bank's activities are divided into eight business lines. Within each business line, gross income is taken as an indicator of the scale of activity. The capital charge is then obtained by multiplying gross income by a fixed percentage, called beta factor, and summing across business lines:

$$ORC^{SA} = \sum_{i=1}^{8} \beta_i \times GI_i \tag{29.15}$$

---

[9] This is taken as the average of positive gross income numbers over the last three years. Negative values are excluded.

The β factors are described in Table 29.12. This approach is still simple but better reflects varying risks across business lines.[10] Trading and sales, for example, carries a higher weight to reflect the possibility of high severity losses due to trader fraud.

The third class of method is the advanced measurement approach (AMA). This allows banks to use their own internal models in the estimation of required capital using quantitative and qualitative criteria set by the Accord. It can only be used if the bank demonstrates effective management and control of operational risk.

The qualitative criteria are similar to those for the use of internal market VAR systems.[11] Once these are satisfied, the risk charge is obtained from the unexpected loss (UL), or VAR at the 99.9% confidence level over a one-year horizon:

$$ORC^{AMA} = UL(1 - year, \ 99.9\% \ confidence) \tag{29.16}$$

Normally, the expected loss (EL) must be included in the capital charge, unless the bank can demonstrate that it adequately captures EL in its internal business practices.

Other quantitative criteria are as follows: (1) Banks must track internal loss data measured over a minimum period of five years; (2) banks must use external data; (3) banks must use scenario analysis to evaluate its exposure to high-severity events; and (4) banks must take into account the business environment and internal control factors. Finally, insurance can be used to offset up to 20% of the operational risk charge.

This approach offers the most refined measurement of operational risk and is expected to be used by more sophisticated institutions.

TABLE 29.12 Beta Factors

Business Line	Beta Factor
Corporate finance	18%
Trading and sales	18%
Retail banking	12%
Commercial banking	15%
Payment, settlement	18%
Agency services	15%
Asset management	12%
Retail brokerage	12%

[10] The formula is actually more complex and allows offsets for some negative GI numbers in a year with positive numbers in other business lines, up to a limit of zero. The exact formula is $ORC^{SA} = \{\sum_{i=1}^{3} Max[\sum_{i=1}^{8}(\beta_i \times GI_i), 0]\}/3$.

[11] Specifically, (1) the bank must have an independent operational risk function, (2) the system must be integrated in day-to-day management, (3) there must be regular reporting, (4) documentation must exist, (5) auditors must perform regular reviews, and (6) there must be external validation.

EXAMPLE 29.13: FRM EXAM 2007—QUESTION 117

Which of the following approaches for calculating operational risk capital charges leads to a higher capital charge for a given accounting income as risk increases?

a. The basic indicator approach
b. The standardized approach
c. The advanced measurement approach
d. All of the above

EXAMPLE 29.14: FRM EXAM 2004—QUESTION 53

Which of the following statements about its methodology for calculating an operational risk capital charge in Basel II is *correct*?

a. Basic indicator approach is suitable for institutions with sophisticated operational risk profile.
b. Under the standardized approach, capital requirement is measured for each of the business line.
c. Advanced measurement approaches will not allow an institution to adopt its own method of assessment of operational risk.
d. AMA is less risk-sensitive than the standardized approach.

EXAMPLE 29.15: FRM EXAM 2004—QUESTION 42

According to the Basel Accord's advanced measurement approach, how are operational capital requirements calculated?

a. As in credit risk, there are formulae specified in the Accord so that only the inputs have to be estimated.
b. Capital requirements have to be estimated using historical data bases but the Accord does not specify which statistical distribution has to be used.
c. The formulae are the same as those used for credit risk capital requirements.
d. Each national supervisor must specify the formulae that the banks have to use.

EXAMPLE 29.16: FRM EXAM 2007—QUESTION 6

Which of the following statements regarding Basel II non-advanced approaches is *incorrect*?

a. The standardized approach makes it advantageous for a bank to book losses early if doing so reduces this year's gross income sufficiently to make it negative.
b. Corporate finance, trading and sales, and payment and settlement are the business lines with the highest regulatory capital requirements.
c. The standardized approach divides the bank into business lines and uses data from the last three years of a business line's gross income and a beta factor to obtain the regulatory capital for that business line.
d. The standardized approach uses data from the last three years of gross income to obtain a bank's operational risk capital charge.

29.4.5 Evaluation

The BCBS has organized a large-scale analysis of the effect of the new capital requirements on the banking system. Table 29.13 reports the results for 228 banks in the G-10 countries. The table shows that the new capital charge will affect banks differentially. Smaller banks, with more retail exposures, will have lower capital requirements than before. Retail risks are indeed more diversified than other types.

Larger banks are more likely to adopt the AIRB approach because it leads to lower capital requirements than the standardized approach. The table lists the

TABLE 29.13   Percentage Changes in Capital Requirements
(Banks in G-10 Countries)

| | Larger Banks | | Smaller Banks | |
| | Method | | Method | |
Portfolio	Standardized	IRB	Standardized	IRB
Corporate	0.9%	−5.0%	−1.0%	−4.5%
Bank	1.5%	0.4%	0.2%	0.1%
Sovereign	0.2%	1.3%	−0.1%	0.6%
SME	−0.2%	−1.3%	−0.1%	−2.2%
Mortgage	−6.3%	−7.6%	−6.2%	−12.6%
Retail	−0.7%	−0.9%	−2.5%	−4.5%
Other	0.8%	2.6%	0.0%	1.5%
Overall credit risk	−3.8%	−10.5%	−9.7%	−21.6%
Operational risk	5.6%	6.1%	8.3%	7.5%
Overall change	1.8%	−4.4%	−1.4%	−14.1%

Source: QIS5 study conducted by the BCBS (2006).

results for the standardized and most likely IRB approach. For instance, large banks will suffer a slightly higher capital charge (by 1.8%) under the standardized approach, which is primarily due to the addition of the operational risk charge. Under the AIRB approach, however, the credit risk charge drops by 10.5%, which leads to a net decrease in capital requirements of −4.4%.

Because the Basel Committee wants to keep the total level of global banking capital unchanged, the new framework introduces a scaling factor, which applies to the credit capital requirements under the IRB approach. This factor, which is incorporated in the table, is currently set at 1.06.

29.5 CONCLUSIONS

The Basel II Accord represents a major step forward for the measurement and management of banking risks. It creates more risk-sensitive capital charges for credit risk and, for the first time, attempts to account for operational risk.

Among winners will be banks that invest in risk management systems, banks with large retail portfolios and with high-grade corporate credits. Indeed, all of these should have lower credit risk than the rest of the industry.

This new framework is certainly not perfect, however. The standardized approach has been criticized for the greater role given to credit rating agencies. The internal approach is viewed as giving too much discretion to banks. These features, however, are certainly much better than the Basel I alternative and are the inevitable result of the move toward risk-sensitive capital charges.

Like any set of formal rules, the Basel II rules leave open some possibilities of regulatory arbitrage, due to discrepancies between economic and regulatory capital for some assets. The framework incorporates "typical" correlations in the construction of the credit risk charge. Institutions that have greater diversification than typical banks cannot enjoy lower capital charges. In theory, this could be corrected by using internal portfolio credit risk models, developed by the banks themselves. In practice, these models are not allowed for setting credit risk charges because they are not viewed as robust enough. Indeed, the losses suffered during the credit crisis that started in 2007 have highlighted major weaknesses in credit risk management.

More importantly, the system of capital charges has missed an important element of banking risk, which is liquidity risk. The need to recapitalize major banks by governments proves that capital levels were not adequate to protect against a major financial crisis. In addition, more risk-sensitive capital charges could have a procyclical effect. In a recession, defaults increase, leading to greater credit risk and greater capital charges. At the same time, the banking system suffers credit losses that erode its actual capital. Caught between lower actual capital and greater regulatory capital requirements, the banking system could respond by reducing its lending activities, thereby aggravating the extent of the recession. The Basel Committee will surely re-examine the adequacy of its capital standards.

EXAMPLE 29.17: FRM EXAM 2002—QUESTION 72

Under the new Basel Accord, which of the following best defines the overall minimum capital ratio?

a. (Total Capital) / (Credit Risk + Market Risk + Operational Risk) = Capital Ratio > 8%
b. (Total Capital) / (Credit Risk + Market Risk + Operational Risk) = Capital Ratio < 8%
c. (Total Capital) / (Credit Risk + Market Risk) = 8%
d. (Tier 1 Capital) / (Market Risk + Operational Risk) = 8%

EXAMPLE 29.18: FRM EXAM 2007—QUESTION 19

Your bank is implementing the AIRB approach for credit risk, the AMA for operational risk, and the internal model approach for market risk. The Chief Risk Officer (CRO) wants to estimate the bank's total risk by adding up the regulatory capital for market risk, credit risk, and operational risk. The CRO asks you to identify the problems with using this approach to estimate the bank's total risk. Which of the following statements about this approach is *incorrect*?

a. It assumes market, credit, and operational risks have zero correlation.
b. It uses a 10-day horizon for market risk.
c. It ignores strategic risks.
d. It ignores the interest risk associated with the bank's loans.

29.8 IMPORTANT FORMULAS

Basel I credit risk charge:
$$CRC = 8\% \times RWA = 8\% \times \left(\sum_i RW_i \times N_i \right)$$

Basel I derivatives credit exposure: Credit Exposure = NRV + Add-On
$$Add\text{-}On = Notional \times Add\text{-}On\ Factor \times (0.4 + 0.6 \times NGR)$$

Basel I total risk charge: TRC = CRC + MRC

Basel II total risk charge: TRC = CRC + MRC + ORC

Basel II credit risk charge: $CRC = 8\% \times \left(\sum_i RW_i \times N_i \right)$
Standardized approach: $RW = f(\text{credit rating})$
FIRB: $RW = f(PD)$
AIRB: $RW = f(PD, LGD)$

Basel II RW including double default effects: $RW_{DD} = RW_0(0.15 + 160 \times PD_g)$

Basel II operational risk charge:

Basic indicator approach: $ORC^{BIA} = \alpha \times GI$

Standardized approach: $ORC^{SA} = \sum_{i=1}^{8} \beta_i \times GI_i$

AMA: $ORC^{AMA} = UL(1 - year, 99.9\% \text{ confidence}) + EL$

29.7 ANSWERS TO CHAPTER EXAMPLES

Example 29.1: Applicable Market Risks

b. In addition to all the risks in the trading book (interest rate, equity, forex, commodity), the market capital charges also include forex and commodity risks in the bank book.

Example 29.2: FRM Exam 2002—Question 71

a. Tier 1 capital includes equity capital, disclosed reserves, and retained earnings. Tier 2 includes undisclosed reserves, hybrid debt, and subordinated debt.

Example 29.3: FRM Exam 2007—Question 53

b. Tier 1 capital consists of equity minus goodwill, or USD 570m. Tier 2 capital includes asset revaluation reserves of $5m and loan loss reserves of $5m. For this question, it is sufficient to find the correct number for tier 1 capital.

Example 29.4: FRM Exam 2004—Question 29

b. Tier 1 capital consists of shareholders' funds plus retained earnings, minus goodwill, which is 671.7. Tier 2 capital consists of subordinated debt plus undisclosed reserves, or 213.5. The ratio is 31.78%. Specific provisions cannot be included in risk capital, because they are likely to be absorbed by specific bad loans.

Example 29.5: FRM Exam 2001—Question 38

b. Under the Basel I rules, the charge is $100 \times 50\% \times 8\% = \$4$ million.

Example 29.6: FRM Exam 2001—Question 45

a. From Table 29.3, the add-on factor is 10%. This gives a credit exposure of $15 + \$50 \times 10\% = \$20$ million, and a credit risk charge of $20 \times 8\% = \$1.6$ million.

Example 29.7: FRM Exam 2000—Question 134

a. Unfunded commitments are off-balance sheet items (unlike funded commitments, which are loans). Below a year, the credit conversion factor is zero, which means zero Basel weight.

Example 29.8: Return on Bank Equity

a. An 8% capital charge applies to this bond. We buy $100 worth of the bond, which is funded at the bank rate, for a net dollar return of $100[(L + 0.15\%) - (L - 0.05\%)] = \$0.20$. We need to keep $8 in capital, which we assume is not invested. The rate of return is then $\$0.20/\$8 = 2.5\%$.

Example 29.9: FRM Exam 2004—Question 67

b. This is because Basel II also covers operational risk. Banks can provide inputs but cannot use their internal models for credit risk, so answer d. is false.

Example 29.10: FRM Exam 2006—Question 79

c. The Basel II rules do not allow cherry-picking.

Example 29.11: FRM Exam 2006—Question 108

a. Banks are never allowed to use their own correlations.

Example 29.12: FRM Exam 2006—Question 90

a. See Equation (29.12). Also, this answer is the only one that makes sense taking units into account because LGD is a unitless ratio.

Example 29.13: FRM Exam 2007—Question 117

c. The basic indicator approach uses a factor of $\alpha = 15\%$. The standardized approach uses a factor ranging from 12% to 18%. For the same level of income, if risk increases for instance by having more exposure to trading, the second approach will correctly pick up an increase in risk. Similarly for the AMA, which is even more risk-sensitive.

Example 29.14: FRM Exam 2004—Question 53

b. The BI approach is suitable for banks with basic risk profiles, so answer a. is incorrect. The AMA approach is an internal model, so answer c. is incorrect. AMA is more risk-sensitive than the SI approach, so answer d. is incorrect.

Example 29.15: FRM Exam 2004—Question 42

c. The AMA approach is an internal models approach, which does not specify a particular loss distribution.

Example 29.16: FRM Exam 2007—Question 6

a. Statement b. is correct, given Table 29.12. Statement a. is incorrect, as only positive income is considered.

Example 29.17: FRM Exam 2002—Question 72

a. The ratio of capital to total risk must be greater than 8%.

Example 29.18: FRM Exam 2007—Question 19

a. The market risk charges uses a 10-day horizon, so statement b. is correct. The Basel capital charges do ignore strategic risk and interest rate in the banking book. Adding up the capital charges assumes perfect correlations (or at least high correlations, implying extreme shocks happen at the same time), not zero correlations.

CHAPTER 30

The Basel Market Risk Charge

After the credit risk charges were instituted in 1988, the Basel Committee turned its attention to market risk in response to the increased proprietary trading activities of commercial banks. The Capital Accord was amended in 1996 to include a capital charge for market risk, which was implemented by January 1998.

The capital charge can be computed using two methods. The first is based on a "standardized" method, similar to the credit risk system with add-ons determined by the Basel rules. Because diversification effects are not fully recognized, this method generates a high market risk charge. The second method is called the internal models approach (IMA) and is based on the banks' own risk management systems, which are more adaptable than the rigid set of standardized rules. This approach must be viewed as a breakthrough in financial regulation. For the first time, regulators relied on the banks' own VAR systems to determine the capital charge. Since banks may have an incentive to understate their market risk, however, the internal models approach also includes a strong system of verification, based on backtesting.

The market risk framework was updated and fully incorporated in Basel II. During the events of 2007 and 2008, however, a number of banks suffered very large losses not captured by the usual VAR numbers. As a result, the Basel Committee has proposed a number of changes to the market risk framework.[1]

This chapter discusses the implementation of capital charges for market risk. Section 30.1 summarizes the standardized method. The application of the internal model approach is described in Section 30.2. Section 30.3 then turns to stress testing. Finally, the framework for backtesting is presented in Section 30.4.

30.1 THE STANDARDIZED METHOD

The objective of the market risk amendment was "to provide an explicit capital cushion for the price risk to which banks are exposed." This was viewed as important in further strengthening the soundness and stability of the international banking system and of financial markers. The original proposal was issued in

---

[1] Basel Committee on Banking Supervision (2008), *Proposed Revisions to the Basel II Market Risk Framework*, Basel: BIS.

April 1993 and was based on a prespecified building block approach. Essentially, this consists of attaching add-ons to all positions, which are added up across the portfolio.

The bank's market risk is first computed for portfolios exposed to interest-rate risk (IR), equity risk (EQ), foreign currency risk (FX), commodity risk (CO), and option risk (OP), using specific guidelines. The bank's total risk is then obtained from the summation of risks across the four categories. Because the construction of the risk charge follows a prespecified process, this approach is sometimes called the standardized method.[2]

The bank's total risk is obtained from the summation of risks across different types of risks, j, on each day, t:

$$\mathrm{MRC}_t^{\mathrm{STD}} = \sum_{j=1}^{5} \mathrm{MRC}_t^j = \mathrm{MRC}_t^{\mathrm{IR}} + \mathrm{MRC}_t^{\mathrm{EQ}} + \mathrm{MRC}_t^{\mathrm{FX}} + \mathrm{MRC}_t^{\mathrm{CO}} + \mathrm{MRC}_t^{\mathrm{OP}}$$

$$(30.1)$$

The interest rate risk charge is the sum of a general market risk charge, which typically increases for longer duration instruments, and a specific risk charge, which covers against issuer-specific risk. For instance, the weight for long-term investment-grade credits is 1.60%. For equity risk, the general market risk charge is 8% of the net positions; the specific risk charge is 8% of the gross positions, unless the portfolio is both liquid and well-diversified, in which case the weight is reduced to 4%. For currency risk, the market risk charge is 8% of the higher of either the net long currency positions or the net short currency positions. For commodity risk, several approaches are possible. In the simplified approach, the risk charge is 15% of the net position in each commodity. Finally, for option risk, several approaches are also possible. In the simplified approach, which applies when banks handle a limited range of purchased options, the capital charge is the lesser of the market risk charge for the underlying security and the option premium.

Thus, the standardized model is relatively straightforward to implement. It is also robust to model misspecification. The building-block approach, however, has been criticized on several grounds. First, the risk classification is arbitrary. For instance, a capital charge of 8% is applied uniformly to equities and currencies without regard for their actual return volatilities. Different currencies have different volatilities relative to the dollar that also can change over time.

Second, the approach leads to high capital requirements because risk charges are systematically added up across different sources of risk, which ignores diversification. For instance, fixed-income charges are computed for each currency separately, then added up across markets. Implicitly, this approach is a worst-case scenario that assumes that the worst loss will occur at the same time across all sources of risk. In practice, these markets are not perfectly correlated, which means that the worst loss will be less than the sum of individual worst losses. Thus, the standardized model fails to recognize the benefits of diversification, which gives

---

[2] See BCBS (1996), *Amendment to the Basel Capital Accord to Incorporate Market Risk*, Basel: BIS.

no incentive for banks to diversify prudently. Recognition of these problems has led to another, more flexible approach based on internal models.

EXAMPLE 30.1: FRM EXAM 2007—QUESTION 63

You are an analyst at Bank Alpha. You were given the task to determine whether under Basel II your bank can use the simplified approach to report option exposure instead of the intermediate approach. Which of the following criteria would your bank have to satisfy in order for it to use the simplified approach?

  a. The bank writes options, but its options trading is insignificant in relation to its overall business activities.
  b. The bank purchases and writes options and has significant options trading.
  c. The bank solely purchases options, and its options trading is insignificant in relation to its overall business activities.
  d. The bank purchases and writes options, but its options trading is insignificant.

30.2 THE INTERNAL MODELS APPROACH

In contrast to the simplistic standardized approach, the internal models approach (IMA) relies on internal risk management systems developed by banks themselves as the basis for the market risk charge.

This approach must be considered as a watershed in financial regulation. For the first time, regulators implicitly recognized that banks had developed sophisticated risk management systems, which should be more precise than simple standardized rules. Indeed, the complexity and speed of development of financial innovations is such that rigid rules can be avoided with new products. Another motivation for the IMA was to provide incentives for banks to develop risk management systems. This is because the IMA approach leads to lower capital charges than the standardized approach.

Regulators, however, have not totally given up their authority. A bank can use internal models only after it has been explicitly approved by the supervisory authority. The bank must satisfy qualitative requirements first. Second, the output is subject to a rigorous backtesting process.

30.2.1 Qualitative Requirements

Not all banks can use internal models, though. Regulators first must have some general reassurance that the bank's risk management system is sound. As a result,

banks first have to satisfy various qualitative standards:

- *Independent risk control unit.* The bank must have a risk control unit that is independent of trading and reports to senior management. This structure minimizes potential conflicts of interest.
- *Backtesting.* The bank must conduct a regular backtesting program, which provides essential feedback on the accuracy of internal VAR models.
- *Involvement.* Senior management and the board need to be involved in the risk control process and devote sufficient resources to risk management.
- *Integration.* The bank's internal risk model must be integrated with day-to-day management. This avoids situations where a bank could compute its VAR simply for regulatory purposes and otherwise ignore it.
- *Use of limits.* The bank should use its risk measurement systems to set internal trading and exposure limits.
- *Stress testing.* The bank should conduct stress tests on a regular basis. Stress tests results should be reviewed by senior management and be reflected in policies and limits set by management and the board of directors.
- *Compliance.* The bank should be compliant with a documented set of policies.
- *Independent review.* An independent review of the trading units and of the risk control unit should be performed regularly, at least once a year. This includes verification with backtesting.

30.2.2 The Market Risk Charge

In addition to these requirements, the bank's risk model must contain a sufficient number of risk factors, where the definition of *sufficient* depends on the extent and complexity of trading activities.[3]

For material exposures to interest rates, there should be at least six factors for yield curve risk plus separate factors to model spread risk. For equity risk, the model should at least consist of beta mapping on an index; a more detailed approach would have industry and even individual risk factor modeling. For active trading in commodities, the risk model should account for movements in spot rates plus convenience yields. Banks should also capture the nonlinear price characteristics of option positions, including vega risk. Correlations *within* broad risk categories are recognized explicitly. Regulators can also recognize correlations *across* risk categories provided the model is sound.

Once these requirements are satisfied, the market risk charge is computed according to these rules:

- *Quantitative parameters.* The computation of daily VAR shall be based on a set of uniform quantitative inputs:
    - A horizon of 10 trading days, or two calendar weeks; banks can, however, scale their daily VAR by the square root of time
    - A 99 percent confidence interval

---

[3] The 2008 proposed revisions would require a bank to justify any factor used in pricing which are left out in the VAR computation.

- An observation period based on at least a year of historical data or, if a non-equal weighting scheme is used, an average time lag of at least six months[4]
- At least quarterly updating, or whenever prices are subject to material changes (so that sudden increases in risk can be picked up)[5]

- *Market risk charge.* The general market capital charge shall be set at the higher of the previous day's VAR, or the average VAR over the last 60 business days times a "multiplicative" factor k. The exact value of this multiplicative factor is to be determined by local regulators, subject to an absolute floor of 3.

  The purpose of this factor is twofold. Without this risk factor, a bank would be expected to have losses that exceed its capital in one 10-day period out of a hundred, or about once in four years. This does not seem prudent. Second, the factor serves as a buffer against model misspecifications, for instance assuming a normal distribution when the distribution really has fatter tails.

- *Plus factor.* A penalty component, called plus factor, shall be added to the multiplicative factor, k, if verification of the VAR forecasts reveals that the bank systematically underestimates its risks. We will discuss this further in the context of backtesting, which will be developed in a further section.

  The purpose of this factor is to penalize a bank that provides an overly optimistic projection of its market risk. It provides a feedback mechanism that rewards truthful internal monitoring and should provide incentives to build sound risk management systems.

In summary, the market risk charge on any day t is

$$\mathrm{MRC}_t^{\mathrm{IMA}} = \mathrm{Max}\left(k\frac{1}{60}\sum_{i=1}^{60}\mathrm{VAR}_{t-i},\ \mathrm{VAR}_{t-1}\right) + \mathrm{SRC}_t, \tag{30.2}$$

where VAR_{t-i} is the bank's VAR over a 10-day horizon at the 99% level of confidence. Here, the factor k reflects both the multiplicative and the plus factors.

The first term consists of a multiplier k times the average VAR over the last 60 days. The second term uses yesterday's VAR, and will be binding only if the positions change dramatically. In practice, this is rarely the case.

Finally, SRC represents the specific risk charge, which represents a buffer against idiosyncratic factors, including basis risk and event risk, related to individual bond and equity issuers. Event risk include the risk of a downgrade and default. Banks that use internal models can incorporate specific risk in their VAR, as long as they (1) satisfy additional criteria and (2) can demonstrate that they can deal with event risk.[6]

---

[4] This is similar to a duration computation. For instance, with equal weights over the last 250 trading days, this average time lag is $\sum_{t=1}^{N} t(1/N) = N(N+1)/2 \ (1/N) = (N+1)/2 = 125.5$ days, or six months. Note that this rules out models such as the GARCH process if the weight on more recent observations is too high. The 2008 proposed revisions would allow to place more weight on more recent data as long as the resulting VAR is higher than with the usual method.

[5] The 2008 proposed revisions would require a minimum monthly update.

[6] The difficulty with event and default risk is that it is typically not reflected in historical data. When a bank cannot satisfy (2), a prudential surcharge is applied to the measure of specific risk.

Consider for instance a corporate bond issued by Ford Motor, a company with a credit rating of "CCC." The general market risk component could capture the effect of movements in yields for an index of CCC-rated corporate bonds. In contrast, the SRC should capture the effect of basis risk between Ford bonds and the index, as well as a downgrade or default.

During the events of 2007 and 2008, a number of banks suffered very large losses not captured by the usual VAR numbers. Banks lost money on credit-related instruments, such as positions in CDOs of ABS, which experienced extreme price moves. Many of these positions had been moved from the banking book to the trading book because of the lower capital charge in the trading book. As a result, the Basel Committee has proposed to expand the specific risk charge to two components.

The first is a more narrow specific risk charge, still based on a 10-day 99% VAR with a multiplier of 3 or more, which covers idiosyncratic risk. This reflects the risk that a security could move by more than the general market factors.

The second is an incremental risk charge calibrated to a 99.9% confidence level over one year that would cover default, credit migration, and credit spread movements for debt instruments. Note that these parameters correspond to the internal ratings-based approach for credit risk in the banking book. For equity instruments, this new charge should cover events such as merger breakups and takeovers. This definition, however, creates some double-counting with risks already incorporated in the usual VAR measures. As of December 2008, this proposal was still open to many questions.

30.2.3 Combination of Approaches

The banks' market risk capital requirement will be either (1) the risk charge obtained by the standardized methodology, obtained from an arithmetic summation across the five risk categories, (2) the risk charge obtained by the internal models approach, or (3) a mixture of (1) and (2) summed arithmetically.

EXAMPLE 30.2: FRM EXAM 2007—QUESTION 91

Under the Basel II Capital Accord, banks that have obtained prior regulatory approval can use the internal models approach to estimate their market risk capital requirement. What approach or methodology is used under the internal models approach to compute capital requirements?

a. Internal rating and vendor models

b. Stress testing and backtesting

c. Expected tail loss, as VAR is not a coherent measure of risk

d. VAR methodology

EXAMPLE 30.3: FRM EXAM 2002—QUESTION 10

Banks have to meet a number of qualitative criteria before they are permitted to use a models-based approach. The qualitative criteria include:

a. The bank should have an independent risk control unit that is responsible for the design and implementation of the bank's risk management system; the unit should conduct a regular backtesting program.

b. The board of directors and senior management should be actively involved in the risk control process; the bank's internal risk measurement model must be closely integrated into the day-to-day risk management process of the bank; an independent review of the risk management system should be carried out regularly; the risk measurement system should be used in conjunction with internal trading and exposure limits.

c. a. and b. together.

d. a. and b., except that the risk measurement system does not have to be used in conjunction with exposure limits.

EXAMPLE 30.4: FRM EXAM 2004—QUESTION 70

Under the market risk amendment to the Basel Accord, a bank can use its internal models to calculate its market risk charge subject to all the following provisions *except*:

a. A time horizon of 10 trading days

b. A 99% confidence level

c. One year of historical observations, which are updated semiannually

d. The market risk charge will be set at the higher of the previous day's VAR or the average VAR over the last 60 days scaled by a multiplicative factor

---

EXAMPLE 30.5: WEIGHTING SCHEME

The 1996 Amendment to the Capital Accord requires that internal models

a. Utilize at least six months of historical data

b. Utilize at least one year of equally weighted historical data

c. Utilize enough historical data so that the weighted average age of the data is at least six months

d. Utilize two years of historical data, unequally weighted

EXAMPLE 30.6: FRM EXAM 2001—QUESTION 42

Which of the following best describes the quantitative parameters of the Internal Models Approach?

a. Ten-day trading horizon, 99% confidence interval, minimum one year of data, minimum quarterly updates

b. One-day trading horizon, 95% confidence interval, five years of data, updated weekly

c. One-day trading horizon, 99% confidence interval, minimum one year of data, updated monthly

d. Ten-day trading horizon, 97.5% confidence interval, minimum five years of data, updated daily

EXAMPLE 30.7: FRM EXAM 2007—QUESTION 114

Your bank is using the internal models approach to estimate its general market risk charge. The multiplication factor k, set by the regulator, is 3 and banks are allowed to use the square root rule to scale daily VAR. The previous day's one-day VAR estimate is EUR 3 million, and the average of the daily VAR over the last 60 days is EUR 2 million. Given the above information, what will be the market risk charge for your bank?

a. EUR 9.49 million

b. EUR 28.46 million

c. EUR 6.32 million

d. EUR 18.97 million

EXAMPLE 30.8: FRM EXAM 2002—QUESTION 62

Specific risk capital charge is designed

a. To protect against credit risk related to the individual issuer of a security

b. To protect against a five standard deviation adverse movement in the price of an individual security

c. To protect against an upward scenario shift in the price of an individual security owing to factors related to the individual issues

d. To protect against credit and liquidity risk related to the individual issuer of a security

30.3 STRESS TESTING

Stress testing is one of the qualitative requirements for a bank to use internal models. The purpose of stress testing is to identify events that could greatly impact the bank but are not captured in VAR measures. A major goal of stress testing is to "evaluate the capacity of the bank's capital to absorb large potential losses."

Stress testing can be described as a process to identify and manage situations that could cause extraordinary losses. This can be done with a set of tools, including (1) scenario analysis, (2) stressing models, volatilities, and correlations, and (3) policy responses.

Scenario analysis consists of evaluating the portfolio under various states of the world. Stress testing also involves evaluating the effect of changes in valuation models, as well as in inputs such as volatilities and correlations. Policy responses consist of identifying steps the bank can take to reduce its risk and conserve capital.

Stress tests fall into three categories:

1. *Scenarios requiring no simulation.* These consist of analyzing large past losses over a recent reporting period to gain a better understanding of the vulnerabilities of the bank. While providing useful information, this approach is backward-looking and does not account for changes in portfolio composition.
2. *Scenarios requiring a simulation.* These consists of running simulations of the current portfolio subject to large historical shocks, for example, the stock market crash of 1987, the Exchange Rate mechanism (ERM) crisis of September 1992, the bond market rout of 1994, the credit crisis that started in 2007, and so on.
3. *Bank-specific scenarios.* These scenarios would be driven by the current position of the bank, instead of historical experience. For instance, a strategy of going long the off-the-run bond while shorting the equivalent on-the-run bond may appear safe based on recent historical patterns of high correlations between these two bonds. With high correlations, a loss on one position will be offset by a gain on the other. This may not be the case, however, if correlations break down. So, the institution should evaluate the effect of a correlation breakdown in this particular example.

The assessment of stress testing is essential to evaluate the risk profile of institutions. Results should be reported routinely to senior management and periodically to the board of directors. When stress-test results reveal a particular vulnerability, corrective action should be taken, by reducing or hedging the position.

In practice, stress testing is much more subjective than VAR measures. The Basel guidelines are suitably vague. First, there is no systematic method to identify scenarios of interest. Second, the process assigns no probability to the extraordinary loss that has been identified. As a result, it is often difficult to know how to follow up on stress-test results. In particular, it would be impractical to guard against every single potential disaster. Overall, however, the most useful aspect of stress testing is that it can help to identify potential weaknesses in the bank's portfolio.

30.4 BACKTESTING

Internal models were allowed by the Basel Committee in large part because they are amenable to verification. Verification is the general process of checking whether the model is adequate. This can be made with a set of tools, including backtesting, stress testing, and independent review and oversight. This section focuses on backtesting techniques for verifying the accuracy of VAR models. Backtesting is a statistical testing framework that consists of checking whether actual trading losses are in line with VAR forecasts. Each exceedence is called an exception.

30.4.1 Measuring Exceptions

But first, we have to define the trading outcome. One definition is the actual profit or loss over the next day. This return, however, does not exactly correspond to the previous day's VAR. All VAR measures assume a *frozen* portfolio from the close of a trading day to the next, and ignore fee income. In practice, trading portfolios do change. Intraday trading will generally increase risk. Fee income is more stable and decreases risk. Although these effects may offset each other, the actual portfolio may have more or less volatility than implied by the VAR.

This is why it is recommended to construct hypothetical portfolios, which are constructed so as to match the VAR measure exactly. Their returns are obtained from fixed positions applied to the actual price changes on all securities, measured from close to close.

The Basel framework recommends using both hypothetical and actual trading outcomes in backtests. The two approaches are likely to provide complementary information on the quality of the risk management system. For instance, suppose the backtest fails using the actual but not the hypothetical portfolio. This indicates that the model is sound but that actual trading increases volatility. On the other hand, if the backtest fails using the hypothetical model, the conclusion should be that the risk model is flawed.

30.4.2 Statistical Decision Rules

The Basel backtesting framework consists of recording *daily* exceptions of the 99% VAR over the last year. Note that even though capital requirements are based on a 10-day period, backtesting uses a daily interval, which entails more observations. On average, we would expect 1% of 250, or 2.5 instances of exceptions over the last year. Too many exceptions indicate that either the model is understating the VAR or the bank is unlucky. How do we decide which explanation is most likely?

Such statistical testing framework must account for two types of error:

- Type 1 errors, which describe the probability of rejecting a correct model, due to bad luck
- Type 2 errors, which describe the probability of not rejecting a model that is false

Ideally, one would want to structure a test that has low type 1 and type 2 error rates. In practice, one has to trade off one type of error against another. Most statistical tests fix the type 1 error rate, say at 5%, and structure the test so as to minimize the type 2 error rate, or to maximize its power. The power of a test is also one minus the type 2 error rate.

Define x as the number of exceptions, n as the total number of observations, and p as the confidence level. The random variable x then has a binomial distribution. Armed with this information, we can find the cutoff point for a type 1 error rate, as seen in Chapter 2 and in Chapter 10.

30.4.3 The Penalty Zones

The Basel Committee has decided that up to four exceptions is acceptable, which defines a "green" light zone. If the number of exceptions is five or more, the bank falls into a "yellow" or "red" zone and incurs a progressive penalty where the multiplicative factor, k, is increased from 3 to 4. The "plus factor" is described in Table 30.1.

An incursion into the red zone generates an *automatic*, nondiscretionary penalty. This is because it would be extremely unlikely to observe 10 or more exceptions if the model was indeed correct.

If the number of exceptions falls within the yellow zone, the supervisor has discretion to apply a penalty, depending on the causes for the exceptions. The Basel Committee uses these categories:

- *Basic integrity of the model*: The deviation occurred because the positions were incorrectly reported or because of an error in the program code. This is a very serious flaw. In this case, a penalty "should" apply and corrective action should be taken.
- *Deficient model accuracy*: The deviation occurred because the model does not measure risk with enough precision (e.g., does not have enough risk factors). This is a serious flaw, too. A penalty "should" apply and the model should be reviewed.
- *Intraday trading*: Positions changed during the day. Here, a penalty "should be considered." If the exception disappears with the hypothetical return, the problem is not in the bank's VAR model.

TABLE 30.1 The Basel Penalty Zones

Zone	Number of Exceptions	Potential Increase in k
Green	0 to 4	0.00
Yellow	5	0.40
	6	0.50
	7	0.65
	8	0.75
	9	0.85
Red	≥ 10	1.00

■ *Bad luck*: Markets were particularly volatile or correlations changed. These exceptions "should be expected to occur at least some of the time" and may not suggest a deficiency of the model but simply bad luck.

EXAMPLE 30.9: FRM EXAM 2002—QUESTION 20

Which of the following procedures is essential in validating the VAR estimates?

a. Stress testing
b. Scenario analysis
c. Backtesting
d. Once approved by regulators no further validation is required.

EXAMPLE 30.10: PENALTY ZONES

The Amendment to the Capital Accord defines the "yellow zone" as the following range of exceptions out of 250 observations

a. 3 to 7
b. 5 to 9
c. 6 to 9
d. 6 to 10

EXAMPLE 30.11: FRM EXAM 2002—QUESTION 23

Backtesting routinely compares daily profits and losses with model-generated risk measures to gauge the quality and accuracy of their risk measurement systems.

The 1996 Market Risk Amendment describes the backtesting framework that is to accompany the internal models capital requirement. This backtesting framework involves

I. The size of outliers
II. The use of risk measure calibrated to a one-day holding period
III. The size of outliers for a risk measure calibrated to a 10-day holding period
IV. Number of outliers

    a. II and III
    b. II only
    c. I and II
    d. II and IV

EXAMPLE 30.12: FRM EXAM 2003—QUESTION 66

Which of the following *correctly* describe the similarities between operational VAR and market VAR?

I. Both VARs, when used for regulatory capital measurement, need to be validated against actual loss experience.
II. Both are built on data (market prices for market VAR and operational loss data for operational VAR) that are readily available.
III. Both are modeled based on a normal distribution.
IV. Extreme Value Theory can be used to model extreme losses at the tail of the distribution for both operational and market VAR.
   a. I and IV
   b. I, II, and III
   c. I, II, and IV
   d. II, III, and IV

30.5 IMPORTANT FORMULAS

Market risk charge, standardized method:
$$MRC_t^{STD} = MRC_t^{IR} + MRC_t^{EQ} + MRC_t^{FX} + MRC_t^{CO} + MRC_t^{OP}$$
Market risk charge, internal models approach:
$$MRC_t^{IMA} = Max\left(k\frac{1}{60}\sum_{i=1}^{60} VAR_{t-i}, VAR_{t-1}\right) + SRC_t$$
Backtests: N, number of exceptions and zones:
Green: $N \le 4$; yellow: $5 < N \le 9$; red: $10 \le N$

30.6 ANSWERS TO CHAPTER EXAMPLES

Example 30.1: FRM Exam 2007—Question 63

c. A bank can only use the simplified approach if it purchases options and its option trading is not significant. Otherwise, it is required to use the intermediate approach. Another way to look at the question is that answer c. contains the weakest conditions, i.e., those least likely to lead to a large loss.

Example 30.2: FRM Exam 2007—Question 61

d. The internal models approach is based on the bank's internal VAR methodology.

Example 30.3: FRM Exam 2002—Question 10

c. The qualitative criteria include those in a., independent risk control, and back-testing, as well as in b., involvement, integration, independent review, and use of limits.

Example 30.4: FRM Exam 2004—Question 70

c. The IMA requires using one year of historical data updated at least quarterly, not semiannually.

Example 30.5: Weighting Scheme

c. Answer b. is correct if the bank uses fixed weights only. Otherwise, the average time lag of the observations cannot be less than six months.

Example 30.6: FRM Exam 2001—Question 42

a. The IMA is based on a 10-day horizon, 99% confidence level, and one year of data, with at least quarterly updates.

Example 30.7: FRM Exam 2007—Question 114

d. The MRC for general risk is $k\text{VAR}\sqrt{10} = 3 \times 2 \times \sqrt{10} = 18.97$. Note that the previous day's VAR is not binding because it is less than 3 times the average VAR.

Example 30.8: FRM Exam 2002—Question 62

a. Specific risk capital is supposed to provide a cushion against idiosyncratic risk, such as a bond default or event risk, in the trading book. It does not cover liquidity risk, which is difficult to ascertain, anyway.

· Example 30.9: FRM Exam 2002—Question 20

c. VAR estimates need to be compared to actual P&L results to be validated, which is called backtesting.

Example 30.10: Penalty Zones

b. See Table 30.1.

Example 30.11: FRM Exam 2002—Question 23

d. The backtesting framework in the IMA only counts the number of times a daily exception occurs, i.e., a loss worse than the VAR. So, this involves the number of outliers and the daily VAR measure.

Example 30.12: FRM Exam 2003—Question 66

a. Operational loss data are not readily available, so II. is wrong. Operational loss distributions are not normally distributed, so III. is wrong. EVT is used for both, so IV. is correct. Both approaches need to be validated with actual data, so I. is correct.
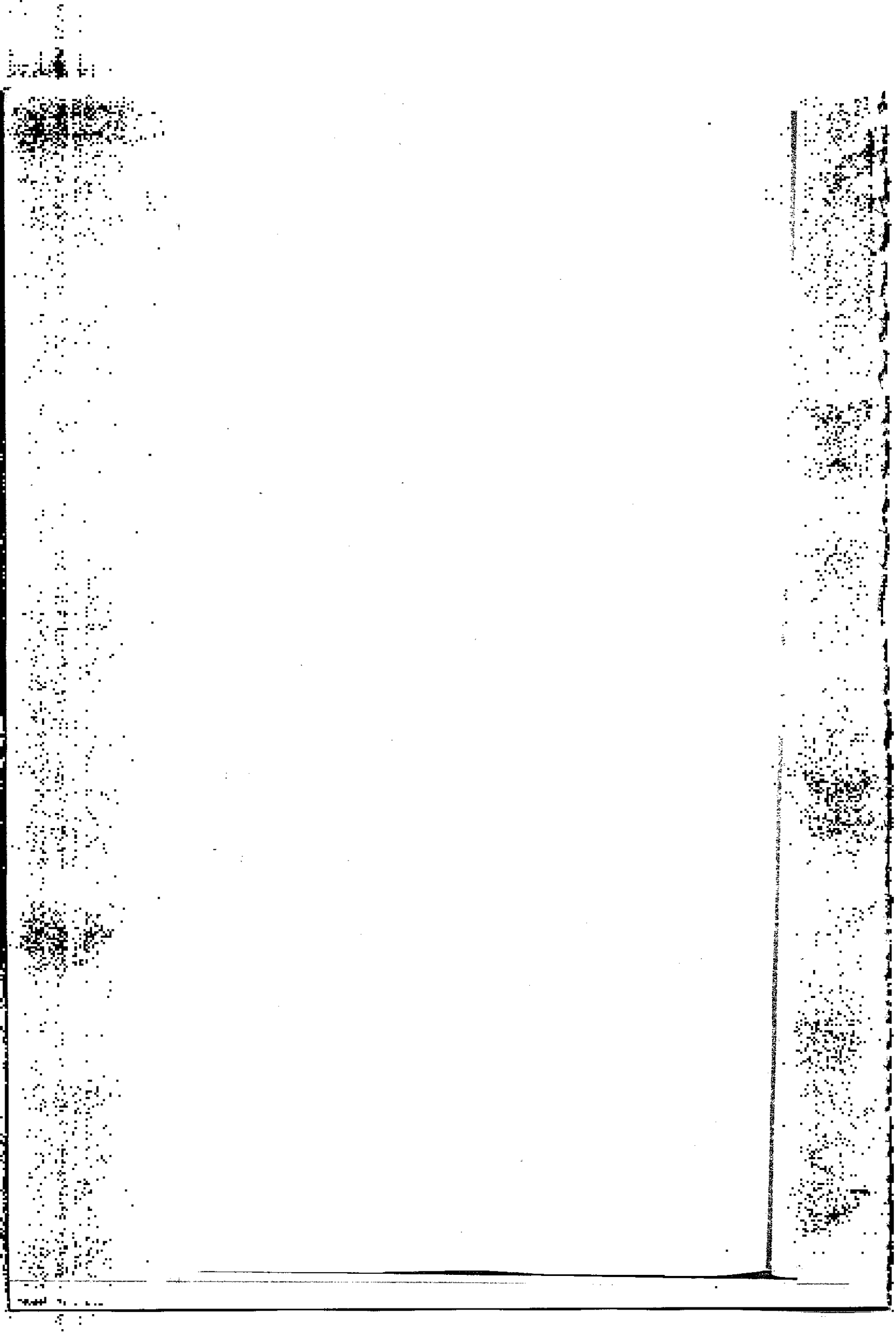
About the CD-ROM

INTRODUCTION

This appendix provides you with information on the contents of the CD that accompanies this book. For the latest and greatest information, please refer to the ReadMe file located at the root of the CD.

SYSTEM REQUIREMENTS

- A computer with a processor running at 120 Mhz or faster
- At least 32 MB of total RAM installed on your computer; for best performance, we recommend at least 64 MB
- A CD-ROM drive

NOTE: Many popular word processing programs are capable of reading Microsoft Word files. However, users should be aware that a slight amount of formatting might be lost when using a program other than Microsoft Word.

USING THE CD WITH WINDOWS

To install the items from the CD to your hard drive, follow these steps:

1. Insert the CD into your computer's CD-ROM drive.
2. The CD-ROM interface will appear. The interface provides a simple point-and-click way to explore the contents of the CD.

If the opening screen of the CD-ROM does not appear automatically, follow these steps to access the CD:

3. Click the Start button on the left end of the taskbar and then choose Run from the menu that pops up.
4. In the dialog box that appears, type d:\start.exe. (If your CD-ROM drive is not drive d, fill in the appropriate letter in place of d.) This brings up the CD Interface described in the preceding set of steps.

WHAT'S ON THE CD

The following sections provide a summary of the software and other materials you'll find on the CD.

CONTENT

The Financial Risk Manager Sample Review Test CD-ROM is a preparatory review for anyone studying for the FRM Exam and for risk professionals interested in self-study to review and improve their knowledge of market, credit, and operational risk management. This interactive CD-ROM contains hundreds of multiple-choice questions from the 2005, 2006, and 2007 FRM exams, with answers and solutions provided.

10 FRM Practice Exams (5 by topic area, 5 by length)

1. Quantitative Analysis Sample Exam (10 Questions)
2. Market Risk Sample Exam (30 Questions)
3. Credit Risk Sample Exam (25 Questions)
4. Operational Risk Sample Exam (25 Questions)
5. Investment Management Sample Exam (10 Questions)
6. Short Practice Exam 1 (20 Questions)
7. Short Practice Exam 2 (20 Questions)
8. Medium Practice Exam 1 (40 Questions)
9. Medium Practice Exam 2 (40 Questions)
10. Long Practice Exam 1 (60 Questions)

APPLICATIONS

Shareware programs are fully functional, trial versions of copyrighted programs. If you like particular programs, register with their authors for a nominal fee and receive licenses, enhanced versions, and technical support.

Freeware programs are copyrighted games, applications, and utilities that are free for personal use. Unlike shareware, these programs do not require a fee or provide technical support.

GNU software is governed by its own license, which is included inside the folder of the GNU product. See the GNU license for more details.

Trial, demo, or evaluation versions are usually limited either by time or functionality (such as being unable to save projects). Some trial versions are very sensitive to system date changes. If you alter your computer's date, the programs will "time out" and no longer be functional.

CUSTOMER CARE

If you have trouble with the CDROM, please call the Wiley Product Technical Support phone number at (800) 762-2974. Outside the United States, call 1(317) 572-3994. You can also contact Wiley Product Technical Support at http://support.wiley.com. John Wiley & Sons will provide technical support only for installation and other general quality control items. For technical support on the applications themselves, consult the program's vendor or author.

To place additional orders or to request information about other Wiley products, please call (877) 762-2974.

Note: Footnote information is indicated by "n" following the page number.