SECOND EDITION

Paul Wilmott

on

QUANTITATIVE FINANCE

WILEY

www.wilmott.com

Paul Wilmott On Quantitative Finance

Second Edition

www.wilmott.com

John Wiley & Sons, Ltd

Paul Wilmott On Quantitative Finance

Paul Wilmott On Quantitative Finance

Second Edition

www.wilmott.com

John Wiley & Sons, Ltd

Published by John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Wilmott, Paul.
Paul Wilmott on quantitative finance.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 13 978-0-470-01870-5 (cloth/cd : alk. paper)
ISBN 10 0-470-01870-4 (cloth/cd : alk. paper)
1. Derivative securities—Mathematical models. 2. Options (Finance)—Mathematical models. 3. Options (Finance)—Prices—Mathematical models. I. Title.
HG6024.A3W555 2006
332.64′53—dc22

2005028317

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13: 978-0-470-01870-5 (HB)
ISBN-10: 0-470-01870-4 (HB)

Typeset in 10/12pt Times by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

In memory of Detlev Vogel

contents of volume one

contents of volume two

contents of volume three

visual basic code

prolog to the second edition

This book is a greatly updated and expanded version of the first edition. The content continues to reflect my own interests and prejudices, based on my skills, such as they are. In the period between the first and second editions, the financial markets have expanded, the tools available to the modeler have expanded, and my girth has expanded. On a personal basis I have spent as much time being a practitioner in a hedge fund as being an independent researcher. Much of the new material therefore represents both my desire as a scientist to build the best, most accurate models, and my need as a practitioner to have models that are fast and robust and simple to understand. As I said, this book is a very personal account of my areas of expertise. Since the subject of quant finance has been galloping apace of late, I advise that you supplement this book with the specialized books that I recommend throughout, and in particular those in the quant library at the end.

I would like to re-thank those people I mentioned in the prolog to the first edition: Arefin Huq, Asli Oztukel, Bafkam Bim, Buddy Holly, Chris McCoy, Colin Atkinson, Daniel Bruno, Dave Thomson, David Bakstein, David Epstein, David Herring, David Wilson, Edna Hepburn-Ruston, Einar Holstad, Eli Lilly, Elisabeth Keck, Elsa Cortina, Eric Cartman, Fouad Khennach, Glen Matlock, Henrik Rassmussen, Hyungsok Ahn, Ingrid Blauer, Jean Laidlaw, Jeff Dewynne, John Lydon, John Ockendon, Karen Mason, Keesup Choe, Malcolm McLaren, Mauricio Bouabci, Patricia Sadro, Paul Cook, Peter Jäckel, Philip Hua, Philipp Schönbucher, Phoebus Theologites, Quentin Crisp, Rich Haber, Richard Arkell, Richard Sherry, Sam Ehrlichman, Sandra Maler, Sara Statman, Simon Gould, Simon Ritchie, Stephen Jefferies, Steve Jones, Truman Capote, Varqa Khadem, and Veronika Guggenbichler.

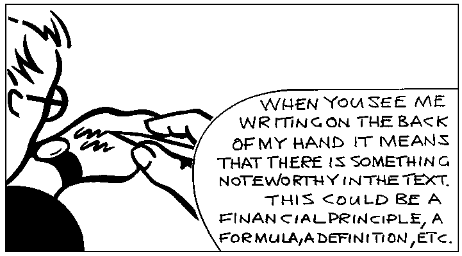
I would also like to thank the following people. My partners in various projects: Paul and Jonathan Shaw at 7city, unequaled in their dedication to training and their imagination for new projects. Also Riaz Ahmad and Seb Lleo who have helped make the Certificate in Quantitative Finance so successful, and for taking some of the pressure off me; Everyone involved in the magazine, especially Aaron Brown, Alan Lewis, Bill Ziemba, Caitlin Cornish, Dan Tudball, Ed Lound, Ed Thorp, Elie Ayache, Espen Gaarder Haug, Graham Russel, Henriette Präst, Jenny McCall, Kent Osband, Liam Larkin, Mike Staunton, Paula Soutinho and Rudi Bogni. I am particularly fortunate and grateful that John Wiley & Sons have been so supportive in what must sometimes seem to them rather wacky schemes; Thanks to Ron Henley, the best hedge fund partner a quant could wish for, 'It's just a jump to the left. And then a step to the right.' And to John Morris of Fulcrum, interesting times; and to Nassim Nicholas Taleb for interesting chats.

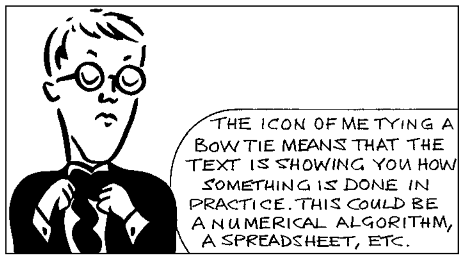
Thanks to, John, Grace, Sel and Stephen, for instilling in me their values: values which have invariably served me well. And to Oscar and Zachary who kept me sane throughout many a series of unfortunate events!

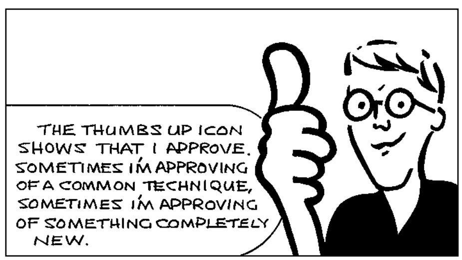
Finally, thanks to my number one fan, Andrea Estrella, from her number one fan, me.

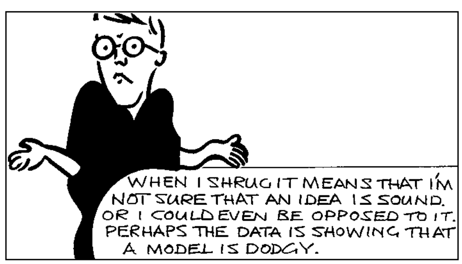
ABOUT THE AUTHOR

Paul Wilmott's professional career spans almost every aspect of mathematics and finance, in both academia and in the real world. He has lectured at all levels, founded a magazine, the leading website for the quant community, and a quant certificate program. He has managed money as a partner in a very successful hedge fund. He lives in London, is married, and has two sons. His only remaining dream is to get some sleep.









More info about the particular meaning of an icon is contained in its 'speech box.'

You will see this icon whenever a method is implemented on the CD.

PART ONE

mathematical and financial foundations; basic theory of derivatives; risk and return

The first part of the book contains the fundamentals of derivatives theory and practice. We look at both equity and fixed income instruments. I introduce the important concepts of hedging and no arbitrage, on which most sophisticated finance theory is based. We draw some insight from ideas first seen in gambling, and we develop those into an analysis of risk and return.

The assumptions, key concepts and results in Part One make up what is loosely known as the 'Black–Scholes world,' named for Fischer Black and Myron Scholes who, together with Robert Merton, first conceived them. Their original work was published in 1973, after some resistance (the famous equation was first written down in 1969). In October 1997 Myron Scholes and Robert Merton were awarded the Nobel Prize for Economics for their work, Fischer Black having died in August 1995. *The New York Times* of Wednesday, 15th October 1997 wrote: 'Two North American scholars won the Nobel Memorial Prize in Economic Science yesterday for work that enables investors to price accurately their bets on the future, a breakthrough that has helped power the explosive growth in financial markets since the 1970's and plays a profound role in the economics of everyday life.'[1]

Part One is self contained, requiring little knowledge of finance or any more than elementary calculus.

Chapter 1: Products and Markets An overview of the workings of the financial markets and their products. A chapter such as this is obligatory. However, my readers will fall into one of two groups. Either they will know everything in this chapter and much, much more besides. Or they will know little, in which case what I write will not be enough.

[1] We'll be hearing more about these two in Chapter 44 on 'Derivatives **** Ups.'

Chapter 2: Derivatives An introduction to options, options markets, market conventions. Definitions of the common terms, simple no arbitrage, put-call parity and elementary trading strategies.

Chapter 3: The Random Behavior of Assets An examination of data for various financial quantities, leading to a model for the random behavior of prices. Almost all of sophisticated finance theory assumes that prices are random, the question is how to model that randomness.

Chapter 4: Elementary Stochastic Calculus We'll need a little bit of theory for manipulating our random variables. I keep the requirements down to the bare minimum. The key concept is Itô's lemma which I will try to introduce in as accessible a manner as possible.

Chapter 5: The Black–Scholes Model I present the classical model for the fair value of options on stocks, currencies and commodities. This is the chapter in which I describe delta hedging and no arbitrage and show how they lead to a unique price for an option. This is the foundation for most quantitative finance theory and I will be building on this foundation for much, but by no means all, of the book.

Chapter 6: Partial Differential Equations Partial differential equations play an important role in most physical applied mathematics. They also play a role in finance. Most of my readers trained in the physical sciences, engineering and applied mathematics will be comfortable with the idea that a partial differential equation is almost the same as 'the answer,' the two being separated by at most some computer code. If you are not sure of this connection I hope that you will persevere with the book. This requires some faith on your part; you may have to read the book through twice: I have necessarily had to relegate the numerics, the real 'answer,' to the last few chapters.

Chapter 7: The Black–Scholes Formulae and the 'Greeks' From the Black–Scholes partial differential equation we can find formulae for the prices of some options. Derivatives of option prices with respect to variables or parameters are important for hedging. I will explain some of the most important such derivatives and how they are used.

Chapter 8: Simple Generalizations of the Black–Scholes World Some of the assumptions of the Black–Scholes world can be dropped or stretched with ease. I will describe several of these. Later chapters are devoted to more extensive generalizations.

Chapter 9: Early Exercise and American Options Early exercise is of particular importance financially. It is also of great mathematical interest. I will explain both of these aspects.

Chapter 10: Probability Density Functions and First-exit Times The random nature of financial quantities means that we cannot say with certainty what the future holds in store. For that reason we need to be able to describe that future in a probabilistic sense.

Chapter 11: Multi-asset Options Another conceptually simple generalization of the basic Black–Scholes world is to options on more than one underlying asset. Theoretically simple, this extension has its own particular problems in practice.

Chapter 12: How to Delta Hedge Not everyone believes in no arbitrage, the absence of free lunches. In this chapter we see how to profit if you have a better forecast for future volatility than the market.

Chapter 13: Fixed-income Products and Analysis: Yield, Duration and Convexity This chapter is an introduction to the simpler techniques and analyses commonly used in the market. In particular I explain the concepts of yield, duration and convexity. In this and the next chapter I assume that interest rates are known, deterministic quantities.

Chapter 14: Swaps Interest-rate swaps are very common and very liquid. I explain the basics and relate the pricing of swaps to the pricing of fixed-rate bonds.

Chapter 15: The Binomial Model One of the reasons that option theory has been so successful is that the ideas can be explained and implemented very easily with no complicated mathematics. The binomial model is the simplest way to explain the basic ideas behind option theory using only basic arithmetic. That's a good thing, right? Yes, but only if you bear in mind that the model is for demonstration purposes only, it is not the real thing. As a model of the financial world it is too simplistic, as a concept for pricing it lacks the elegance that makes other methods preferable, and as a numerical scheme it is prehistoric. Use once and then throw away, that's my recommendation.

Chapter 16: How Accurate is the Normal Approximation? One of the major assumptions of finance theory is that returns are Normally distributed. In this chapter we take a look at why we make this assumption, and how good it really is.

Chapter 17: Investment Lessons from Blackjack and Gambling We draw insights and inspiration from the not-unrelated world of gambling to help us in the treatment of risk, return, and money/risk management.

Chapter 18: Portfolio Management If you are willing to accept some risk how should you invest? I explain the classical ideas of Modern Portfolio Theory and the Capital Asset Pricing Model.

Chapter 19: Value at Risk How risky is your portfolio? How much might you conceivably lose if there is an adverse market move? These are the topics of this chapter.

Chapter 20: Forecasting the Markets? Although almost all sophisticated finance theory assumes that assets move randomly, many traders rely on technical indicators to predict the future direction of assets. These indicators may be simple geometrical constructs of the asset price path or quite complex algorithms. The hypothesis is that information about short-term future asset price movements are contained within the past history of prices. All traders use technical indicators at some time. In this chapter I describe some of the more common techniques.

Chapter 21: A Trading Game Many readers of this book will never have traded anything more sophisticated than baseball cards. To get them into the swing of the subject from a practical point of view I include some suggestions on how to organize your own trading game based on the buying and selling of derivatives. I had a lot of help with this chapter from David Epstein who has been running such games for several years.

CHAPTER 1
products and markets

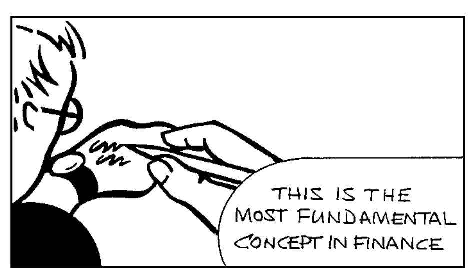
In this Chapter...

- the time value of money
- an introduction to equities, commodities, currencies and indices
- fixed and floating interest rates
- futures and forwards
- no-arbitrage, one of the main building blocks of finance theory

1.1 INTRODUCTION

This first chapter is a very gentle introduction to the subject of finance, and is mainly just a collection of definitions and specifications concerning the financial markets in general. There is little technical material here, and the one technical issue, the 'time value of money,' is extremely simple. I will give the first example of 'no arbitrage.' This is important, being one part of the foundation of derivatives theory. Whether you read this chapter thoroughly or just skim it will depend on your background; mathematicians new to finance may want to spend more time on it than practitioners, say.

1.2 THE TIME VALUE OF MONEY



The simplest concept in finance is that of the **time value of money**; \$1 today is worth more than \$1 in a year's time. This is because of all the things we can do with \$1 over the next year. At the very least, we can put it under the mattress and take it out in one year. But instead of putting it under the mattress we could invest it in a gold mine, or a new company. If those are too risky, then lend the money to someone who is willing to take the risks and will give you back the dollar with a little bit extra, the **interest**. That is what banks do, they borrow your money and invest it in various risky ways, but by spreading their risk over many investments they reduce their overall risk. And by borrowing money from many people they can invest in ways that the average individual cannot. The banks compete for your money by offering high interest rates. Free markets and the ability to change banks quickly and cheaply ensure that interest rates are fairly consistent from one bank to another.

I am going to denote interest rates by r. Although rates vary with time I am going to assume for the moment that they are constant. We can talk about several types of interest. First of all there is **simple** and **compound interest**. Simple interest is when the interest you receive is based only on the amount you invest initially, whereas compound interest is when you also get interest on your interest. Compound interest is the main case of relevance. And compound interest comes in two forms, **discretely compounded** and **continuously compounded**. Let me illustrate how they each work.

Suppose I invest \$1 in a bank at a discrete interest rate of r paid once *per annum*. At the end of one year my bank account will contain

$$\$1 \times (1 + r).$$

If the interest rate is 10% I will have one dollar and ten cents. After two years I will have

$$\$1 \times (1 + r) \times (1 + r) = (1 + r)^2,$$

or one dollar and twenty-one cents. After n years I will have $(1 + r)^n$ dollars. That is an example of discrete compounding.

Now suppose I receive m interest payments at a rate of r/m *per annum*. After one year I will have

$$\left(1 + \frac{r}{m}\right)^m. \tag{1.1}$$

(I have dropped the \$ sign, taking it as read from now on.)

I am going to imagine that these interest payments come at increasingly frequent intervals, but at an increasingly smaller interest rate: I am going to take the limit $m \to \infty$. This will lead to a rate of interest that is paid continuously. Expression (1.1) becomes

$$\left(1 + \frac{r}{m}\right)^m = e^{m \log\left(1 + \frac{r}{m}\right)} \sim e^r.$$

This is a simple application of Taylor series when r/m is small. And that is how much money I will have in the bank after one year if the interest is continuously compounded. Similarly, after a time t I will have an amount

$$e^{rt} \tag{1.2}$$

in the bank. Almost everything in this book assumes that interest is compounded continuously.

Another way of deriving the result (1.2) is via a differential equation. Suppose I have an amount $M(t)$ in the bank at time t, how much does this increase in value from one day to the next? If I look at my bank account at time t and then again a short while later, time $t + dt$, the amount will have increased by

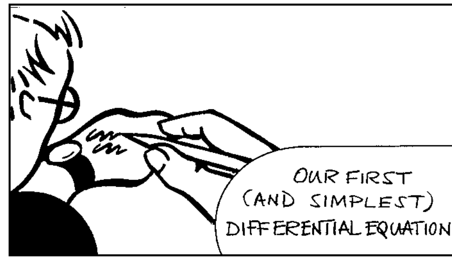
$$M(t + dt) - M(t) \approx \frac{dM}{dt}\, dt + \cdots,$$

where the right-hand side comes from a Taylor series expansion of $M(t + dt)$. But I also know that the interest I receive must be proportional to the amount I have, M, the interest rate, r, and the time step, dt. Thus

$$\frac{dM}{dt}\, dt = rM(t)\, dt.$$

Dividing by dt gives the ordinary differential equation

$$\boxed{\frac{dM}{dt} = rM(t)}$$



the solution of which is

$$M(t) = M(0)\, e^{rt}.$$

If the initial amount at $t = 0$ was $1 then I get (1.2) again.

This equation relates the value of the money I have now to the value in the future. Conversely, if I know I will get one dollar at time T in the future, its value at an earlier time t is simply

$$e^{-r(T-t)}.$$

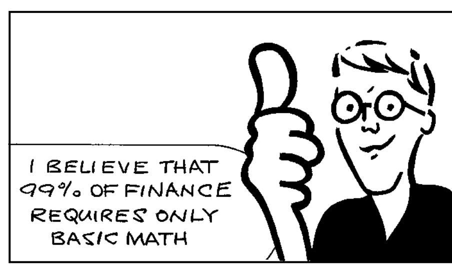
I can relate cashflows in the future to their **present value** by multiplying by this factor. As an example, suppose that r is 5% i.e. $r = 0.05$, then the present value of $1,000,000 to be received in two years is

$$\$1{,}000{,}000 \times e^{-0.05\times 2} = \$904{,}837.$$

The present value is clearly less than the future value.

Interest rates are a very important factor determining the present value of future cashflows. For the moment I will only talk about one interest rate, and that will be constant. In later chapters I will generalize.

Important Aside



What mathematics have we seen so far? To get to (1.2) all we needed to know about are the two functions e (or exp) and log, and Taylor series. Believe it or not, you can appreciate almost all finance theory by knowing these three things together with 'expectations.' I'm going to build up to the basic Black–Scholes and derivatives theory assuming that you know all four of these. Don't worry if you don't know about these things yet, take a look at Appendix A where I review these requisites and show how to interpret finance theory and practice in terms of the most elementary mathematics.

Just because you *can* understand derivatives theory in terms of basic math doesn't mean that you *should*. I hope that there's enough in the book to please the Ph.D.s[1] as well.

1.3 EQUITIES

The most basic of financial instruments is the **equity**, **stock** or **share**. This is the ownership of a small piece of a company. If you have a bright idea for a new product or service

[1] And Nobel laureates.

then you could raise capital to realize this idea by selling off future profits in the form of a stake in your new company. The investors may be friends, your Aunt Joan, a bank, or a venture capitalist. The investor in the company gives you some cash, and in return you give him a contract stating how much of the company he owns. The **shareholders** who own the company between them then have some say in the running of the business, and technically the directors of the company are meant to act in the best interests of the shareholders. Once your business is up and running, you could raise further capital for expansion by issuing new shares.

This is how small businesses begin. Once the small business has become a large business, your Aunt Joan may not have enough money hidden under the mattress to invest in the next expansion. At this point shares in the company may be sold to a wider audience or even the general public. The investors in the business may have no link with the founders. The final point in the growth of the company is with the quotation of shares on a regulated stock exchange so that shares can be bought and sold freely, and capital can be raised efficiently and at the lowest cost.

Figures 1.1 and 1.2 show screens from Bloomberg giving details of Microsoft stock, including price, high and low, names of key personnel, weighting in various indices (see below) etc. There is much, much more info available on Bloomberg for this and all other stocks. We'll be seeing many Bloomberg screens throughout this book.

```
MSFT US $  C 95+15/16 Q  Q↑94 15/16/95Q               DL18 Equity DES
As of Sep10 DELAYED Vol 17,227,500 Op 95 1/16 Q  Hi 95 5/16 Q  Lo 94 Q
                          DESCRIPTION                       Page  1 /10
MSFT US                     MICROSOFT CORP          12) CN All News/Research
Computer Software                                   13) CWP Company Web Page
CUSIP 594918104                                     14) HH Hoover's Handbook
Microsoft Corporation develops, manufactures, licenses, sells, and supports
software products.  The Company offers operating system software, server
application software, business and consumer applications software, software
development tools, and Internet and intranet software.  Microsoft also develops
the MSN network of Internet products and services.

       STOCK DATA      Round Lot      100 | 8)DVD     DIVIDENDS  - None
1)GPO  Current Price         USD       95 |      Indicated Gross Yld
       52Wk High  7/19/1999 USD   100 3/4 |      Dividend Growth
       52Wk Low  10/ 8/1998 USD    43 7/8 |       Ex-Date       Type        Amt
       YTD Chng ( 37.00%)   USD  25 21/32 |
2)TRA  1 Yr Total Return           82.25% |       3/29/99       Split     2 for 1
3)CH1  Shares Out as of  4/30   5103.859M | EARNINGS  - Ann Date  10/20/99 (Est)
       Market Cap          USD 484866.63M | 9)ERN Trailing 12mo EPS   USD   1.395
Float    3576.78M  Short Int     24.823M  | 10)EE  Est EPS      6/2000 USD   1.566
5)BETA Beta    vs. SPX               1.25 | 11)GE  P/E    68.10  Est P/E     60.66
6)OCM  Options avail & Stk Marginable     | LT Growth     25.21 Est PEG      2.41
Par Value = .0000125

Copyright 1999 BLOOMBERG L.P.   Frankfurt:69-920410  Hong Kong:2-977-6000  London:171-330-7500  New York:212-318-2000
Princeton:609-279-3000     Singapore:226-3000     Sydney:2-9777-8686     Tokyo:3-3201-8900     Sao Paulo:11-3048-4500
                                                                          I741-53-0 11-Sep-99 15:35:34
Bloomberg
PROFESSIONAL
```

Figure 1.1 Details of Microsoft stock. Source: Bloomberg L.P.

```
Page                                                        DL18 Equity DES
Hit 1 <GO> for a more detailed company management profile (MGMT).
MSFT US              MICROSOFT CORP                              Page  2 /10
 One Microsoft Way                     T:425-882-8080        F:425-936-8000
 Bldg 8 Southwest                        2) http://www.microsoft.com/msft/
 Redmond,WA   98052-6399               TR AG ChaseMellon Shareholder Services
 United States                         # OF EMPLOYEES         27,055
 WILLIAM H GATES III                      CHAIRMAN/CEO
 STEVEN A BALLMER                         PRESIDENT
 ROBERT J HERBOLD                         EXEC VP/COO
 GREGORY B MAFFEI                         SENIOR VP/CFO
 TIM HALLADAY                             INVESTOR RELATIONS CONTACT
 STEVE SCHIRO                             VP:CONSUMER CUSTOMER UNIT
 Type   Common Stock     PAR  $   .00001 3)WGT MEMBER            TICKER    WEIGHT
 PRIMARY EXCHANGE    NASDAQ N-Mkt         S&P 500 INDEX          SPX        4.368%
 COUNTRY             United States        NASDAQ 100 STOCK       NDX       14.287%
 FISCAL YEAR END     JUNE                 S&P 100 INDEX          OEX        8.752%
 SIC Code  7372   PREPAKG SOFTW           TRIB WORLD INDEX       TRIB       5.245%
 VALOREN    000951692                     AMEX INSTITUTION       XII        6.540%
 WPK Number 870747                        AMEX COMPUTER TE       XCI       23.453%
 SEDOL      2588173                       PHILA NATIONAL O       XOC       21.223%
 Sicovam    903099                        CBOE TECHNOLOGY        TXX        4.157%
 ISIN       US5949181045                  S&P INDUSTRIALS        SPXI       5.316%
                                          S&P CAPITAL GOOD       SPCAPC    15.140%
Copyright 1999 BLOOMBERG L.P.  Frankfurt:69-920410  Hong Kong:2-977-6000  London:171-330-7500  New York:212-318-2000
Princeton:609-279-3000    Singapore:226-3000    Sydney:2-9777-8686    Tokyo:3-3201-8900    Sao Paulo:11-3048-4500
                                                                  I741-53-0 11-Sep-99 15:35:41
Bloomberg
PROFESSIONAL
```

Figure 1.2 Details of Microsoft stock continued. Source: Bloomberg L.P.

In Figure 1.3 I show an excerpt from *The Wall Street Journal Europe* of 14th April 2005. This shows a small selection of the many stocks traded on the New York Stock Exchange. The listed information includes highs and lows for the day as well as the change since the previous day's close.

The behavior of the quoted prices of stocks is far from being predictable. In Figure 1.4 I show the Dow Jones Industrial Average over the period January 1950 to March 2004. In Figure 1.5 is a time series of the Glaxo–Wellcome share price, as produced by Bloomberg.

If we could predict the behavior of stock prices in the future then we could become very rich. Although many people have claimed to be able to predict prices with varying degrees of accuracy, no one has yet made a completely convincing case. In this book I am going to take the point of view that prices have a large element of randomness. This does *not* mean that we cannot model stock prices, but it does mean that the modeling must be done in a probabilistic sense. No doubt the reality of the situation lies somewhere between complete predictability and perfect randomness, not least because there have been many cases of market manipulation where large trades have moved stock prices in a direction that was favorable to the person doing the moving.

To whet your appetite for the mathematical modeling later, I want to show you a simple way to simulate a random walk that looks something like a stock price. One of the simplest random processes is the tossing of a coin. I am going to use ideas related to coin tossing as a model for the behavior of a stock price. As a simple experiment start with the number 100 which you

Publisher's Note:
Permission to reproduce this image online was not granted by the copyright holder. Readers are kindly requested to refer to the printed version of this chapter.

Figure 1.3 *The Wall Street Journal Europe* of 14th April 2005. Reproduced by permission of Dow Jones & Company, Inc.

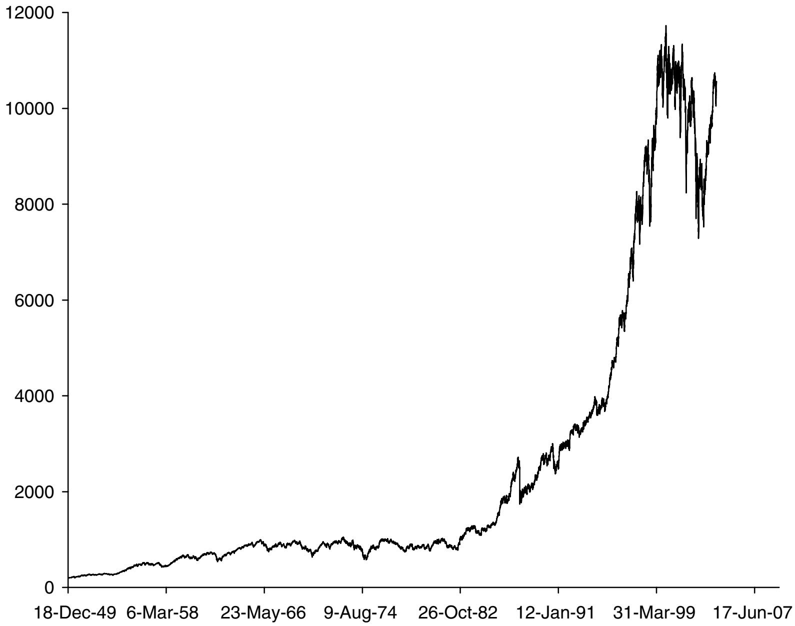


Figure 1.4 A time series of the Dow Jones Industrial Average from January 1950 to March 2004.

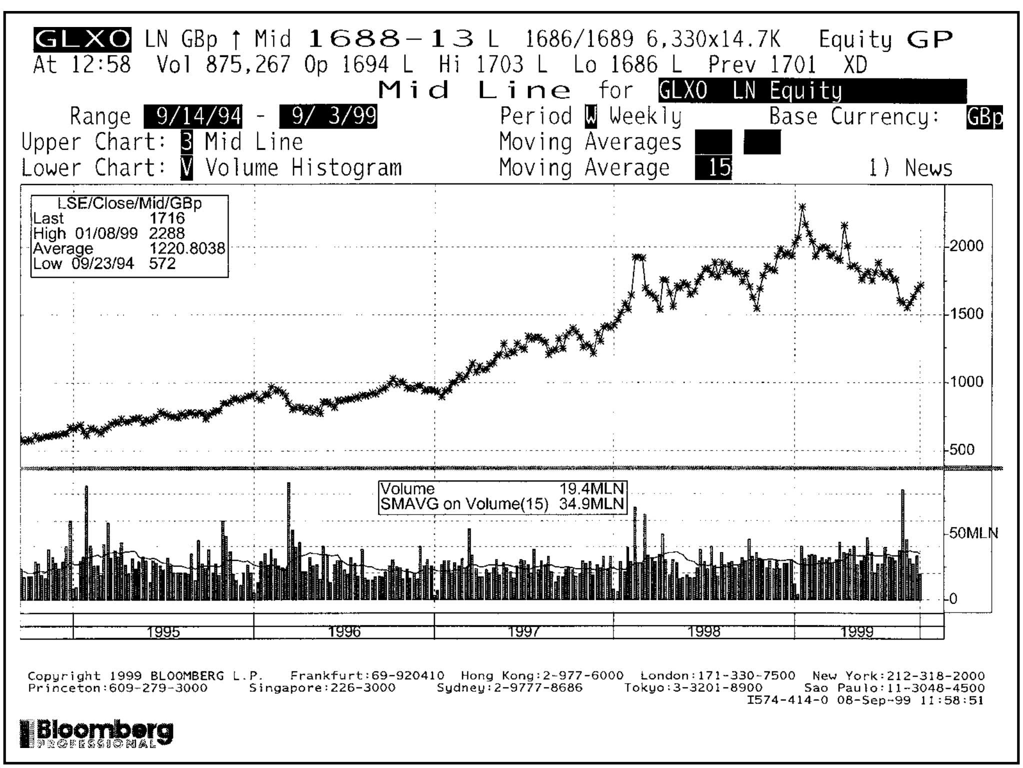


Figure 1.5 Glaxo−Wellcome share price (volume below). Source: Bloomberg L.P.

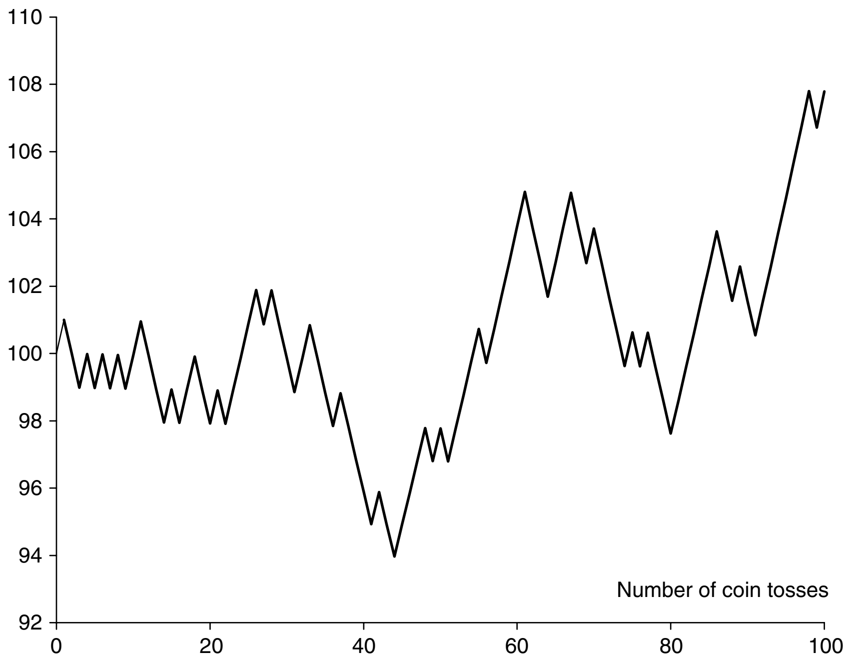


Figure 1.6 A simulation of an asset price path?

should think of as the price of your stock, and toss a coin. If you throw a head multiply the number by 1.01, if you throw a tail multiply by 0.99. After one toss your number will be either 99 or 101. Toss again. If you get a head multiply your *new* number by 1.01 or by 0.99 if you throw a tail. You will now have either $1.01^2 \times 100$, $1.01 \times 0.99 \times 100 = 0.99 \times 1.01 \times 100$ or $0.99^2 \times 100$. Continue this process and plot your value on a graph each time you throw the coin. Results of one particular experiment are shown in Figure 1.6. Instead of physically tossing a coin, the series used in this plot was generated on a spreadsheet like that in Figure 1.7. This uses the Excel spreadsheet function `RAND()` to generate a uniformly distributed random number between 0 and 1. If this number is greater than one half it counts as a 'head' otherwise a 'tail.'

1.3.1 Dividends

The owner of the stock theoretically owns a piece of the company. This ownership can only be turned into cash if he owns so many of the stock that he can take over the company and keep all the profits for himself. This is unrealistic for most of us. To the average investor the value in holding the stock comes from the **dividends** and any growth in the stock's value. Dividends are lump sum payments, paid out every quarter or every six months, to the holder of the stock.

The amount of the dividend varies from year to year depending on the profitability of the company. As a general rule companies like to try to keep the level of dividends about the same each time. The amount of the dividend is decided by the board of directors of the company and is usually set a month or so before the dividend is actually paid.

When the stock is bought it either comes with its entitlement to the next dividend (**cum**) or not (**ex**). There is a date at around the time of the dividend payment when the stock goes

	A	B	C	D	E
1	**Initial stock price**	**100**		**Stock**	
2	**Up move**	**1.01**		100	
3	**Down move**	**0.99**		101	
4	**Probability of up**	**0.5**		99.99	
5				98.9901	
6				99.98	
7				98.9802	
8				99.97	
9				[illegible]	
10				99.96001	
11				98.96041	
12				99.95001	
13				100.9495	
14				99.94001	
15				98.94061	
16				97.95121	
17				98.93072	
18				97.94141	
19				98.92083	
20				99.91004	
21				98.91094	
22				97.92183	
23				98.90104	
24				97.91203	
25				98.89115	
26				99.88007	
27				100.8789	
28				101.8877	
29				100.8688	
30				101.8775	
31				100.8587	

=B1

=D6*IF(RAND()>1-$B$4,$B$2,$B$3)

Figure 1.7 Simple spreadsheet to simulate the coin-tossing experiment.

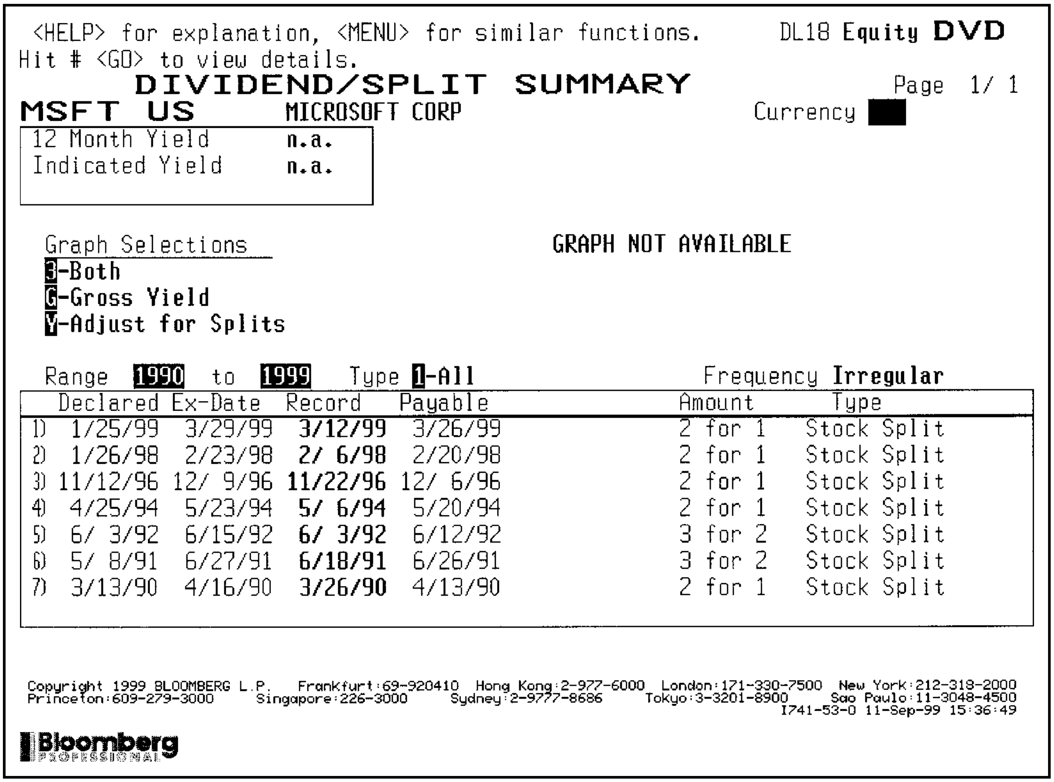
from cum to ex. The original holder of the stock gets the dividend but the person who buys it obviously does not. All things being equal a stock that is cum dividend is better than one that is ex dividend. Thus at the time that the dividend is paid and the stock goes ex dividend there will be a drop in the value of the stock. The size of this drop in stock value offsets the disadvantage of not getting the dividend.

This jump in stock price is in practice more complex than I have just made out. Often capital gains due to the rise in a stock price are taxed differently from a dividend, which is often treated as income. Some people can make a lot of risk-free money by exploiting tax 'inconsistencies.'

I discuss dividends in depth in Chapter 8 and again in Chapter 64.

1.3.2 Stock Splits

Stock prices in the US are usually of the order of magnitude of $100. In the UK they are typically around £1. There is no real reason for the popularity of the number of digits, after all, if I buy a stock I want to know what percentage growth I will get, the absolute level of the stock is irrelevant to me, it just determines whether I have to buy tens or thousands



<HELP> for explanation, <MENU> for similar functions. DL18 Equity DVD
Hit # <GO> to view details.
DIVIDEND/SPLIT SUMMARY Page 1/ 1
MSFT US MICROSOFT CORP Currency

12 Month Yield n.a.
Indicated Yield n.a.

Graph Selections
B-Both
G-Gross Yield
Y-Adjust for Splits

GRAPH NOT AVAILABLE

Range 1990 to 1999 Type 1-All Frequency Irregular

	Declared	Ex-Date	Record	Payable	Amount	Type
1)	1/25/99	3/29/99	3/12/99	3/26/99	2 for 1	Stock Split
2)	1/26/98	2/23/98	2/ 6/98	2/20/98	2 for 1	Stock Split
3)	11/12/96	12/ 9/96	11/22/96	12/ 6/96	2 for 1	Stock Split
4)	4/25/94	5/23/94	5/ 6/94	5/20/94	2 for 1	Stock Split
5)	6/ 3/92	6/15/92	6/ 3/92	6/12/92	3 for 2	Stock Split
6)	5/ 8/91	6/27/91	6/18/91	6/26/91	3 for 2	Stock Split
7)	3/13/90	4/16/90	3/26/90	4/13/90	2 for 1	Stock Split

Copyright 1999 BLOOMBERG L.P. Frankfurt:69-920410 Hong Kong:2-977-6000 London:171-330-7500 New York:212-318-2000
Princeton:609-279-3000 Singapore:226-3000 Sydney:2-9777-8686 Tokyo:3-3201-8900 Sao Paulo:11-3048-4500
I741-53-0 11-Sep-99 15:36:49

Bloomberg PROFESSIONAL

Figure 1.8 Stock split info for Microsoft. Source: Bloomberg L.P.

of the stock to invest a given amount. Nevertheless there is some psychological element to the stock size. Every now and then a company will announce a **stock split** (see Figure 1.8). For example, the company with a stock price of $900 announces a three-for-one stock split. This simply means that instead of holding one stock valued at $900, I hold three valued at $300 each.[2]

1.4 COMMODITIES

Commodities are usually raw products such as precious metals, oil, food products etc. The prices of these products are unpredictable but often show seasonal effects. Scarcity of the product results in higher prices. Commodities are usually traded by people who have no need of the raw material. For example they may just be speculating on the direction of gold without wanting to stockpile it or make jewelry. Most trading is done on the futures market, making deals to buy or sell the commodity at some time in the future. The deal is then closed out before the commodity is due to be delivered. Futures contracts are discussed below.

Figure 1.9 shows a time series of the price of pulp, used in paper manufacture.

[2] In the UK this would be called a two-for-one split.

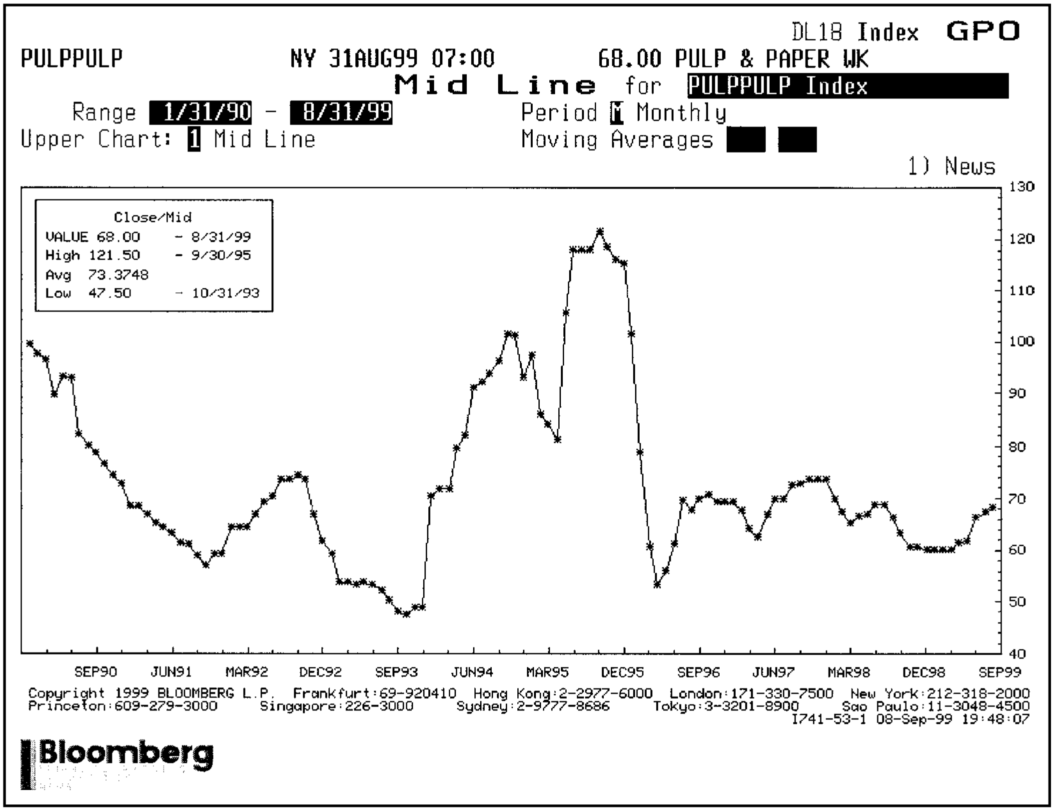


Figure 1.9 Pulp price. Source: Bloomberg L.P.

1.5 CURRENCIES

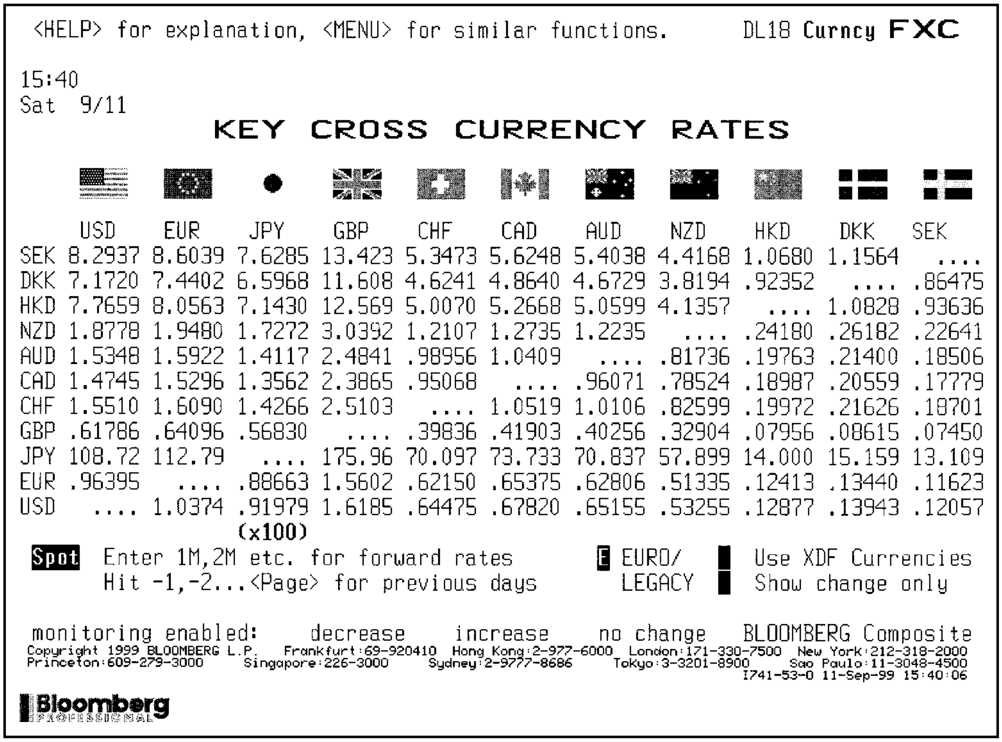
Another financial quantity we shall discuss is the **exchange rate**, the rate at which one currency can be exchanged for another. This is the world of **foreign exchange**, or **Forex** or **FX** for short. Some currencies are pegged to one another, and others are allowed to float freely. Whatever the exchange rates from one currency to another, there must be consistency throughout. If it is possible to exchange dollars for pounds and then the pounds for yen, this implies a relationship between the dollar/pound, pound/yen and dollar/yen exchange rates. If this relationship moves out of line it is possible to make **arbitrage profits** by exploiting the mispricing.

Figure 1.10 is an excerpt from *The Wall Street Journal Europe* of 14th April 2005. At the top of this excerpt is a matrix of exchange rates. A similar matrix is shown in Figure 1.11 from Bloomberg.

Although the fluctuation in exchange rates is unpredictable, there is a link between exchange rates and the interest rates in the two countries. If the interest rate on dollars is raised while the interest rate on pounds sterling stays fixed we would expect to see sterling depreciating against the dollar for a while. Central banks can use interest rates as a tool for manipulating exchange rates, but only to a degree.

At the start of 1999 Euroland currencies were fixed at the rates shown in Figure 1.12.

Figure 1.10 *The Wall Street Journal Europe* of 14th April 2005, currency exchange rates. Reproduced by permission of Dow Jones & Company, Inc.



<HELP> for explanation, <MENU> for similar functions. DL18 Curncy **FXC**

15:40
Sat 9/11

KEY CROSS CURRENCY RATES

	USD	EUR	JPY	GBP	CHF	CAD	AUD	NZD	HKD	DKK	SEK
SEK	8.2937	8.6039	7.6285	13.423	5.3473	5.6248	5.4038	4.4168	1.0680	1.1564	....
DKK	7.1720	7.4402	6.5968	11.608	4.6241	4.8640	4.6729	3.8194	.92352	....	.86475
HKD	7.7659	8.0563	7.1430	12.569	5.0070	5.2668	5.0599	4.1357	....	1.0828	.93636
NZD	1.8778	1.9480	1.7272	3.0392	1.2107	1.2735	1.2235	....	.24180	.26182	.22641
AUD	1.5348	1.5922	1.4117	2.4841	.98956	1.0409	....	.81736	.19763	.21400	.18506
CAD	1.4745	1.5296	1.3562	2.3865	.95068	....	.96071	.78524	.18987	.20559	.17779
CHF	1.5510	1.6090	1.4266	2.5103	....	1.0519	1.0106	.82599	.19972	.21626	.18701
GBP	.61786	.64096	.56830	....	.39836	.41903	.40256	.32904	.07956	.08615	.07450
JPY	108.72	112.79	....	175.96	70.097	73.733	70.837	57.899	14.000	15.159	13.109
EUR	.96395	....	.88663	1.5602	.62150	.65375	.62806	.51335	.12413	.13440	.11623
USD	....	1.0374	.91979	1.6185	.64475	.67820	.65155	.53255	.12877	.13943	.12057

(x100)

Spot Enter 1M,2M etc. for forward rates E EURO/ LEGACY Use XDF Currencies
Hit -1,-2...<Page> for previous days Show change only

monitoring enabled: decrease increase no change BLOOMBERG Composite

Copyright 1999 BLOOMBERG L.P. Frankfurt:69-920410 Hong Kong:2-977-6000 London:171-330-7500 New York:212-318-2000
Princeton:609-279-3000 Singapore:226-3000 Sydney:2-9777-8686 Tokyo:3-3201-8900 Sao Paulo:11-3048-4500
I741-53-0 11-Sep-99 15:40:06

Bloomberg PROFESSIONAL

Figure 1.11 Key cross currency rates. Source: Bloomberg L.P.

1.6 INDICES

For measuring how the stock market/economy is doing as a whole, there have been developed the stock market **indices**. A typical index is made up from the weighted sum of a selection or **basket** of representative stocks. The selection may be designed to represent the whole market, such as the Standard & Poor's 500 (S&P500) in the US or the Financial Times Stock Exchange index (FTSE100) in the UK, or a very special part of a market. In Figure 1.4 we saw the DJIA, representing major US stocks. In Figure 1.13 is shown JP Morgan's Emerging Market Bond Index. The EMBI+ is an index of emerging market debt instruments, including external-currency-denominated Brady bonds, Eurobonds and US dollar local markets instruments. The main components of the index are the three major Latin American countries, Argentina, Brazil and Mexico. Bulgaria, Morocco, Nigeria, the Philippines, Poland, Russia and South Africa are also represented.

Figure 1.14 shows a time series of the MAE All Bond Index which includes Peso and US dollar denominated bonds sold by the Argentine Government.

1.7 FIXED-INCOME SECURITIES

In lending money to a bank you may get to choose for how long you tie your money up and what kind of interest rate you receive. If you decide on a fixed-term deposit the bank will offer to lock in a fixed rate of interest for the period of the deposit, a month, six months, a year, say.

<HELP> for explanation, <MENU> for similar functions. DL18 **Curncy ERM**

EURO FIXING RATES

Official Fixing Rates vs. Euro

German Mark	DEM	1.955830
Belgian Franc	BEF	40.339900
Luxembourg Franc	LUF	40.339900
Spanish Peseta	ESP	166.386000
French Franc	FRF	6.559570
Irish Punt	IEP	0.787564
Italian Lira	ITL	1936.270000
Dutch Guilder	NLG	2.203710
Austrian Schilling	ATS	13.760300
Portuguese Escudo	PTE	200.482000
Finnish Markka	FIM	5.945730

The Danish Krone is linked at a parity of 7.46038 per EUR +/- 2.25 %
The Greek Drachma is linked at a parity of 353.109 per EUR +/- 15.0 %

Figure 1.12 Euro fixing rates. Source: Bloomberg L.P.

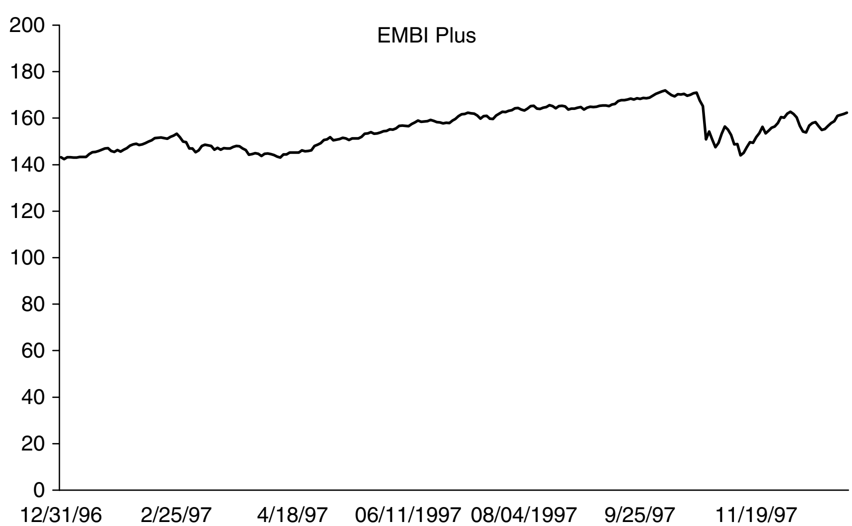


Figure 1.13 JP Morgan's EMBI Plus.

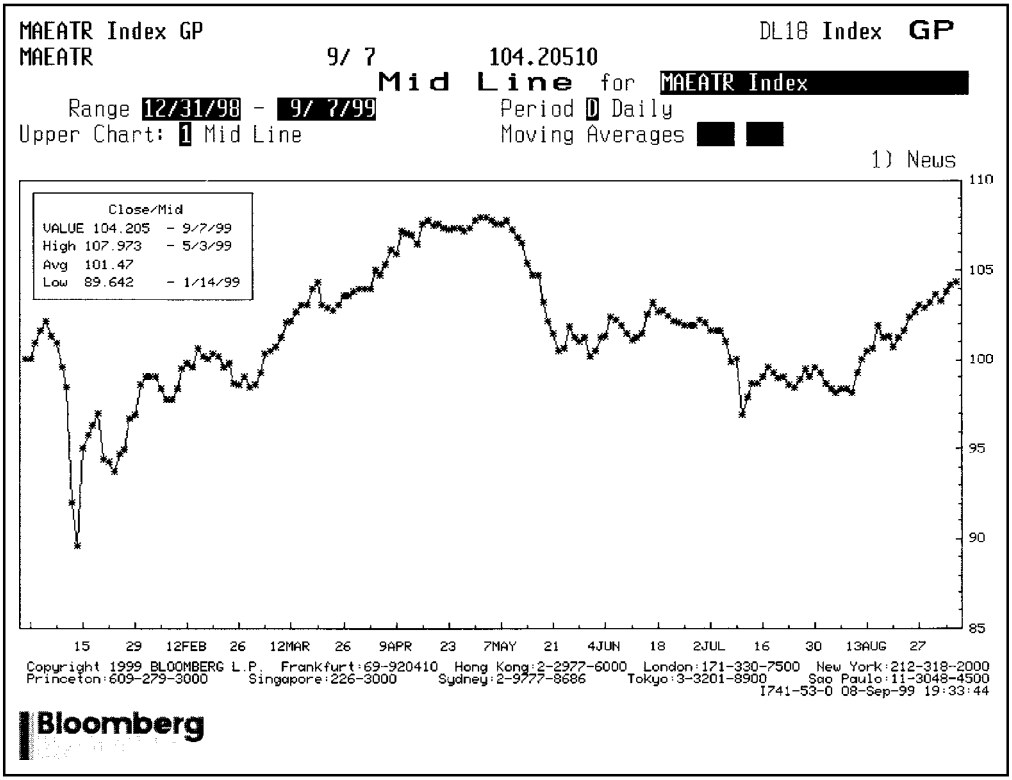


Figure 1.14 A time series of the MAE All Bond Index. Source: Bloomberg L.P.

The rate of interest will not necessarily be the same for each period, and generally the longer the time that the money is tied up the higher the rate of interest, although this is not always the case. Often, if you want to have immediate access to your money then you will be exposed to interest rates that will change from time to time, as interest rates are not constant.

These two types of interest payments, **fixed** and **floating**, are seen in many financial instruments. **Coupon-bearing bonds** pay out a known amount every six months or year etc. This is the **coupon** and would often be a fixed rate of interest. At the end of your fixed term you get a final coupon and the return of the **principal**, the amount on which the interest was calculated. **Interest rate swaps** are an exchange of a fixed rate of interest for a floating rate of interest. Governments and companies issue bonds as a form of borrowing. The less creditworthy the issuer, the higher the interest that they will have to pay out. Bonds are actively traded, with prices that continually fluctuate.

Fixed-income modeling and products are the subject of Chapters 13 and 14 and the whole of Part Three.

1.8 INFLATION-PROOF BONDS

A recent addition to the list of bonds issued by the US government is the **index-linked bond**. These have been around in the UK since 1981, and have provided a very successful way of ensuring that income is not eroded by inflation.

Figure 1.15 UK gilts prices from *The Financial Times* of 14th April 2005. Reproduced by permission of *The Financial Times.*

In the UK inflation is measured by the **Retail Price Index** or **RPI**. This index is a measure of year-on-year inflation, using a 'basket' of goods and services including mortgage interest payments. The index is published monthly. The coupons and principal of the index-linked bonds are related to the level of the RPI. Roughly speaking, the amounts of the coupon and principal are scaled with the increase in the RPI over the period from the issue of the bond to the time of the payment. There is one slight complication in that the actual RPI level used in these calculations is set back *eight months*. Thus the base measurement is eight months before issue and the scaling of any coupon is with respect to the increase in the RPI from this base measurement to the level of the RPI eight months before the coupon is paid. One of the reasons for this complexity is that the initial estimate of the RPI is usually corrected at a later date.

Figure 1.15 shows the UK gilts prices published in *The Financial Times* of 14th April 2005. The index-linked bonds are on the right.

In the US the inflation index is the **Consumer Price Index** (**CPI**). A time series of this index is shown in Figure 1.16.

The dynamics of the relationship between inflation and short-term interest rates is particularly interesting. Clearly the level of interest rates will affect the rate of inflation directly through mortgage repayments, but also interest rates are often used by central banks as a tool for keeping inflation down.

We look at the modeling of inflation in Chapter 71.

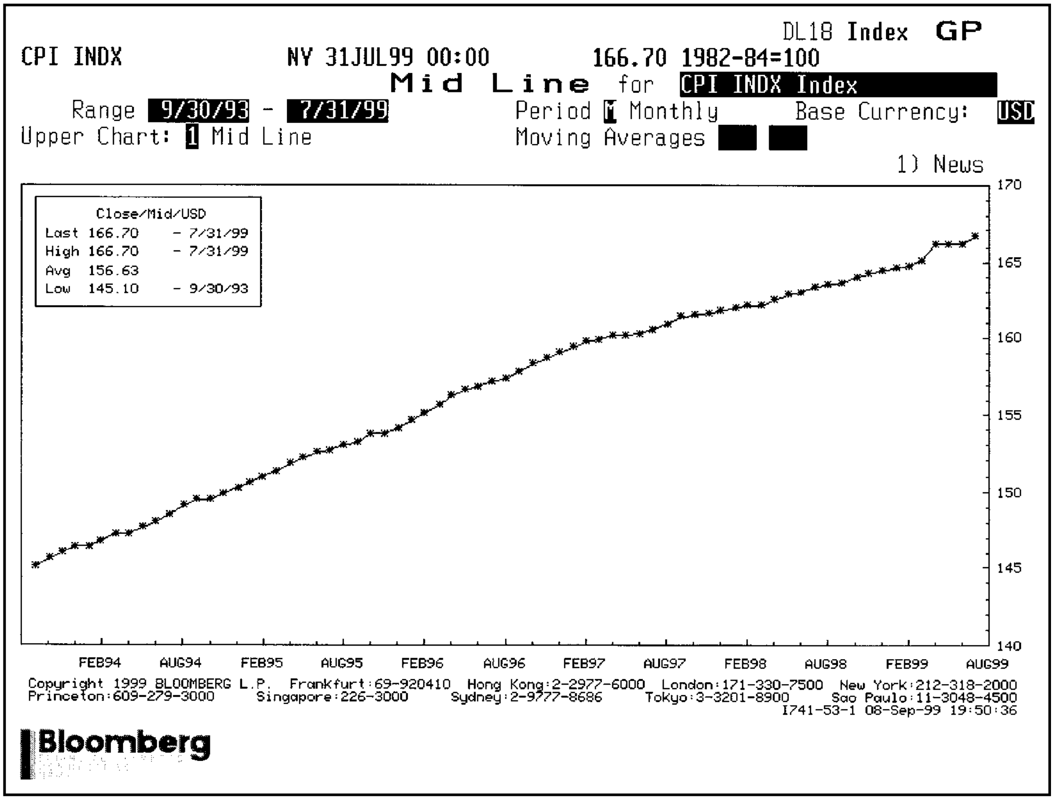
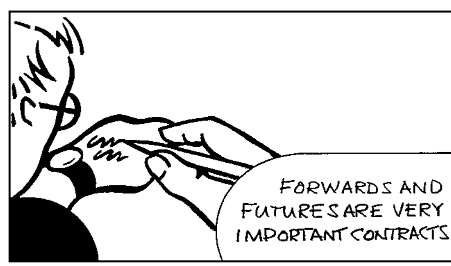


Figure 1.16 The CPI index. Source: Bloomberg L.P.

1.9 FORWARDS AND FUTURES



A **forward contract** is an agreement where one party promises to buy an asset from another party at some specified time in the future and at some specified price. No money changes hands until the **delivery date** or **maturity** of the contract. The terms of the contract make it an obligation to buy the asset at the delivery date, there is no choice in the matter. The asset could be a stock, a commodity or a currency.

A **futures contract** is very similar to a forward contract. Futures contracts are usually traded through an exchange, which standardizes the terms of the contracts. The profit or loss from the futures position is calculated every day and the change in this value is paid from one party to the other. Thus with futures contracts there is a gradual payment of funds from initiation until maturity.

Forwards and futures have two main uses, in speculation and in hedging. If you believe that the market will rise you can benefit from this by entering into a forward or futures contract.

If your market view is right then a lot of money will change hands (at maturity or every day) in your favor. That is speculation and is very risky. Hedging is the opposite, it is avoidance of risk. For example, if you are expecting to get paid in yen in six months' time, but you live in America and your expenses are all in dollars, then you could enter into a futures contract to lock in a guaranteed exchange rate for the amount of your yen income. Once this exchange rate is locked in you are no longer exposed to fluctuations in the dollar/yen exchange rate. But then you won't benefit if the yen appreciates.

1.9.1 A First Example of No Arbitrage

Futures and forwards provide us with our first example of the **no-arbitrage** principle.

Consider a forward contract that obliges us to hand over an amount $\$F$ at time T to receive the underlying asset. Today's date is t and the price of the asset is currently $\$S(t)$, this is the **spot price**, the amount for which we could get immediate delivery of the asset. When we get to maturity we will hand over the amount $\$F$ and receive the asset, then worth $\$S(T)$. How much profit we make cannot be known until we know the value $\$S(T)$, and we can't know this until time T. From now on I am going to drop the '\$' sign from in front of monetary amounts.

We know all of F, $S(t)$, t and T. But is there any relationship between them? You might think not, since the forward contract entitles us to receive an amount $S(T) - F$ at expiry and this is unknown. However, by entering into a special portfolio of trades *now* we can eliminate all randomness in the future. This is done as follows.

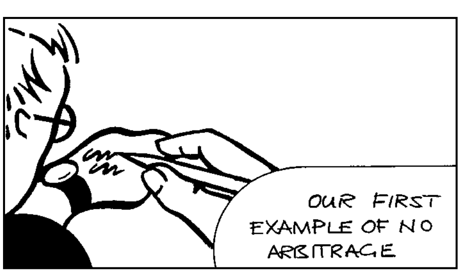
Enter into the forward contract. This costs us nothing up front but exposes us to the uncertainty in the value of the asset at maturity. Simultaneously sell the asset. It is called **going short** when you sell something you don't own. This is possible in many markets, but with some timing restrictions. We now have an amount $S(t)$ in cash due to the sale of the asset, a forward contract, and a short asset position. But our net position is zero. Put the cash in the bank, to receive interest.

When we get to maturity we hand over the amount F and receive the asset, this cancels our short asset position regardless of the value of $S(T)$. At maturity we are left with a guaranteed $-F$ in cash as well as the bank account. The word 'guaranteed' is important because it emphasizes that it is independent of the value of the asset. The bank account contains the initial investment of an amount $S(t)$ with added interest, which has a value at maturity of

$$S(t)e^{r(T-t)}.$$

Our net position at maturity is therefore

$$S(t)e^{r(T-t)} - F.$$



Since we began with a portfolio worth zero and we end up with a predictable amount, that predictable amount should also be zero. We can conclude that

$$\boxed{F = S(t)e^{r(T-t)}.} \tag{1.3}$$

Table 1.1 Cashflows in a hedged portfolio of asset and forward.

Holding	Worth today (t)	Worth at maturity (T)
Forward	0	$S(T) - F$
$-$Stock	$-S(t)$	$-S(T)$
Cash	$S(t)$	$S(t)e^{r(T-t)}$
Total	0	$S(t)e^{r(T-t)} - F$

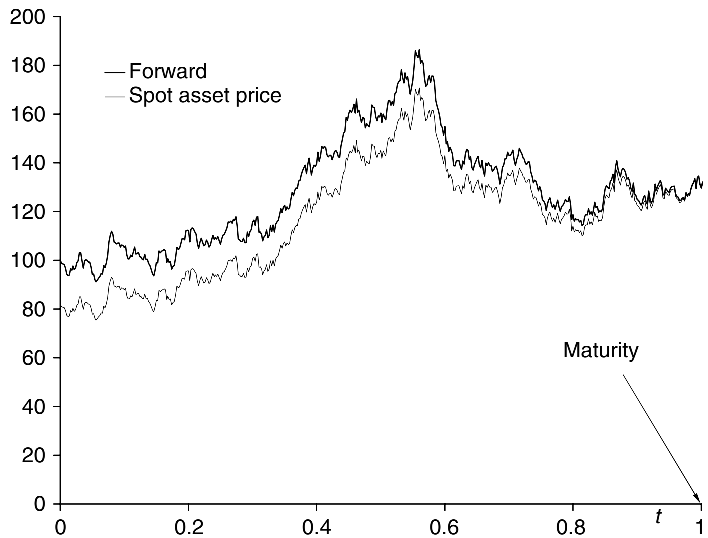


Figure 1.17 A time series of a spot asset price and its forward price.

This is the relationship between the spot price and the forward price. It is a linear relationship, the forward price is proportional to the spot price.

The cashflows in this special hedged portfolio are shown in Table 1.1.

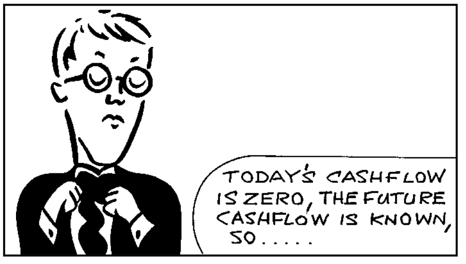


Figure 1.17 shows a path taken by the spot asset price and its forward price. As long as interest rates are constant, these two are related by (1.3).

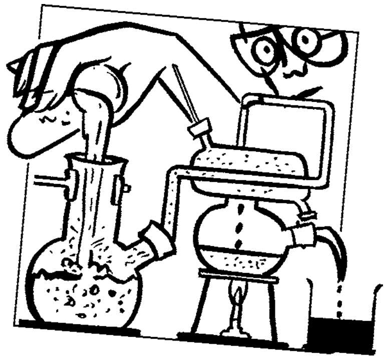
If this relationship is violated then there will be an arbitrage opportunity. To see what is meant by this, imagine that F is less than $S(t)e^{r(T-t)}$. To exploit this and make a riskless arbitrage profit, enter into the deals as explained above. At maturity you will have $S(t)e^{r(T-t)}$ in the bank, a short asset and a long forward. The asset position cancels when you hand over the amount F, leaving you with a profit of $S(t)e^{r(T-t)} - F$. If F is greater than that given by (1.3) then you enter into the opposite positions, going short the forward. Again you make a riskless profit. The standard economic argument then says that investors will act quickly to exploit the opportunity, and in the process prices will adjust to eliminate it.

1.10 SUMMARY

The above descriptions of financial markets are enough for this introductory chapter. Perhaps the most important point to take away with you is the idea of no arbitrage. In the example here, relating spot prices to futures prices, we saw how we could set up a very simple portfolio which completely eliminated any dependence on the future value of the stock. When we come to value derivatives, in the way we just valued a future, we will see that the same principle can be applied albeit in a far more sophisticated way.

FURTHER READING

- For general financial news visit `www.bloomberg.com` and `www.reuters.com`. CNN has online financial news at `www.cnnfn.com`. There are also online editions of *The Wall Street Journal*, `www.wsj.com`, *The Financial Times*, `www.ft.com` and *Futures and Options World*, `www.fow.com`.
- For more information about futures see the Chicago Board of Trade website `www.cbot.com`.
- Many, many financial links can be found at Wahoo!, `www.io.com/˜gibbonsb/wahoo.html`.
- In the main, we'll be assuming that markets are random. For insight about alternative hypotheses see Schwager (1990, 1992).
- See Brooks (1967) for how the raising of capital for a business might work in practice.
- Cox, Ingersoll & Ross (1981) discuss the relationship between forward and future prices.



CHAPTER 2
derivatives

In this Chapter...

- the definitions of basic derivative instruments
- option jargon
- how to draw payoff diagrams
- no arbitrage and put-call parity
- simple option strategies

2.1 INTRODUCTION

The previous chapter dealt with some of the basics of financial markets. I didn't go into any detail, just giving the barest outline and setting the scene for this chapter. Here I introduce the theme that is central to the book, the subject of options, a.k.a. derivatives or contingent claims. This chapter is nontechnical, being a description of some of the most common option contracts, and explanation of the market-standard jargon. It is in later chapters that I start to get technical.

Options have been around for many years, but it was only on 26th April 1973 that they were first traded on an exchange. It was then that The Chicago Board Options Exchange (CBOE) first created standardized, listed options. Initially there were just calls on 16 stocks. Puts weren't even introduced until 1977. In the US options are traded on CBOE, the American Stock Exchange, the Pacific Exchange and the Philadelphia Stock Exchange. Worldwide, there are over 50 exchanges on which options are traded.

2.2 OPTIONS

If you are reading the book in a linear fashion, from start to finish, then the last topics you read about will have been futures and forwards. The holder of future or forward contracts is *obliged* to trade at the maturity of the contract. Unless the position is closed before maturity the holder must take possession of the commodity, currency or whatever is the subject of the contract, regardless of whether the asset has risen or fallen. Wouldn't it be nice if we only had to take possession of the asset if it had risen in value?

The simplest **option** gives the holder the *right* to trade in the future at a previously agreed price but takes away the obligation. So if the stock falls, we don't have to buy it after all.

> A **call option** is the right to buy a particular asset for an agreed amount at a specified time in the future

As an example, consider the following call option on Microsoft stock. It gives the holder the right to buy one of Microsoft stock for an amount \$25 in one month's time. Today's stock price is \$24.5. The amount '25' which we can pay for the stock is called the **exercise price** or **strike price**. The date on which we must **exercise** our option, if we decide to, is called the **expiry** or **expiration date**. The stock on which the option is based is known as the **underlying asset**.

Let's consider what may happen over the next month, up until expiry. Suppose that nothing happens, that the stock price remains at \$24.5. What do we do at expiry? We could exercise the option, handing over \$25 to receive the stock. Would that be sensible? No, because the stock is only worth \$24.5, either we wouldn't exercise the option or if we really wanted the stock we would buy it in the stock market for the \$24.5. But what if the stock price rises to \$29? Then we'd be laughing, we would exercise the option, paying \$25 for a stock that's worth \$29, a profit of \$4.

We would exercise the option at expiry if the stock is above the strike and not if it is below. If we use S to mean the stock price and E the strike then at expiry the option is worth

$$\max(S - E, 0).$$

This function of the underlying asset is called the **payoff function**. The 'max' function represents the optionality.

Why would we buy such an option? Clearly, if you own a call option you want the stock to rise as much as possible. The higher the stock price the greater will be your profit. I will discuss this below, but our decision whether to buy it will depend on how much it costs; the option is valuable, there is no downside to it unlike a future. In our example the option was valued at \$1.875. Where did this number come from? The valuation of options is one of the subjects of this book, and I'll be showing you how to find this value later on.

What if you believe that the stock is going to fall, is there a contract that you can buy to benefit from the fall in a stock price?

> A **put option** is the right to *sell* a particular asset for an agreed amount at a specified time in the future

The holder of a put option wants the stock price to fall so that he can sell the asset for more than it is worth. The payoff function for a put option is

$$\max(E - S, 0).$$

Now the option is only exercised if the stock falls below the strike price.

Publisher's Note:
Permission to reproduce this image online was not granted by the copyright holder. Readers are kindly requested to refer to the printed version of this chapter.

Figure 2.1 *The Wall Street Journal Europe* of 14th April 2005, Stock Options. Reproduced by permission of Dow Jones & Company, Inc.

Publisher's Note:
Permission to reproduce this image online was not granted by the copyright holder. Readers are kindly requested to refer to the printed version of this chapter.

Figure 2.2 *The Wall Street Journal Europe* of 5th January 2000, Index Options. Reproduced by permission of Dow Jones & Company, Inc.

Figure 2.1 is an excerpt from *The Wall Street Journal Europe* of 14th April 2005 showing options on various stocks. The table lists closing prices of the underlying stocks and the last traded prices of the options on the stocks. To understand how to read this let us examine the prices of options on Apple. Go to 'AppleC' in the list, there are several instances. The closing price on 13th April 2005 was \$41.35, (the LAST column, second from the right). Calls and puts are quoted here with strikes of \$37.50, \$40, ..., \$47.50, \$50, others may exist but are not included in the newspaper. The expiries mentioned are April, May and July. Part of the information included here is the volume of the transactions in each series, we won't worry about that but some people use option volume as a trading indicator. From the data, we can see that the April calls with a strike of \$40 were worth \$2.40. The puts with same strike and expiry were worth \$1.20. The April calls with a strike of \$42.50 were worth \$1.20 and the puts with same strike and expiry were worth \$2.45. Note that the higher the strike, the lower the value of the calls but the higher the value of the puts. This makes sense when you remember that the call allows you to buy the underlying for the strike, so that the lower the strike price the more this right is worth to you. The opposite is true for a put since it allows you to sell the underlying for the strike price.

There are more strikes and expiries available for options on indices, so let's now look at the Index Options section of *The Wall Street Journal Europe* 5th January 2000, this is shown in Figure 2.2.

In Figure 2.3 are the quoted prices of the March and June DJIA calls against the strike price. Also plotted is the payoff function *if the underlying were to finish at its current value at expiry*, the closing price of the DJIA was 10997.93 on the day the option prices were quoted.

This plot reinforces the fact that the higher the strike the lower the value of a call option. It also appears that the longer the time to maturity the higher the value of the call. Is it obvious

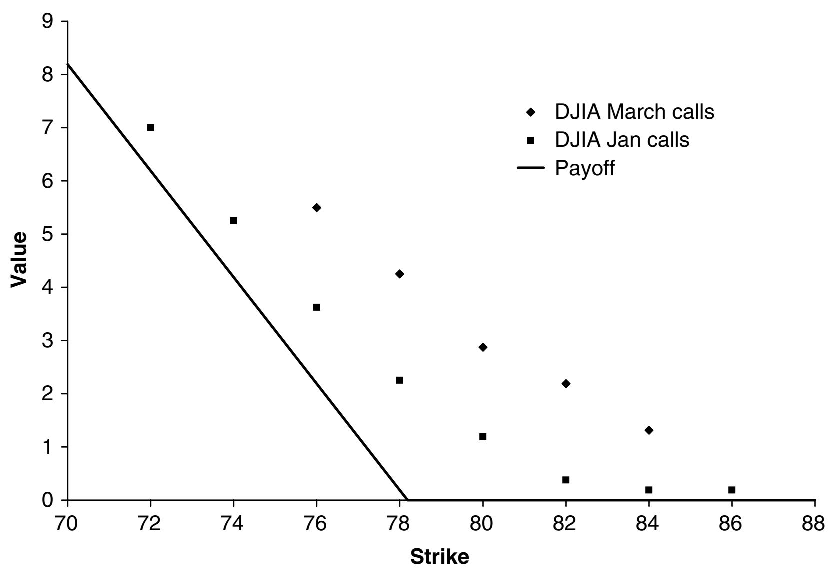


Figure 2.3 Option prices versus strike, March and June series of DJIA.

that this should be so? As the time to expiry decreases what would we see happen? As there is less and less time for the underlying to move, so the option value must converge to the payoff function.

One of the most interesting features of calls and puts is that they have a non-linear dependence on the underlying asset. This contrasts with futures which have a linear dependence on the underlying. This non-linearity is very important in the pricing of options since the randomness in the underlying asset and the curvature of the option value with respect to the asset are intimately related.

Calls and puts are the two simplest forms of option. For this reason they are often referred to as **vanilla** because of the ubiquity of that flavor. There are many, many more kinds of options, some of which will be described and examined later on. Other terms used to describe contracts with some dependence on a more fundamental asset are **derivatives** or **contingent claims**.

Figure 2.4 shows the prices of call options on Glaxo–Wellcome for a variety of strikes as of January. All these options are expiring in October. The table shows many other quantities that we will be seeing later on.

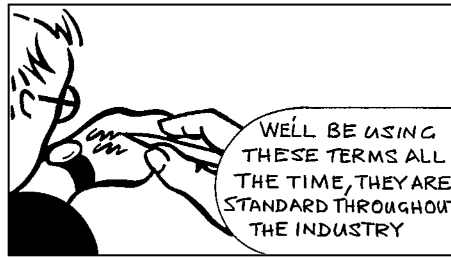
GLXO LN GBp ↑ 1688 −13 L 5s L 1686/1689 L Trd Equity OCM
At 12:50 Vol 854,194 Op 1694 L Hi 1703 L Lo 1686 L Prev 1701
OPTION MONITOR 3 COMP Center: 1687 1 <GO> to Edit Spreadsheet

	BID m	ASK m	LAST m	1CHG m	IVBD	IVAS	BEST	DEBS	GABS	VEBS	THEO	7DEC
GLXO LN CALLS	Bid Price	Ask Price	Last Trade	1 Day Net Change	Imp Volat Bid	Imp Volat Ask	Best Price	Delta Best Price	Gamma Best Price	Vega Best Price	Theo. Value	7 Day Decay
GLXOCT99	1686.0	1689.0	1688.0	-13.0			1687					
1) 1200	489.50	504.50	509.50	unch	N.A.	69.97	504.50	.942	.0003	.674	494.09	4.6870
2) 1250	440.00	455.00	460.00	unch	N.A.	63.58	455.00	.936	.0003	.689	444.92	4.6853
3) 1300	390.50	405.50	410.50	unch	N.A.	57.36	405.50	.928	.0004	.837	396.33	4.6828
4) 1350	342.00	357.00	362.00	unch	N.A.	52.29	357.00	.915	.0005	.853	348.72	4.8888
5) 1400	294.50	309.50	314.50	unch	N.A.	48.07	309.50	.895	.0007	1.018	302.62	5.2385
6) 1450	249.00	264.00	268.50	unch	29.45	45.11	264.00	.864	.0008	1.194	258.66	5.8316
7) 1500	203.00	218.00	224.00	unch	30.67	42.27	220.00	.823	.0011	1.538	217.53	6.3538
8) 1600	125.00	137.50	136.00	-6.00	29.86	37.59	136.00	.706	.0017	2.013	146.02	7.0423
9) 1700	69.00	76.00	80.00	unch	30.95	34.02	76.00	.516	.0020	2.280	90.956	7.4785
10) 1800	32.00	38.00	40.00	unch	30.62	33.12	37.00	.319	.0019	2.005	52.713	6.2390
11) 1900	16.00	20.00	21.50	unch	32.84	35.47	20.00	.190	.0013	1.552	28.386	4.8611
12) 2000	6.00	9.00	9.00	unch	32.53	35.83	9.00	.099	.0008	1.041	14.273	2.9660
13) 2100	2.00	4.00	3.50	unch	32.32	36.52	3.50	.044	.0005	.581	6.728	1.4568
14) 2200		2.00	1.00	unch	N.A.	38.08	1.00	.015	.0002	.232	2.968	.5272
15) 2300		1.50	.50	unch	N.A.	41.58	.50	.008	.0001	.132	1.262	.2929
16) 2400		1.00	.50	unch	N.A.	43.98	.50	.007	.0001	.126	.502	.2977
17) 2500		1.00	.50	unch	N.A.	48.40	.50	.007	.0001	.101	.195	.3010

Bloomberg PROFESSIONAL

Figure 2.4 Prices for Glaxo–Wellcome calls expiring in October. Source: Bloomberg L.P.

2.3 DEFINITION OF COMMON TERMS



The subjects of mathematical finance and derivatives theory are filled with jargon. The jargon comes from both the mathematical world and the financial world. Generally speaking the jargon from finance is aimed at simplifying communication, and to put everyone on the same footing.[1] Here are a few loose definitions to be going on with, some you have already seen and there will be many more throughout the book.

- **Premium:** The amount paid for the contract initially. How to find this value is the subject of much of this book.
- **Underlying (asset):** The financial instrument on which the option value depends. Stocks, commodities, currencies and indices are going to be denoted by S. The option payoff is defined as some function of the underlying asset at expiry.
- **Strike (price)** or **exercise price:** The amount for which the underlying can be bought (call) or sold (put). This will be denoted by E. This definition only really applies to the simple calls and puts. We will see more complicated contracts in later chapters and the definition of strike or exercise price will be extended.
- **Expiration (date)** or **expiry (date):** Date on which the option can be exercised or date on which the option ceases to exist or give the holder any rights. This will be denoted by T.
- **Intrinsic value:** The payoff that would be received if the underlying is at its current level when the option expires.
- **Time value:** Any value that the option has above its intrinsic value. The uncertainty surrounding the future value of the underlying asset means that the option value is generally different from the intrinsic value.
- **In the money:** An option with positive intrinsic value. A call option when the asset price is above the strike, a put option when the asset price is below the strike.
- **Out of the money:** An option with no intrinsic value, only time value. A call option when the asset price is below the strike, a put option when the asset price is above the strike.
- **At the money:** A call or put with a strike that is close to the current asset level.
- **Long position:** A positive amount of a quantity, or a positive exposure to a quantity.
- **Short position:** A negative amount of a quantity, or a negative exposure to a quantity. Many assets can be sold short, with some constraints on the length of time before they must be bought back.

[1] I have serious doubts about the purpose of most of the math jargon.

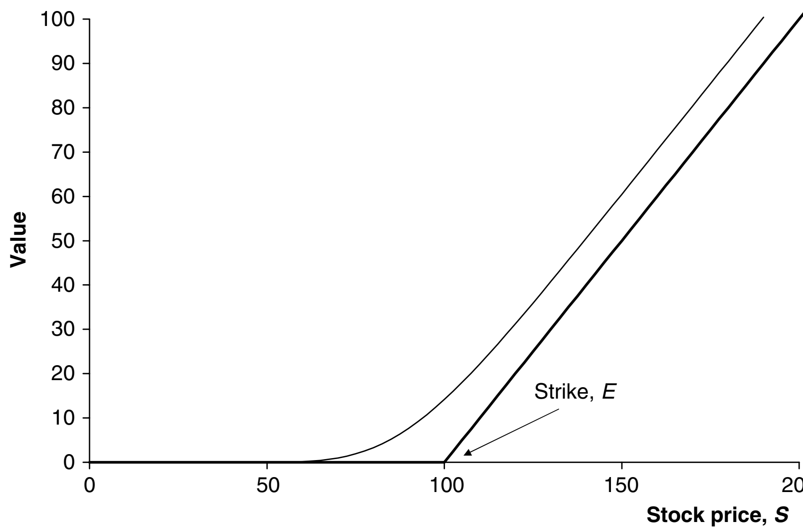


Figure 2.5 Payoff diagram for a call option.

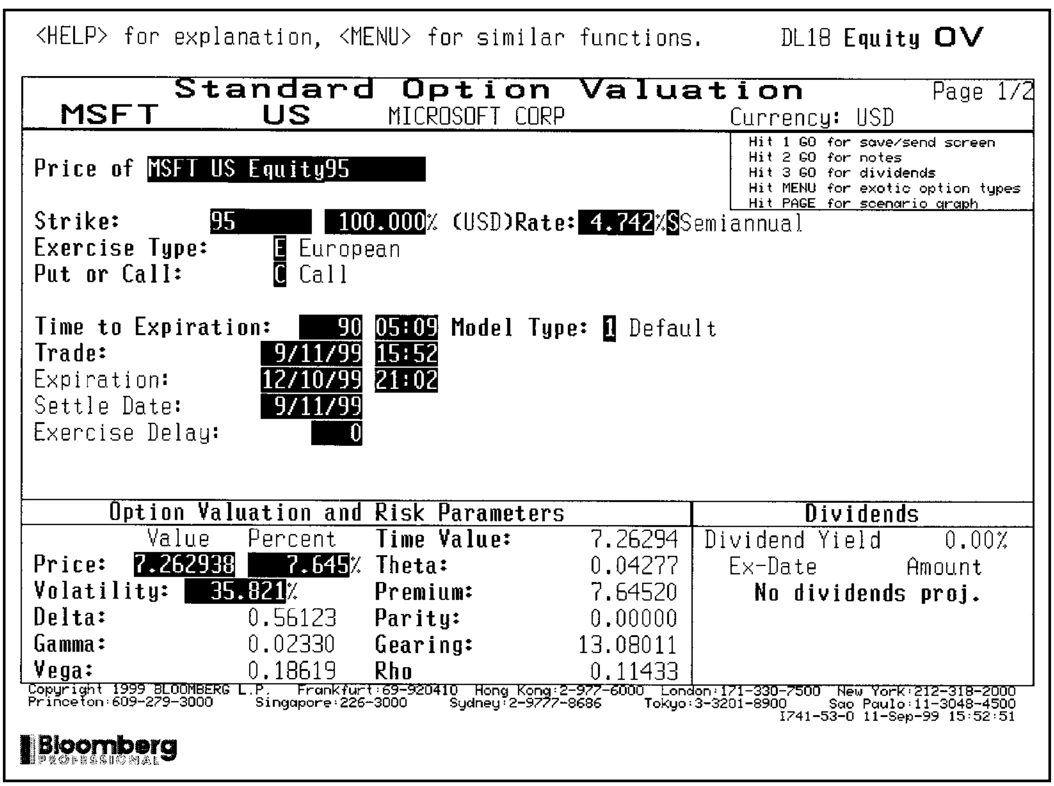


Figure 2.6 Bloomberg option valuation screen, call. Source: Bloomberg L.P.

2.4 PAYOFF DIAGRAMS

The understanding of options is helped by the visual interpretation of an option's value at expiry. We can plot the value of an option at expiry as a function of the underlying in what is known as a **payoff diagram**. At expiry the option is worth a known amount. In the case of a call option the contract is worth $\max(S - E, 0)$. This function is the bold line in Figure 2.5.

Figure 2.6 shows Bloomberg's standard option valuation screen and Figure 2.7 shows the value against the underlying and the payoff.

The payoff for a put option is $\max(E - S, 0)$, this is the bold line plotted in Figure 2.8. Figure 2.9 shows Bloomberg's option valuation screen and Figure 2.10 shows the value against the underlying and the payoff.

These payoff diagrams are useful since they simplify the analysis of complex strategies involving more than one option.

Make a note of the thin lines in all of these figures. The meaning of these will be explained very shortly.

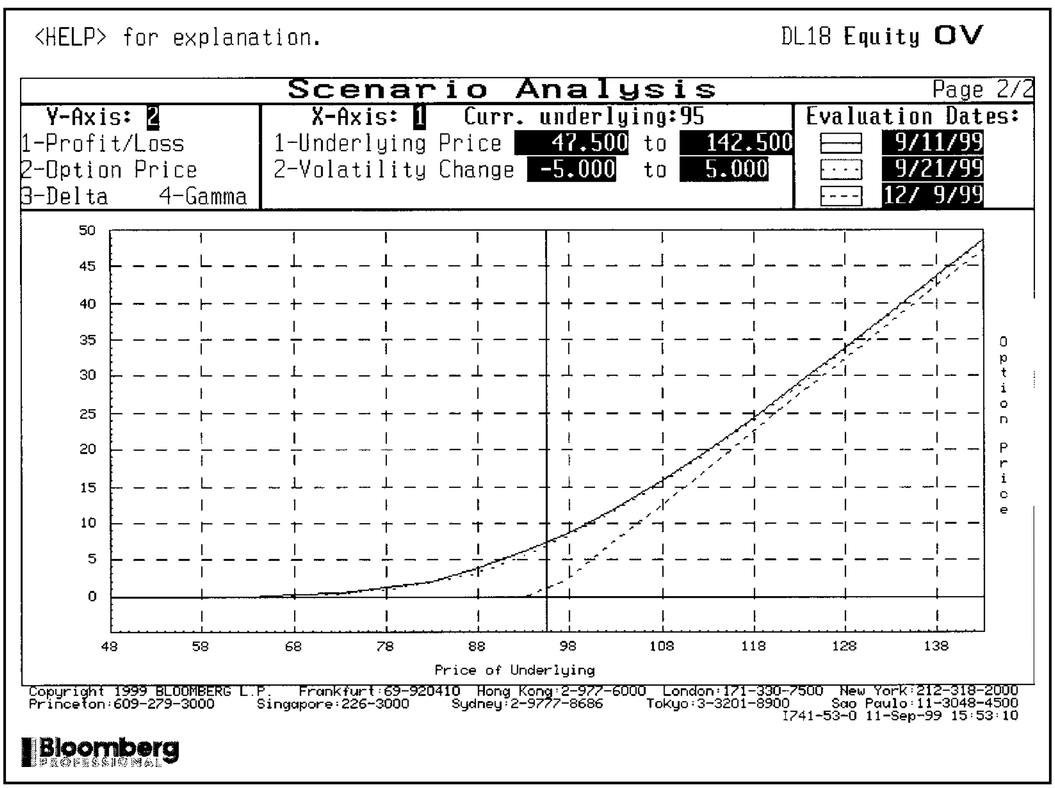


Figure 2.7 Bloomberg scenario analysis, call. Source: Bloomberg L.P.

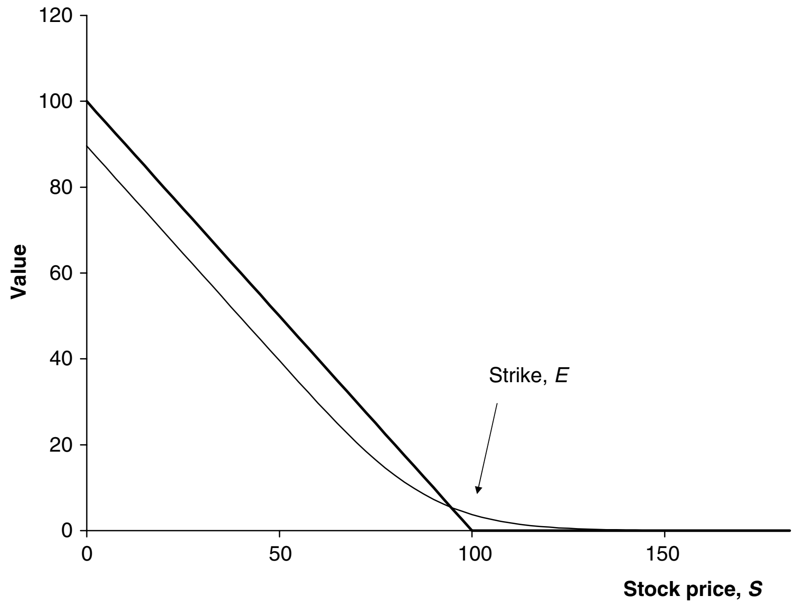


Figure 2.8 Payoff diagram for a put option.

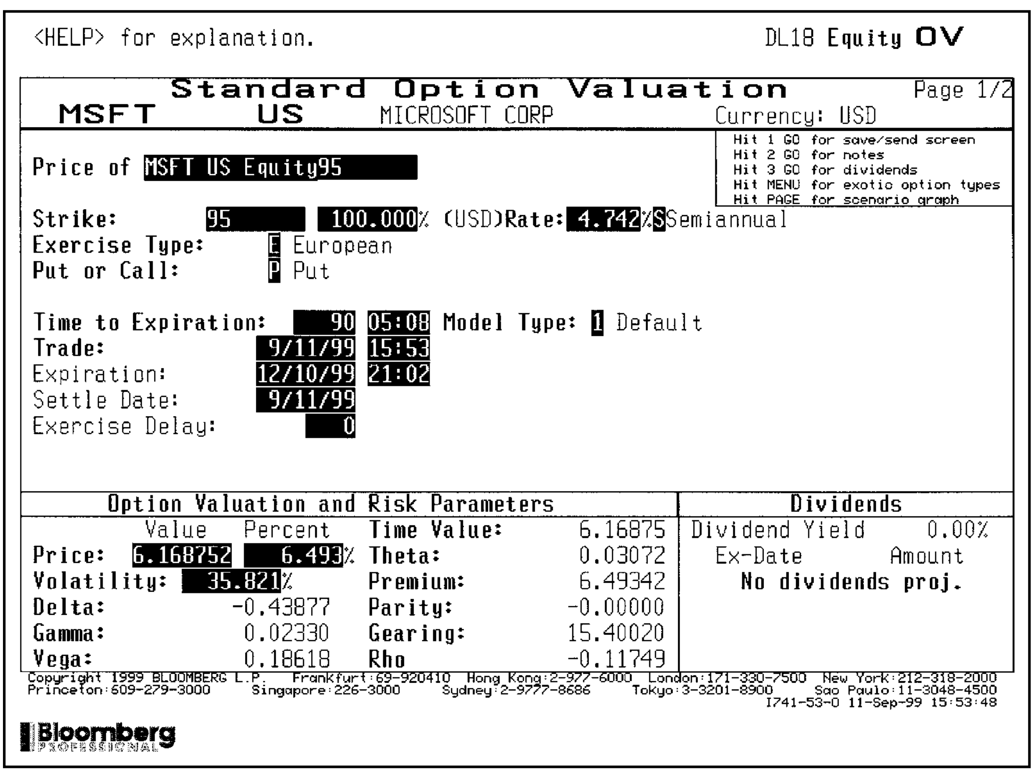


Figure 2.9 Bloomberg option valuation screen, put. Source: Bloomberg L.P.

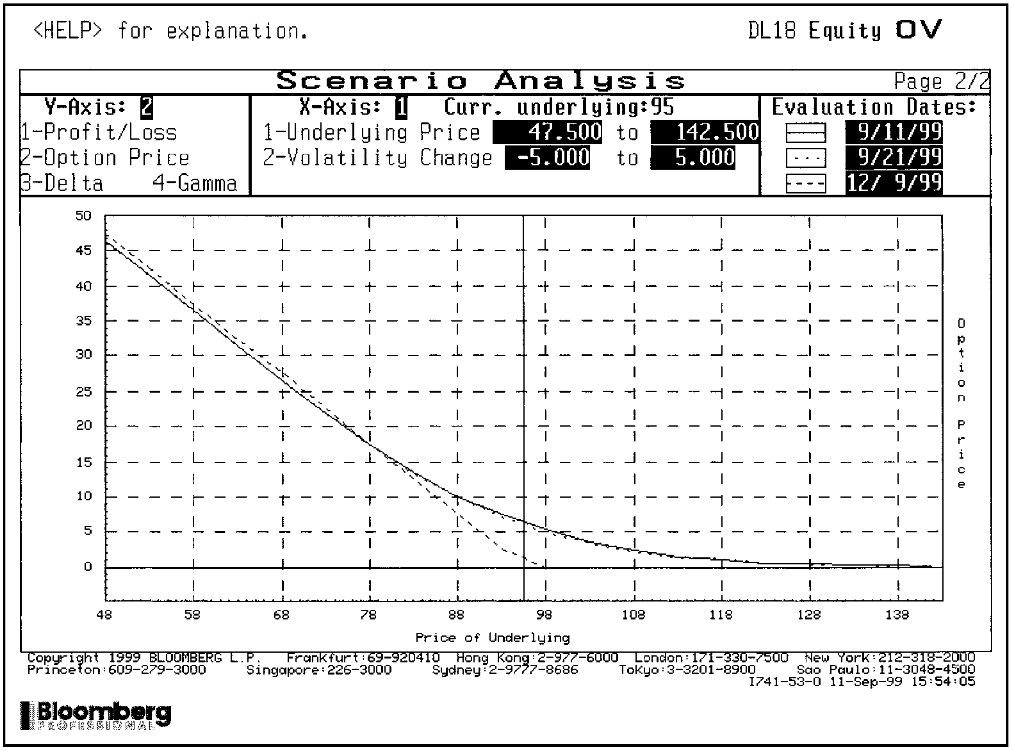


Figure 2.10 Bloomberg scenario analysis, put. Source: Bloomberg L.P.

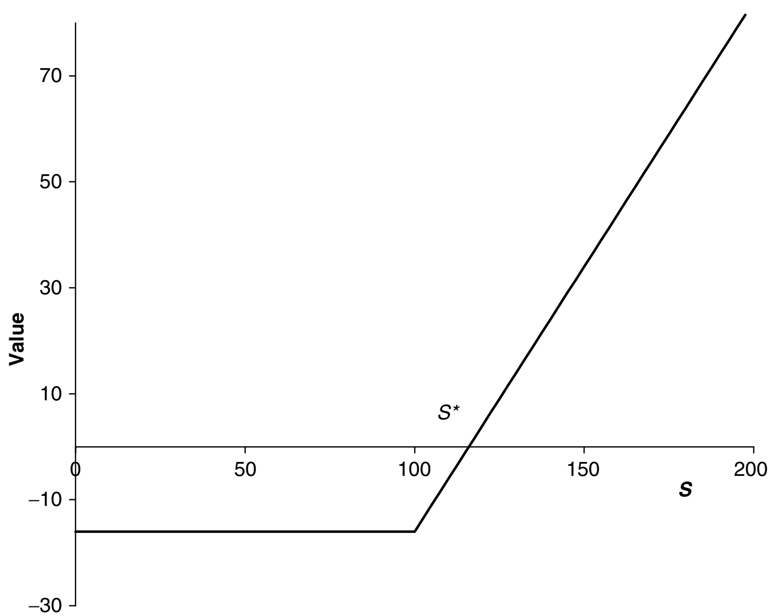


Figure 2.11 Profit diagram for a call option.

2.4.1 Other Representations of Value

The payoff diagrams shown above only tell you about what happens at expiry, how much money your option contract is worth at that time. It makes no allowance for how much premium you had to pay for the option. To adjust for the original cost of the option, sometimes one plots a diagram such as that shown in Figure 2.11. In this **profit diagram** for a call option I have subtracted from the payoff the premium originally paid for the call option. This figure is helpful because it shows how far into the money the asset must be at expiry before the option becomes profitable. The asset value marked S^* is the point which divides profit from loss; if the asset at expiry is above this value then the contract has made a profit, if below the contract has made a loss.

As it stands, this profit diagram takes no account of the time value of money. The premium is paid up front but the payoff, if any, is only received at expiry. To be consistent one should either discount the payoff by multiplying by $e^{-r(T-t)}$ to value everything at the present, or multiply the premium by $e^{r(T-t)}$ to value all cashflows at expiry.

Figure 2.12 shows Bloomberg's call option profit diagram. Note that the profit today is zero; if we buy the option and immediately sell it we make neither a profit nor a loss (this is subject to issues of transaction costs).

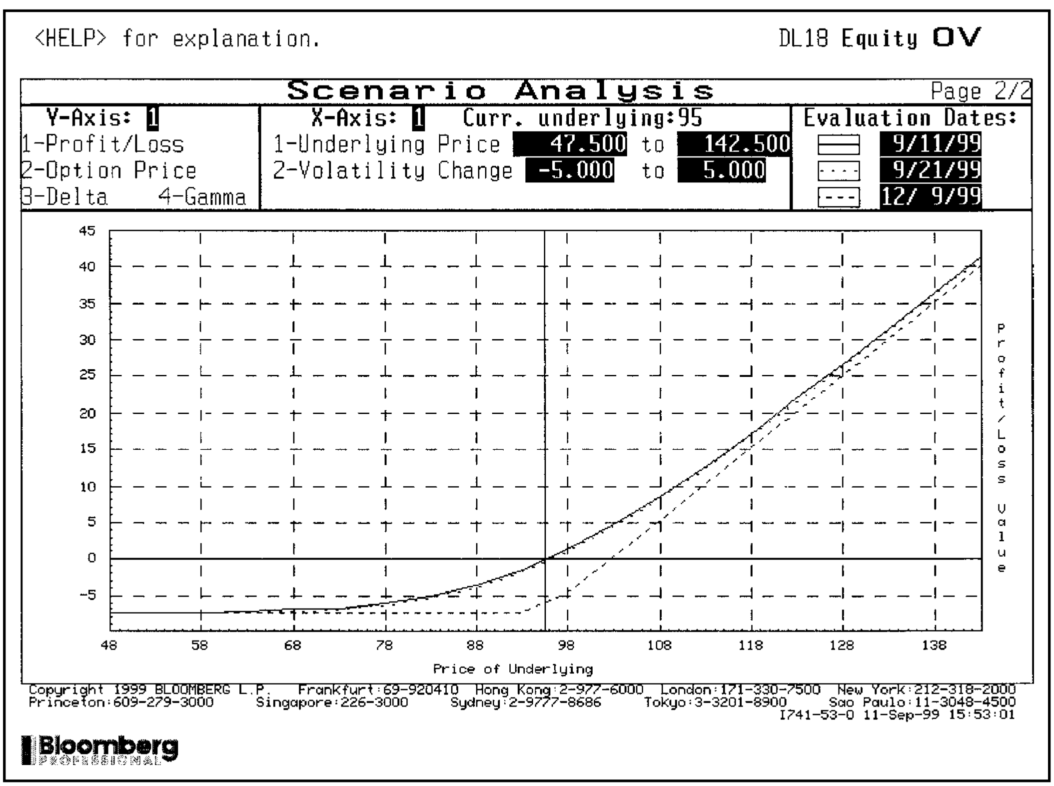
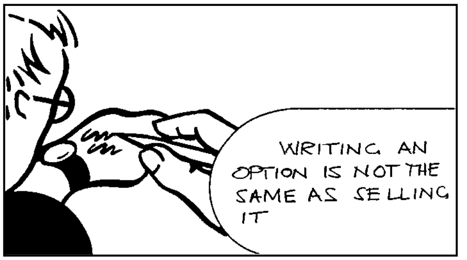


Figure 2.12 Profit diagram for a call. Source: Bloomberg L.P.

2.5 WRITING OPTIONS

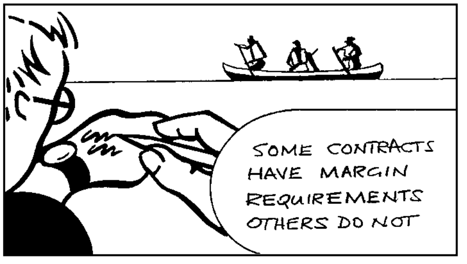


I have talked above about the rights of the purchaser of the option. But for every option that is sold, someone somewhere must be liable if the option is exercised. If I hold a call option entitling me to buy a stock some time in the future, who do I buy this stock from? Ultimately, the stock must be delivered by the person who **wrote** the option. The **writer** of an option is the person who promises to deliver the underlying asset, if the option is a call, or buy it, if the option is a put. The writer is the person who receives the premium.

In practice, most simple option contracts are handled through an exchange so that the purchaser of an option does not know who the writer is. The holder of the option can even sell the option on to someone else via the exchange to close his position. However, regardless of who holds the option, or who has handled it, the writer is the person who has the obligation to deliver or buy the underlying.

The asymmetry between owning and writing options is now clear. The purchaser of the option hands over a premium in return for special rights, and an uncertain outcome. The writer receives a guaranteed payment up front, but then has obligations in the future.

2.6 MARGIN



Writing options is very risky. The downside of *buying* an option is just the initial premium, the upside may be unlimited. The upside of *writing* an option is limited, but the downside could be huge. For this reason, to cover the risk of default in the event of an unfavorable outcome, the **clearing houses** that register and settle options insist on the deposit of a **margin** by the writers of options. Clearing houses act as counterparty to each transaction.

Margin comes in two forms, the **initial margin** and the **maintenance margin**. The initial margin is the amount deposited at the initiation of the contract. The total amount held as margin must stay above a prescribed maintenance margin. If it ever falls below this level then more money (or equivalent in bonds, stocks etc.) must be deposited. The levels of these margins vary from market to market.

Margin has been much neglected in the academic literature. But a poor understanding of the subject has led to a number of famous financial disasters, most notably Metallgesellschaft and Long Term Capital Management. We'll discuss the details of these cases in Chapter 44, and we'll also be seeing how to model margin and how to margin hedge.

2.7 MARKET CONVENTIONS

Most of the simpler options contracts are bought and sold through exchanges. These exchanges make it easier and more efficient to match buyers with sellers. Part of this simplification involves the conventions about such features of the contracts as the available strikes and expiries. For example, vanilla calls and puts come in **series**. This refers to the strike and expiry dates.

Typically a stock has three choices of expiries trading at any time. Having standardized contracts traded through an exchange promotes liquidity of the instruments.

Some options are an agreement between two parties, not through an exchange, but often brought together by an intermediary. These agreements can be very flexible and the contract details do not need to satisfy any conventions. Such contracts are known as **over the counter** or **OTC** contracts. I give an example at the end of this chapter.

2.8 THE VALUE OF THE OPTION BEFORE EXPIRY

We have seen how much calls and puts are worth at expiry, and drawn these values in payoff diagrams. The question that we can ask, and the question that is central to this book, is 'How much is the contract worth *now*, before expiry?' How much would you pay for a contract, a piece of paper, giving you rights in the future? You may have no idea what the stock price will do between now and expiry in six months, say, but clearly the contract has value. At the very least you know that there is no downside to owning the option, the contract gives you specific rights but no *obligations*. Two things are clear about the contract value before expiry: the value will depend on how high the asset price is today and how long there is before expiry.

The higher the underlying asset today, the higher we might expect the asset to be at expiry of the option and therefore the more valuable we might expect a call option to be. On the other hand a put option might be cheaper by the same reasoning.

The dependence on time to expiry is more subtle. The longer the time to expiry, the more time there is for the asset to rise or fall. Is that good or bad if we own a call option? Furthermore, the longer we have to wait until we get any payoff, the less valuable will be that payoff simply because of the time value of money.

I will ask you to suspend disbelief for the moment (it won't be the first time in the book) and trust me that we will be finding a 'fair value' for these options contracts. The aspect of finding the 'fair value' that I want to focus on now is the dependence on the asset price and time. I am going to use V to mean the value of the option, and it will be a function of the value of the underlying asset S at time t. Thus we can write $V(S,t)$ for the value of the contract.

We know the value of the contract *at expiry*. If I use T to denote the expiry date then at $t = T$ the function V is known, it is just the payoff function. For example if we have a call option then

$$V(S,T) = \max(S - E, 0).$$

This is the function of S that I plotted in the earlier payoff diagrams. Now I can tell you what the fine lines are in Figures 2.5 and 2.8: they are the values of the contracts $V(S,t)$ *at some time before expiry*, plotted against S. I have not specified how long before expiry, since the plot is for explanatory purposes only. I will spend a lot of time showing you how to find these values for a wide variety of contracts.

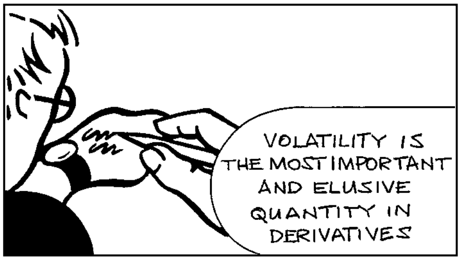
2.9 FACTORS AFFECTING DERIVATIVES PRICES

The two most important factors affecting the prices of options are the value of the underlying asset S and the time to expiry t. These quantities are **variables** meaning that they inevitably change during the life of the contract; if the underlying did not change then the pricing would be trivial. This contrasts with the **parameters** that affect the price of options.

Examples of parameters are the interest rate and strike price. The interest rate will have an effect on the option value via the time value of money since the payoff is received in the future. The interest rate also plays another role which we will see later. Clearly the strike price is important; the higher the strike in a call, the lower the value of the call.

If we have an equity option then its value will depend on any dividends that are paid on the asset during the option's life. If we have an FX option then its value will depend on the interest rate received by the foreign currency.

There is one important parameter that I have not mentioned, and which has a major impact on the option value. That parameter is the **volatility**. Volatility is a measure of the amount of fluctuation in the asset price, a measure of the randomness. Figure 2.13 shows two asset price paths, the more jagged of the two has the higher volatility. The technical definition of volatility is the 'annualized standard deviation of the asset returns.' I will show how to measure this parameter in Chapter 3.



Volatility is a particularly interesting parameter because it is so hard to estimate. And having estimated it, one finds that it never stays constant and is unpredictable. Once you start to think of the volatility as varying in a random fashion then it becomes natural to treat it as a variable also. We will see examples of this later in the book.[2]

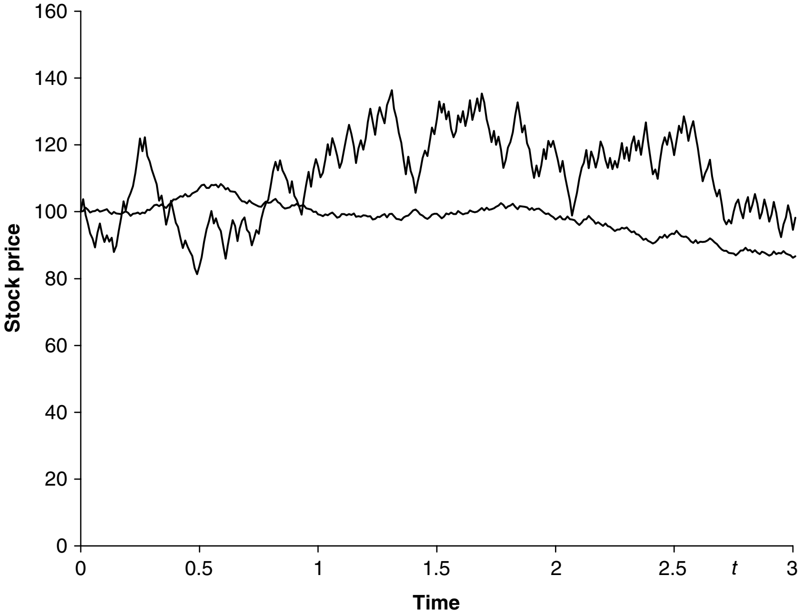


Figure 2.13 Two (simulated) asset price paths, one is much more volatile than the other.

[2] In finance we are overloaded with data, but it is never clear how useful these data will be for helping us to model the future. Contrast this with problems in the hard sciences. For example, the average number of sunspots on the sun has probably been quite stable for millions of years, but data going that far back are impossible to get, obviously.

The distinction between parameters and variables is very important. I shall be deriving equations for the value of options, partial differential equations. These equations will involve differentiation with respect to the variables, but the parameters, as their name suggests, remain as parameters in the equations.

2.10 SPECULATION AND GEARING

If you buy a far out-of-the-money option it may not cost very much, especially if there is not very long until expiry. If the option expires worthless, then you also haven't lost very much. However, if there is a dramatic move in the underlying, so that the option expires in the money, you may make a large profit relative to the amount of the investment. Let me give an example.

Example

Today's date is 14th April and the price of Wilmott Inc. stock is \$666. The cost of a 680 call option with expiry 22nd August is \$39. I expect the stock to rise significantly between now and August, how can I profit if I am right?

Buy the stock

Suppose I buy the stock for \$666. And suppose that by the middle of August the stock has risen to \$730. I will have made a profit of \$64 per stock. More importantly my investment will have risen by

$$\frac{730 - 666}{666} \times 100 = 9.6\%.$$

Buy the call

If I buy the call option for \$39, then at expiry I can exercise the call, paying \$680 to receive something worth \$730. I have paid \$39 and I get back \$50. This is a profit of \$11 per option, but in percentage terms I have made

$$\frac{\text{value of asset at expiry} - \text{strike} - \text{cost of call}}{\text{cost of call}} \times 100 = \frac{730 - 680 - 39}{39} \times 100 = 28\%.$$

This is an example of **gearing** or **leverage**. The out-of-the-money option has a high gearing, a possible high payoff for a small investment. The downside of this leverage is that the call option is more likely than not to expire completely worthless and you will lose all of your investment. If Wilmott Inc. remains at \$666 then the stock investment has the same value but the call option experiences a 100% loss.

Highly-leveraged contracts are very risky for the writer of the option. The buyer is only risking a small amount; although he is very likely to lose, his downside is limited to his initial premium. But the writer is risking a large loss in order to make a probable small profit. The writer is likely to think twice about such a deal unless he can offset his risk by buying other contracts. This offsetting of risk by buying other related contracts is called **hedging**.

Gearing explains one of the reasons for buying options. If you have a strong view about the direction of the market then you can exploit derivatives to make a better return, if you are right, than buying or selling the underlying.

2.11 EARLY EXERCISE

The simple options described above are examples of **European options** because exercise is only permitted *at expiry*. Some contracts allow the holder to exercise *at any time* before expiry, and these are called **American options**. American options give the holder more rights than their European equivalent and can therefore be more valuable, and they can never be less valuable. The main point of interest with American-style contracts is deciding *when* to exercise. In Chapter 9 I will discuss American options in depth, and show how to determine when it is *optimal* to exercise, so as to give the contract the highest value.

Note that the terms 'European' and 'American' do not in any way refer to the continents on which the contracts are traded.

Finally, there are **Bermudan options**. These allow exercise on specified dates, or in specified periods. In a sense they are half way between European and American since exercise is allowed on some days and not on others.

2.12 PUT-CALL PARITY

Imagine that you buy one European call option with a strike of E and an expiry of T and that you write a European put option with the same strike and expiry. Today's date is t. The payoff you receive at T for the call will look like the line in the first plot of Figure 2.14. The payoff for the put is the line in the second plot in the figure. Note that the sign of the payoff is negative; you *wrote* the option and are liable for the payoff. The payoff for the portfolio of the two options is the sum of the individual payoffs, shown in the third plot. The payoff for this portfolio of options is

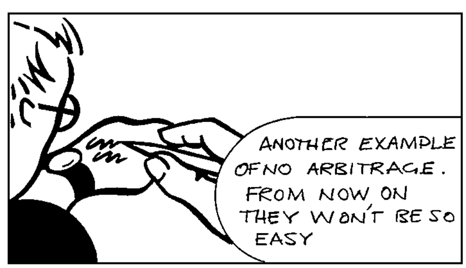
$$\max(S(T) - E, 0) - \max(E - S(T), 0) = S(T) - E,$$

where $S(T)$ is the value of the underlying asset at time T.

The right-hand side of this expression consists of two parts, the asset and a fixed sum E. Is there another way to get exactly this payoff? If I buy the asset today it will cost me $S(t)$ and be worth $S(T)$ at expiry. I don't know what the value $S(T)$ will be but I do know how to guarantee to get that amount, and that is to buy the asset. What about the E term? To lock in a payment of E at time T involves a cash flow of $Ee^{-r(T-t)}$ at time t. The conclusion is that the portfolio of a long call and a short put gives me exactly the same payoff as a long asset, short cash position. The equality of these cashflows is independent of the future behavior of the stock and is model independent:

$$C - P = S - Ee^{-r(T-t)},$$

where C and P are today's values of the call and the put respectively. This relationship holds at any time up to expiry and is known as **put-call parity**. If this relationship did not hold then there would be riskless arbitrage opportunities.



In Table 2.1 I show the cashflows in the perfectly hedged portfolio. In this table I have set up the cashflows to have a guaranteed value of zero at expiry.

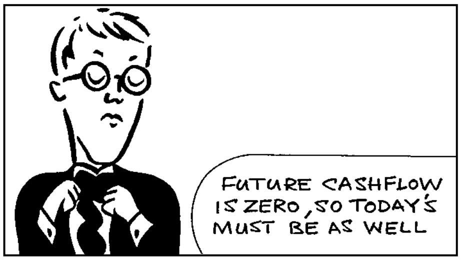


Table 2.1 Cashflows in a hedged portfolio of options and asset.

Holding	Worth today (t)	Worth at expiry (T)
Call	C	$\max(S(T)-E,0)$
$-$Put	$-P$	$-\max(E-S(T),0)$
$-$Stock	$-S(t)$	$-S(T)$
Cash	$Ee^{-r(T-t)}$	E
Total	$C-P-S(t)+Ee^{-r(T-t)}$	0

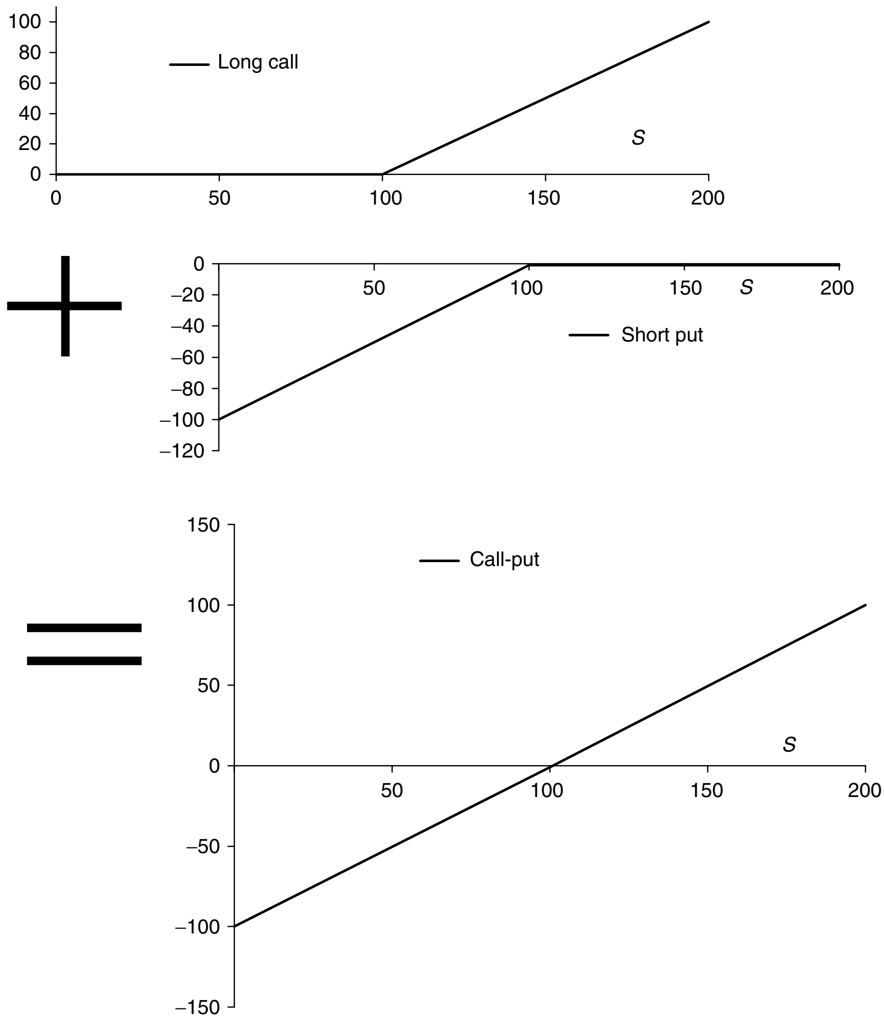


Figure 2.14 Schematic diagram showing put-call parity.

2.13 BINARIES OR DIGITALS

The original and still most common contracts are the vanilla calls and puts. Increasingly important are the **binary** or **digital options**. These contracts have a payoff at expiry that is discontinuous in the underlying asset price. An example of the payoff diagram for one of these options, a **binary call**, is shown in Figure 2.15. This contract pays \$1 at expiry, time T, if the asset price is then greater than the exercise price E. Again, and as with the rest of the figures in this chapter, the bold line is the payoff and the fine line is the contract value some time before expiry.

Why would you invest in a binary call? If you think that the asset price will rise by expiry, to finish above the strike price then you might choose to buy either a vanilla call or a binary call. The vanilla call has the best upside potential, growing linearly with S beyond the strike. The binary call, however, can never pay off more than the \$1. If you expect the underlying will rise dramatically then it may be best to buy the vanilla call. If you believe that the asset rise will be less dramatic then buy the binary call. The gearing of the vanilla call is greater than that for a binary call if the move in the underlying is large.

Figure 2.16 shows the payoff diagram for a **binary put**, the holder of which receives \$1 if the asset is *below* E at expiry. The binary put would be bought by someone expecting a modest fall in the asset price.

There is a particularly simple binary put-call parity relationship. What do you get at expiry if you hold both a binary call and a binary put with the same strikes and expiries? The answer is that you will always get \$1 regardless of the level of the underlying at expiry. Thus

$$\text{Binary call} + \text{Binary put} = e^{-r(T-t)}.$$

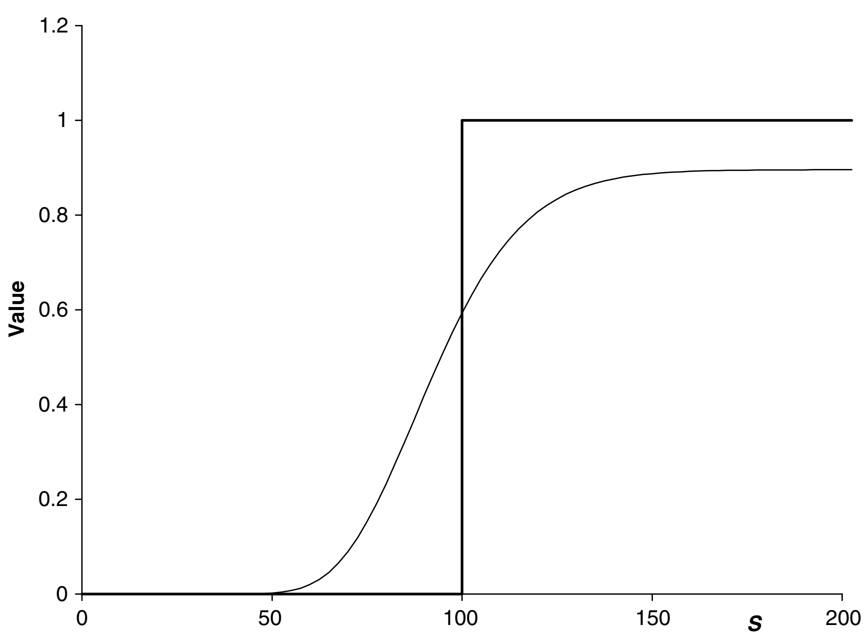


Figure 2.15 Payoff diagram for a binary call option.

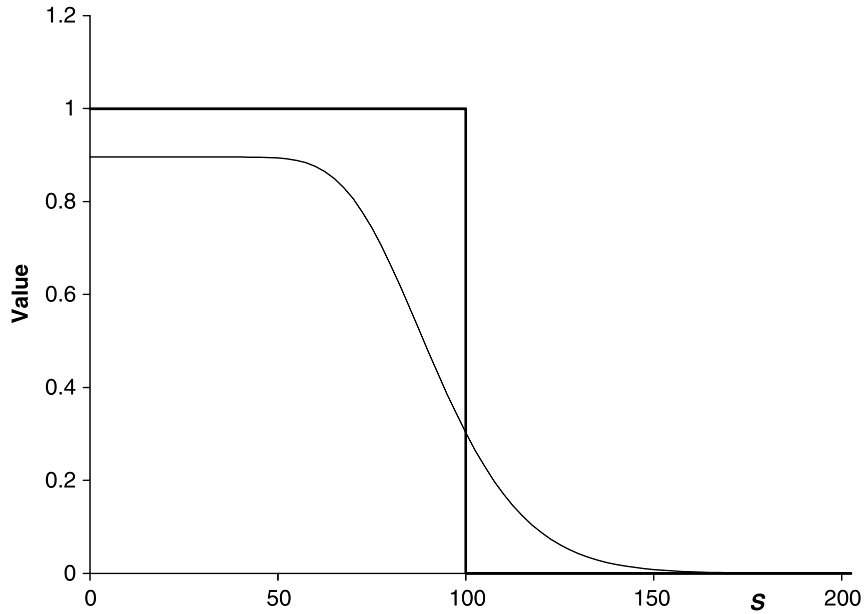


Figure 2.16 Payoff diagram for a binary put option.

What would the table of cashflows look like for the perfectly hedged digital portfolio?

2.14 BULL AND BEAR SPREADS

A payoff that is similar to a binary option can be made up with vanilla calls. This is our first example of a **portfolio of options** or an **option strategy**.

Suppose I buy one call option with a strike of 100 and write another with a strike of 120 and with the same expiration as the first then my resulting portfolio has a payoff that is shown in Figure 2.17. This payoff is zero below 100, 20 above 120 and linear in between. The payoff is continuous, unlike the binary call, but has a payoff that is superficially similar. This strategy is called a **bull spread** or a **call spread**; it benefits from a bull, i.e. rising, market.

The payoff for a general bull spread, made up of calls with strikes E_1 and E_2, is given by

$$\frac{1}{E_2 - E_1}(\max(S - E_1, 0) - \max(S - E_2, 0)),$$

where $E_2 > E_1$. Here I have bought/sold $(E_2 - E_1)^{-1}$ of each of the options so that the maximum payoff is scaled to 1.

If I write a put option with strike 100 and buy a put with strike 120 I get the payoff shown in Figure 2.18. This is called a **bear spread** or a **put spread**, benefitting from a bear, i.e. falling, market. Again, it is very similar to a binary put except that the payoff is continuous.

Because of put-call parity it is possible to build up these payoffs using other contracts.

A strategy involving options of the same type (i.e. calls or puts) is called a **spread**.

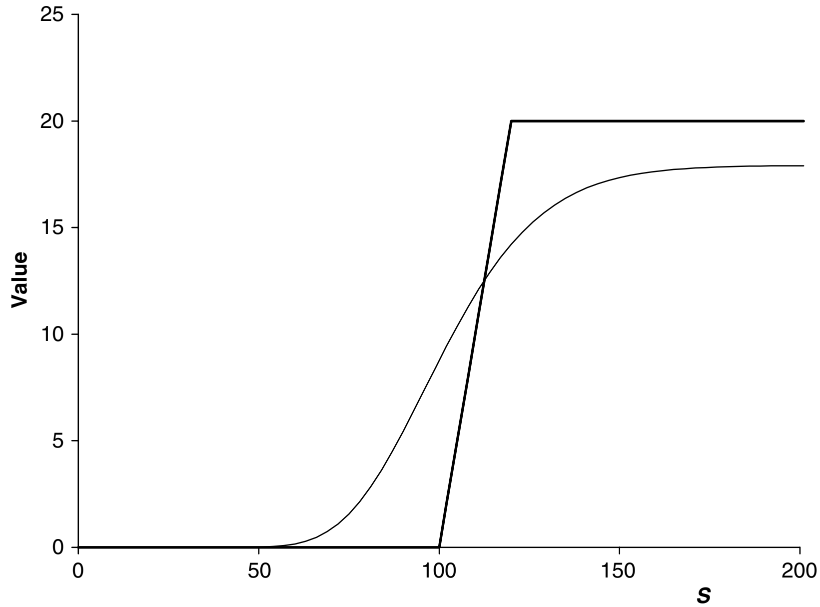


Figure 2.17 Payoff diagram for a bull spread.

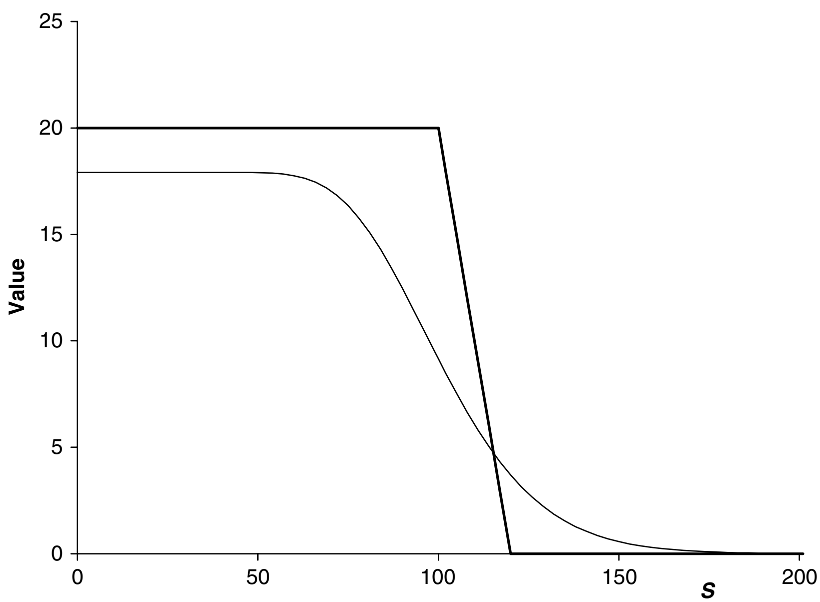


Figure 2.18 Payoff diagram for a bear spread.

2.15 STRADDLES AND STRANGLES

If you have a precise view on the behavior of the underlying asset you may want to be more precise in your choice of option; simple calls, puts, and binaries may be too crude.

The **straddle** consists of a call and a put with the same strike. The payoff diagram is shown in Figure 2.19. Such a position is usually bought at the money by someone who expects the underlying to either rise or fall, but not to remain at the same level. For example, just before an anticipated major news item stocks often show a 'calm before the storm.' On the announcement the stock suddenly moves either up or down depending on whether or not the news was favorable to the company. They may also be bought by technical traders who see the stock at a key support or resistance level and expect the stock to either break through dramatically or bounce back.

The straddle would be sold by someone with the opposite view, someone who expects the underlying price to remain stable.

Figure 2.20 shows the Bloomberg screen for setting up a straddle. Figure 2.21 shows the profit and loss for this position at various times before expiry. The profit/loss is the option value less the upfront premium.

The **strangle** is similar to the straddle except that the strikes of the put and the call are different. The contract can be either an **out-of-the-money strangle** or **an in-the-money strangle**. The payoff for an out-of-the money strangle is shown in Figure 2.22. The motivation behind the purchase of this position is similar to that for the purchase of a straddle. The difference is that the buyer expects an even larger move in the underlying one way or the other. The contract is usually bought when the asset is around the middle of the two strikes and is cheaper than a straddle. This cheapness means that the gearing for the out-of-the-money strangle is higher than that for the straddle. The downside is that there is a much greater range over which the

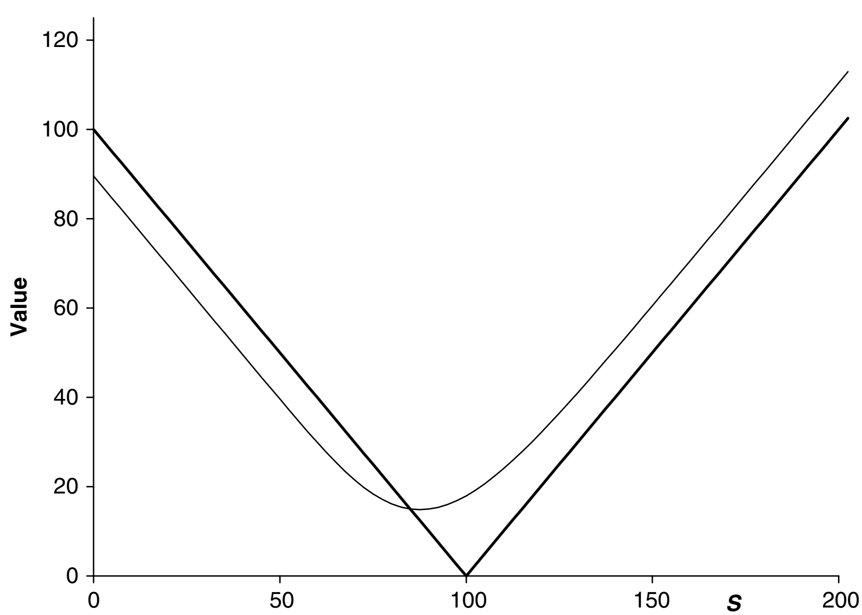


Figure 2.19 Payoff diagram for a straddle.

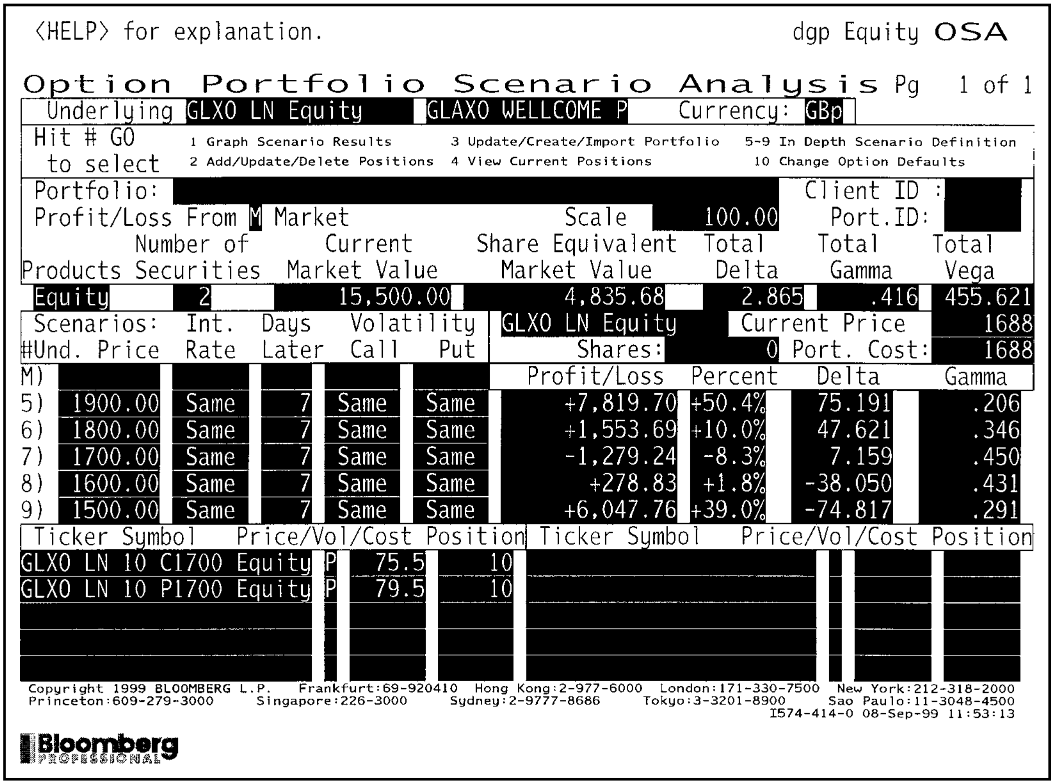


Figure 2.20 A portfolio of two options making up a straddle. Source: Bloomberg L. P.

strangle has no payoff at expiry; for the straddle there is only the one point at which there is no payoff.

There is another reason for a straddle or strangle trade that does not involve a view on the direction of the underlying. These contracts are bought or sold by those with a view on the direction of volatility, they are one of the simplest **volatility trades**. Because of the relationship between the price of an option and the volatility of the asset one can speculate on the direction of volatility. Do you expect the volatility to rise? If so, how can you benefit from this? Until we know more about this relationship, we cannot go into this in more detail.

Straddles and strangles are rarely held until expiry.

A strategy involving options of different types (i.e. both calls and puts) is called a **combination**.

2.16 RISK REVERSAL

The **risk reversal** is a combination of a long call, with strike above the current spot, and a short put with a strike below the current spot. Both have the same expiry. The payoff is shown in Figure 2.23.

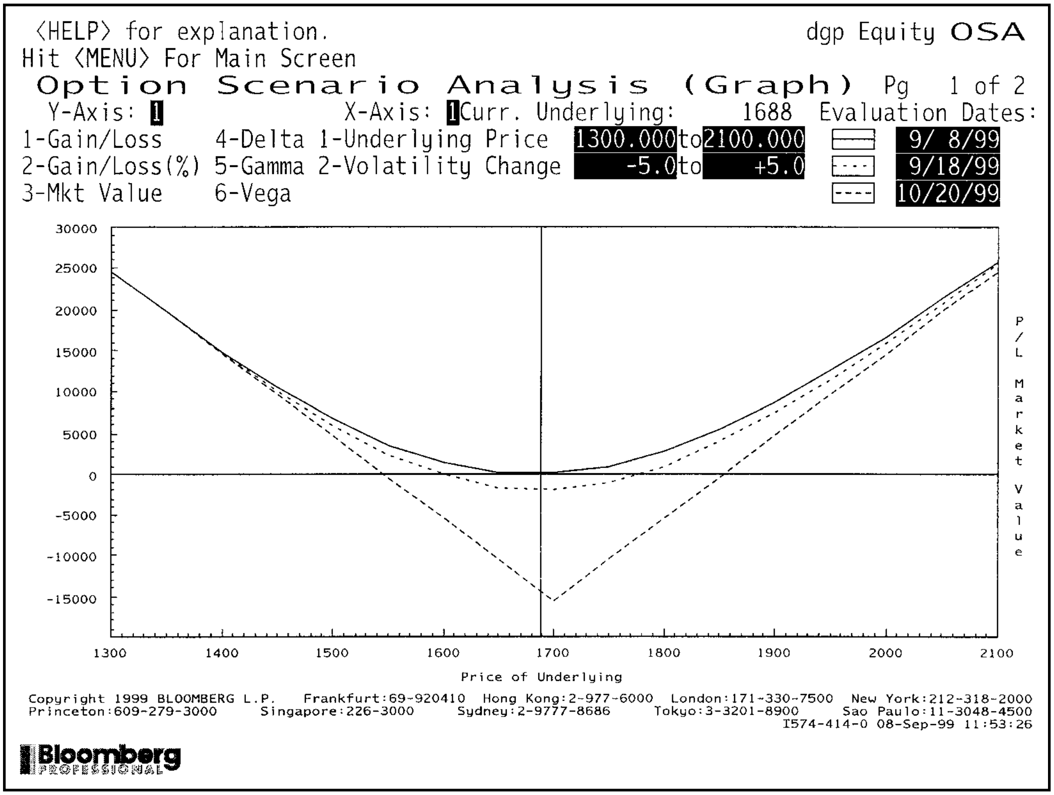


Figure 2.21 Profit/loss for the straddle at several times before expiry. Source: Bloomberg L. P.

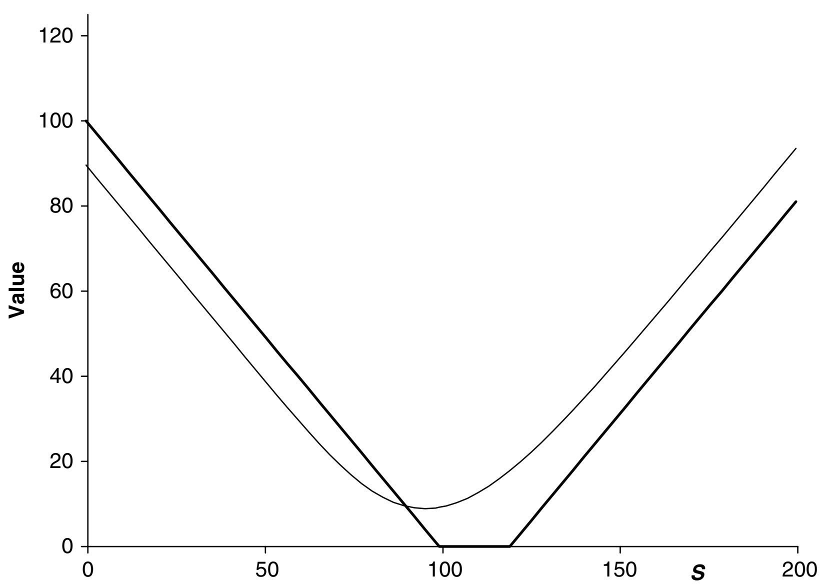


Figure 2.22 Payoff diagram for a strangle.

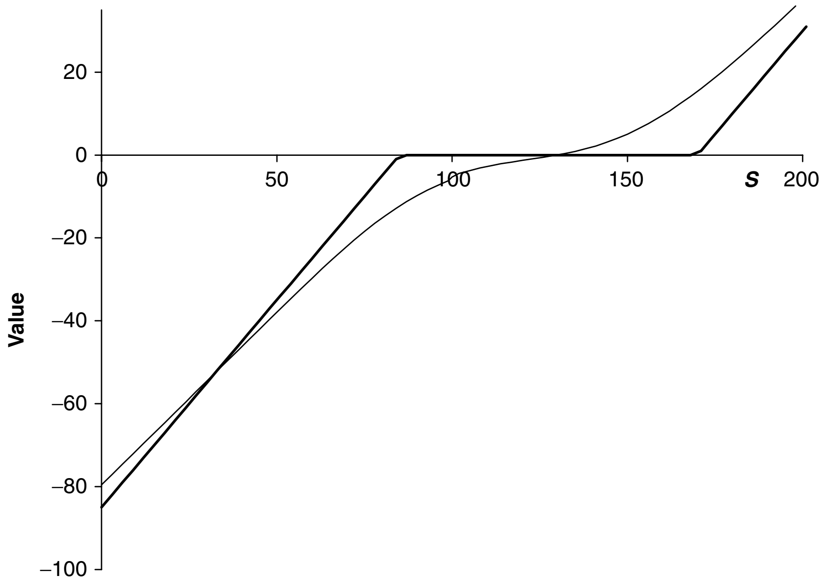


Figure 2.23 Payoff diagram for a risk reversal.

The risk reversal is a very special contract, popular with practitioners. Its value is usually quite small and related to the market's expectations of the behavior of volatility. This is too complex to go into now, but will be explained in Chapter 50.

2.17 BUTTERFLIES AND CONDORS

A more complicated strategy involving the purchase and sale of options with *three* different strikes is a **butterfly spread**. Buying a call with a strike of 90, writing two calls struck at 100 and buying a 110 call gives the payoff in Figure 2.24. This is the kind of position you might enter if you believe that the asset is not going anywhere, either up or down. Because it has no large upside potential (in this case the maximum payoff is 10) the position will be relatively cheap. With options, cheap is good.

The **condor** is like a butterfly except that four strikes, and four call options, are used. The payoff is shown in Figure 2.25.

2.18 CALENDAR SPREADS

All of the strategies I have described above have involved buying or writing calls and puts with different strikes *but all with the same expiration*. A strategy involving options with different expiry dates is called a **calendar spread**. You may enter into such a position if you have a precise view on the timing of a market move as well as the direction of the move. As always the motive behind such a strategy is to reduce the payoff at asset values and times which you believe are irrelevant, while increasing the payoff where you think it will matter. Any reduction in payoff will reduce the overall value of the option position.

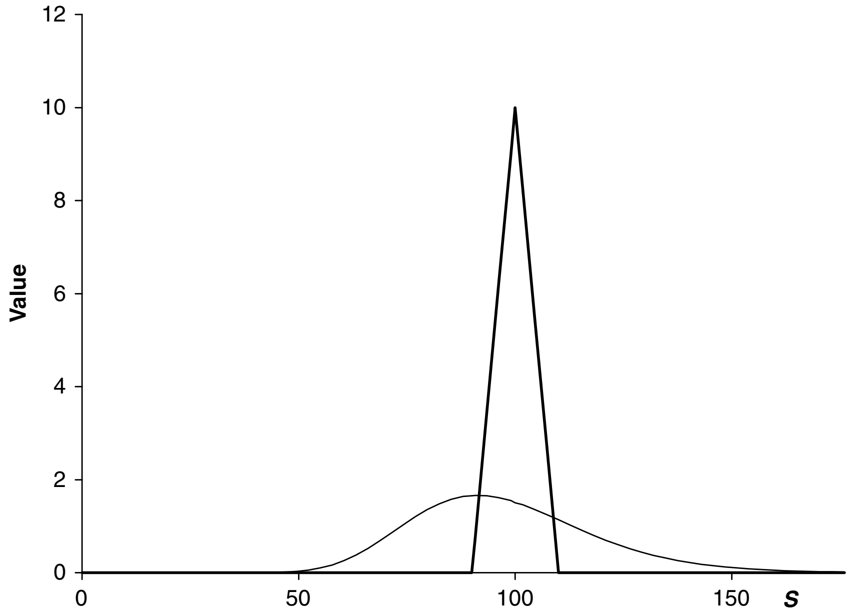


Figure 2.24 Payoff diagram for a butterfly spread.

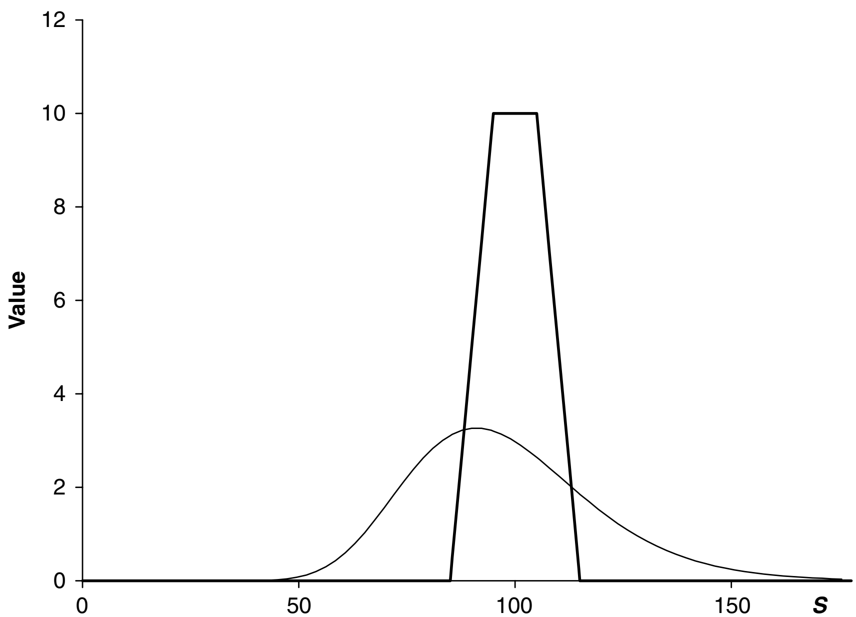


Figure 2.25 Payoff diagram for a condor.

2.19 LEAPS AND FLEX

LEAPS or **Long-term equity anticipation securities** are longer-dated exchange-traded calls and puts. They began trading on the CBOE in the late 1980s. They are standardized so that they are available with expiries up to three years. They come with three strikes, corresponding to at the money and approximately 20% in and out of the money with respect to the underlying asset price when issued.

In 1993 the CBOE created **FLEX** or **FLexible EXchange-traded options** on several indices. These allow a degree of customization, in the expiry date (up to five years), the strike price and the exercise style.

2.20 WARRANTS

A contract that is very similar to an option is a **warrant**. Warrants are call options issued by a company on its own equity. The main differences between traded options and warrants are the timescales involved, warrants usually have a longer lifespan, and on exercise the company issues new stock to the warrant holder. On exercise, the holder of a *traded* option receives stock that has already been issued. Exercise is usually allowed any time before expiry, but after an initial waiting period.

The typical lifespan of a warrant is five or more years. Occasionally **perpetual warrants** are issued; these have no maturity.

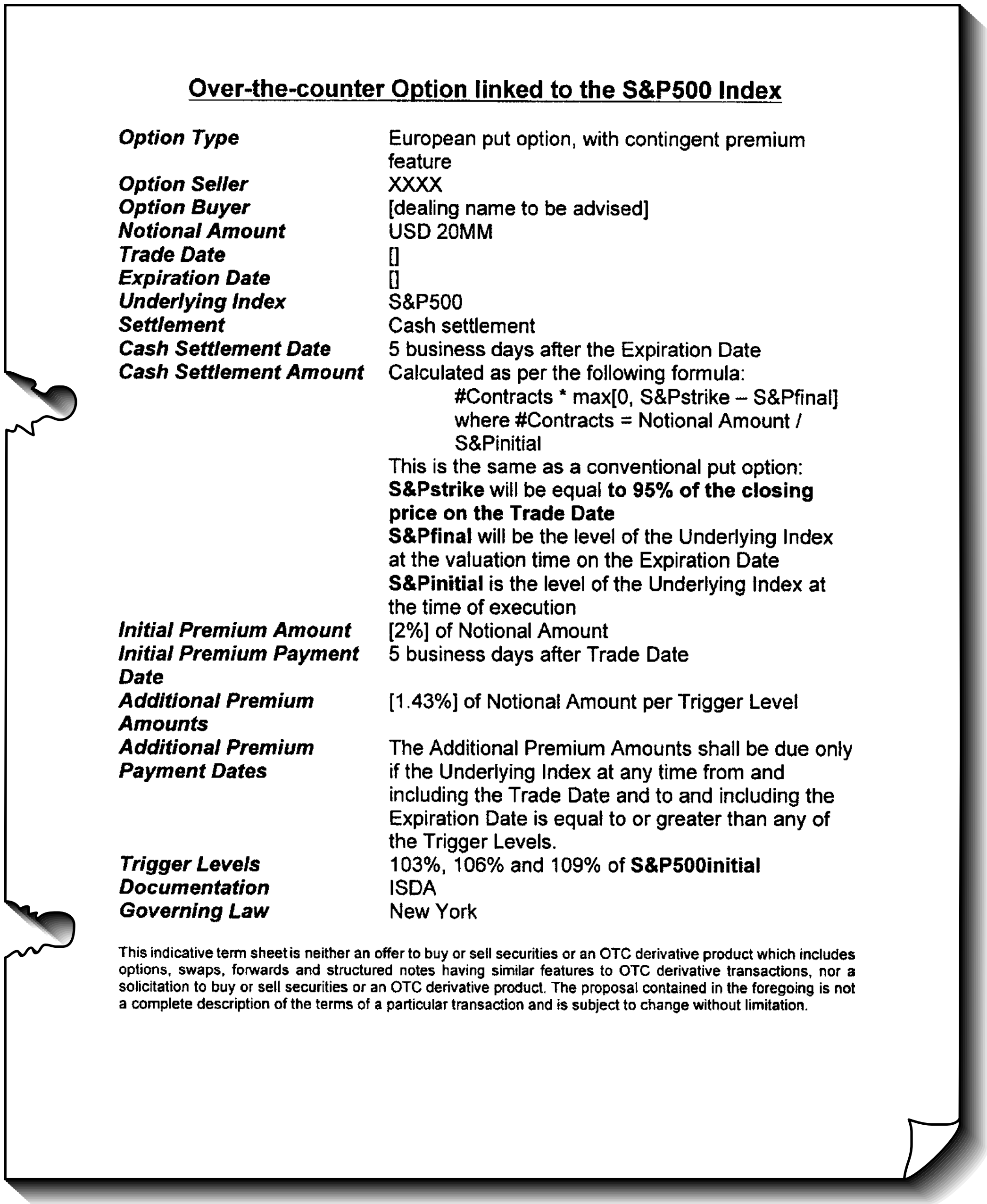
2.21 CONVERTIBLE BONDS

Convertible bonds or **CB**s have features of both bonds and warrants. They pay a stream of coupons with a final repayment of principal at maturity, but they can be converted into the underlying stock before expiry. On conversion rights to future coupons are lost. If the stock price is low then there is little incentive to convert to the stock; the coupon stream is more valuable. In this case the CB behaves like a bond. If the stock price is high then conversion is likely and the CB responds to the movement in the asset. Because the CB can be converted into the asset, its value has to be at least the value of the asset. This makes CBs similar to American options; early exercise and conversion are mathematically the same.

There are other interesting features of convertible bonds, callback, resetting, etc. and the whole of Chapter 33 is devoted to their description and analysis.

2.22 OVER THE COUNTER OPTIONS

Not all options are traded on an exchange. Some, known as over the counter or OTC options are sold privately from one counterparty to another. In Figure 2.26 is the term sheet for an OTC put option, having some special features. A **term sheet** specifies the precise details of an OTC contract. In this OTC put the holder gets a put option on S&P500, but more cheaply than a vanilla put option. This contract is cheap because part of the premium does not have to be paid until and unless the underlying index trades above a specified level. Each time that a new level is reached an extra payment is triggered. This feature means that the contract is not



Over-the-counter Option linked to the S&P500 Index

Option Type	European put option, with contingent premium feature
Option Seller	XXXX
Option Buyer	[dealing name to be advised]
Notional Amount	USD 20MM
Trade Date	[]
Expiration Date	[]
Underlying Index	S&P500
Settlement	Cash settlement
Cash Settlement Date	5 business days after the Expiration Date
Cash Settlement Amount	Calculated as per the following formula: #Contracts * max[0, S&Pstrike – S&Pfinal] where #Contracts = Notional Amount / S&Pinitial This is the same as a conventional put option: **S&Pstrike** will be equal **to 95% of the closing price on the Trade Date** **S&Pfinal** will be the level of the Underlying Index at the valuation time on the Expiration Date **S&Pinitial** is the level of the Underlying Index at the time of execution
Initial Premium Amount	[2%] of Notional Amount
Initial Premium Payment Date	5 business days after Trade Date
Additional Premium Amounts	[1.43%] of Notional Amount per Trigger Level
Additional Premium Payment Dates	The Additional Premium Amounts shall be due only if the Underlying Index at any time from and including the Trade Date and to and including the Expiration Date is equal to or greater than any of the Trigger Levels.
Trigger Levels	103%, 106% and 109% of **S&P500initial**
Documentation	ISDA
Governing Law	New York

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 2.26 Term sheet for an OTC 'Put.'

vanilla, and makes the pricing more complicated. We will be discussing special features like the ones in this contract in later chapters.

All of the term sheets in this book are real, in the sense that someone at a bank wrote them for possible commercial purposes. However, many of them were given to me before they were finalized. For that reason you will see that often there are bits 'missing,' and these would have been set at the time that the deal was finally struck.

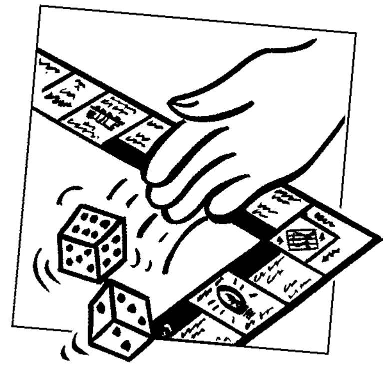
2.23 SUMMARY

We now know the basics of options and markets, and a few of the simplest trading strategies. We know some of the jargon and the reasons why people might want to buy an option. We've also seen another example of no arbitrage in put-call parity. This is just the beginning. We don't know how much these instruments are worth, how they are affected by the price of the underlying, how much risk is involved in the buying or writing of options. And we have only seen the very simplest of contracts, there are many, many more complex products to examine. All of these issues are going to be addressed in later chapters.

FURTHER READING

- McMillan (1996) and Options Institute (1995) describe many option strategies used in practice.
- Most exchanges have websites. The London International Financial Futures Exchange website contains information about the money markets, bonds, equities, indices and commodities. See `www.liffe.com`. For information about options and derivatives generally, see `www.cboe.com`, the Chicago Board Options Exchange website. The American Stock Exchange is on `www.amex.com` and the New York Stock Exchange on `www.nyse.com`.
- Derivatives have often had bad press (and there's probably more to come). See Miller (1997) for a discussion of the pros and cons of derivatives.
- The best books on options are Hull (2005) and Cox & Rubinstein (1985), modesty forbids me from mentioning others.

CHAPTER 3
the random behavior of assets



In this Chapter. . .

- Jensen's inequality
- more notation commonly used in mathematical finance
- how to examine time-series data to model returns
- the Wiener process, a mathematical model of randomness
- a simple model for equities, currencies, commodities and indices

3.1 INTRODUCTION

In this chapter I describe a simple continuous-time model for equities and other financial instruments, inspired by our earlier coin-tossing experiment. This takes us into the world of stochastic calculus and Wiener processes. Although there is a great deal of theory behind the ideas I describe, I am going to explain everything in as simple and accessible manner as possible. We will be modeling the behavior of equities, currencies and commodities, but the ideas are applicable to the fixed-income world as we shall see in Part Three.

3.2 THE POPULAR FORMS OF 'ANALYSIS'

There are three forms of 'analysis' commonly used in the financial world:

- Fundamental
- Technical
- Quantitative

Fundamental analysis is all about trying to determine the 'correct' worth of a company by an in-depth study of balance sheets, management teams, patent applications, competitors, lawsuits, etc. In other words, getting to the heart of the firm, doing lots of accounting and projections and what-not. This sounds like a really sensible way to model a company and hence its stock price. Well, it is and we will talk about this later in the subject of Real Options.

However, there are unfortunately two difficulties with this approach. First it is very, very hard. You need a degree in accounting and plenty of patience. And even then all the most important stuff can be hidden 'off balance sheet.' Second, and more importantly, 'The market can stay irrational longer than you can stay solvent' (possibly said by Keynes). In other words, even if you have the perfect model for the value of a firm it doesn't mean you can make money. You have to find some mispricing and then hope that the rest of the world starts to see your point of view. And this may never happen. If fundamental analysis is hard, then the next form of analysis is the exact opposite, because it is so easy.

Technical analysis is when you don't care anything about the company other than the information contained within its stock price history. You draw trendlines, look for specific patterns in the share price and make predictions accordingly. This is the subject of Chapter 20. Most academic evidence suggests that most technical analysis is bunk.

The final form of analysis is the one we are really concerned with in this book, and is the form that has been most successful over the last 50 years, forming a solid foundation for portfolio theory, derivatives pricing and risk management. It is **quantitative analysis**. Quantitative analysis is all about treating financial quantities such as stock prices or interest rates as random, and then choosing the best models for that randomness. Let's see why randomness is important and then build up a simple, random, stock price model.

3.3 WHY WE NEED A MODEL FOR RANDOMNESS: JENSEN'S INEQUALITY

Why is 'randomness' so crucial to modeling the world of derivatives? Why can't we just try to forecast the future stock price as best we can and figure out the option's payoff? To best see the importance of randomness in option theory let's take a look at some very simple mathematics, called **Jensen's Inequality**.

The stock price today is \$100. Let's suppose that in one year's time it could be \$50 or \$150, with both equally likely (see Figure 3.1). How can we value an option on this stock, a call option with a strike of 100 expiring in one year, say?

Two ways spring to mind.

With those two possible scenarios we could say that we expect the stock price to be at \$100 in one year, this being the average of the possible future values. The payoff for the call option would then be 0, since it is exactly at the money. And the present value of this is zero. Could this be the way to value an option?

Probably not. You expect the value to be greater than zero, since half the time there is some payoff.

Alternatively we could look at the two possible payoffs and then calculate that expectation. If the stock falls to 50 then the payoff is zero, if it rises to 150 then the payoff is 50. The average payoff is therefore 25, which we could present value to give us some idea of the option's value.

It turns out that the second calculation is closer to what we do in practice to value options (although it turns out that the probabilities don't come into the calculation) and we'll see lots of this throughout the book. But that calculation also illustrates another point of great importance, that the order in which we do the payoff calculation and the expectation matters. In this example we had

$$\text{Payoff (Expected stock price)} = 0$$

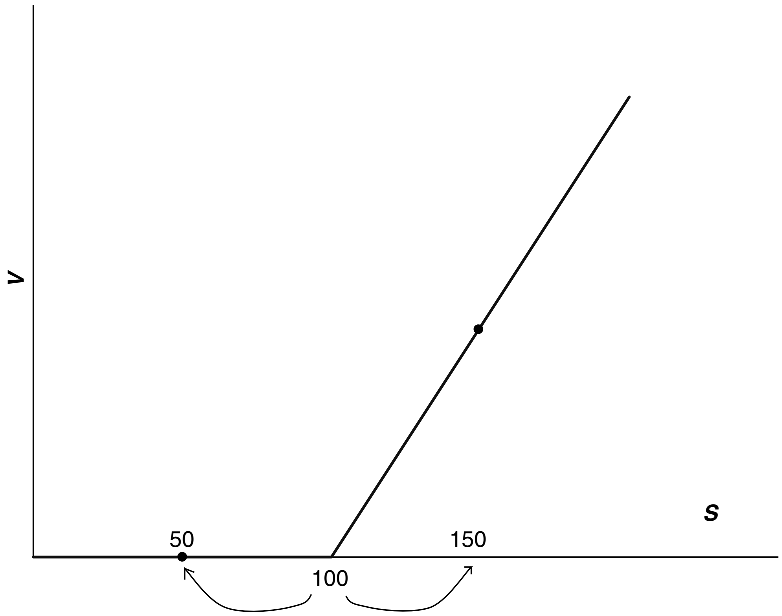


Figure 3.1 Future scenarios.

whereas

$$\text{Expected}\left[\text{Payoff(Stock price)}\right] = 25.$$

This is an example of Jensen's inequality. Let's use some symbols. If we have a convex function $f(S)$ (in our example the payoff function for a call) of a random variable S (in our example the stock price) then

$$E\left[f(S)\right] \geq f(E[S]). \tag{3.1}$$

We can even get an idea of how much greater the left-hand side is than the right-hand side by using a Taylor series approximation around the mean of S. Write

$$S = \overline{S} + \epsilon,$$

where $\overline{S} = E[S]$, so the $E[\epsilon] = 0$. Then

$$\begin{aligned} E\left[f(S)\right] = E\left[f(\overline{S}+\epsilon)\right] &= E\left[f(\overline{S}) + \epsilon f'(\overline{S}) + \tfrac{1}{2}\epsilon^2 f''(\overline{S}) + \cdots\right] \\ &\approx f(\overline{S}) + \tfrac{1}{2} f''(\overline{S}) E\left[\epsilon^2\right] \\ &= f(E[S]) + \tfrac{1}{2} f''(E[S]) E[\epsilon^2]. \end{aligned}$$

So the left-hand side of (3.1) is greater than the right by approximately

$$\tfrac{1}{2} f''(E[S])\ E[\epsilon^2].$$

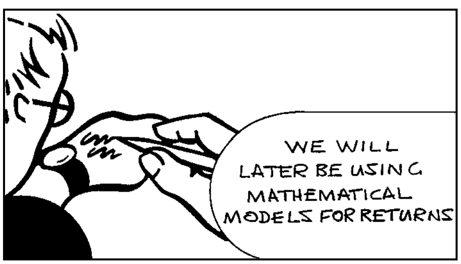
This shows the importance of two concepts:

- $f''(E[S])$: The **convexity** of an option. As a rule this adds value to an option. It also means that any intuition we may get from linear contracts (forwards and futures) might not be helpful with non-linear instruments such as options.
- $E\left[\epsilon^2\right]$: Randomness in the underlying, and its **variance**. As stated above, modeling randomness is the key to modeling options.

Now that we have seen a hint as to why randomness is so important, let's start modeling some assets!

3.4 SIMILARITIES BETWEEN EQUITIES, CURRENCIES, COMMODITIES AND INDICES

When you invest in something, whether it is a stock, commodity, work of art or a racehorse, your main concern is that you will make a comfortable return on your investment. By **return** we tend to mean the percentage growth in the value of an asset, together with accumulated dividends, over some period:



$$\text{Return} = \frac{\text{Change in value of the asset} \atop + \text{accumulated cashflows}}{\text{Original value of the asset}}.$$

I want to distinguish here between the percentage or relative growth and the absolute growth. Suppose we could invest in either of two stocks, both of which grow on average by \$10 *per annum*. Stock A has a value of \$100 and stock B is currently worth \$1000. Clearly the former is a better investment, at the end of the year stock A will probably be worth around \$110 (if the past is anything to go by) and stock B \$1010. Both have gone up by \$10, but A has risen by 10% and B by only 1%. If we have \$1000 to invest we would be better off investing in ten of asset A than one of asset B. This illustrates that when we come to model assets, it is the return that we should concentrate on. In this respect, all equities, currencies, commodities and stock market indices can be treated similarly. What return do we expect to get from them?

Part of the business of estimating returns for each asset is to estimate how much unpredictability there is in the asset value. In the next section I am going to show that randomness plays a large part in financial markets, and start to build up a model for asset returns incorporating this randomness.

3.5 EXAMINING RETURNS

In Figure 3.2 I show the quoted price of Perez Companc, an Argentinian conglomerate, over the period February 1995 to November 1996. This is a very typical plot of a financial asset. The asset shows a general upward trend over the period but this is far from guaranteed. If you bought and sold at the wrong times you would lose a lot of money. The unpredictability that is seen in this figure is the main feature of financial modeling. Because there is so much randomness, any mathematical model of a financial asset must acknowledge the randomness and have a probabilistic foundation.

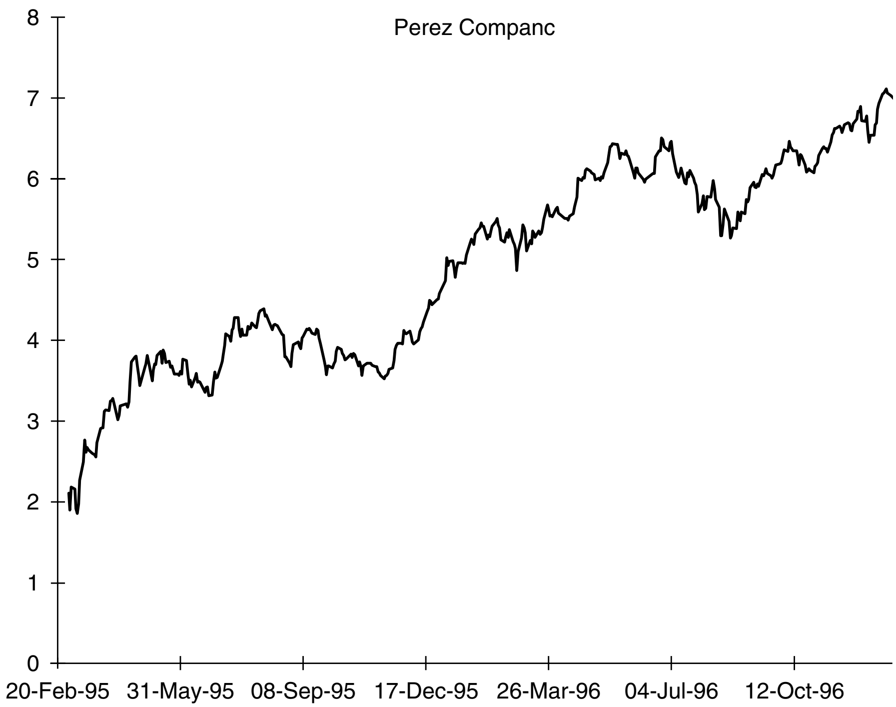


Figure 3.2 Perez Companc from February 1995 to November 1996.

Remembering that the returns are more important to us than the absolute level of the asset price, I show in Figure 3.3 how to calculate returns on a spreadsheet. Denoting the asset value on the ith day by S_i, then the return from day i to day $i+1$ is given by

$$\frac{S_{i+1} - S_i}{S_i} = R_i.$$

(I've ignored dividends here, they are easily allowed for, especially since they only get paid two or four times a year typically.) Of course, I didn't need to use data spaced at intervals of a day, I will comment on this later.

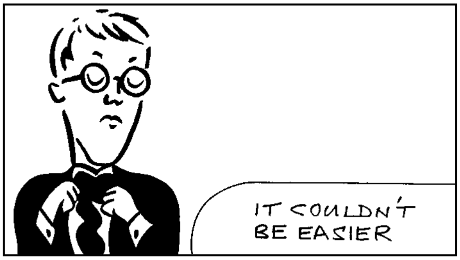
In Figure 3.4 I show the daily returns for Perez Companc. This looks very much like 'noise,' and that is exactly how we are going to model it.

The mean of the returns distribution is

$$\overline{R} = \frac{1}{M}\sum_{i=1}^{M} R_i \tag{3.2}$$

and the sample standard deviation is

$$\sqrt{\frac{1}{M-1}\sum_{i=1}^{M}(R_i - \overline{R})^2}, \tag{3.3}$$



Date	Perez	Return		
01-Mar-95	2.11		**Average return**	0.002916
02-Mar-95	1.90	−0.1	**Standard deviation**	0.024521
03-Mar-95	2.18	0.149906		
06-Mar-95	2.16	−0.01081		
07-Mar-95	1.91	−0.11258	=AVERAGE(C3:C463)	
08-Mar-95	1.86	−0.02985		
09-Mar-95	1.97	0.061538		
10-Mar-95	2.27	0.15		=STDEVP(C3:C463)
13-Mar-95	2.49	0.099874		
14-Mar-95	2.76	0.108565		
15-Mar-95	2.61	−0.05426		
16-Mar-95	2.67	0.021858		
17-Mar-95	2.64	−0.0107		
20-Mar-95	2.60	−0.01622	=(B13-B12)/B12	
21-Mar-95	2.59	−0.00275		
22-Mar-95	2.59	−0.00275		
23-Mar-95	2.55	−0.01232		
24-Mar-95	2.73	0.069307		
27-Mar-95	2.91	0.064815		
28-Mar-95	2.92	0.002899		
29-Mar-95	2.92	0		
30-Mar-95	3.12	0.069364		
31-Mar-95	3.14	0.005405		
03-Apr-95	3.13	−0.00269		
04-Apr-95	3.24	0.037736		
05-Apr-95	3.25	0.002597		
06-Apr-95	3.28	0.007772		
07-Apr-95	3.21	−0.02057		
10-Apr-95	3.02	−0.06037		
11-Apr-95	3.08	0.019553		
12-Apr-95	3.19	0.035616		
17-Apr-95	3.21	0.007936		
18-Apr-95	3.17	−0.01312		
19-Apr-95	3.24	0.021277		

Figure 3.3 Spreadsheet for calculating asset returns.

where M is the number of returns in the sample (one fewer than the number of asset prices). From the data in this example we find that the mean is 0.002916 and the standard deviation is 0.024521.

Notice how the mean daily return is much smaller than the standard deviation. This is very typical of financial quantities over short timescales. On a day-by-day basis you will tend to see the noise in the stock price, and will have to wait months perhaps before you can spot the trend.

The frequency distribution of this time series of daily returns is easily calculated, and very instructive to plot. In Excel use Tools | Data Analysis | Histogram. In Figure 3.5 is shown the frequency distribution of daily returns for Perez Companc. This distribution has been scaled and translated to give it a mean of zero, a standard deviation of one and an area under the curve of one. On the same plot is drawn the probability density function for the standardized Normal distribution function

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\phi^2},$$

where ϕ is a standardized Normal variable. The two curves are not identical but are fairly close.

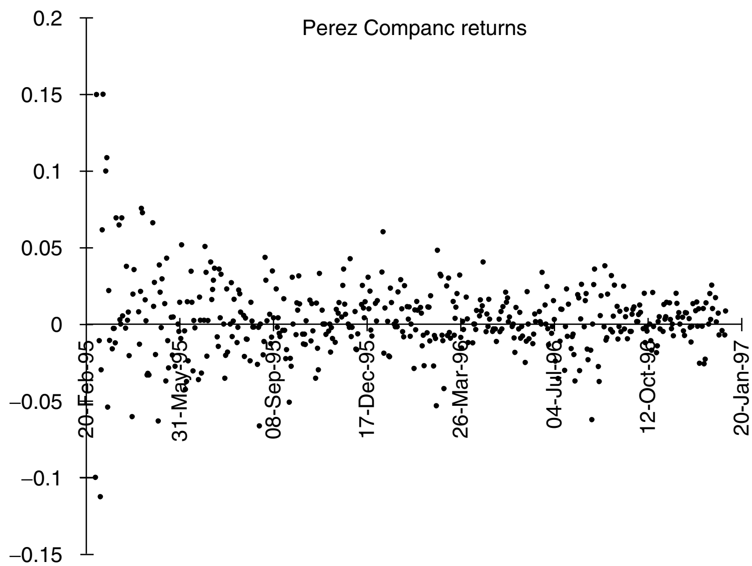


Figure 3.4 Daily returns of Perez Companc.

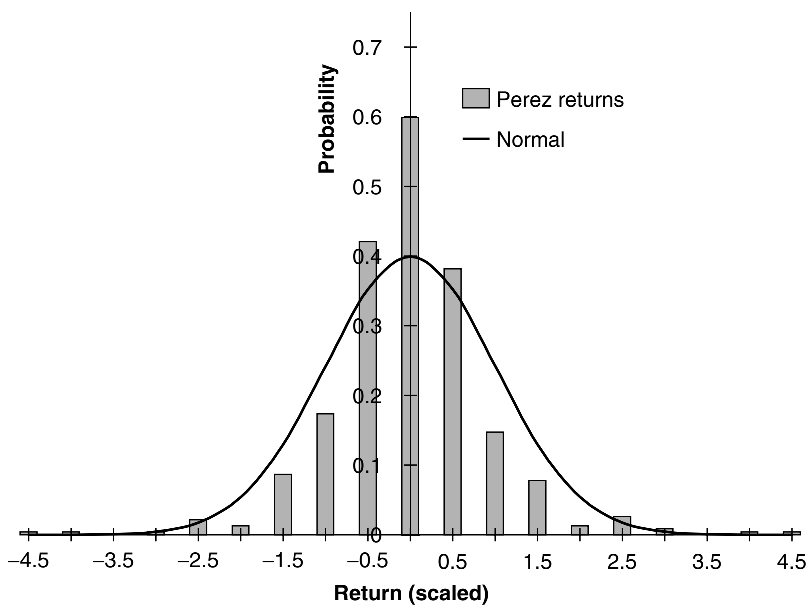


Figure 3.5 Normalized frequency distribution of Perez Companc and the standardized Normal distribution.

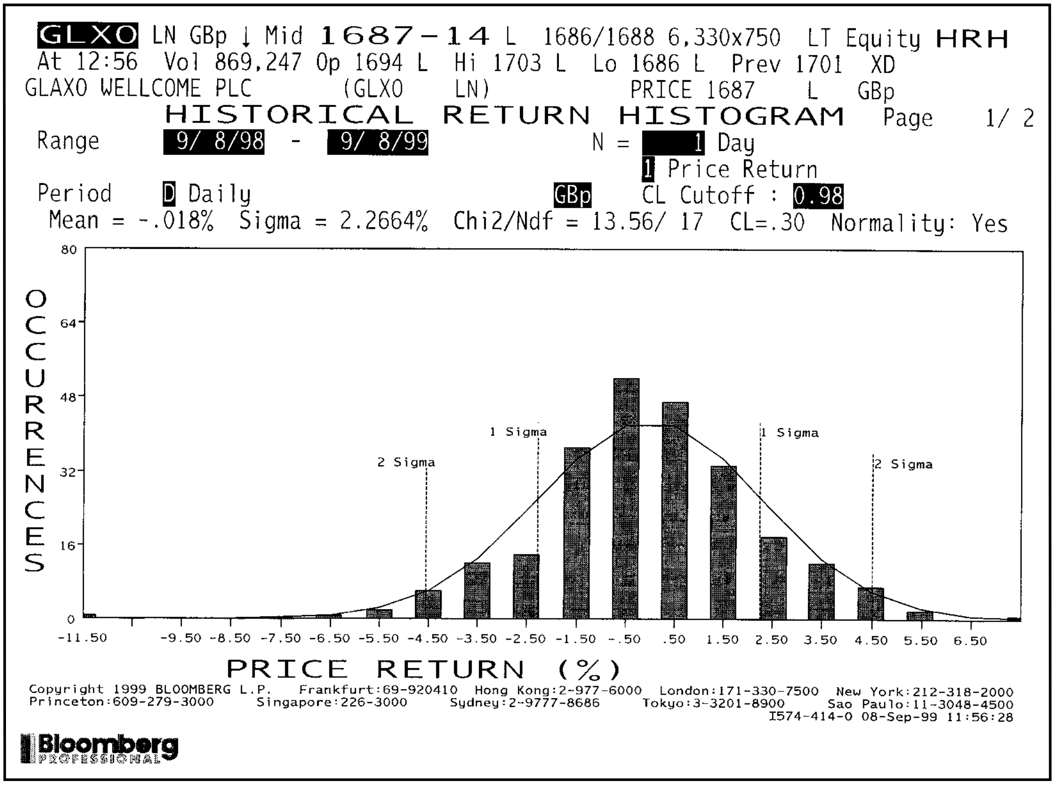


Figure 3.6 Glaxo–Wellcome returns histogram. Source: Bloomberg L.P.

Supposing that we believe that the empirical returns are close enough to Normal for this to be a good approximation, then we have come a long way towards a model. I am going to write the returns as a random variable, drawn from a Normal distribution with a known, constant, non-zero mean and a known, constant, non-zero standard deviation:

$$R_i = \frac{S_{i+1} - S_i}{S_i} = \text{mean} + \text{standard deviation} \times \phi.$$

Figure 3.6 shows the returns distribution of Glaxo–Wellcome as calculated by Bloomberg. This has not been normalized.

3.6 TIMESCALES

How do the mean and standard deviation of the returns' time series, as estimated by (3.2) and (3.3), scale with the time step between asset price measurements? In the example the time step is one day, but suppose I sampled at hourly intervals or weekly, how would this affect the distribution?

Call the time step δt. The mean of the return scales with the size of the time step. That is, the larger the time between sampling the more the asset will have moved in the meantime, *on average*. I can write

$$\text{mean} = \mu\ \delta t,$$

for some μ which we will assume to be constant. In the Perez Companc example we had a mean of 0.002916 over a timescale of one day, $\delta t = 1/252$ so that

$$\mu = 252 \times 002916 = 0.735 = 73.5\%.$$

Ignoring randomness for the moment, our model is simply

$$\frac{S_{i+1} - S_i}{S_i} = \mu\ \delta t.$$

Rearranging, we get

$$S_{i+1} = S_i(1 + \mu\ \delta t).$$

If the asset begins at S_0 at time $t = 0$ then after one time step $t = \delta t$ and

$$S_1 = S_0(1 + \mu\,\delta t).$$

After two time steps $t = 2\ \delta t$ and

$$S_2 = S_1(1 + \mu\ \delta t) = S_0(1 + \mu\ \delta t)^2,$$

and after M time steps $t = M\ \delta t = T$ and

$$S_M = S_0(1 + \mu\ \delta t)^M.$$

This is just

$$S_M = S_0\,(1 + \mu\ \delta t)^M = S_0 e^{M \log(1+\mu\ \delta t)} \approx S_0 e^{\mu M\ \delta t} = S_0 e^{\mu T}.$$

In the limit as the time step tends to zero with the total time T fixed, this approximation becomes exact. This result is important for two reasons.

First, in the absence of any randomness the asset exhibits exponential growth, just like cash in the bank.

Second, the model is meaningful in the limit as the time step tends to zero. If I had chosen to scale the mean of the returns distribution with any other power of δt it would have resulted in either a trivial model ($S_T = S_0$) or infinite values for the asset ($S_T = \pm\infty$).

The second point can guide us in the choice of scaling for the random component of the return. How does the standard deviation of the return scale with the time step δt? (Recall that you add *variances* not standard deviations.) Again, consider what happens after $T/\delta t$ time steps each of size δt (i.e. after a total time of T). Inside the square root in expression (3.3) there are a large number of terms, $T/\delta t$ of them. In order for the standard deviation to remain finite as we let δt tend to zero, the individual terms in the expression must each be of $O(\delta t)$. Since each term is a square of a return, the standard deviation of the asset return over a time step δt must be $O(\delta t^{1/2})$:

$$\text{standard deviation} = \sigma\ \delta t^{1/2},$$

where σ is some parameter measuring the amount of randomness; the larger this parameter the more uncertain is the return. For the moment let's assume that it is constant. For Perez Companc we have a standard deviation of 0.024521 over one day so that

$$\sigma = \sqrt{252} \times 0.024521 = 0.389 = 38.9\%.$$

Putting these scalings explicitly into our asset return model

$$R_i = \frac{S_{i+1} - S_i}{S_i} = \mu\, \delta t + \sigma \phi\ \delta t^{1/2}. \tag{3.4}$$

I can rewrite Equation (3.4) as

$$S_{i+1} - S_i = \mu S_i\, \delta t + \sigma S_i \phi\ \delta t^{1/2}. \tag{3.5}$$

The left-hand side of this equation is the change in the asset price from time step i to time step $i + 1$. The right-hand side is the 'model.' We can think of this equation as a model for a **random walk** of the asset price. This is shown schematically in Figure 3.7. We know exactly where the asset price is today but tomorrow's value is unknown. It is distributed about today's value according to (3.5).

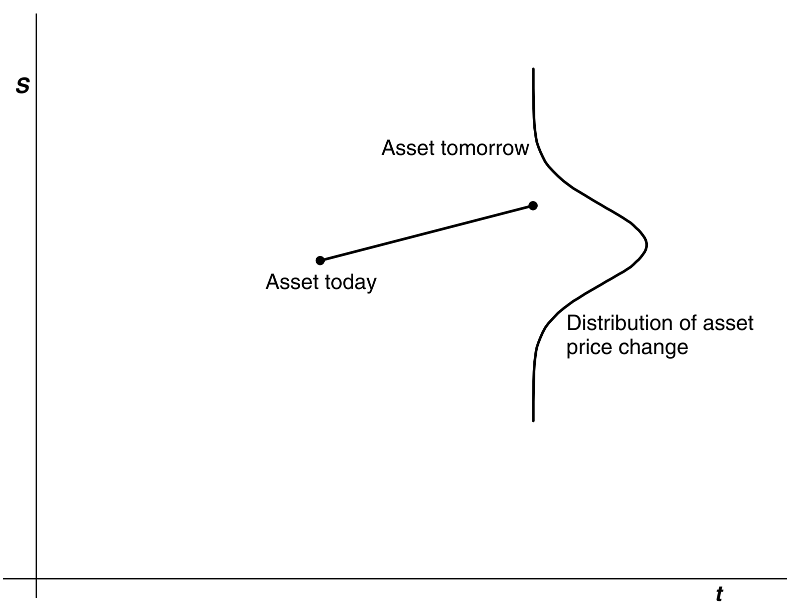


Figure 3.7 A representation of the random walk.

3.6.1 The Drift

The parameter μ is called the **drift rate**, the **expected return** or the **growth rate** of the asset. Statistically it is very hard to measure since the mean scales with the usually small parameter δt. It can be estimated by

$$\mu = \frac{1}{M\,\delta t}\sum_{i=1}^{M} R_i.$$

The unit of time that is usually used is the year, in which case μ is quoted as an *annualized* growth rate.

In the classical option pricing theory the drift plays almost no role. So even though it is hard to measure, this doesn't matter too much.[1]

3.6.2 The Volatility

The parameter σ is called the **volatility** of the asset. It can be estimated by

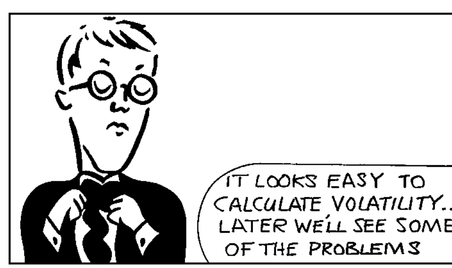
$$\sqrt{\frac{1}{(M-1)\,\delta t}\sum_{i=1}^{M}(R_i - \overline{R})^2}.$$

Again, this is almost always quoted in annualized terms.

The volatility is the most important and elusive quantity in the theory of derivatives. I will come back again and again to its estimation and modeling.

Because of their scaling with time, the drift and volatility have different effects on the asset path. The drift is not apparent over short timescales for which the volatility dominates. Over long timescales, for instance decades, the drift becomes important. Figure 3.8 is a realized path of the logarithm of an asset, together with its expected path and a 'confidence interval.' In this example the confidence interval represents one standard deviation. With the assumption of Normality this means that 68% of the time the asset should be within this range. The mean path is growing linearly in time and the confidence interval grows like the square root of time. Thus over short timescales the volatility dominates.[2]

3.7 ESTIMATING VOLATILITY



The most common estimate of volatility is simply

$$\sqrt{\frac{1}{(M-1)\,\delta t}\sum_{i=1}^{M}(R_i - \overline{R})^2}.$$

If δt is sufficiently small the mean return $\overline{R}$ term can be ignored. For small δt

$$\sqrt{\frac{1}{(M-1)\,\delta t}\sum_{i=1}^{M}(\log S(t_i) - \log S(t_{i-1}))^2}$$

can also be used, where $S(t_i)$ is the closing price on day t_i.

[1] In non-classical theories and in portfolio management, it *does* often matter, very much.

[2] Why did I take the logarithm? Because changes in the logarithm are related to the return on the asset.

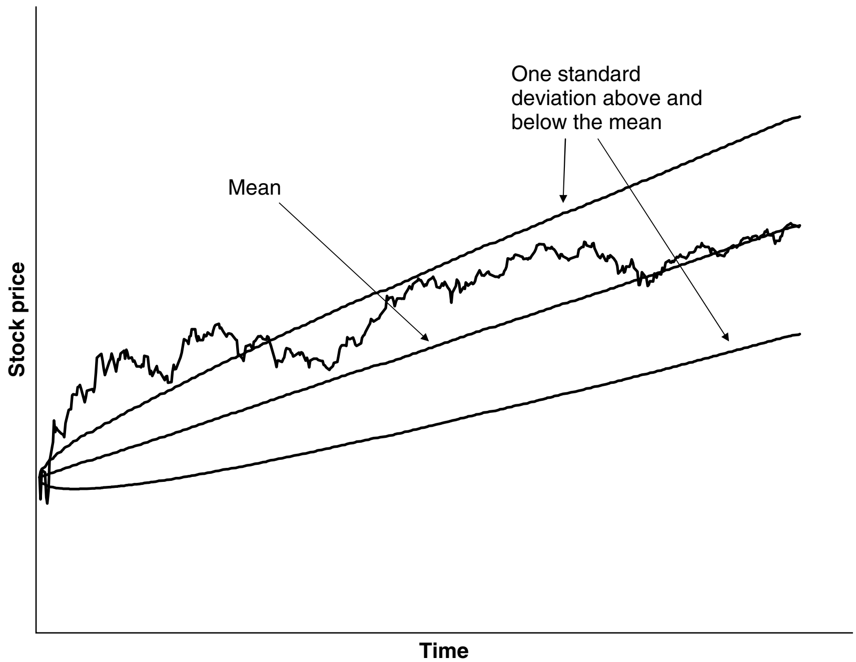


Figure 3.8 Path of the logarithm of an asset, its expected path and one standard deviation above and below.

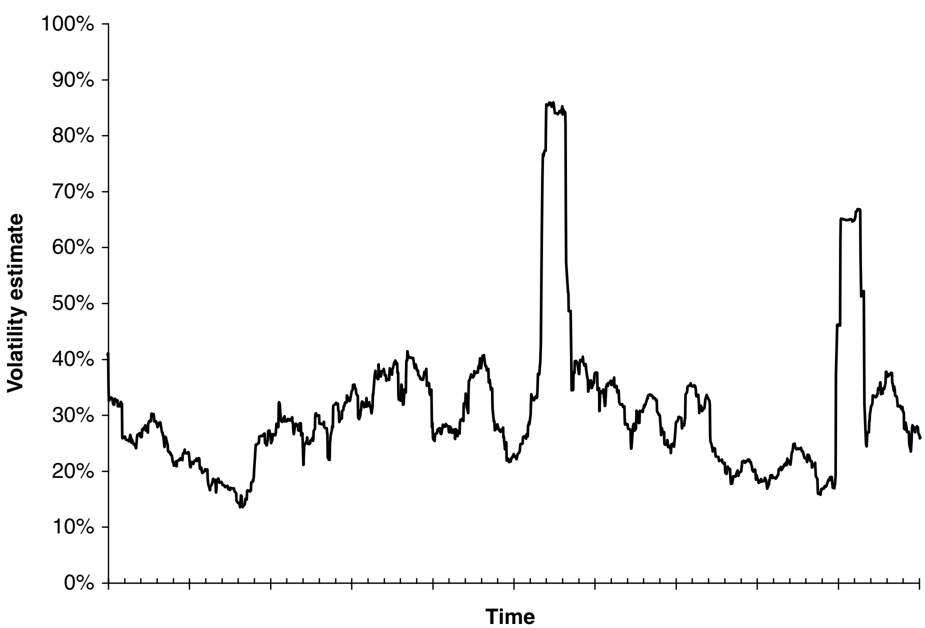


Figure 3.9 The plateauing effect when using a moving window volatility estimate.

It is highly unlikely that volatility is constant in time. Changing economic circumstances, seasonality etc. will inevitably result in volatility changing with time. If you want to know the volatility today you must use some past data in the calculation. Unfortunately, this means that there is no guarantee that you are actually calculating *today's* volatility.

Typically you would use daily closing prices to work out daily returns and then use the past $10, 30, 100, \ldots$ daily returns in the formula above. Or you could use returns over longer or shorter periods. Since all returns are equally weighted, while they are in the estimate of volatility, any large return will stay in the estimate of volatility until the 10 (or 30 or 100) days have passed. This gives rise to a plateauing of volatility, and is totally spurious.

In Figure 3.9 is shown the spurious plateauing effect associated with a sudden large drop in a stock price.

Since volatility is not directly observable, and because of the plateauing effect in the simple measure of volatility, you might want to use other volatility estimates. We'll see some more in Chapter 49.

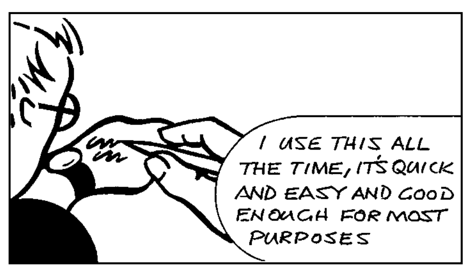
3.8 THE RANDOM WALK ON A SPREADSHEET

The random walk (3.5) can be written as a 'recipe' for generating S_{i+1} from S_i:

$$S_{i+1} = S_i \left(1 + \mu\, \delta t + \sigma \phi\, \delta t^{1/2}\right). \tag{3.6}$$

We can easily simulate the model using a spreadsheet. In this simulation we must input several parameters, a starting value for the asset, a time step δt, the drift rate μ, the volatility σ and the total number of time steps. Then, at each time step, we must choose a random number ϕ from a Normal distribution. I will talk about simulations in depth in Chapter 80dir , for the moment let me just say that an approximation to a Normal variable that is fast in a spreadsheet, and quite accurate, is simply to add up twelve random variables drawn from a uniform distribution over zero to one, and subtract six:

$$\left(\sum_{i=1}^{12} \mathtt{RAND()}\right) - 6.$$



The Excel spreadsheet function `RAND()` gives a uniformly-distributed random variable.

In Figure 3.10 I show the details of a spreadsheet used for simulating the asset price random walk.

3.9 THE WIENER PROCESS

So far we have a model that allows the asset to take any value after a time step. This is some progress but we have still not reached our goal of continuous time, we still have a discrete time step. This section is a brief introduction to the continuous-time limit of equations like (3.4). I will start to introduce ideas from the world of stochastic modeling and Wiener processes, delving more deeply in Chapter 4.

I am now going to use the notation $d\cdot$ to mean 'the change in' some quantity. Thus dS is the 'change in the asset price.' But this change will be in *continuous time*. Thus we will go to

	A	B	C	D	E	F	G	H
1	**Asset**	**100**		**Time**	**Asset**			
2	**Drift**	**0.15**		0	100			
3	**Volatility**	**0.25**		0.01	98.38844			
4	**Timestep**	**0.01**		0.02	94.28005			
5				0.03	95.40441			
6				0.04	92.79735			
7				0.05	93.45168			
8				0.06	93.99664			
9				0.07	97.66597			
10				0.08	96.52319			
11								
12								
13				0.11	99.60075			
14				0.12	99.01974			
15				0.13	100.8729			
16				0.14	101.2378			
17				0.15	102.4736			
18				0.16	102.7694			
19				0.17	100.7347			
20				0.18	102.7021			
21				0.19	107.3493			
22				0.2	109.887			
23				0.21	108.688			
24				0.22	110.7826			
25				0.23	112.8932			
26				0.24	111.0625			
27				0.25	111.6157			
28				0.26	112.5443			
29				0.27	111.9805			
30				0.28	115.6002			
31				0.29	117.9831			
32				0.3	115.2694			
33				0.31	117.4374			

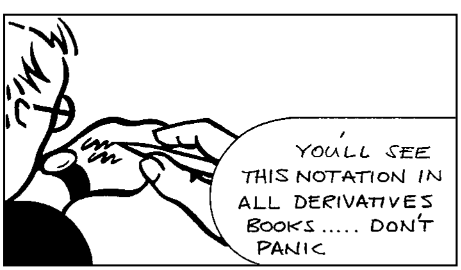
=D4+$B$4

=E7*(1+$B$2*$B$4+$B$3*SQRT($B$4)*(RAND()+RAND()+RAND()+RAND()
+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()-6))

Figure 3.10 Simulating the random walk on a spreadsheet.

the limit $\delta t = 0$. The first δt on the right-hand side of (3.5) becomes dt but the second term is more complicated.

I cannot straightforwardly write $dt^{1/2}$ instead of $\delta t^{1/2}$. If I do go to the zero-time step limit then any random $dt^{1/2}$ term will dominate any deterministic dt term. Yet in our problem the factor in front of $dt^{1/2}$ has a mean of zero, so maybe it does not outweigh the drift after all. Clearly something subtle is happening in the limit.



It turns out, and we will see this in Chapter 4, that because the *variance* of the random term is $O(\delta t)$ we *can* make a sensible continuous-time limit of our discrete-time model. This brings us into the world of Wiener processes.

I am going to write the term $\phi\ \delta t^{1/2}$ as

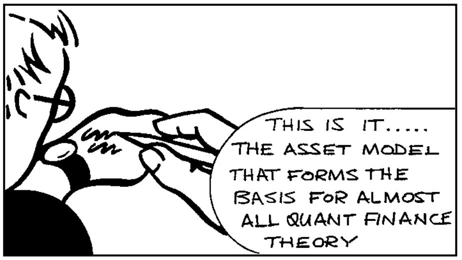
$$dX.$$

You can think of dX as being a random variable, drawn from a Normal distribution with mean zero and variance dt:

$$E[dX] = 0 \quad \text{and} \quad E[dX^2] = dt.$$

This is not exactly what it is, but it is close enough to give the right idea. This is called a **Wiener process**. The important point is that we can build up a continuous-time theory using Wiener processes instead of Normal distributions and discrete time.

3.10 THE WIDELY ACCEPTED MODEL FOR EQUITIES, CURRENCIES, COMMODITIES AND INDICES



Our asset price model in the continuous-time limit, using the Wiener process notation, can be written as

$$dS = \mu S\, dt + \sigma S\, dX. \tag{3.7}$$

This is our first **stochastic differential equation**. It is a continuous-time model of an asset price. It is the most widely accepted model for equities, currencies, commodities and indices, and the foundation of so much finance theory.

We've now built up a simple model for equities that we are going to be using quite a lot. You could ask, if the stock market is so random how can fund managers justify their fee? Do they manage to outsmart the market? Are they clairvoyant or aren't the markets random? Well, I won't swear that markets are random but I can say with confidence that fund managers don't outperform the market. In Figure 3.11 is shown the percentage of funds that outperform an index of all UK stocks. Whether we look at a one-, three-, five- or 10-year horizon we can see that the vast majority of funds can't even keep up with the market. And statistically speaking, there are bound to be a few that beat the market, but only by chance. Maybe one should invest in a fund that does the opposite of all other funds. Great idea except that the management fee and transaction costs probably mean that that would be a poor investment too. This doesn't prove that markets are random, but it's sufficiently suggestive that most of my personal share exposure is via an index-tracker fund.

3.11 SUMMARY

In this chapter I introduced a simple model for the random walk of asset. Initially I built the model up in discrete time, showing what the various terms mean, how they scale with the time step and showing how to implement the model on a spreadsheet.

Most of this book is about continuous-time models for assets. The continuous-time version of the random walk involves concepts such as stochastic calculus and Wiener processes. I introduced these briefly in this chapter and will now go on to explain the underlying theory of stochastic calculus to give the necessary background for the rest of the book.

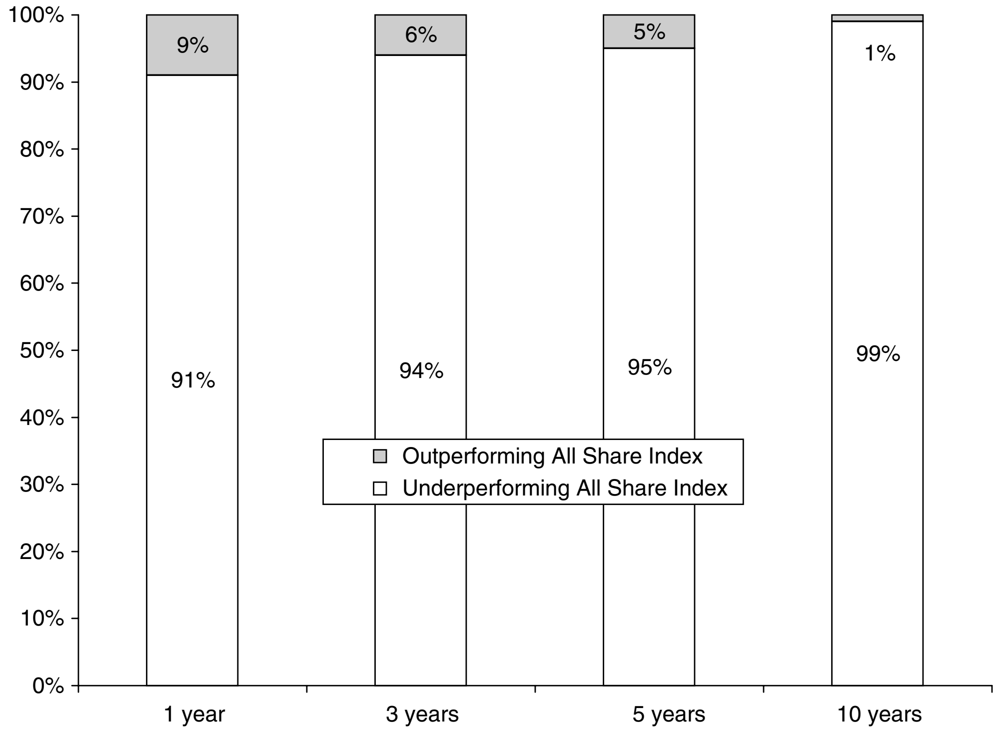


Figure 3.11 Fund performances compared with UK All Share Index. To end December 1998. Data supplied by Virgin Direct.

FURTHER READING

- Mandelbrot (1963) and Fama (1965) did some of the early work on the analysis of financial data.
- For an introduction to random walks and Wiener processes, see Øksendal (1992) and Schuss (1980).
- Some high frequency data can be ordered through Olsen Associates, `www.olsen.ch`. It's not free, but nor is it expensive.
- The famous book by Malkiel (1990) is well worth reading for its insights into the behavior of the stock market. Read what he has to say about chimpanzees, blindfolds and darts. In fact, if you haven't already read Malkiel's book make sure that it is the next book you read after finishing mine.

CHAPTER 4
elementary stochastic calculus

In this Chapter...

- all the stochastic calculus you need to know, and no more
- the meaning of Markov and martingale
- Brownian motion
- stochastic integration
- stochastic differential equations
- Itô's lemma in one and more dimensions

4.1 INTRODUCTION

Stochastic calculus is very important in the mathematical modeling of financial processes. This is because of the assumed underlying random nature of financial markets. Because stochastic calculus is such an important tool I want to ensure that it can be used by everyone. To that end, I am going to try to make this chapter as accessible and intuitive as possible. By the end, I hope that the reader will know what various technical terms mean (and rarely are they very complicated), but, more importantly, will also know how to use the techniques with the minimum of fuss.

Most academic articles in finance have a 'pure' mathematical theme. The mathematical rigor in these works is occasionally justified, but more often than not it only succeeds in obscuring the content. When a subject is young, as is mathematical finance (young*ish*), there is a tendency for technical rigor to feature very prominently in research. This is due to lack of confidence in the methods and results. As the subject ages, researchers will become more cavalier in their attitudes and we will see much more rapid progress.

4.2 A MOTIVATING EXAMPLE

Toss a coin. Every time you throw a head I give you \$1, every time you throw a tail you give me \$1. Figure 4.1 shows how much money you have after six tosses. In this experiment the sequence was THHTHT, and we finished even.

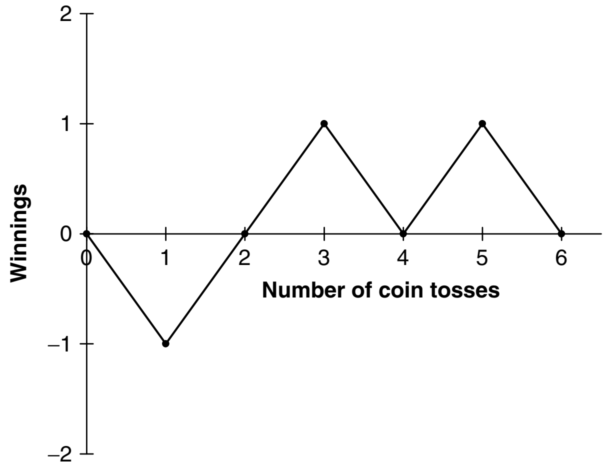


Figure 4.1 The outcome of a coin tossing experiment.

If I use R_i to mean the random amount, either \$1 or −\$1, you make on the ith toss then we have

$$E[R_i] = 0, \quad E[R_i^2] = 1 \quad \text{and} \quad E[R_i R_j] = 0.$$

In this example it doesn't matter whether or not these expectations are conditional on the past. In other words, if I threw five heads in a row it does not affect the outcome of the sixth toss. To the gamblers out there, this property is also shared by a fair die, a balanced roulette wheel, but not by the deck of cards in Blackjack. In Blackjack the same deck is used for game after game, the odds during one game depend on what cards were dealt out from the same deck in previous games. That is why you can in the long run beat the house at Blackjack but not roulette.

Introduce S_i to mean the total amount of money you have won up to and including the ith toss so that

$$S_i = \sum_{j=1}^{i} R_j.$$

Later on it will be useful if we have $S_0 = 0$, i.e., you start with no money.

If we now calculate expectations of S_i it *does* matter what information we have. If we calculate expectations of future events before the experiment has even begun then

$$E[S_i] = 0 \quad \text{and} \quad E[S_i^2] = E[R_1^2 + 2R_1R_2 + \cdots] = i.$$

On the other hand, suppose there have been five tosses already, can I use this information and what can we say about expectations for the sixth toss? This is the **conditional expectation**. The expectation of S_6 conditional upon the previous five tosses gives

$$E[S_6|R_1, \ldots, R_5] = S_5.$$

4.3 THE MARKOV PROPERTY

This result is special, the distribution of the value of the random variable S_i conditional upon all of the past events *only depends on the previous value* S_{i-1}. This is the **Markov property**. We say that the random walk has no memory beyond where it is now. Note that it doesn't have to be the case that the expected value of the random variable S_i is the same as the previous value.

This can be generalized to say that, given information about S_j for some values of $1 \leq j < i$, then the only information that is of use to us in estimating S_i is the value of S_j for the largest j for which we have information.

Almost all of the financial models that I will show you have the Markov property. This is of fundamental importance in modeling in finance. I will also show you examples where the system has a small amount of memory, meaning that one or two other pieces of information are important. And I will also give a couple of examples where *all* of the random walk path contains relevant information.

4.4 THE MARTINGALE PROPERTY

The coin-tossing experiment possesses another property that can be important in finance. You know how much money you have won after the fifth toss. Your expected winnings after the sixth toss, and indeed after any number of tosses if we keep playing, is just the amount you already hold. That is, the conditional expectation of your winnings at any time in the future is just the amount you already hold:

$$E[S_i|S_j, j < i] = S_j.$$

This is called the **martingale property**.

4.5 QUADRATIC VARIATION

I am now going to define the **quadratic variation** of the random walk. This is defined by

$$\sum_{j=1}^{i} \left(S_j - S_{j-1}\right)^2.$$

Because you either win or lose an amount \$1 after each toss, $|S_j - S_{j-1}| = 1$. Thus the quadratic variation is always i:

$$\sum_{j=1}^{i} \left(S_j - S_{j-1}\right)^2 = i.$$

I want to use the coin-tossing experiment for one more demonstration. And that will lead us to a continuous-time random walk.

4.6 BROWNIAN MOTION

I am going to change the rules of my coin-tossing experiment. First of all I am going to restrict the time allowed for the six tosses to a period t, so each toss will take a time $t/6$. Second, the size of the bet will not be \$1 but $\sqrt{t/6}$.

This new experiment clearly still possesses both the Markov and martingale properties, and its quadratic variation measured over the whole experiment is

$$\sum_{j=1}^{6} \left(S_j - S_{j-1}\right)^2 = 6 \times \left(\sqrt{\frac{t}{6}}\right)^2 = t.$$

I have set up my experiment so that the quadratic variation is just the time taken for the experiment.

There is nothing special about the choice of '6' tosses of the coin, so I will change the rules again, to speed up the game even more. We will have n tosses in the allowed time t, with an amount $\sqrt{t/n}$ riding on each throw. Again, the Markov and martingale properties are retained and the quadratic variation is still

$$\sum_{j=1}^{n} \left(S_j - S_{j-1}\right)^2 = n \times \left(\sqrt{\frac{t}{n}}\right)^2 = t.$$

I am now going to make n larger and larger. All I am doing with my rule changes is to speed up the game, decreasing the time between tosses, with a smaller amount for each bet. But I have chosen my new scalings very carefully; the time step is decreasing like n^{-1} but the bet size only decreases by $n^{-1/2}$.

In Figure 4.2 I show a series of experiments, each lasting for a time 1, with increasing number of tosses per experiment.

As I go to the limit $n = \infty$, the resulting random walk stays finite. It has an expectation, conditional on a starting value of zero, of

$$E[S(t)] = 0$$

and a variance

$$E[S(t)^2] = t.$$

I use $S(t)$ to denote the amount you have won or the value of the random variable after a time t. The limiting process for this random walk as the time steps go to zero is called **Brownian motion**, and I will denote it by $X(t)$.

The important properties of Brownian motion are as follows.

- *Finiteness:* Any other scaling of the bet size or 'increments' with time step would have resulted in either a random walk going to infinity in a finite time, or a limit in which there was no motion at all. It is important that the increment scales with the square root of the time step.
- *Continuity:* The paths are continuous, there are no discontinuities. Brownian motion is the continuous-time limit of our discrete time random walk.

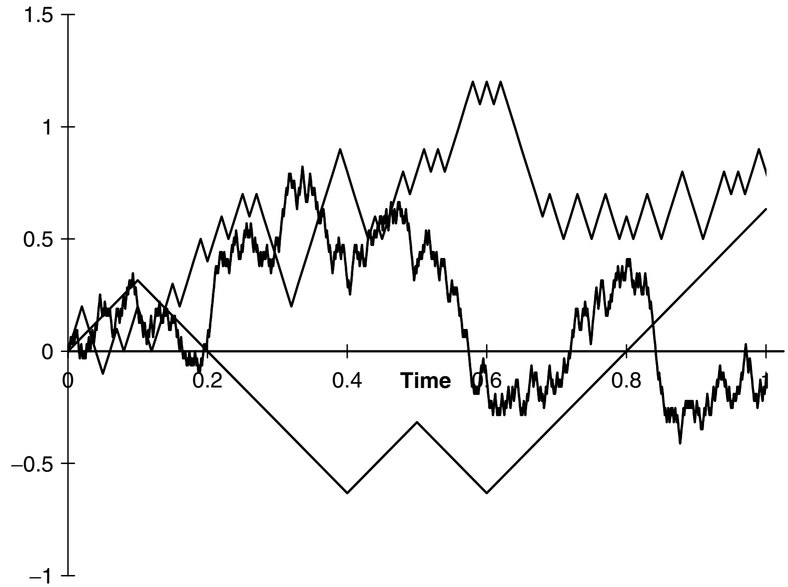


Figure 4.2 A series of coin-tossing experiments, the limit of which is Brownian motion.

- *Markov:* The conditional distribution of $X(t)$ given information up until $\tau < t$ depends only on $X(\tau)$.
- *Martingale:* Given information up until $\tau < t$ the conditional expectation of $X(t)$ is $X(\tau)$.
- *Quadratic variation:* If we divide up the time 0 to t in a partition with $n+1$ partition points $t_i = i\,t/n$ then

$$\sum_{j=1}^{n} \left(X(t_j) - X(t_{j-1})\right)^2 \to t. \quad \text{(Technically ‘almost surely.’)}$$

- *Normality:* Over finite time increments t_{i-1} to t_i, $X(t_i) - X(t_{i-1})$ is Normally distributed with mean zero and variance $t_i - t_{i-1}$.

Having built up the idea and properties of Brownian motion from a series of experiments, we can discard the experiments, to leave the Brownian motion that is defined by its properties. These properties will be very important for our financial models.

4.7 STOCHASTIC INTEGRATION

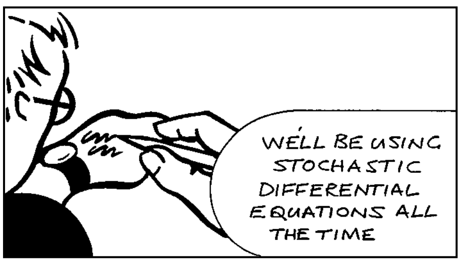
I am going to define a **stochastic integral** by

$$W(t) = \int_0^t f(\tau)\,dX(\tau) = \lim_{n\to\infty} \sum_{j=1}^{n} f(t_{j-1})\left(X(t_j) - X(t_{j-1})\right)$$

with

$$t_j = \frac{jt}{n}.$$

Before I manipulate this in any way or discuss its properties, I want to stress that the function $f(t)$ which I am integrating is evaluated in the summation at the *left-hand point* t_{j-1}. It will be crucially important that each function evaluation does not know about the random increment that multiplies it, i.e. the integration is **non-anticipatory**. In financial terms, we will see that we take some action such as choosing a portfolio and only then does the stock price move. This choice of integration is natural in finance, ensuring that we use no information about the future in our current actions.



4.8 STOCHASTIC DIFFERENTIAL EQUATIONS

Stochastic integrals are important for any theory of stochastic calculus since they can be meaningfully defined. (And in the next section I show how the definition leads to some important properties.) However, it is very common to use a shorthand notation for expressions such as

$$W(t) = \int_0^t f(\tau)\, dX(\tau). \tag{4.1}$$

That shorthand comes from 'differentiating' (4.1) and is

$$dW = f(t)\, dX. \tag{4.2}$$

Think of dX as being an increment in X, i.e. a Normal random variable with mean zero and standard deviation $dt^{1/2}$.

Equations (4.1) and (4.2) are meant to be equivalent. One of the reasons for this shorthand is that the equation (4.2) looks a lot like an ordinary differential equation. We *do not* go the further step of dividing by dt to make it look exactly like an ordinary differential equation because then we would have the difficult task of defining $\frac{dX}{dt}$.

Pursuing this idea further, imagine what might be meant by

$$dW = g(t)\, dt + f(t)\, dX. \tag{4.3}$$

This is simply shorthand for

$$W(t) = \int_0^t g(\tau)\, d\tau + \int_0^t f(\tau)\, dX(\tau).$$

Equations like (4.3) are called **stochastic differential equations**. Their precise meaning comes, however, from the technically more accurate equivalent stochastic integral. In this book I will use the shorthand versions almost everywhere, so no confusion should arise.

You might find it helpful to think of stochastic differential equations as being of the form

$$d\,\underline{\text{Something}} = \underline{\text{Deterministic}}\; dt + \underline{\text{Random}}\; dX.$$

The interpretation is that 'Something' is the thing that you are modeling (stock price, option value, interest rate, volatility, ...), with 'Deterministic' being a function representing the growth in 'Something' when you switch randomness off, and 'Random' being another function representing how random the 'Something' is. Hope that helps!

4.9 THE MEAN SQUARE LIMIT

I am going to describe the technical term **mean square limit**. This is useful in the precise definition of stochastic integration. I will explain the idea by way of the simplest example.

Examine the quantity

$$E\left[\left(\sum_{j=1}^{n}(X(t_j)-X(t_{j-1}))^2-t\right)^2\right] \tag{4.4}$$

where

$$t_j=\frac{jt}{n}.$$

This can be expanded as

$$E\left[\sum_{j=1}^{n}(X(t_j)-X(t_{j-1}))^4+2\sum_{i=1}^{n}\sum_{j<i}(X(t_i)-X(t_{i-1}))^2(X(t_j)-X(t_{j-1}))^2\right.$$
$$\left.-2t\sum_{j=1}^{n}(X(t_j)-X(t_{j-1}))^2+t^2\right].$$

Since $X(t_j)-X(t_{j-1})$ is Normally distributed with mean zero and variance t/n we have

$$E\left[(X(t_j)-X(t_{j-1}))^2\right]=\frac{t}{n}$$

and

$$E\left[(X(t_j)-X(t_{j-1}))^4\right]=\frac{3t^2}{n^2}.$$

Thus (4.4) becomes

$$n\frac{3t^2}{n^2}+n(n-1)\frac{t^2}{n^2}-2tn\frac{t}{n}+t^2=O\left(\frac{1}{n}\right).$$

As $n\to\infty$ this tends to zero. We therefore say that

$$\sum_{j=1}^{n}(X(t_j)-X(t_{j-1}))^2=t$$

in the 'mean square limit.' This is often written, for obvious reasons, as

$$\int_0^t (dX)^2 = t.$$

I am not going to use this result, nor will I use the mean square limit technique. However, when I talk about 'equality' in the following 'proof' I mean equality in the mean square sense.

4.10 FUNCTIONS OF STOCHASTIC VARIABLES AND ITÔ'S LEMMA

I am now going to introduce the idea of a function of a stochastic variable. In Figure 4.3 is shown a realization of a Brownian motion $X(t)$ and the function $F(X) = X^2$.

Clearly X is random (we made it so) and therefore $F(X)$ is random, 'stochastic.' If F is stochastic then what is the stochastic differential equation it satisfies?

If $F = X^2$ is it true that $dF = 2X\,dX$?

No. The ordinary rules of calculus do not generally hold in a stochastic environment. Then what are the rules of calculus?

I am going to 'derive' the most important rule of stochastic calculus, **Itô's lemma**. My derivation is more heuristic than rigorous, but at least it is transparent. I will do this for an arbitrary function $F(X)$.

In this derivation I will need to introduce various timescales. The first timescale is very, very small. I will denote it by

$$\frac{\delta t}{n} = h.$$

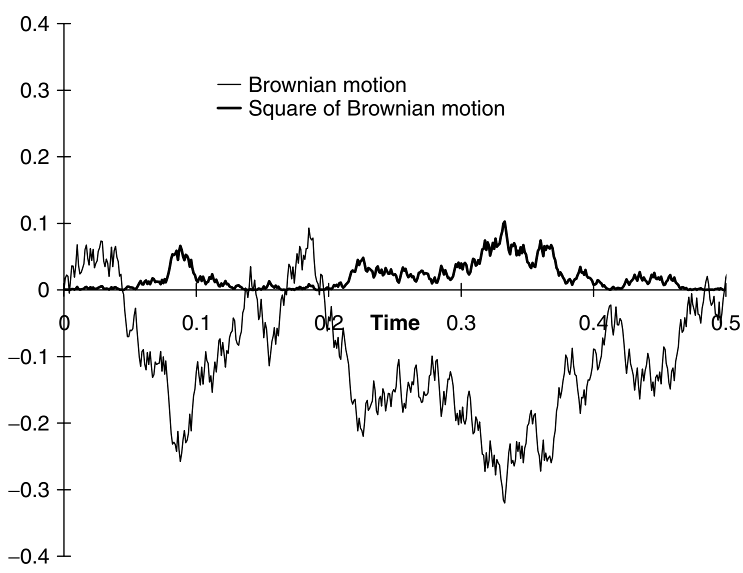


Figure 4.3 A realization of a Brownian motion and its square.

This timescale is so small that the function $F(X(t+h))$ can be approximated by a Taylor series:

$$F(X(t+h)) - F(X(t)) = (X(t+h) - X(t))\frac{dF}{dX}(X(t)) + \tfrac{1}{2}(X(t+h) - X(t))^2\frac{d^2F}{dX^2}(X(t)) + \cdots.$$

From this it follows that

$$\begin{aligned}&(F(X(t+h)) - F(X(t))) + (F(X(t+2h)) - F(X(t+h))) + \cdots + (F(X(t+nh)) \\ &\quad -F(X(t+(n-1)h))) \\ &= \sum_{j=1}^{n}(X(t+jh) - X(t+(j-1)h))\frac{dF}{dX}(X(t+(j-1)h)) \\ &\quad + \tfrac{1}{2}\frac{d^2F}{dX^2}(X(t))\sum_{j=1}^{n}(X(t+jh) - X(t+(j-1)h))^2 + \cdots.\end{aligned}$$

In this I have used the approximation

$$\frac{d^2F}{dX^2}(X(t+(j-1)h)) = \frac{d^2F}{dX^2}(X(t)).$$

This is consistent with the order of accuracy I require.

The first line in this becomes simply

$$F(X(t+nh)) - F(X(t)) = F(X(t+\delta t)) - F(X(t)).$$

The second is just the definition of

$$\int_t^{t+\delta t}\frac{dF}{dX}\,dX$$

and the last is

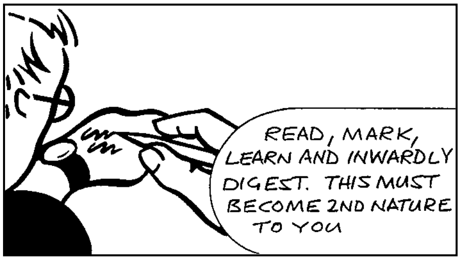
$$\tfrac{1}{2}\frac{d^2F}{dX^2}(X(t))\,\delta t,$$

in the *mean square sense*. Thus we have

$$F(X(t+\delta t)) - F(X(t)) = \int_t^{t+\delta t}\frac{dF}{dX}(X(\tau))\,dX(\tau) + \tfrac{1}{2}\int_t^{t+\delta t}\frac{d^2F}{dX^2}(X(\tau))\,d\tau.$$

I can now extend this result over longer timescales, from zero up to t, over which F *does* vary substantially to get

$$\boxed{F(X(t)) = F(X(0)) + \int_0^t\frac{dF}{dX}(X(\tau))\,dX(\tau) + \tfrac{1}{2}\int_0^t\frac{d^2F}{dX^2}(X(\tau))\,d\tau.}$$



This is the integral version of **Itô's lemma**, which is usually written as

$$\boxed{dF = \frac{dF}{dX}\,dX + \tfrac{1}{2}\frac{d^2F}{dX^2}\,dt.} \tag{4.5}$$

We can now answer the question: If $F = X^2$ what stochastic differential equation does F satisfy? In this example

$$\frac{dF}{dX} = 2X \quad \text{and} \quad \frac{d^2F}{dX^2} = 2.$$

Therefore Itô's lemma tells us that

$$dF = 2X\,dX + dt.$$

This is *not* what we would get if X were a deterministic variable. In integrated form

$$X^2 = F(X) = F(0) + \int_0^t 2X\,dX + \int_0^t 1\,d\tau = \int_0^t 2X\,dX + t.$$

Therefore

$$\int_0^t X\,dX = \tfrac{1}{2}X^2 - \tfrac{1}{2}t.$$

4.11 INTERPRETATION OF ITÔ'S LEMMA

Itô's lemma is going to be of great importance to us when we start to look at pricing options. If we can get comfortable with manipulating random quantities via simple rules of stochastic calculus then we will find most option theory quite straightforward.

To help in that regard, and to give you some insight into the role that Itô's lemma will be playing, take a look at Figure 4.4.

In this figure you will see at the top a realization of a stock price, just a basic lognormal random walk. Below this is the value of an option on this stock.[1] What you will notice about these plots is that both have a direction to them (both are rising overall) and both have a random element (the bouncing around of the values).

Both look stochastic and we know that the stock price satisfies a stochastic differential equation

$$dS = \mu S\,dt + \sigma S\,dX,$$

so maybe the option value (call it $V(S, t)$) also satisfies a stochastic differential equation

$$dV = \underline{\qquad}\,dt + \underline{\qquad}\,dX.$$

[1] It doesn't much matter whether it is a call, a put or something more exotic, the concept is relevant to all options.

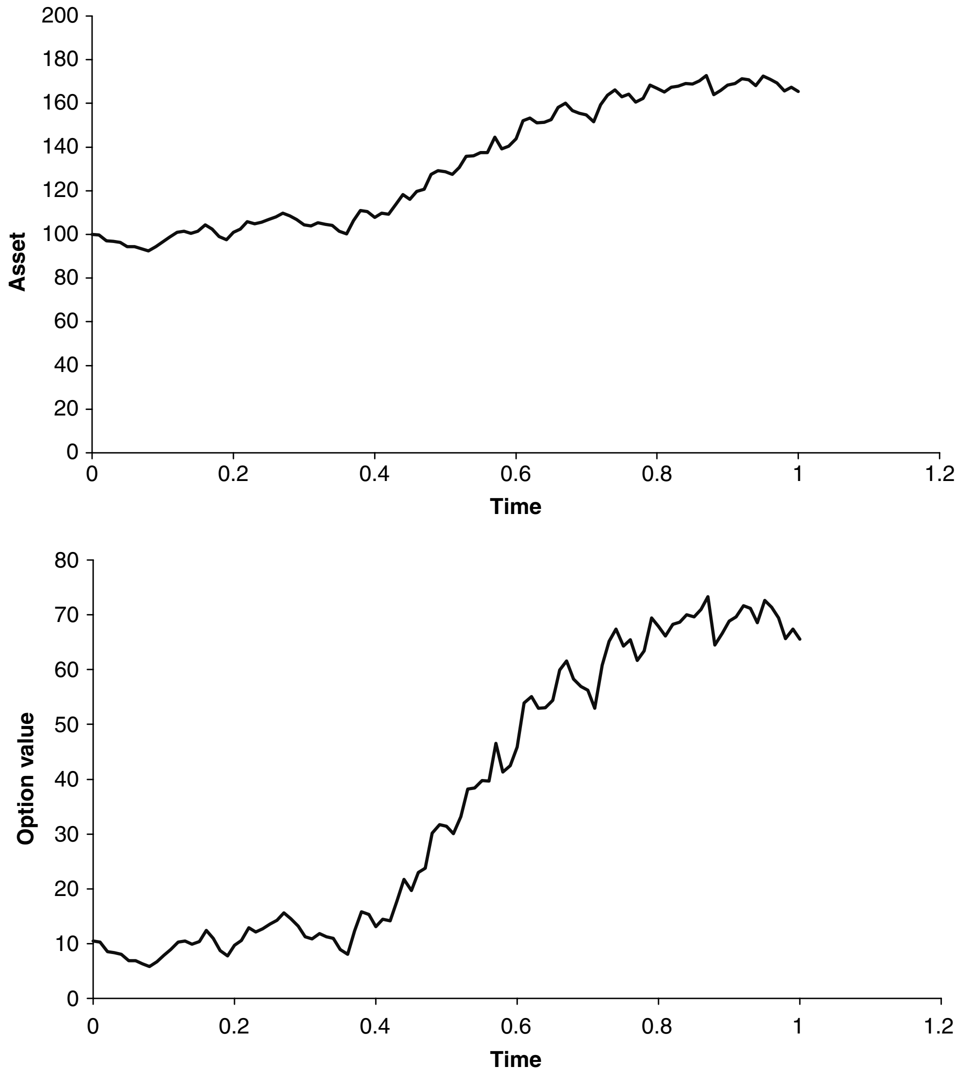


Figure 4.4 A realization of a stock price and the value of an option on that stock.

The question is then: What are the underlined bits? And that is precisely what Itô tells us.

This will be important later when we do the Black–Scholes theory, because knowing how much randomness there is in an option's value relative to a stock's value will give us a recipe for eliminating that randomness by buying an option and selling short a special quantity of the stock.

4.12 ITÔ AND TAYLOR

Having derived Itô's lemma, I am going to give some further intuition behind the result, to make it really easy to use/derive, and then slightly generalize it.

If we were to do a naive Taylor series expansion of F, completely disregarding the nature of X, and treating dX as a small increment in X, we would get

$$F(X + dX) = F(X) + \frac{dF}{dX}\,dX + \tfrac{1}{2}\frac{d^2F}{dX^2}\,dX^2,$$

ignoring higher-order terms. We could argue that $F(X + dX) - F(X)$ was just the 'change in' F and so

$$dF = \frac{dF}{dX}\,dX + \tfrac{1}{2}\frac{d^2F}{dX^2}\,dX^2.$$

This is very similar to (4.5) (and Taylor series *is* very similar to Itô), with the only difference being that there is a dX^2 instead of a dt. However, since in a sense

$$\int_0^t (dX)^2 = t$$

I could perhaps write

$$dX^2 = dt. \tag{4.6}$$

Although this lacks any rigor (because it's wrong) it does give the correct result. However, on a positive note you can, with little risk of error, use Taylor series with the 'rule of thumb' (4.6) and in practice you will get the right result. Although this is technically incorrect, you almost certainly[2] won't get the wrong result. I will use this rule of thumb almost every time I want to differentiate a function of a random variable. In Chapter 48 I will show when it *is* correct, and better, to use Taylor series.

4.12.1 The Intuition

To aid in the understanding of Itô's lemma it may be helpful if you think of it as follows.

Stochastic integrals, which are the 'proper' way of manipulating random quantities, involve adding up an infinite number of random variables. We saw this in Section 4.10.

There are several contributions to the sum: Those that are a sum of random variables and those that are the sum of the squares of random variables, and then there are higher-order terms.

Add up a large number of random variables and the Central Limit Theorem kicks in, the end result being a Normally distributed random variable. But what is its mean and standard deviation? This is the key.

When we add up N terms that are Normal, each with a mean of zero and a standard deviation of $\sqrt{\delta t/N}$, we end up with another Normal, with a mean of zero and a standard deviation of $\sqrt{\delta t}$. This is our dX. Notice how the N disappears in this limit.

Then if we add up the N *squares* of the same Normal numbers (mean of zero and a standard deviation of $\sqrt{\delta t/N}$) then we get something which is Normally distributed with a mean of

$$N \times \left(\sqrt{\frac{\delta t}{N}}\right)^2 = \delta t$$

[2] Or should that be 'almost surely'?

and a standard deviation which is

$$\delta t\sqrt{\frac{2}{N}}.$$

This tends to zero as N gets larger. In this limit we end up with, in a sense, our $dX^2 = dt$, because the randomness as measured by the standard deviation disappears leaving us with just the mean, dt.

The higher-order terms have means and standard deviations that are too small, disappearing rapidly in the limit as N tends to infinity.

4.12.2 Simple Generalization

To end this section I will generalize slightly. Suppose my stochastic differential equation is

$$dS = a(S)\,dt + b(S)\,dX, \tag{4.7}$$

say, for some functions $a(S)$ and $b(S)$. Here dX is the usual Brownian increment. Now if I have a function of S, $V(S)$, what stochastic differential equation does it satisfy? The answer is

$$dV = \frac{dV}{dS}\,dS + \tfrac{1}{2}b^2\frac{d^2V}{dS^2}\,dt.$$

We could derive this properly or just cheat by using Taylor series with $dX^2 = dt$. I could, if I wanted, substitute for dS from (4.7) to get an equation for dV in terms of the pure Brownian motion X:

$$dV = \left(a(S)\frac{dV}{dS} + \tfrac{1}{2}b(S)^2\frac{d^2V}{dS^2}\right)dt + b(S)\frac{dV}{dS}\,dX.$$

4.13 ITÔ IN HIGHER DIMENSIONS

In financial problems we often have functions of one stochastic variable S and a deterministic variable t, time: $V(S, t)$. If

$$dS = a(S, t)\,dt + b(S, t)\,dX,$$

then the increment dV is given by

$$dV = \frac{\partial V}{\partial t}\,dt + \frac{\partial V}{\partial S}\,dS + \tfrac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}\,dt. \tag{4.8}$$

Again, this is shorthand notation for the correct integrated form. This result is obvious, as is the use of partial instead of ordinary derivatives.

Occasionally, we have a function of two, or more, random variables, and time as well: $V(S_1, S_2, t)$. An example would be the value of an option to buy the more valuable out of Nike and Reebok. I will write the behavior of S_1 and S_2 in the general form

$$dS_1 = a_1(S_1, S_2, t)\,dt + b_1(S_1, S_2, t)\,dX_1$$

and

$$dS_2 = a_2(S_1, S_2, t)\, dt + b_2(S_1, S_2, t)\, dX_2.$$

Note that I have *two* Brownian increments dX_1 and dX_2. We can think of these as being Normally distributed with variance dt, but *they are correlated*. The correlation between these two random variables I will call ρ. This can also be a function of S_1, S_2 and t but must satisfy

$$-1 \le \rho \le 1.$$

The 'rules of thumb' can readily be imagined:

$$dX_1^2 = dt, \quad dX_2^2 = dt \quad \text{and} \quad dX_1 dX_2 = \rho\, dt.$$

Itô's lemma becomes

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_1} dS_1 + \frac{\partial V}{\partial S_2} dS_2 + \tfrac{1}{2} b_1^2 \frac{\partial^2 V}{\partial S_1^2} dt + \rho b_1 b_2 \frac{\partial^2 V}{\partial S_1 \partial S_2} dt + \tfrac{1}{2} b_2^2 \frac{\partial^2 V}{\partial S_2^2} dt. \qquad (4.9)$$

4.14 SOME PERTINENT EXAMPLES

In this section I am going to introduce a few common random walks and talk about their properties.

Remember that a stochastic differential equation model for variable S is something of the form

$$dS = \underline{\qquad}\, dt + \underline{\qquad}\, dX.$$

The bit in front of the dt is deterministic and the bit in front of the dX tells us how much randomness there is. Modeling is very much about choosing functions to go where the underlining is; it is about choosing the functional form for the deterministic part and the functional form for the amount of randomness. We will now look at some examples.

4.14.1 Brownian Motion with Drift

The first example is like the simple Brownian motion but with a drift:

$$dS = \mu\, dt + \sigma\, dX.$$

A realization of this is shown in Figure 4.5. The point to note about this realization is that S has gone negative, near the start. This random walk would therefore not be a good model for many financial quantities, such as interest rates or equity prices. This stochastic differential equation can be integrated exactly to get

$$S(t) = S(0) + \mu t + \sigma (X(t) - X(0)).$$

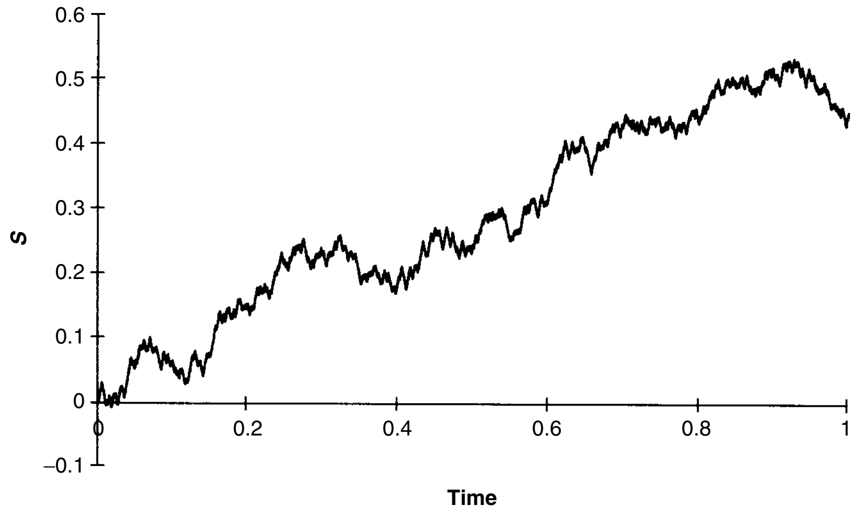


Figure 4.5 A realization of $dS = \mu\, dt + \sigma\, dX$.

4.14.2 The Lognormal Random Walk

My second example is similar to the above but the drift and randomness scale with S:

$$dS = \mu S\, dt + \sigma S\, dX. \tag{4.10}$$

A realization of this is shown in Figure 4.6. If S starts out positive it can never go negative; the closer that S gets to zero the smaller the increments dS. For this reason I have had to start the simulation with a non-zero value for S. This property of this random walk is clearly seen if we examine the function $F(S) = \log S$ using Itô's lemma. From Itô we have

$$dF = \frac{dF}{dS}\, dS + \tfrac{1}{2}\sigma^2 S^2 \frac{d^2F}{dS^2}\, dt = \frac{1}{S}\left(\mu S\, dt + \sigma S\, dX\right) - \tfrac{1}{2}\sigma^2\, dt = \left(\mu - \tfrac{1}{2}\sigma^2\right) dt + \sigma\, dX.$$

This shows us that $\log S$ can range between minus and plus infinity but cannot reach these limits in a finite time, therefore S cannot reach zero or infinity in a finite time.

How does the time series in Figure 4.6 which was generated on a spreadsheet using random returns compare qualitatively with the time series in Figure 4.7 which is the real series for Glaxo–Wellcome?

The integral form of this stochastic differential equation follows simply from the stochastic differential equation for $\log S$:

$$S(t) = S(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma(X(t) - X(0))}.$$

The stochastic differential equation (4.10) will be particularly important in the modeling of many asset classes. And if we have some function $V(S, t)$ then from Itô it follows that

$$\boxed{dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dt.} \tag{4.11}$$

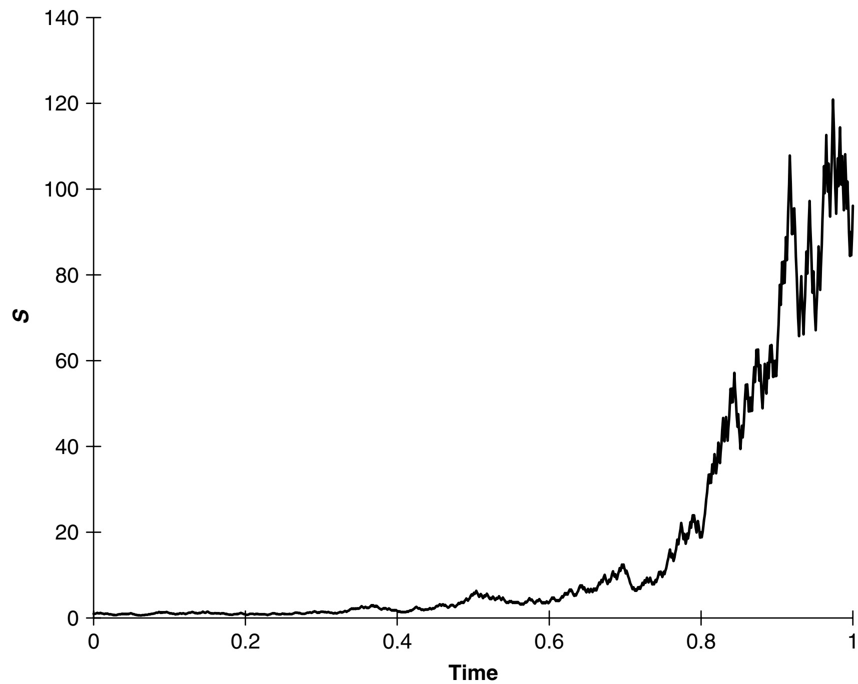


Figure 4.6 A realization of $dS = \mu S\,dt + \sigma S\,dX$.

4.14.3 A Mean-reverting Random Walk

The third example is

$$dS = (\nu - \mu S)\,dt + \sigma\,dX.$$

A realization of this is shown in Figure 4.8.

This random walk is an example of a **mean-reverting** random walk. If S is large, greater than ν/μ, the negative coefficient in front of dt means that S will move down on average; if S is small, less than ν/μ, it rises on average. There is still no incentive for S to stay positive in this random walk. With r instead of S this random walk is the Vasicek model for the short-term interest rate.

Mean-reverting models are used for modeling a random variable that 'isn't going anywhere.' That's why they are often used for interest rates; Figure 4.9 shows the yield on a Japanese Government Bond.

Let's take a look at the Vasicek model for the spot interest rate r

$$dr = (\nu - \gamma r)\,dt + \sigma\,dX$$

where γ is the reversion rate and ν/γ is the mean rate.

By setting $W = r - \nu$, W is a solution of

$$dW = -\gamma W\,dt + \sigma\,dX.$$

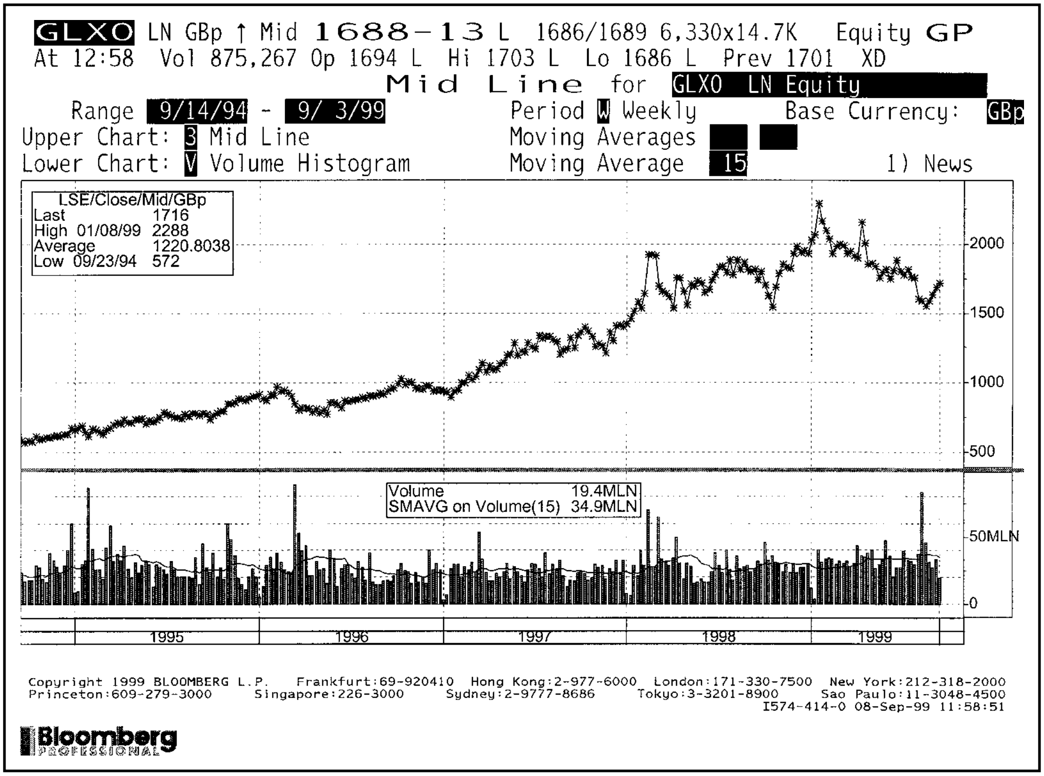


Figure 4.7 Glaxo–Wellcome share price (volume below). Source: Bloomberg L.P.

This random walk for W is an **Ornstein–Uhlenbeck** process. An analytic solution for this equation exists, and we shall derive it now.

Introduce the integrating factor $I = e^{\gamma t}$. Write

$$d(IW) = I\,dW + W\,dI = e^{\gamma t}\left(-\gamma W\,dt + \sigma\,dX\right) + \gamma W e^{\gamma t}\,dt$$
$$= \sigma e^{\gamma t}\,dX.$$

Integrating over $[0, t]$ gives

$$IW = \text{constant} + \sigma \int_0^t e^{\gamma s}\,dX(s),$$

so that

$$W = W(0)\,e^{-\gamma t} + \sigma \int_0^t e^{\gamma(s-t)}\,dX(s). \tag{4.12}$$

By using integration by parts we can simplify (4.12).

$$\int_0^t e^{\gamma(s-t)}\,dX(s) = X - \gamma \int_0^t e^{\gamma(s-t)} X(s)\,ds.$$

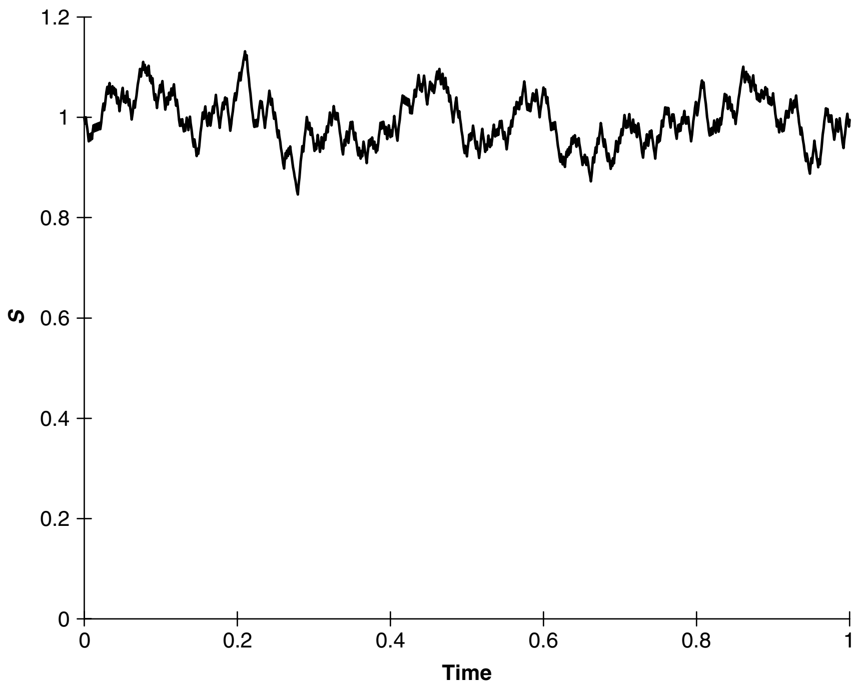


Figure 4.8 A realization of $dS = (\nu - \mu S)dt + \sigma\, dX$.

And we can write (4.12) as

$$W(t) = W(0)\, e^{-\gamma t} + \sigma \left(X(t) - \gamma \int_0^t e^{\gamma(s-t)} X(s)\, ds \right).$$

Hence

$$r = \nu + W = \nu + (r(0) - \nu) \exp(-\gamma t) + \sigma \left(X(t) - \gamma \int_0^t e^{\gamma(s-t)} X(s)\, ds \right).$$

4.14.4 And Another Mean-reverting Random Walk

The final example is similar to the third, and I will again write it in terms of r, but I am going to adjust the random term slightly:

$$dr = (\nu - \mu r)\, dt + \sigma r^{1/2}\, dX.$$

Now if r ever gets close to zero the randomness decreases, perhaps this will stop r from going negative? Let's play around with this example for a while. And we'll see Itô in practice.

Write $F = r^{1/2}$. What stochastic differential equation does F satisfy? Since

$$\frac{dF}{dr} = \tfrac{1}{2} r^{-1/2} \quad \text{and} \quad \frac{d^2F}{dr^2} = -\tfrac{1}{4} r^{-3/2}$$

Figure 4.9 Time series of the yield on a JGB. Source: Bloomberg L.P.

we have

$$dF = \left(\frac{4\nu - \sigma^2}{8F} - \tfrac{1}{2}\mu F \right) dt + \tfrac{1}{2}\sigma\, dX.$$

I have just turned the original stochastic differential equation with a variable coefficient in front of the random term into a stochastic differential equation with a constant random term. In so doing I have made the drift term nastier. In particular, the drift is now singular at $F = r = 0$. Something special is happening at $r = 0$.

Instead of examining $F(r) = r^{1/2}$, can I find a function $F(r)$ such that its stochastic differential equation has a zero drift term? For this I will need

$$(\nu - \mu r)\frac{dF}{dr} + \tfrac{1}{2}\sigma^2 r \frac{d^2F}{dr^2} = 0.$$

This is easily integrated once to give

$$\frac{dF}{dr} = Ar^{-2\nu/\sigma^2} e^{2\mu r/\sigma^2} \tag{4.13}$$

for any constant A. I won't take this any further but just make one observation. If

$$\frac{2\nu}{\sigma^2} \geq 1$$

we cannot integrate (4.13) at $r = 0$. This makes the origin **non-attainable**. In other words, if the parameter ν is sufficiently large it forces the random walk to stay away from zero.

This particular stochastic differential equation for r will be important later on, it is the Cox, Ingersoll & Ross model for the short-term interest rate.

These are just four of the many random walks we will be seeing.

4.15 SUMMARY

This chapter introduced the most important tool of the trade, Itô's lemma. Itô's lemma allows us to manipulate functions of a random variable. If we think of S as the value of an asset for which we have a stochastic differential equation, a 'model,' then we can handle functions of the asset, and ultimately value contracts such as options.

If we use Itô as a tool we do not need to know why or how it works, only how to use it. Essentially all we require to use the lemma successfully is a rule of thumb, as explained in the text. Unless we are using Itô in highly unusual situations, then we are unlikely to make any errors.

FURTHER READING

- Neftci (1996) is the only readable book on stochastic calculus for beginners. It does not assume any knowledge about anything. It takes the reader very slowly through the basics as applied to finance.
- Once you have got beyond the basics, move on to Øksendal (1992) and Schuss (1980).

CHAPTER 5

the Black–Scholes model

In this Chapter...

- the foundations of derivatives theory: delta hedging and no arbitrage
- the derivation of the Black–Scholes partial differential equation
- the assumptions that go into the Black–Scholes equation
- how to modify the equation for commodity and currency options

5.1 INTRODUCTION

This is, without doubt, the most important chapter in the book. In it I describe and explain the basic building blocks of derivatives theory. These building blocks are delta hedging and no arbitrage. They form a moderately sturdy foundation to the subject and have performed well since 1973 when the ideas became public.

In this chapter I begin with the stochastic differential equation model for equities and exploit the correlation between this asset and an option on this asset to make a perfectly risk-free portfolio. I then appeal to no arbitrage to equate returns on all risk-free portfolios to the risk-free interest rate, the so called 'no free lunch' argument.

The arguments are trivially modified to incorporate dividends on the underlying and also to price commodity and currency options and options on futures.

This chapter is quite theoretical, yet all of the ideas contained here are regularly used in practice. Even though all of the assumptions can be shown to be wrong to a greater or lesser extent, the Black–Scholes model is profoundly important both in theory and in practice.

5.2 A VERY SPECIAL PORTFOLIO

In Chapter 2 I described some of the characteristics of options and options markets. I introduced the idea of call and put options, amongst others. The value of a call option is clearly going to be a function of various parameters in the contract, such as the strike price E and the time to expiry $T - t$, where T is the date of expiry and t is the current time. The value will also depend on properties of the asset itself, such as its price, its drift and its volatility, as well as the risk-free rate of interest.[1] We can write the option value as

$$V(S, t; \sigma, \mu; E, T; r).$$

Notice that the semicolons separate different types of variables and parameters:

- S and t are variables;
- σ and μ are parameters associated with the asset price;
- E and T are parameters associated with the details of the particular contract;
- r is a parameter associated with the currency in which the asset is quoted.

I'm not going to carry all the parameters around, except when it is important. For the moment I'll just use $V(S, t)$ to denote the option value as a function of its variables.

One simple observation is that a call option will rise in value if the underlying asset rises, and will fall if the asset falls. This is clear since a call has a larger payoff the greater the value of the underlying at expiry. This is an example of **correlation** between two financial instruments, in this case the correlation is positive. A put and the underlying have a negative correlation. We can exploit these correlations to construct a very special portfolio.

Use Π to denote the value of a portfolio of one long option position and a short position in some quantity Δ, **delta**, of the underlying:

$$\Pi = V(S, t) - \Delta S. \tag{5.1}$$

The first term on the right is the option and the second term is the short asset position. Notice the minus sign in front of the second term. The quantity Δ will for the moment be some constant quantity of our choosing. We will assume that the underlying follows a lognormal random walk

$$dS = \mu S\, dt + \sigma S\, dX.$$

It is natural to ask how the value of the portfolio changes from time t to $t + dt$. The change in the portfolio value is due partly to the change in the option value and partly to the change in the underlying:

$$d\Pi = dV - \Delta\, dS.$$

Notice that Δ has not changed during the time step; we have not anticipated the change in S. From Itô we have

$$dV = \frac{\partial V}{\partial t}\, dt + \frac{\partial V}{\partial S}\, dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\, dt.$$

Thus the portfolio changes by

$$d\Pi = \frac{\partial V}{\partial t}\, dt + \frac{\partial V}{\partial S}\, dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\, dt - \Delta\, dS. \tag{5.2}$$

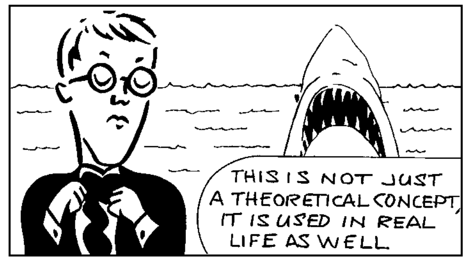
[1] Actually, I'm lying. One of these parameters does not affect the option value.

5.3 ELIMINATION OF RISK: DELTA HEDGING

The right-hand side of (5.2) contains two types of terms, the deterministic and the random. The deterministic terms are those with the dt, and the random terms are those with the dS. Pretending for the moment that we know V and its derivatives then we know everything about the right-hand side of (5.2) *except for the value of* dS. And this quantity we can never know in advance.

These random terms are the risk in our portfolio. Is there any way to reduce or even eliminate this risk? This can be done in theory (and *almost* in practice) by carefully choosing Δ. The random terms in (5.2) are

$$\left(\frac{\partial V}{\partial S} - \Delta\right) dS.$$



If we choose

$$\Delta = \frac{\partial V}{\partial S} \tag{5.3}$$

then the randomness is reduced to zero.

Any reduction in randomness is generally termed **hedging**, whether that randomness is due to fluctuations in the stock market or the outcome of a horse race. The perfect elimination of risk, by exploiting correlation between two instruments (in this case an option and its underlying) is generally called **delta hedging**.

Delta hedging is an example of a **dynamic hedging** strategy. From one time step to the next the quantity $\frac{\partial V}{\partial S}$ changes, since it is, like V, a function of the ever-changing variables S and t. This means that the perfect hedge must be continually rebalanced. In later chapters we will see examples of static hedging, where a hedging position is *not* changed as the variables evolve.

Delta hedging was effectively first described by Thorp & Kassouf (1967). (We will see more of Thorp when we look at casino Blackjack as an investment in Chapter 17.)

5.4 NO ARBITRAGE

After choosing the quantity Δ as suggested above, we hold a portfolio whose value changes by the amount

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt. \tag{5.4}$$

This change is completely *riskless*. If we have a completely risk-free change $d\Pi$ in the portfolio value Π then it must be the same as the growth we would get if we put the equivalent amount of cash in a risk-free interest-bearing account:

$$d\Pi = r\Pi\, dt. \tag{5.5}$$

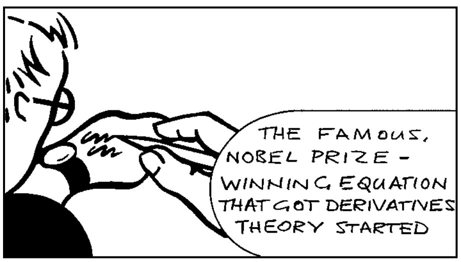
This is an example of the **no arbitrage** principle.

To see why this should be so, consider in turn what might happen if the return on the portfolio were, first, greater and, second, less than the risk-free rate. If we were guaranteed to get a return of greater than r from the delta-hedged portfolio then what we could do is borrow from the bank, paying interest at the rate r, invest in the risk-free option/stock portfolio and make a profit. If, on the other hand, the return were less than the risk-free rate we should go short the option, delta hedge it, and invest the cash in the bank. Either way, we make a riskless profit in excess of the risk-free rate of interest. At this point we say that, all things being equal, the action of investors buying and selling to exploit the arbitrage opportunity will cause the market price of the option to move in the direction that eliminates the arbitrage.

5.5 THE BLACK–SCHOLES EQUATION

Substituting (5.1), (5.3) and (5.4) into (5.5) we find that

$$\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt = r\left(V - S\frac{\partial V}{\partial S}\right) dt.$$



On dividing by dt and rearranging we get

$$\boxed{\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.} \tag{5.6}$$

This is the **Black–Scholes equation**. The equation was first written down in 1969, but a few years passed, with Fischer Black and Myron Scholes justifying the model, before it was published. The derivation of the equation was finally published in 1973, although the call and put formulae had been published a year earlier.[2]

The Black–Scholes equation is a **linear parabolic partial differential equation**. In fact, almost all partial differential equations in finance are of a similar form. They are almost always linear, meaning that if you have two solutions of the equation then the sum of these is itself also a solution. Or at least they tended to be linear until recently. In Part Five I will show you some examples of recent models which lead to non-linear equations. Financial equations are also usually parabolic, meaning that they are related to the heat or diffusion equation of mechanics. One of the good things about this is that such equations are relatively easy to solve numerically.

The Black–Scholes equation contains all the obvious variables and parameters such as the underlying, time, and volatility, but there is no mention of the drift rate μ. Why is this? Any dependence on the drift dropped out at the same time as we eliminated the dS component of the portfolio. The economic argument for this is that since we can perfectly hedge the option with the underlying we should not be rewarded for taking unnecessary risk; only the risk-free rate of return is in the equation. This means that if you and I agree on the volatility of an asset we will agree on the value of its derivatives *even if we have differing estimates of the drift*.

Another way of looking at the hedging argument is to ask what happens if we hold a portfolio consisting of just the stock, in a quantity Δ, and cash. If Δ is the partial derivative of some option value then such a portfolio will yield an amount at expiry that is simply that option's

[2] The pricing formulae were being used even earlier by Ed Thorp to make money.

payoff. In other words, we can use the same Black–Scholes argument to **replicate** an option just by buying and selling the underlying asset. This leads to the idea of a **complete market**. In a complete market an option can be replicated with the underlying, thus making options redundant. Why buy an option when you can get the same payoff by trading in the asset? Many things conspire to make markets incomplete and we will discuss some of these, such as transaction costs and stochastic volatility, in later chapters.

Since the Black–Scholes equation is of such importance I'm going to spend a moment relating the equation to ideas and models in other parts of the book.

For the rest of this part of the book I'm going to explain what the Black–Scholes equation means as a differential equation and show how to solve it in a few special cases. I will generalize the model slightly but not too much. I will also show what happens to the equation when early exercise is allowed. Superficially, early exercise does not make much difference, but on a closer inspection it changes the whole nature of the problem.

The second part of the book is devoted to more serious generalizations to accommodate the pricing of exotic derivative contracts. Still the extensions are what might be termed 'classical,' there is nothing too outrageous here and there is nothing outside the 'Black–Scholes world.'

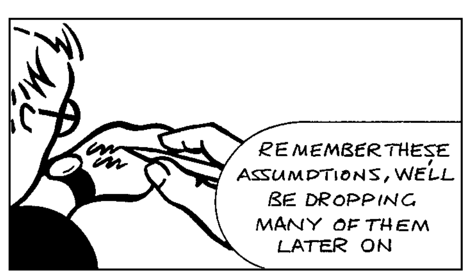
Many of the Black–Scholes ideas, plus a few new ones, will be found in Part Three on the pricing of interest rate products. There are some technical reasons that make the fixed-income world harder than the equity, currency, commodity world. Nevertheless delta hedging and no arbitrage play the same role.

Credit risk management is the subject of Part Four. Here's a situation where delta hedging is typically not possible.

Part Five is devoted to more serious generalizations and some advanced topics. Some of the generalizations are minor modifications to the Black–Scholes world and can still be thought of as classical. These are just attempts to broaden the Black–Scholes world, to relax some of the assumptions. Other models take us a long way from the Black–Scholes world into fairly new and uncharted waters. Delta hedging and no arbitrage are seen to be not quite as straightforward as I had led you to understand. You are free to believe or disbelieve any of these models, but I do ask that you appreciate them.

In Part Six I show how to solve the Black–Scholes equation, and related equations, numerically. I describe the two main numerical methods: finite-difference methods and Monte Carlo simulations. I include Visual Basic code to illustrate many of the methods. None of these methods are hard; we are lucky in this subject that the governing equation, usually being parabolic, is simple to solve numerically.

5.6 THE BLACK–SCHOLES ASSUMPTIONS



What are the 'assumptions' that I've just been referring to? Here is a partial list, together with some discussion.

- *The underlying follows a lognormal random walk:* This is not entirely necessary. To find explicit solutions we will need the random term in the stochastic differential equation for S to be proportional to S. The 'factor' σ does not need to be constant to find solutions, but it must only be time dependent, see Chapter 7. As far as the validity of the equation is concerned it doesn't matter if the volatility is also asset-price dependent, but then the equation will either have very messy explicit solutions, if it has any at all, or have to be solved numerically. Then there is the question of the drift term μS. Do we need this term to take this form, after all it doesn't even appear in the equation? There is a technicality here that whatever the stochastic differential equation for the asset S, the domain over which the asset can range must be zero to infinity. This is a technicality I am not going into, but it amounts to another elimination of arbitrage. It is possible to choose the drift so that the asset is restricted to lie within a range; such a drift could leave us with arbitrage opportunities.

- *The risk-free interest rate is a known function of time:* This restriction is just to help us find explicit solutions again. If r were constant this job would be even easier. In practice, the interest rate is often taken to be time dependent but known in advance. Explicit formulae still exist for the prices of simple contracts and I discuss this issue in Chapter 8. In reality the rate r is not known in advance and is itself stochastic, or so it seems from data. I will discuss stochastic interest rates in Part Three. We've also assumed that lending and borrowing rates are the same. It is not difficult to relax this assumption, and it is related to ideas in Chapter 52.
- *There are no dividends on the underlying:* I will drop this restriction in a moment, and discuss the subject more generally in Chapter 8.
- *Delta hedging is done continuously:* This is definitely impossible. Hedging must be done in discrete time. Often the time between rehedges will depend on the level of transaction costs in the market for the underlying; the lower the costs, the more frequent the rehedging. This subject is covered in depth in Chapter 47.
- *There are no transaction costs on the underlying:* The dynamic business of delta hedging is in reality expensive since there is a bid-offer spread on most underlyings. In some markets this matters and in some it doesn't. Chapter 48 is devoted to a discussion of these issues.
- *There are no arbitrage opportunities:* This is a beauty. Of course there are arbitrage opportunities, a lot of people make a lot of money finding them.[3] It is extremely important to stress that we are ruling out model-dependent arbitrage. This is highly dubious since it depends on us having the correct model in the first place, and that is unlikely. I am happier ruling out model-independent arbitrage, i.e. arbitrage arising when two identical cashflows have different values. But even that can be criticized.

There are many more assumptions but the above are the most important. In other parts of the book I will drop these assumptions or, if I don't drop them, I will at least loosen them a bit.

5.7 FINAL CONDITIONS

The Black–Scholes equation (5.6) knows nothing about what kind of option we are valuing, whether it is a call or a put, nor what is the strike and the expiry. These points are dealt with by the **final condition**. We must specify the option value V as a function of the underlying at the expiry date T. That is, we must prescribe $V(S, T)$, the payoff.

[3] Life, and everything in it, is based on arbitrage opportunities and their exploitation. I always wonder how, after learning about the absence of free lunches, students can join a bank and then expect to get six or seven figure bonuses. Where do they expect this money to come from?

For example, if we have a call option then we know that

$$V(S,T)=\max(S-E,0).$$

For a put we have

$$V(S,T)=\max(E-S,0),$$

for a binary call

$$V(S,T)=\mathcal{H}(S-E)$$

and for a binary put

$$V(S,T)=\mathcal{H}(E-S),$$

where $\mathcal{H}(\cdot)$ is the **Heaviside function**, which is zero when its argument is negative and one when it is positive.

The imposition of the final condition will be explained in Chapters 6 and 7, and implemented numerically in Part Six.

As an aside, observe that both the asset, S, and 'money in the bank,' e^{rt} satisfy the Black–Scholes equation.

5.8 OPTIONS ON DIVIDEND-PAYING EQUITIES

The first generalization we discuss is how to value options on stocks paying dividends. This is just about the simplest generalization of the Black–Scholes model. To keep things easy (I will complicate matters in Chapter 8) let's assume that the asset receives a continuous and constant dividend yield, D. Thus in a time dt each asset receives an amount $DS\,dt$. This must be factored into the derivation of the Black–Scholes equation. I take up the Black–Scholes argument at the point where we are looking at the change in the value of the portfolio:

$$d\Pi=\frac{\partial V}{\partial t}\,dt+\frac{\partial V}{\partial S}\,dS+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}\,dt-\Delta\,dS-D\Delta S\,dt.$$

The last term on the right-hand side is simply the amount of the dividend per asset, $DS\,dt$, multiplied by the number of the asset held, $-\Delta$. The Δ is still given by the rate of change of the option value with respect to the underlying, but after some simple substitutions we now get

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}+(r-D)S\frac{\partial V}{\partial S}-rV=0. \qquad (5.7)$$

5.9 CURRENCY OPTIONS

Options on currencies are handled in exactly the same way. In holding the foreign currency we receive interest at the foreign rate of interest r_f. This is just like receiving a continuous dividend. I will skip the derivation but we readily find that

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}+(r-r_f)S\frac{\partial V}{\partial S}-rV=0. \qquad (5.8)$$

5.10 COMMODITY OPTIONS

The relevant feature of commodities requiring that we adjust the Black–Scholes equation is that they have a **cost of carry**. That is, the storage of commodities is not without cost. Let us introduce q as the fraction of the value of a commodity that goes towards paying the cost of carry. This means that just holding the commodity will result in a gradual loss of wealth even if the commodity price remains fixed. To be precise, for each unit of the commodity held an amount $qS\,dt$ will be required during short time dt to finance the holding. This is just like having a negative dividend and so we get

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r+q)S\frac{\partial V}{\partial S} - rV = 0. \tag{5.9}$$

5.11 OPTIONS ON FUTURES

The final modification to the Black–Scholes model in this chapter is to value options on futures. Recalling that the future price of a non-dividend paying equity F is related to the spot price by

$$F = e^{r(T_F - t)} S$$

where T_F is the maturity date of the futures contract. We can easily change variables, and look for a solution $V(S,t) = \mathcal{V}(F,t)$. We find that

$$\frac{\partial \mathcal{V}}{\partial t} + \tfrac{1}{2}\sigma^2 F^2 \frac{\partial^2 \mathcal{V}}{\partial F^2} - r\mathcal{V} = 0. \tag{5.10}$$

The equation for an option on a future is actually simpler than the Black–Scholes equation.

5.12 SOME OTHER WAYS OF DERIVING THE BLACK–SCHOLES EQUATION

The derivation of the Black–Scholes equation above is the classical one, and similar to the original Black and Scholes derivation. There are other ways of getting to the same result. Here are a few, without any of the details. The details, and more examples, are contained in the final reference in the Further Reading.

5.12.1 The Martingale Approach

The value of an option can be shown to be an expectation, not a real expectation but a special, risk-neutral one. We'll be seeing lots of this subject later. This is a useful result, since it forms the basis for pricing by simulation, see Chapter 80. The partial differential equation can be derived from the expectation, see Chapter 10. The concepts of hedging and no arbitrage are obviously still used in this derivation.

5.12.2 The Binomial Model

The binomial model is a discrete time, discrete asset price model for underlyings and again uses hedging and no arbitrage to derive a pricing algorithm for options. We shall see this in detail in Chapter 15. In taking the limit as the time step shrinks to zero we get the continuous-time Black–Scholes equation.

5.12.3 CAPM/Utility

Again, we'll be seeing the Capital Asset Pricing Model later, for the moment you just need to know that it is a model for the behavior of risky assets and a principle and algorithm for defining and finding optimal ways to allocate wealth among the assets. Portfolios are described in terms of their risk (standard deviation of returns) and reward (expected growth). If you include options in this framework then the possible combinations of risk and reward are not increased. This is because options are, in a sense, just functions of their underlyings. This is market completeness. The risk and reward on an option and on its underlying are related and the Black–Scholes equation follows.

5.13 SUMMARY

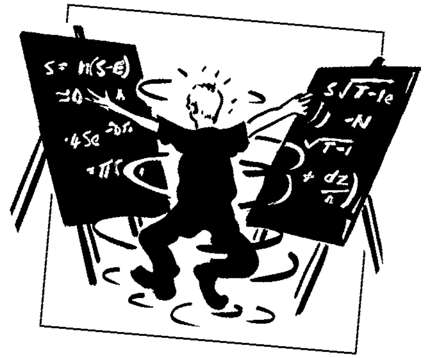
This was an important but not too difficult chapter. In it I introduced some very powerful and beautiful concepts such as delta hedging and no arbitrage. These two fundamental principles led to the Black–Scholes option pricing equation. Everything from this point on is based on, or is inspired by, these ideas.

FURTHER READING

- The history of option theory, leading up to Black–Scholes is described in Briys, Mai, Bellalah & de Varenne (1998).
- The story of the derivation of the Black–Scholes equation, written by Bob Whaley, can be found in the 10th anniversary issue of *Risk* magazine, published in December 1997.
- Of course, you must read the original work, Black & Scholes (1973) and Merton (1973).
- See Black (1976) for the details of the pricing of options on futures, and Garman & Kohlhagen (1983) for the pricing of FX options.
- For details of other ways to derive the Black–Scholes equation, see Andreasen, Jensen & Poulson (1998).

CHAPTER 6

partial differential equations



In this Chapter...

- properties of the parabolic partial differential equation
- the meaning of terms in the Black–Scholes equation
- some solution techniques

6.1 INTRODUCTION

The analysis and solution of partial differential equations is a BIG subject. We can only skim the surface in this book. If you don't feel comfortable with the subject, then the list of books at the end should be of help. However, to understand finance, and even to solve partial differential equations numerically, does not require any great depth of understanding. The aim of this chapter is to give just enough background to the subject to permit any reasonably numerate person to follow the rest of the book; I want to keep the entry requirements to the subject as low as possible.

6.2 PUTTING THE BLACK–SCHOLES EQUATION INTO HISTORICAL PERSPECTIVE

The Black–Scholes partial differential equation is in two dimensions, S and t. It is a parabolic equation, meaning that it has a second derivative with respect to one variable, S, and a first derivative with respect to the other, t. Equations of this form are more colloquially known as **heat** or **diffusion equations**.

The equation, in its simplest form, goes back to almost the beginning of the 19th century. Diffusion equations have been successfully used to model

- diffusion of one material within another, such as smoke particles in air
- flow of heat from one part of an object to another

- chemical reactions, such as the Belousov–Zhabotinsky reaction which exhibits fascinating wave structure
- electrical activity in the membranes of living organisms, such as the Hodgkin–Huxley model
- dispersion of populations, such as individuals moving both randomly and to avoid overcrowding
- pursuit and evasion in predator-prey systems
- pattern formation in animal coats, such as the formation of zebra stripes
- dispersion of pollutants in a running stream

In most of these cases the resulting equations are more complicated than the Black–Scholes equation.

The simplest heat equation for the temperature in a bar is usually written in the form

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{6.1}$$

where u is the temperature, x is a spatial coordinate and t is time. This equation comes from a heat balance. Consider the flow into and out of a small section of the bar. The flow of heat along the bar is proportional to the spatial gradient of the temperature

$$\frac{\partial u}{\partial x}$$

and thus the derivative of this, the *second* derivative of the temperature, $\partial^2 u/\partial x^2$, is the heat retained by the small section. This retained heat is seen as a rise in the temperature, represented mathematically by

$$\frac{\partial u}{\partial t}.$$

The balance of the second x-derivative and the first t-derivative results in the heat equation, Equation (6.1). (There would be a coefficient in the equation, depending on the properties of the bar, but I have set this to one.)

6.3 THE MEANING OF THE TERMS IN THE BLACK–SCHOLES EQUATION

The Black–Scholes equation can be accurately interpreted as a reaction-convection-diffusion equation.

The basic diffusion equation is a balance of a first-order t derivative and a second-order S derivative:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}.$$

If these were the only terms in the Black–Scholes equation it would still exhibit the smoothing-out effect, that any discontinuities in the payoff would be instantly diffused away. The only difference between these terms and the terms as they appear in the basic heat or diffusion equation, is that the diffusion coefficient is a function of one of the variables S. Thus we really have diffusion in a nonhomogeneous medium.

The first-order S-derivative term

$$rS\frac{\partial V}{\partial S}$$

can be thought of as a convection term. If this equation represented some physical system, such as the diffusion of smoke particles in the atmosphere, then the convective term would be due to a breeze, blowing the smoke in a preferred direction.

The final term

$$-rV$$

is a reaction term. Balancing this term and the time derivative would give a model for decay of a radioactive body, with the half-life being related to r.

Figure 6.1 shows how a call option payoff would evolve in three cases: (i) diffusion only, (ii) convection only and (iii) reaction only.

Putting these terms together we get a reaction-convection-diffusion equation. An almost identical equation would be arrived at for the dispersion of pollutant along a flowing river with absorption by the sand. In this, the dispersion is the diffusion, the flow is the convection, and the absorption is the reaction.

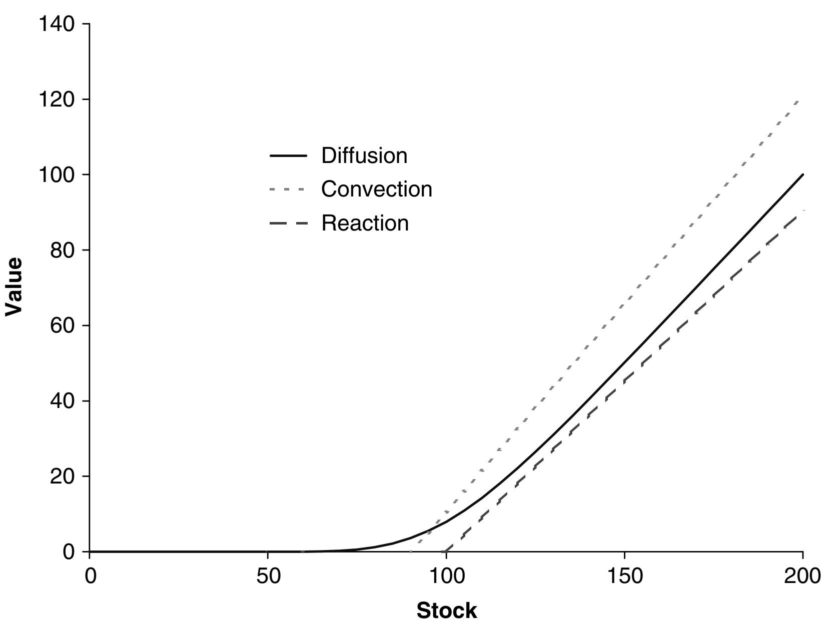


Figure 6.1 Call option value before expiration: (i) diffusion only, (ii) convection only and (iii) reaction only.

6.4 BOUNDARY AND INITIAL/FINAL CONDITIONS

To specify a problem uniquely we must prescribe **boundary conditions** and an **initial** or **final condition**. Boundary conditions tell us how the solution must behave for all time at certain values of the asset. In financial problems we usually specify the behavior of the solution at $S = 0$ and as $S \to \infty$. We must also tell the problem how the solution begins. The Black–Scholes equation is a backward equation, meaning that the signs of the t derivative and the second S derivative in the equation are the same when written on the same side of the equals sign. We therefore have to impose a final condition. This is usually the payoff function at expiry.

The Black–Scholes equation in its basic form is linear and homogeneous, and therefore satisfies the superposition principle; add together two solutions of the equation and you will get a third. This is not true of non-linear equations. Linear diffusion equations have some very nice properties. Even if we start out with a discontinuity in the final data, due to a discontinuity in the payoff, this *immediately* gets smoothed out; this is due to the diffusive nature of the equation. Another nice property is the uniqueness of the solution. Provided that the solution is not allowed to grow too fast as S tends to infinity the solution will be unique. This precise definition of 'too fast' need not worry us, as we will not have to worry about uniqueness for any problems we encounter.

6.5 SOME SOLUTION METHODS

We are not going to spend much time on the exact solution of the Black–Scholes equation. Such a solution is important, but current market practice is such that models have features which preclude the exact solution. The few explicit, closed-form solutions that are used by practitioners will be covered in the next two chapters.

6.5.1 Transformation to Constant Coefficient Diffusion Equation

It can sometimes be useful to transform the basic Black–Scholes equation into something a little bit simpler by a change of variables. If we write

$$V(S, t) = e^{\alpha x + \beta \tau} U(x, \tau),$$

where

$$\alpha = -\tfrac{1}{2}\left(\frac{2r}{\sigma^2} - 1\right), \quad \beta = -\tfrac{1}{4}\left(\frac{2r}{\sigma^2} + 1\right)^2, \quad S = e^x \quad \text{and} \quad t = T - \frac{2\tau}{\sigma^2},$$

then $U(x, \tau)$ satisfies the basic diffusion equation

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2}. \tag{6.2}$$

This simpler equation is easier to handle than the Black–Scholes equation. Sometimes that can be important, for example when seeking closed-form solutions, or in some simple numerical schemes. We shall not pursue this any further.

6.5.2 Green's Functions

One solution of the Black–Scholes equation is

$$V'(S,t) = \frac{e^{-r(T-t)}}{\sigma S'\sqrt{2\pi(T-t)}} e^{-\left(\log(S/S')+(r-\frac{1}{2}\sigma^2)(T-t)\right)^2/2\sigma^2(T-t)} \tag{6.3}$$

for any S'. (You can verify this by substituting back into the equation, but we'll also be seeing it derived in the next chapter.) This solution is special because as $t \to T$ it becomes zero everywhere, except at $S = S'$. In this limit the function becomes what is known as a **Dirac delta function**. Think of this as a function that is zero everywhere except at one point where it is infinite, in such a way that its integral is one. How is this be of help to us?

Expression (6.3) is a solution of the Black–Scholes equation for any S'. Because of the superposition principle we can multiply (6.3) by any constant, and we get another solution. But then we can also get another solution by adding together expressions of the form (6.3) but with different values for S'. Putting this together, and thinking of an integral as just a way of adding together many solutions, we find that

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_0^\infty e^{-\left(\log(S/S')+(r-\frac{1}{2}\sigma^2)(T-t)\right)^2/2\sigma^2(T-t)} f(S')\frac{dS'}{S'}$$

is also a solution of the Black–Scholes equation for any function $f(S')$. (If you don't believe me, substitute it into the Black–Scholes equation.)

Because of the nature of the integrand as $t \to T$ (i.e. that it is zero everywhere except at S' and has integral one), if we choose the arbitrary function $f(S')$ to be the payoff function then this expression becomes the solution of the problem:

$$V(S,t) = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_0^\infty e^{-\left(\log(S/S')+(r-\frac{1}{2}\sigma^2)(T-t)\right)^2/2\sigma^2(T-t)} \,\mathrm{Payoff}(S')\frac{dS'}{S'}.$$

The function $V'(S,t)$ given by (6.3) is called the **Green's function**.

6.5.3 Series Solution

Sometimes we have boundary conditions at two finite (and non-zero) values of S, S_u and S_d, say (we see examples in Chapter 23). For this type of problem, we postulate that the required solution of the Black–Scholes equation can be written as an infinite sum of special functions. First of all, transform to the nicer basic diffusion equation in x and τ. Now write the solution as

$$e^{\alpha x+\beta\tau}\sum_{i=0}^{\infty} a_i(\tau)\sin(i\omega x) + b_i(\tau)\cos(i\omega x),$$

for some ω and some functions a and b to be found. The linearity of the equation suggests that a sum of solutions might be appropriate. If this is to satisfy the Black–Scholes equation then we must have

$$\frac{da_i}{d\tau} = -i^2\omega^2 a_i(\tau) \quad\text{and}\quad \frac{db_i}{d\tau} = -i^2\omega^2 b_i(\tau).$$

You can easily show this by substitution. The solutions are thus

$$a_i(\tau) = A_i e^{-i^2\omega^2\tau} \quad \text{and} \quad b_i(\tau) = B_i e^{-i^2\omega^2\tau}.$$

The solution of the Black–Scholes equation is therefore

$$e^{\alpha x+\beta\tau} \sum_{i=0}^{\infty} e^{-i^2\omega^2\tau} \left(A_i \sin(i\omega x) + B_i \cos(i\omega x)\right). \tag{6.4}$$

We have solved the equation; all that we need to do now is to satisfy boundary and initial conditions.

Consider the example where the payoff at time $\tau = 0$ is $f(x)$ (although it would be expressed in the original variables, of course) but the contract becomes worthless if ever $x = x_d$ or $x = x_u$.[1]

Rewrite the term in large brackets in (6.4) as

$$C_i \sin\left(i\omega' \frac{x - x_d}{x_u - x_d}\right) + D_i \cos\left(i\omega' \frac{x - x_d}{x_u - x_d}\right).$$

To ensure that the option is worthless on these two x values, choose $D_i = 0$ and $\omega' = \pi$. The boundary conditions are thereby satisfied. All that remains is to choose the C_i to satisfy the final condition:

$$e^{\alpha x} \sum_{i=0}^{\infty} C_i \sin\left(i\omega' \frac{x - x_d}{x_u - x_d}\right) = f(x).$$

This also is simple. Multiplying both sides by

$$\sin\left(j\omega' \frac{x - x_d}{x_u - x_d}\right),$$

and integrating between x_d and x_u we find that

$$C_j = \frac{2}{x_u - x_d} \int_{x_d}^{x_u} f(x) e^{-\alpha x} \sin\left(j\omega' \frac{x - x_d}{x_u - x_d}\right) dx.$$

This technique, which can be generalized, is the **Fourier series method**. There are some problems with the method if you are trying to represent a discontinuous function with a sum of trigonometrical functions. The oscillatory nature of an approximate solution with a finite number of terms is known as **Gibbs phenomenon**.

6.6 SIMILARITY REDUCTIONS

Apart from the Green's function, we're not going to use any of the above techniques in this book; rarely will we even find explicit solutions. But one technique that we will find useful is the **similarity reduction**. I will demonstrate the idea using the simple diffusion equation; we will later use it in many other, more complicated problems.

[1] This is an example of a double knock-out option, see Chapter 23.

The basic diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{6.5}$$

is an equation for the function u which depends on the two variables x and t. Sometimes, in very, very special cases we can write the solution as a function of just one variable. Let me give an example.

The function

$$u(x,t) = \int_0^{x/t^{1/2}} e^{-\frac{1}{4}\xi^2} d\xi$$

satisfies (6.5). This is easy to verify. But in this function x and t only appear in the combination

$$\frac{x}{t^{1/2}}.$$

Thus, in a sense, u is a function of only one variable.

A slight generalization, but also demonstrating the idea of similarity solutions, is to look for a solution of the form

$$u = t^{-1/2} f(\xi) \tag{6.6}$$

where

$$\xi = \frac{x}{t^{1/2}}.$$

Substitute (6.6) into (6.5) to find that a solution for f is

$$f = e^{-\frac{1}{4}\xi^2},$$

so that

$$t^{-1/2} e^{-(1/4)(x^2/t)}$$

is also a special solution of the diffusion equation.

Be warned, though. You can't always find similarity solutions; not only must the equation have a particularly nice structure but also the similarity form must be consistent with any initial condition or boundary conditions.

6.7 OTHER ANALYTICAL TECHNIQUES

The other two main solution techniques for linear partial differential equations are Fourier and Laplace transforms. These are such large and highly technical subjects that I really cannot begin to give an idea of how they work, space is far too short. Recently these techniques have become useful in some higher dimensional models such as those incorporating stochastic volatility.

6.8 NUMERICAL SOLUTION

Even though there are several techniques that we can use for finding solutions, in the vast majority of cases we must solve the Black–Scholes equation numerically. But we are lucky. Parabolic differential equations are just about the easiest equations to solve numerically. Obviously, there are any number of really sophisticated techniques, but if you stick with the simplest then you can't go far wrong. In Chapters 15, 77 and 78 we discuss these methods in detail. I want to stress that I am going to derive many partial differential equations from now on, and I am going to assume you trust me that we will at the end of the book see how to solve them.

6.9 SUMMARY

This short chapter is only intended as a primer on partial differential equations. If you want to study this subject in depth, see the books and articles mentioned below.

FURTHER READING

- Grindrod (1991) is all about reaction-diffusion equations, where they come from and their analysis. The book includes many of the physical models described above.
- Murray (1989) also contains a great deal on reaction-diffusion equations, but concentrates on models of biological systems.
- Wilmott & Wilmott (1990) describe the diffusion of pollutant along a river with convection and absorption by the river bed.
- The classical reference works for diffusion equations are Crank (1989) and Carslaw & Jaeger (1989). But also see the book on partial differential equations by Sneddon (1957) and the book on general applied mathematical methods by Strang (1986).

CHAPTER 7

the Black–Scholes formulae and the 'greeks'

In this Chapter...

- the derivation of the Black–Scholes formulae for calls, puts and simple digitals
- the meaning and importance of the 'greeks,' delta, gamma, theta, vega and rho
- the difference between differentiation with respect to variables and to parameters
- formulae for the greeks for calls, puts and simple digitals

7.1 INTRODUCTION

The Black–Scholes equation has simple solutions for calls, puts and some other contracts. In this chapter I'm going to walk you through the derivation of these formulae step by step. This is one of the few places in the book where I do derive formulae. The reason that I don't often derive formulae is that the majority of contracts do not have explicit solutions for their theoretical value. Instead much of my emphasis will be placed on finding numerical solutions of the Black–Scholes equation.

We've seen how the quantity 'delta,' the first derivative of the option value with respect to the underlying, occurs as an important quantity in the derivation of the Black–Scholes equation. In this chapter I describe the importance of other derivatives of the option price, with respect to the variables (the underlying asset and time) and with respect to some of the parameters. These derivatives are important in the hedging of an option position, playing key roles in risk management. It can be argued that it is more important to get the hedging correct than to be precise in the pricing of a contract. The reason for this is that if you are accurate in your hedging you will have reduced or eliminated future uncertainty. This leaves you with a profit (or loss) that is set the moment that you buy or sell the contract. But if your hedging is inaccurate, then it doesn't matter, within reason, what you sold the contract for initially; future uncertainty could easily dominate any initial profit. Of course, life is not so simple, in reality we are exposed to model error, which can make a mockery of anything we do. However, this illustrates the importance of good hedging, and that's where the 'greeks' come in.

7.2 DERIVATION OF THE FORMULAE FOR CALLS, PUTS AND SIMPLE DIGITALS

The Black–Scholes equation is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{7.1}$$

This equation must be solved with final condition depending on the payoff; each contract will have a different functional form prescribed at expiry $t = T$, depending on whether it is a call, a put or something more fancy. This is the final condition that must be imposed to make the solution unique. We'll worry about final conditions later, for the moment concentrate on manipulating (7.1) into something we can easily solve.

The first step in the manipulation is to change from present value to future value terms. Recalling that the payoff is received at time T but that we are valuing the option at time t this suggests that we write

$$V(S, t) = e^{-r(T-t)}U(S, t).$$

This takes our differential equation to

$$\frac{\partial U}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S} = 0.$$

The second step is really trivial. Because we are solving a backward equation, discussed in Chapter 6, we'll write

$$\tau = T - t.$$

This now takes our equation to

$$\frac{\partial U}{\partial \tau} = \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + rS\frac{\partial U}{\partial S}.$$

When we first started modeling equity prices we used intuition about the asset price *return* to build up the stochastic differential equation model. Let's go back to examine the return and write

$$\xi = \log S.$$

With this as the new variable, we find that

$$\frac{\partial}{\partial S} = e^{-\xi}\frac{\partial}{\partial \xi} \quad \text{and} \quad \frac{\partial^2}{\partial S^2} = e^{-2\xi}\frac{\partial^2}{\partial \xi^2} - e^{-2\xi}\frac{\partial}{\partial \xi}.$$

Now the Black–Scholes equation becomes

$$\frac{\partial U}{\partial \tau} = \tfrac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial \xi^2} + \left(r - \tfrac{1}{2}\sigma^2\right)\frac{\partial U}{\partial \xi}.$$

What has this done for us? It has taken the problem defined for $0 \le S < \infty$ to one defined for $-\infty < \xi < \infty$. But more importantly, the coefficients in the equation are now all constant, independent of the underlying. This is a big step forward, made possible by the lognormality of the underlying asset. We are nearly there.

The last step is simple, but the motivation is not so obvious. Write

$$x = \xi + \left(r - \tfrac{1}{2}\sigma^2\right)\tau,$$

and $U = W(x, \tau)$. This is just a 'translation' of the coordinate system. It's a bit like using the forward price of the asset instead of the spot price as a variable. After this change of variables the Black–Scholes becomes the simpler

$$\frac{\partial W}{\partial \tau} = \tfrac{1}{2}\sigma^2 \frac{\partial^2 W}{\partial x^2}. \tag{7.2}$$

To summarize,

$$\begin{aligned} V(S, t) &= e^{-r(T-t)}U(S, t) = e^{-r\tau}U(S, T-\tau) = e^{-r\tau}U(e^{\xi}, T-\tau) \\ &= e^{-r\tau}U\left(e^{x-\left(r-\frac{1}{2}\sigma^2\right)\tau}, T-\tau\right) = e^{-r\tau}W(x, \tau). \end{aligned}$$

To those of you who already know the Black–Scholes formulae for calls and puts the variable x will ring a bell:

$$x = \xi + \left(r - \tfrac{1}{2}\sigma^2\right)\tau = \log S + \left(r - \tfrac{1}{2}\sigma^2\right)(T-t).$$

Having turned the original Black–Scholes equation into something much simpler, let's take a break for a moment while I explain where we are headed.

I'm going to derive an expression for the value of any option whose payoff is a known function of the asset price at expiry. This includes calls, puts and digitals. This expression will be in the form of an integral. For special cases, I'll show how to rewrite this integral in terms of the cumulative distribution function for the Normal distribution. This is particularly useful since the function can be found on spreadsheets, calculators and in the backs of books. But there are two steps before I can write down this integral.

The first step is to find a special solution of (7.2), called the fundamental solution. This solution has useful properties. The second step is to use the linearity of the equation and the useful properties of the special solution to find the *general solution* of the equation. Here we go.

I'm going to look for a special solution of (7.2) of the following form

$$W(x, \tau) = \tau^{\alpha} f\left(\frac{(x - x')}{\tau^{\beta}}\right), \tag{7.3}$$

where x' is an arbitrary constant. And I'll call this special solution $W_f(x, \tau; x')$. Note that the unknown function depends on only *one* variable $(x - x')/\tau^{\beta}$. As well as finding the function f we must find the constant parameters α and β. We can expect that if this approach works, the equation for f will be an ordinary differential equation since the function only has one variable. This reduction of dimension is an example of a similarity reduction, discussed in Chapter 6.

Substituting expression (7.3) into (7.2) we get

$$\tau^{\alpha-1}\left(\alpha f - \beta\eta \frac{df}{d\eta}\right) = \tfrac{1}{2}\sigma^2 \tau^{\alpha-2\beta}\frac{d^2 f}{d\eta^2}, \tag{7.4}$$

where

$$\eta = \frac{x - x'}{\tau^{\beta}}.$$

Examining the dependence of the two terms in (7.4) on both τ and η we see that we can only have a solution if

$$\alpha - 1 = \alpha - 2\beta \quad \text{i.e.} \quad \beta = \tfrac{1}{2}.$$

I want to ensure that my 'special solution' has the property that its integral over all ξ is independent of τ, for reasons that will become apparent. To ensure this, I require

$$\int_{-\infty}^{\infty} \tau^{\alpha} f((x - x')/\tau^{\beta})\, dx$$

to be constant. I can write this as

$$\int_{-\infty}^{\infty} \tau^{\alpha+\beta} f(\eta)\, d\eta$$

and so I need

$$\alpha = -\beta = -\tfrac{1}{2}.$$

The function f now satisfies

$$-f - \eta \frac{df}{d\eta} = \sigma^2 \frac{d^2 f}{d\eta^2}.$$

This can be written

$$\sigma^2 \frac{d^2 f}{d\eta^2} + \frac{d(\eta f)}{d\eta} = 0,$$

which can be integrated once to give

$$\sigma^2 \frac{df}{d\eta} + \eta f = a,$$

where a is a constant. For my special solution I'm going to choose $a = 0$. This equation can be integrated again to give

$$f(\eta) = b e^{-\eta^2/(2\sigma^2)}.$$

I will choose the constant b such that the integral of f from minus infinity to plus infinity is one:

$$f(\eta) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\eta^2/(2\sigma^2)}.$$

This is the special solution I have been seeking:[1]

$$W_f(x, \tau) = \frac{1}{\sqrt{2\pi\tau}\,\sigma} e^{-((x-x')^2/2\sigma^2\tau)}.$$

Now I will explain why it is useful in our quest for the Black–Scholes formulae.

[1] It is just the probability density function for a Normal random variable with mean zero and standard deviation σ.

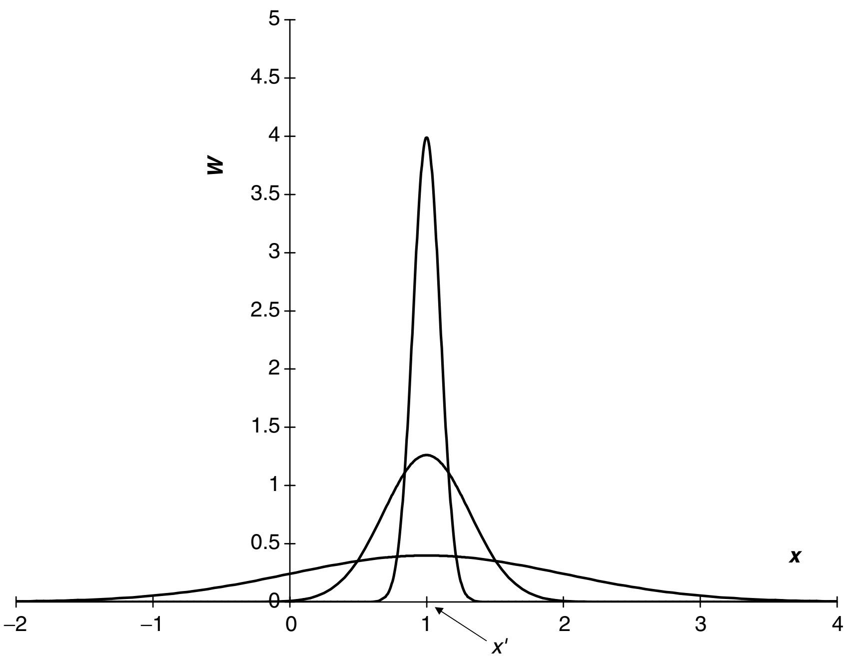


Figure 7.1 The fundamental solution.

In Figure 7.1 is plotted W_f as a function of x' for several values of τ. Observe how the function rises in the middle but decays at the sides. As $\tau \to 0$ this becomes more pronounced. The 'middle' is the point $x' = x$. At this point the function grows unboundedly, and away from this point the function decays to zero as $\tau \to 0$. Although the function is increasingly confined to a narrower and narrower region its area remains fixed at one. These properties of decay away from one point, unbounded growth at that point and constant area, result in a **Dirac delta function** $\delta(x' - x)$ as $\tau \to 0$. The delta function has one important property, namely

$$\int \delta(x' - x)g(x')\,dx' = g(x)$$

where the integration is from any point below x to any point above x. Thus the delta function 'picks out' the value of g at the point where the delta function is singular i.e. at $x' = x$. In the limit as $\tau \to 0$ the function W_f becomes a delta function at $x = x'$. This means that

$$\lim_{\tau \to 0} \frac{1}{\sigma\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} e^{-\frac{(x'-x)^2}{2\sigma^2\tau}} g(x')\,dx' = g(x).$$

This property of the special solution, together with the linearity of the Black–Scholes equation, are all that are needed to find some explicit solutions.

Now is the time to consider the payoff. Let's call it

$$\text{Payoff}(S).$$

This is the value of the option at time $t = T$. It is the final condition for the function V, satisfying the Black–Scholes equation:

$$V(S, T) = \text{Payoff}(S).$$

With our new variables, this final condition is

$$W(x, 0) = \text{Payoff}(e^x). \tag{7.5}$$

I claim that the solution of this for $\tau > 0$ is

$$W(x, \tau) = \int_{-\infty}^{\infty} W_f(x, \tau; x')\text{Payoff}(e^{x'})\, dx'. \tag{7.6}$$

To show this, I just have to demonstrate that the expression satisfies the equation (7.2) and the final condition (7.5). Both of these are straightforward. The integration with respect to x' is similar to a summation, and since each individual component satisfies the equation so does the sum/integral. Alternatively, differentiate (7.6) under the integral sign to see that it satisfies the partial differential equation. That it satisfies the condition (7.5) follows from the special properties of the fundamental solution W_f.

Retracing our steps to write our solution in terms of the original variables, we get

$$V(S, t) = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_0^{\infty} e^{-\left(\log(S/S')+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2 \Big/ 2\sigma^2(T-t)}\,\text{Payoff}(S')\frac{dS'}{S'}, \tag{7.7}$$

where I have written $x' = \log S'$.

This is the exact solution for the option value in terms of the arbitrary payoff function. In the next sections I will manipulate this expression for special payoff functions.

7.2.1 Formula for a Call

The call option has the payoff function

$$\text{Payoff}(S) = \max(S - E, 0).$$

Expression (7.7) can then be written as

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_E^{\infty} e^{-\left(\log(S/S')+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2 / 2\sigma^2(T-t)}(S' - E)\frac{dS'}{S'}.$$

Return to the variable $x' = \log S'$, to write this as

$$\begin{aligned}
&\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^{\infty} e^{-\left(-x'+\log S+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2 / 2\sigma^2(T-t)}(e^{x'} - E)\, dx' \\
&\quad = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^{\infty} e^{-\left(-x'+\log S+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2 / 2\sigma^2(T-t)} e^{x'}\, dx' \\
&\qquad - E\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_{\log E}^{\infty} e^{-\left(-x'+\log S+\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2 / 2\sigma^2(T-t)}\, dx'.
\end{aligned}$$

Both integrals in this expression can be written in the form

$$\int_d^{\infty} e^{-\frac{1}{2}x'^2}\,dx'$$

for some d (the second is just about in this form already, and the first just needs a completion of the square).

Apart from a couple of minor differences, this integral is just like the cumulative distribution function for the standardized Normal distribution[2] defined by

$$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}\phi^2}\,d\phi.$$

This function, plotted in Figure 7.2, is the probability that a Normally distributed variable is less than x.

Thus the option price can be written as two separate terms involving the cumulative distribution function for a Normal distribution:

$$\text{Call option value} = SN(d_1) - Ee^{-r(T-t)}N(d_2)$$

where

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

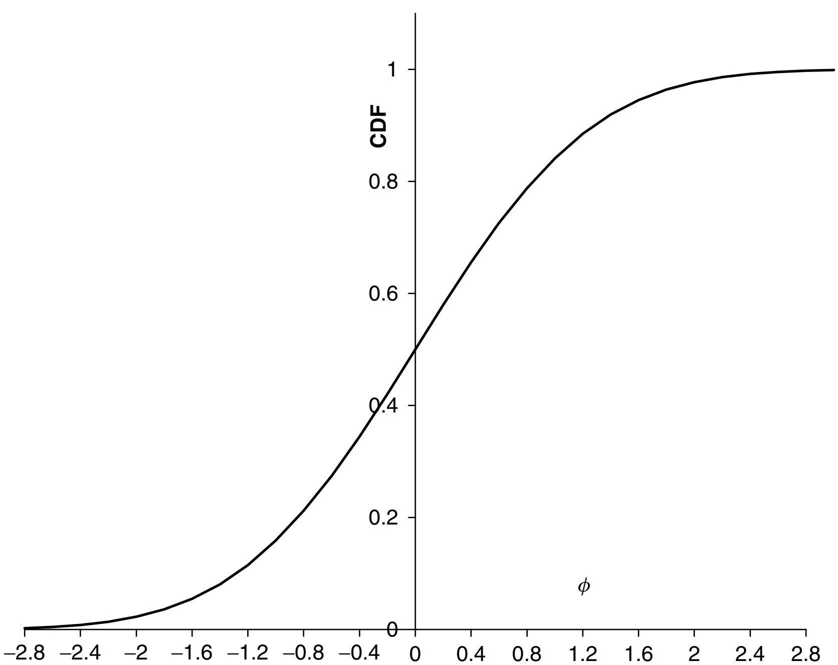


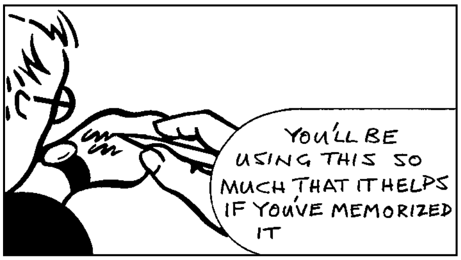
Figure 7.2 The cumulative distribution function for a standardized Normal random variable, $N(x)$.

[2] i.e. having zero mean and unit standard deviation.

and

$$d_2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

When there is continuous dividend yield on the underlying, or it is a currency, then



Call option value

$$Se^{-D(T-t)}N(d_1) - Ee^{-r(T-t)}N(d_2)$$

$$d_1 = \frac{\log(S/E) + (r - D + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\log(S/E) + (r - D - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$= d_1 - \sigma\sqrt{T-t}$$

The option value is shown in Figure 7.3 as a function of the underlying asset at a fixed time to expiry. In Figure 7.4 the value of the at-the-money option is shown as a function of time, and expiry is $t = 1$. In Figure 7.5 is the call value as a function of both the underlying and time.

When the asset is 'at-the-money forward,' i.e. $S = Ee^{-(r-D)(T-t)}$, then there is a simple approximation for the call value (Brenner & Subrahmanyam, 1994):

$$\text{Call} \approx 0.4\, Se^{-D(T-t)}\sigma\sqrt{T-t}.$$

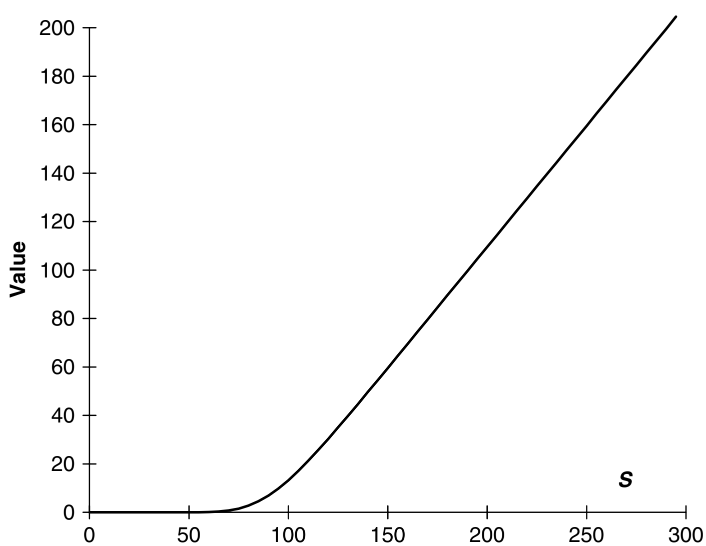


Figure 7.3 The value of a call option as a function of the underlying asset price at a fixed time to expiry.

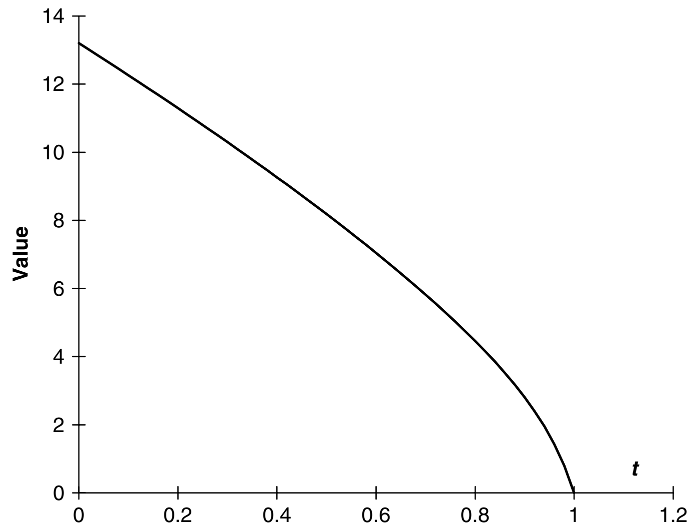


Figure 7.4 The value of an at-the-money call option as a function of time.

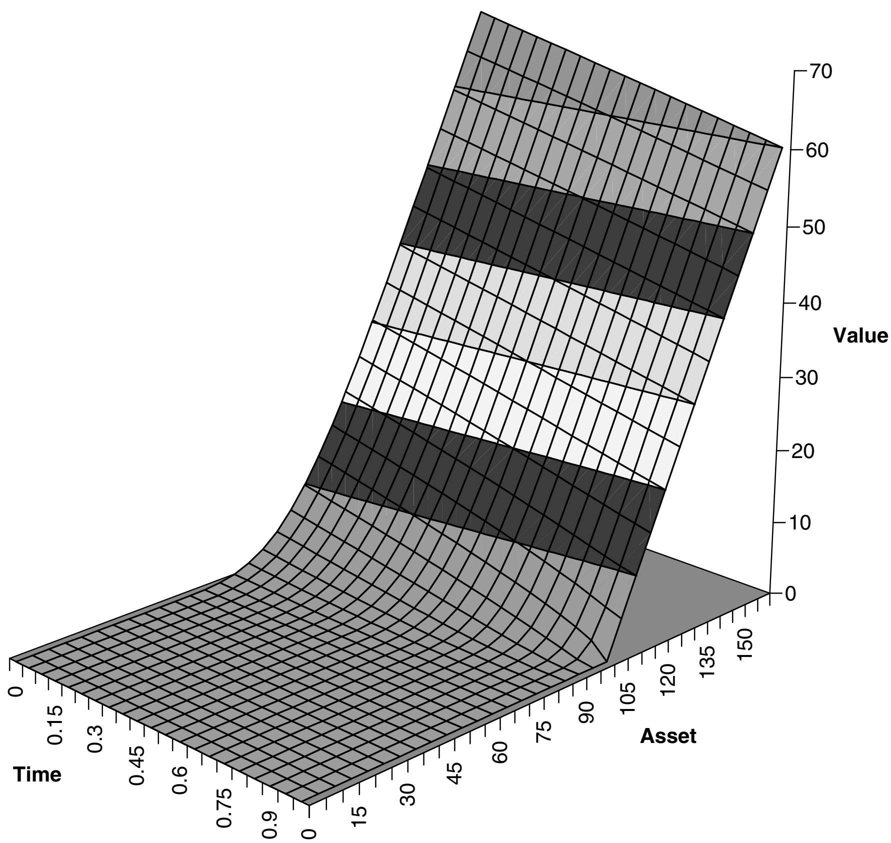


Figure 7.5 The value of a call option as a function of asset and time.

7.2.2 Formula for a Put

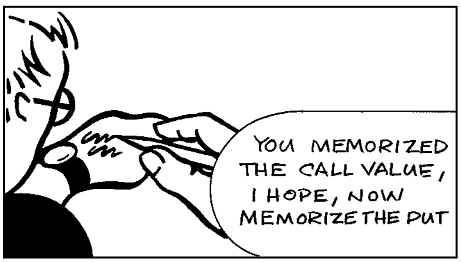
The put option has payoff

$$\text{Payoff}(S) = \max(E - S, 0).$$

The value of a put option can be found in the same way as above, or using put-call parity

$$\text{Put option value} = -SN(-d_1) + Ee^{-r(T-t)}N(-d_2),$$

with the same d_1 and d_2.



When there is continuous dividend yield on the underlying, or it is a currency, then

Put option value

$$-Se^{-D(T-t)}N(-d_1) + Ee^{-r(T-t)}N(-d_2)$$

The option value is shown in Figure 7.6 against the underlying asset and in Figure 7.7 against time. In Figure 7.8 is the option value as a function of both the underlying asset and time.

When the asset is at-the-money forward the simple approximation for the put value (Brenner & Subrahmanyam, 1994) is

$$\text{Put} \approx 0.4\ Se^{-D(T-t)}\sigma\sqrt{T-t}.$$

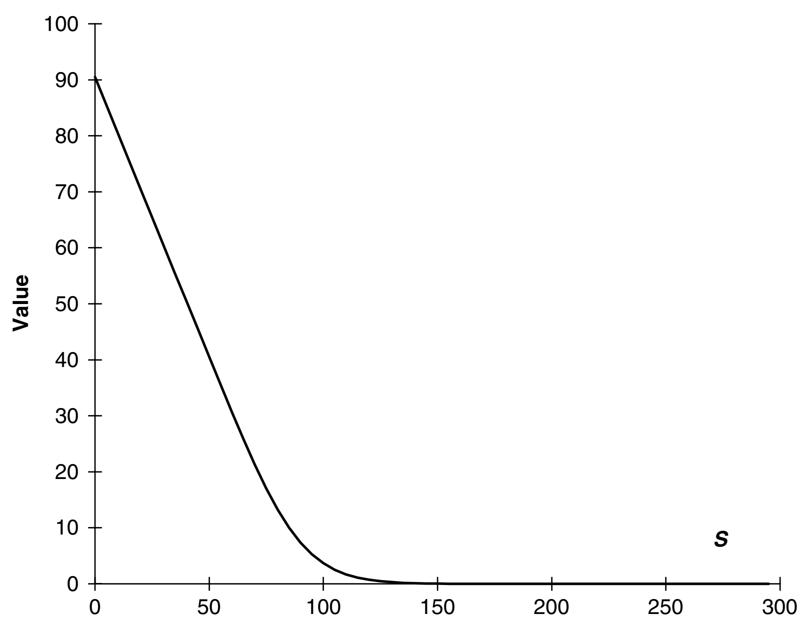


Figure 7.6 The value of a put option as a function of the underlying asset at a fixed time to expiry.

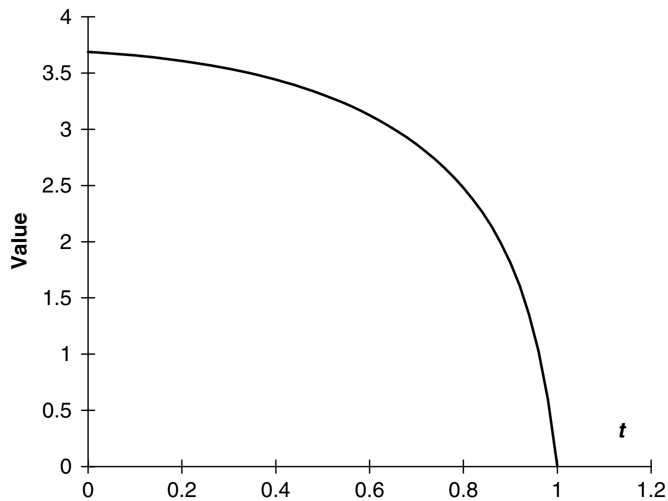


Figure 7.7 The value of an at-the money put option as a function of time.

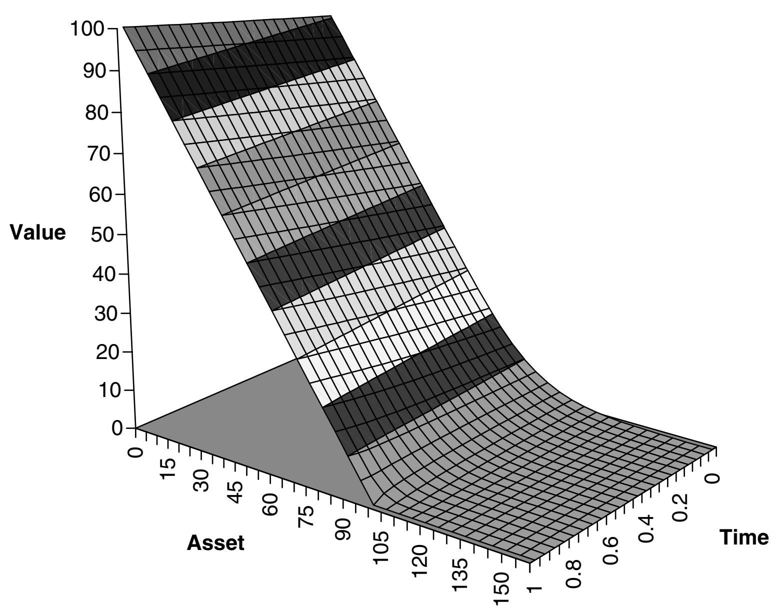


Figure 7.8 The value of a put option as a function of asset and time.

7.2.3 Formula for a Binary Call

The binary call has payoff

$$\text{Payoff}(S) = \mathcal{H}(S - E),$$

where $\mathcal{H}$ is the Heaviside function taking the value one when its argument is positive and zero otherwise.

Incorporating a dividend yield, we can write the option value as

$$\frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}}\int_{\log E}^{\infty} e^{-\left(x'-\log S-\left(r-D-\frac{1}{2}\sigma^2\right)(T-t)\right)^2/2\sigma^2(T-t)}\,dx'.$$

This term is just like the second term in the call option equation and so

Binary call option value

$$e^{-r(T-t)}N(d_2)$$

The option value is shown in Figure 7.9.

7.2.4 Formula for a Binary Put

The binary put has a payoff of one if $S < E$ at expiry. It has a value of

Binary put option value

$$e^{-r(T-t)}(1-N(d_2))$$

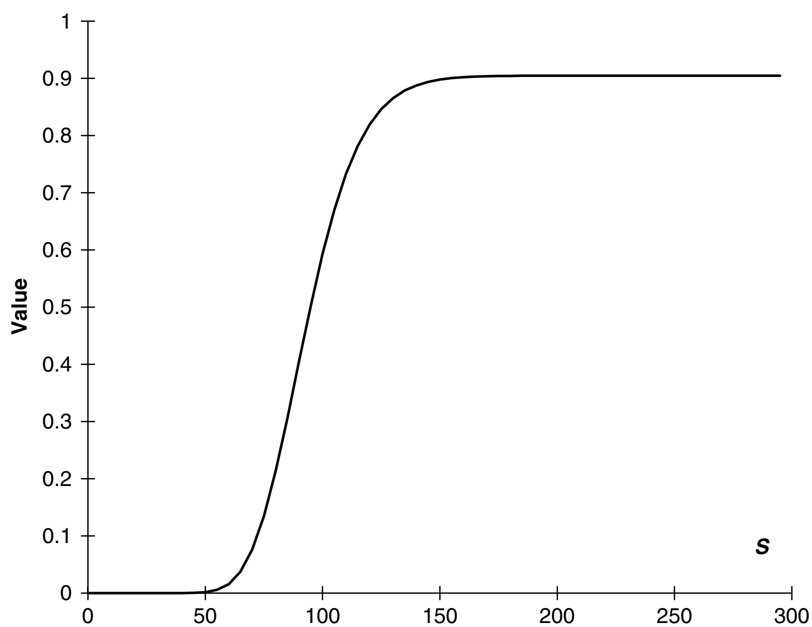


Figure 7.9 The value of a binary call option.

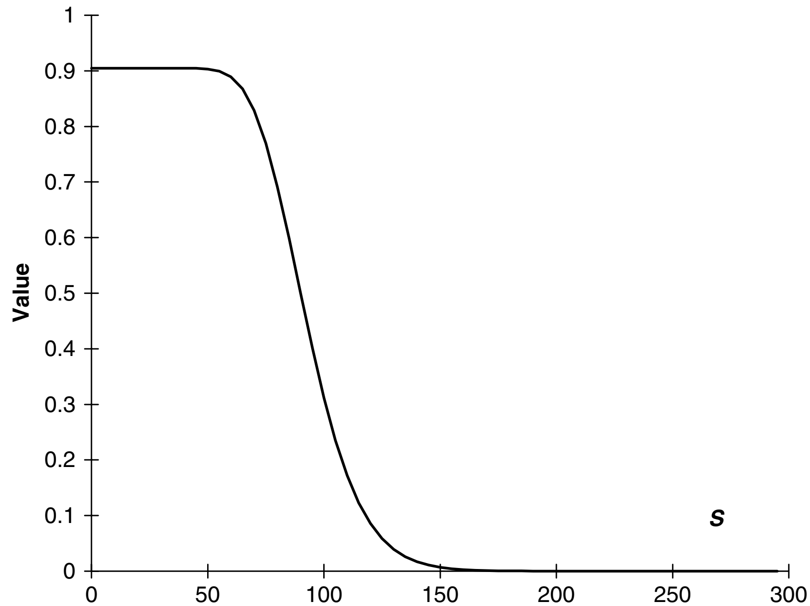


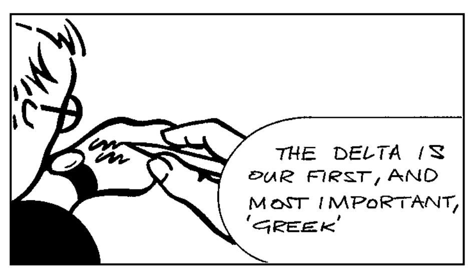
Figure 7.10 The value of a binary put option.

since a binary call and a binary put must add up to the present value of \$1 received at time T. The option value is shown in Figure 7.10.

7.3 DELTA

The **delta**, Δ, of an option or a portfolio of options is the sensitivity of the option or portfolio to the underlying. It is the rate of change of value with respect to the asset:

$$\boxed{\Delta = \frac{\partial V}{\partial S}}$$



Here V can be the value of a single contract or of a whole portfolio of contracts. The delta of a portfolio of options is just the sum of the deltas of all the individual positions.

The theoretical device of delta hedging, introduced in Chapter 5, for eliminating risk is far more than that, it is a very important practical technique.

Roughly speaking, the financial world is divided up into speculators and hedgers. The speculators take a view on the direction of some quantity such as the asset price (or more abstract quantities such as volatility) and implement a strategy to take advantage of their view. Such people may not hedge at all.

Then there are the hedgers. There are two kinds of hedger: the ones who hold a position already and want to eliminate some very specific risk (usually using options) and the ones selling (or buying) the options because they believe they have a better price and can make money by hedging away *all* risk. It is the latter type of hedger that is delta hedging. They can only guarantee to make a profit by selling a contract for a high value if they can eliminate all of the risk due to the random fluctuation in the underlying.

Delta hedging means holding one of the option and short a quantity Δ of the underlying. Delta can be expressed as a function of S and t, I'll give some formulae later in this section. This function varies as S and t vary. This means that the number of assets held must be continuously changed to maintain a **delta-neutral** position; this procedure is called **dynamic hedging**. Changing the number of assets held requires the continual purchase and/or sale of the stock. This is called **rehedging** or **rebalancing** the portfolio.

This delta hedging may take place very frequently in highly liquid markets where it is relatively costless to buy and sell. Thus the Black–Scholes assumption of continuous hedging may be quite accurate. In less liquid markets, you lose a lot on bid-offer spread and will therefore hedge less frequently. Moreover, you may not even be able to buy or sell in the quantities you want. Even in the absence of costs, you cannot be sure that your model for the underlying is accurate. There will certainly be some risk associated with the model. These issues make delta hedging less than perfect and in practice the risk in the underlying cannot be hedged away perfectly. Issues of discrete hedging and transaction costs are covered in depth in Chapters 47 and 48.

Some contracts (see especially Chapter 23) have a delta that becomes very large at special times or asset values. The size of the delta makes delta hedging impossible; what can you do if you find yourself with a theoretical delta requiring you to buy more stock than exists? In such a situation the basic foundation of the Black–Scholes world has collapsed and you would be right to question the validity of any pricing formula. This happens at expiry close to the strike for binary options. Although I've given a formula for their price above and a formula for their delta below, I'd be careful using them if I were you.

Here are some formulae for the deltas of common contracts (all formulae assume that the underlying pays dividends or is a currency):

Deltas of common contracts

Call	$e^{-D(T-t)}N(d_1)$
Put	$e^{-D(T-t)}(N(d_1)-1)$
Binary call	$\dfrac{e^{-r(T-t)}N'(d_2)}{\sigma S\sqrt{T-t}}$
Binary put	$-\dfrac{e^{-r(T-t)}N'(d_2)}{\sigma S\sqrt{T-t}}$

$$N'(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$$

Examples of these functions are plotted in Figure 7.11, with some scaling of the binaries.

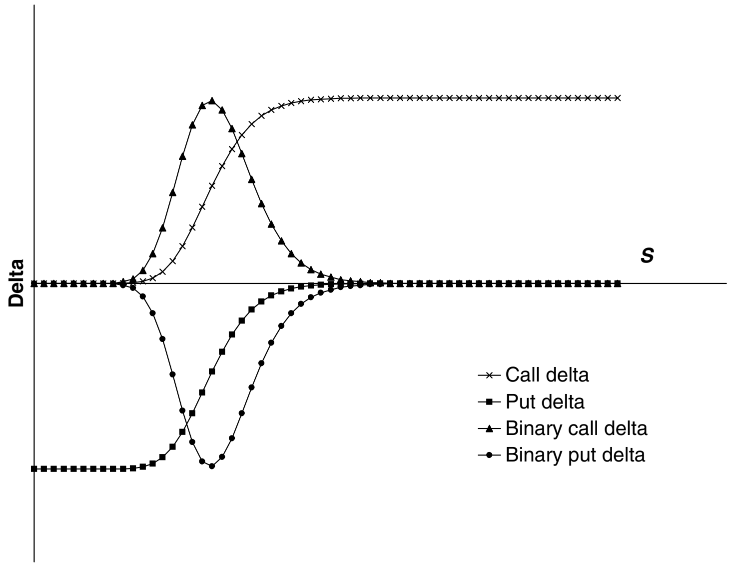


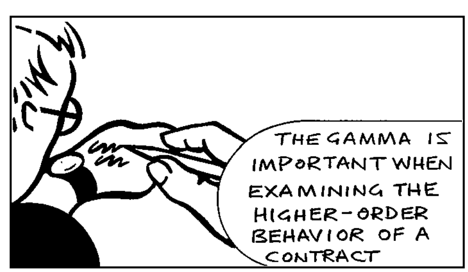
Figure 7.11 The deltas of a call, a put, a binary call and a binary put option. (Binary values scaled to a maximum value of one.)

7.4 GAMMA

The **gamma**, Γ, of an option or a portfolio of options is the second derivative of the position with respect to the underlying:

$$\boxed{\Gamma = \frac{\partial^2 V}{\partial S^2}}$$

Since gamma is the sensitivity of the delta to the underlying it is a measure of by how much or how often a position must be rehedged in order to maintain a delta-neutral position. Although the delta also varies with time this effect is dominated by the Brownian nature of the movement in the underlying.



In a delta-neutral position the gamma is partly responsible for making the return on the portfolio equal to the risk-free rate, the no-arbitrage condition of Chapter 5. The rest of this task falls to the time-derivative of the option value, discussed below. Actually, the situation is far more complicated than this because of the necessary discreteness in the hedging, there is a finite time between rehedges. In any delta-hedged position you make money on some hedges and lose some on others. In a long gamma position ($\Gamma > 0$)

you make money on the large moves in the underlying and lose it on the small moves. To be precise, you make money 32% of the time and lose it 68%. But when you make it, you make more. The net effect is to get the risk-free rate of return on the portfolio. You won't have a clue where this fact came from, but all will be made clear in Chapter 47.

Gamma also plays an important role when there is a mismatch between the market's view of volatility and the actual volatility of the underlying, again this is discussed in Chapter 47.

Since costs can be large and because one wants to reduce exposure to model error it is natural to try to minimize the need to rebalance the portfolio too frequently. Since gamma is a measure of sensitivity of the hedge ratio Δ to the movement in the underlying, the hedging requirement can be decreased by a gamma-neutral strategy. This means buying or selling more *options*, not just the underlying. Because the gamma of the underlying (its second derivative) is zero, we cannot add gamma to our position just with the underlying. We can have as many options in our position as we want; we choose the quantities of each such that both delta and gamma are zero. The minimal requirement is to hold two different types of option and the underlying. In practice, the option position is not readjusted too often because, if the cost of transacting in the underlying is large, then the cost of transacting in its derivatives is even larger.

Here are some formulae for the gammas of common contracts:

Gammas of common contracts

Call	$\dfrac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$
Put	$\dfrac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$
Binary call	$-\dfrac{e^{-r(T-t)}d_1 N'(d_2)}{\sigma^2 S^2(T-t)}$
Binary put	$\dfrac{e^{-r(T-t)}d_1 N'(d_2)}{\sigma^2 S^2(T-t)}$

Examples of these functions are plotted in Figure 7.12, with some scaling for the binaries.

7.5 THETA

Theta, Θ, is the rate of change of the option price with time.

$$\Theta = \frac{\partial V}{\partial t}$$

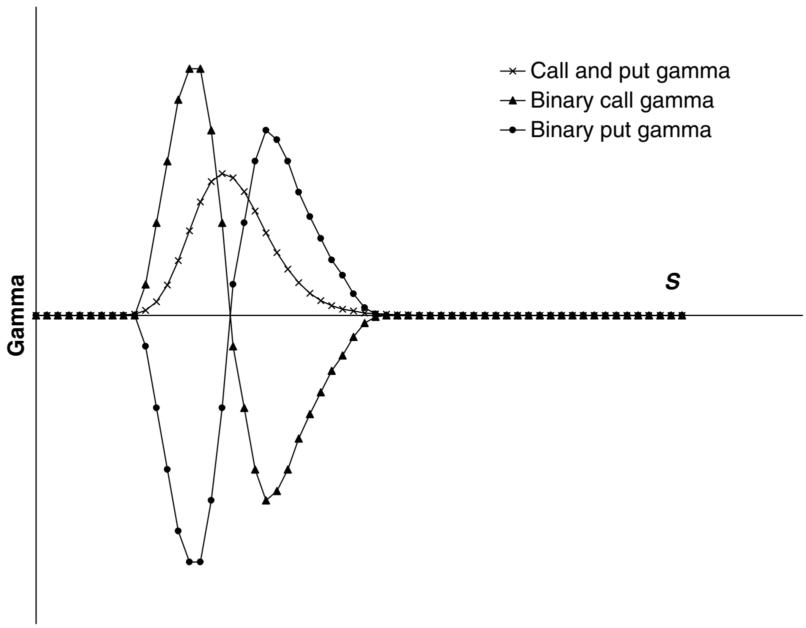


Figure 7.12 The gammas of a call, a put, a binary call and a binary put option.

The theta is related to the option value, the delta and the gamma by the Black–Scholes equation. In a delta-hedged portfolio the theta contributes to ensuring that the portfolio earns the risk-free rate. But it contributes in a completely certain way, unlike the gamma which contributes the right amount *on average*.

Here are some formulae for the thetas of common contracts:

Thetas of common contracts

Call $$-\frac{\sigma S e^{-D(T-t)} N'(d_1)}{2\sqrt{T-t}} + DSN(d_1)e^{-D(T-t)} - rEe^{-r(T-t)}N(d_2)$$

Put $$-\frac{\sigma S e^{-D(T-t)} N'(-d_1)}{2\sqrt{T-t}} - DSN(-d_1)e^{-D(T-t)} + rEe^{-r(T-t)}N(-d_2)$$

Binary call $$re^{-r(T-t)}N(d_2) + e^{-r(T-t)}N'(d_2)\left(\frac{d_1}{2(T-t)} - \frac{r-D}{\sigma\sqrt{T-t}}\right)$$

Binary put $$re^{-r(T-t)}(1-N(d_2)) - e^{-r(T-t)}N'(d_2)\left(\frac{d_1}{2(T-t)} - \frac{r-D}{\sigma\sqrt{T-t}}\right)$$

These functions are plotted in Figure 7.13.

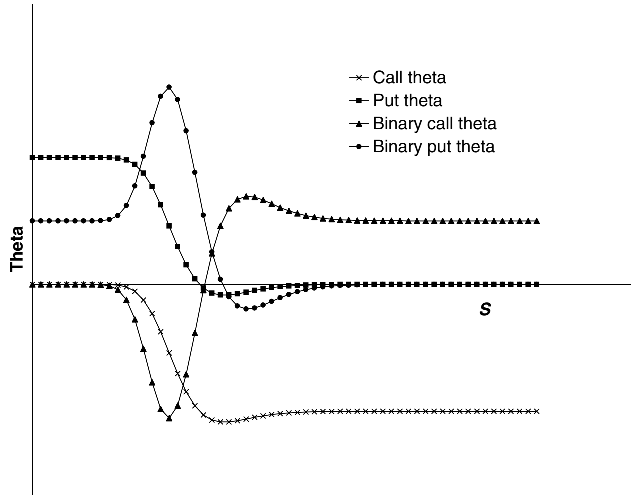


Figure 7.13 The thetas of a call, a put, a binary call and a binary put option.

7.6 SPEED

The **speed** of an option is the rate of change of the gamma with respect to the stock price.

$$\text{Speed} = \frac{\partial^3 V}{\partial S^3}$$

Traders use the gamma to estimate how much they will have to rehedge by if the stock moves. The stock moves by \$1 so the delta changes by whatever the gamma is. But that's only an approximation. The delta may change by more or less than this, especially if the stock moves by a larger amount, or the option is close to the strike and expiration. Hence the use of speed in a higher-order Taylor series expansion.

Here are some formulae for the speed of common contracts:

Speed of common contracts

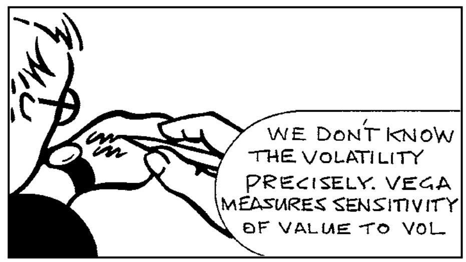
Call	$-\dfrac{e^{-D(T-t)}N'(d_1)}{\sigma^2 S^2(T-t)}\left(d_1+\sigma\sqrt{T-t}\right)$
Put	$-\dfrac{e^{-D(T-t)}N'(d_1)}{\sigma^2 S^2(T-t)}\left(d_1+\sigma\sqrt{T-t}\right)$
Binary call	$-\dfrac{e^{-r(T-t)}N'(d_2)}{\sigma^2 S^3(T-t)}\left(-2d_1+\dfrac{1-d_1d_2}{\sigma\sqrt{T-t}}\right)$
Binary put	$\dfrac{e^{-r(T-t)}N'(d_2)}{\sigma^2 S^3(T-t)}\left(-2d_1+\dfrac{1-d_1d_2}{\sigma\sqrt{T-t}}\right)$

7.7 VEGA

Vega, a.k.a. zeta and kappa, is a very important but confusing quantity. It is the sensitivity of the option price to volatility.

$$\text{Vega} = \frac{\partial V}{\partial \sigma}$$

This is completely different from the other greeks[3] since it is a derivative with respect to a parameter and not a variable. This makes something of a difference when we come to finding numerical solutions for such quantities.



In practice, the volatility of the underlying is not known with certainty. Not only is it very difficult to measure at any time, it is even harder to predict what it will do in the future. Suppose that we put a volatility of 20% into an option pricing formula, how sensitive is the price to that number? That's the vega.

As with gamma hedging, one can vega hedge to reduce sensitivity to the volatility. This is a major step towards eliminating some model risk, since it reduces dependence on a quantity that, to be honest, is not known very accurately.

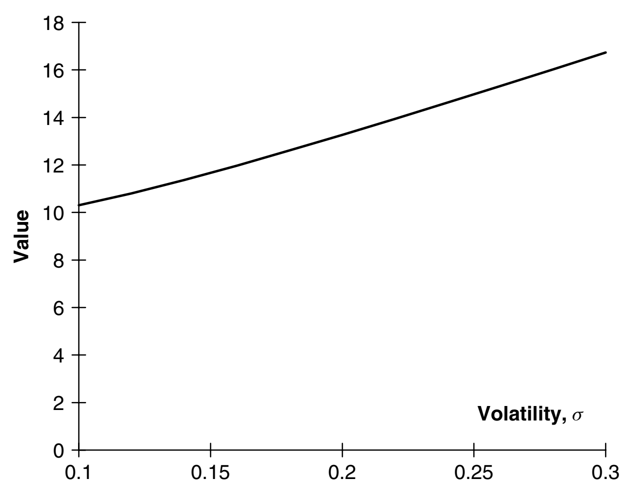
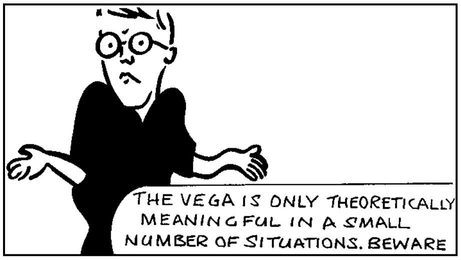


Figure 7.14 The value of an at-the-money call option as a function of volatility.

---

[3] It's not even Greek. Among other things it is an American car, a star (Alpha Lyrae), the real name of Zorro, there are a couple of 16th century Spanish authors called Vega, an Op art painting by Vasarely and a character in the computer game ‘Street Fighter.’ And who could forget Vincent, and his brother?

The second derivative with respect to σ has been called ‘vomma’ and the second-order derivative with respect to the asset and the volatility has been called ‘kabanga.’ I doubt that they represent what their fans think they represent, and I'm going to make no further mention of them.



There is a downside to the measurement of vega. It is only really meaningful for options having single-signed gamma everywhere. For example it makes sense to measure vega for calls and puts but not binary calls and binary puts. I have included the formulae for the vega of such contracts below, but they should be used with care, if at all. The reason for this is that call and put values (and options with single-signed gamma) have values that are monotonic in the volatility: increase the volatility in a call and its value increases everywhere. Contracts with a gamma that changes sign may have a vega measured at zero because as we increase the volatility the price may rise somewhere and fall somewhere else. Such a contract is very exposed to volatility risk but that risk is not measured by the vega. See Chapter 52 for more details.

Here are formulae for the vegas of common contracts:

Vegas of common contracts

Call	$S\sqrt{T-t}\,e^{-D(T-t)}N'(d_1)$
Put	$S\sqrt{T-t}\,e^{-D(T-t)}N'(d_1)$
Binary call	$-e^{-r(T-t)}N'(d_2)\left(\sqrt{T-t}+\dfrac{d_2}{\sigma}\right)$
Binary put	$e^{-r(T-t)}N'(d_2)\left(\sqrt{T-t}+\dfrac{d_2}{\sigma}\right)$

In Figure 7.14 is shown the value of an at-the-money call option as a function of the volatility. There is one year to expiry, the strike is 100, the interest rate is 10% and there are no dividends. No matter how far in or out of the money this curve is always monotonically increasing for call options and put options; uncertainty adds value to the contract. The slope of this curve is the vega.

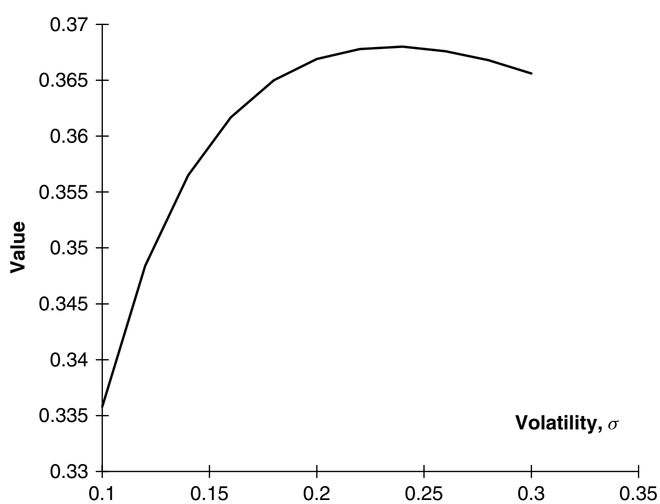


Figure 7.15 The value of an out-of-the-money binary call option as a function of volatility.

In Figure 7.15 is shown the value of an out-of-the-money binary call option as a function of the volatility. There is one year to expiry, the asset value is 88, strike is 100, the interest rate is 10% and there are no dividends. Observe that there is maximum at a volatility of about 24%. The value of the option is not monotonic in the volatility. We will see later why this makes the meaning of vega somewhat suspect.

7.8 RHO

Rho, ρ, is the sensitivity of the option value to the interest rate used in the Black–Scholes formulae:

$$\rho = \frac{\partial V}{\partial r}$$

In practice one often uses a whole term structure of interest rates, meaning a time-dependent rate $r(t)$. Rho would then be the sensitivity to the level of the rates assuming a parallel shift in rates at all times. Again, you must be careful for which contracts you measure rho; see Chapter 52 for more details.

Here are some formulae for the rhos of common contracts:

Rhos of common contracts

Call	$E(T-t)e^{-r(T-t)}N(d_2)$
Put	$-E(T-t)e^{-r(T-t)}N(-d_2)$
Binary call	$-(T-t)e^{-r(T-t)}N(d_2) + \dfrac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$
Binary put	$-(T-t)e^{-r(T-t)}(1-N(d_2)) - \dfrac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$

The sensitivities of common contract to the dividend yield or foreign interest rate are given by the following formulae:

Sensitivity to dividend for common contracts

Call	$-(T-t)Se^{-D(T-t)}N(d_1)$
Put	$(T-t)Se^{-D(T-t)}N(-d_1)$
Binary call	$-\dfrac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$
Binary put	$\dfrac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$

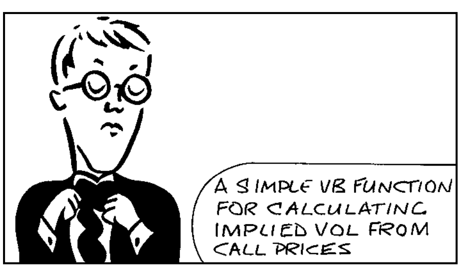
7.9 IMPLIED VOLATILITY

The Black–Scholes formula for a call option takes as input the expiry, the strike, the underlying and the interest rate *together with the volatility* to output the price. All but the volatility are easily measured. How do we know what volatility to put into the formulae? A trader can see on his screen that a certain call option with four months until expiry and a strike of 100 is trading at 6.51 with the underlying at 101.5 and a short-term interest rate of 8%. Can we use this information in some way?

Turn the relationship between volatility and an option price on its head. If we can see the price at which the option is trading, we can ask 'What volatility must I use to get the correct market price?' This is called the **implied volatility**. The implied volatility is the volatility of the underlying which when substituted into the Black–Scholes formula gives a theoretical price equal to the market price. In a sense it is the market's view of volatility over the life of the option. Assuming that we are using call prices to estimate the implied volatility then provided the option price is less than the asset and greater than zero then we can find a unique value for the implied volatility. (If the option price is outside these bounds then there's a very extreme arbitrage opportunity.) Because there is no simple formula for the implied volatility as a function of the option value we must solve the equation

$$V_{BS}(S_0, t_0; \sigma, r; E, T) = \text{known value}$$

for σ, where V_{BS} is the Black–Scholes formula. Today's asset price is S_0, the date is t_0 and everything is known in this equation except for σ. Below is an algorithm for finding the implied volatility from the market price of a call option to any required degree of accuracy. The method used is **Newton–Raphson** which uses the derivative of the option price with respect to the volatility (the vega) in the calculation. This method is particularly good for such a well-behaved function as a call value.




```
Function ImpVolCall(MktPrice As Double, Strike As
         Double, Expiry As Double, _ Asset As Double,
         IntRate As Double, error As Double)
Volatility = 0.2
dv = error + 1
While Abs(dv) > error
    d1 = Log(Asset / Strike) + (IntRate + 0.5 *
         Volatility * Volatility) * Expiry
    d1 = d1 / (Volatility * Sqr(Expiry))
    d2 = d1 - Volatility * Sqr(Expiry)
    PriceError = Asset * cdf(d1) - Strike *
                 Exp(-IntRate * Expiry) _
                 * cdf(d2) - MktPrice
    Vega = Asset * Sqr(Expiry / 3.1415926 / 2) *
           Exp(-0.5 * d1 * d1)
    dv = PriceError / Vega
    Volatility = Volatility - dv
Wend
ImpVolCall = Volatility
End Function
```

In this we need the cumulative distribution function for the Normal distribution. The following is a simple algorithm which gives an accurate, and fast, approximation to the cumulative distribution function of the standardized Normal:

$$\text{For} \quad x \geq 0 \quad N(x) \approx 1 - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \left(a_1 d + a_2 d^2 + a_3 d^3 + a_4 d^4 + a_5 d^5\right)$$

where

$$d = \frac{1}{1 + 0.0.2316419x}$$

and

$a1 = 0.31938153$, $a2 = -0.356563782$, $a3 = 1.781477937$, $a4 = -1.821255978$ and $a5 = 1.330274429$.

For $x < 0$ use the fact that $N(x) + N(-x) = 1$.

```
Function cdf(x As Double) As Double
Dim d As Double
Dim temp as Double
Dim a1 As Double
Dim a2 As Double
Dim a3 As Double
Dim a4 As Double
Dim a5 As Double
d = 1 / (1 + 0.2316419 * Abs(x))
a1 = 0.31938153
a2 = -0.356563782
a3 = 1.781477937
a4 = -1.821255978
a5 = 1.330274429
temp = a5
temp = a4 + d * temp
temp = a3 + d * temp
temp = a2 + d * temp
temp = a1 + d * temp
temp = d * temp
cdf = 1 - 1 / Sqr(2 * 3.1415926) * Exp(-0.5 * x * x) * temp
If x < 0 Then cdf = 1 - cdf
End Function
```

In practice if we calculate the implied volatility for many different strikes and expiries on the same underlying then we find that *the volatility is not constant*. A typical result is that of Figure 7.16 which shows the implied volatilities for the S&P500 on 9th September 1999 for options expiring later in the month. The implied volatilities for the calls and puts should be identical, because of put-call parity. The differences seen here could be due to bid-offer spread or calculations performed at slightly different times.

This shape is commonly referred to as the **skew,** or the smile, if it is turned up at both ends, but it could also be in the shape of a **frown**. In this example it's a rather lopsided wry grin, a negative skew since it slopes downwards. Whatever the shape, it tends to persist with time, with certain shapes being characteristic of certain markets.

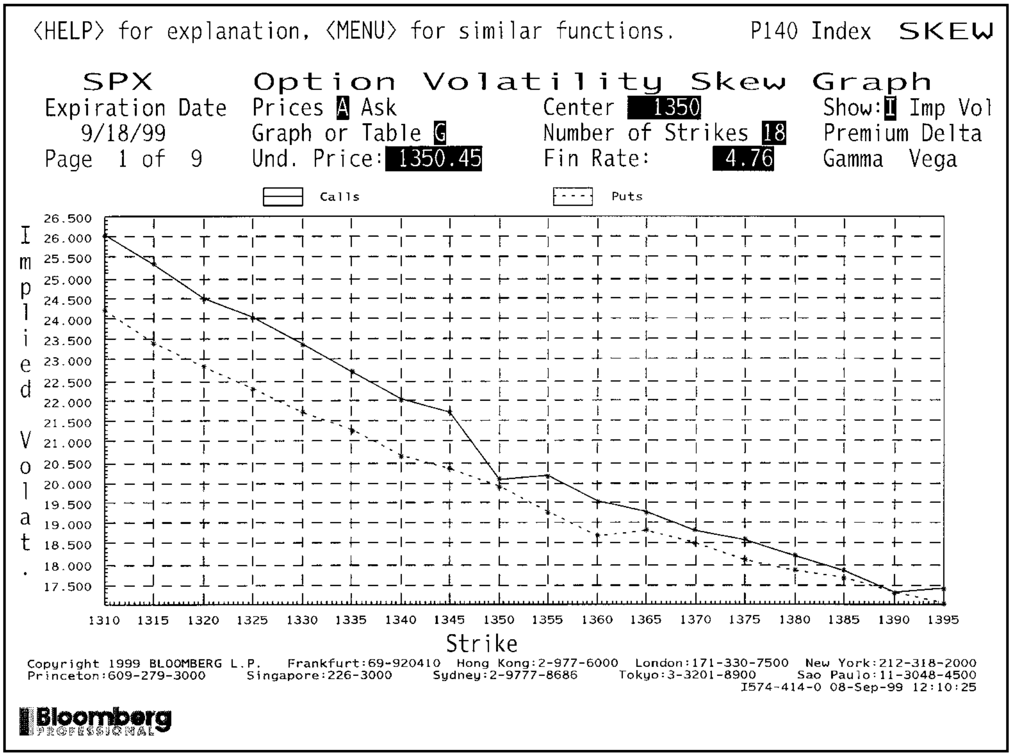


Figure 7.16 Implied volatilities for the S&P500. Source: Bloomberg L.P.

The dependence of the implied volatility on strike and expiry can be interpreted in many ways. The easiest interpretation is that it represents the market's view of future volatility in some complex way. This issue is covered in depth in Chapter 50. Another possibility is that it reflects the uncertainty in volatility; perhaps volatility is also a stochastic variable, see Chapter 51.

In the foreign exchange markets they tend to plot implied volatility versus the delta of an option. Figures 7.17 and 7.18 show plots of implied volatility versus strike, the picture usually looked at by equity derivatives traders, and implied volatility versus delta, the picture usually looked at by FX derivatives traders, respectively.

Although these two ways of looking at the data are effectively the same (and the same as simply looking at option price versus strike as well) they do lead to different dynamics in equity and FX markets. In practice, when the value of the underlying changes we will see the option prices change and also often see the implied volatilities change. Now this latter is totally inconsistent with the theory we've developed so far, but it's what happens and we'll be looking at this more closely later on. But the important point is that in equity markets it is usually said that implied volatility stays unchanged *as a function of strike* whereas in FX markets the implied volatilities stay unchanged *as a function of delta*. These are called the sticky strike and sticky delta models. This is probably due in part simply to the pictures that the different types of traders look at rather than due to any fundamental difference between equities and exchange rates.

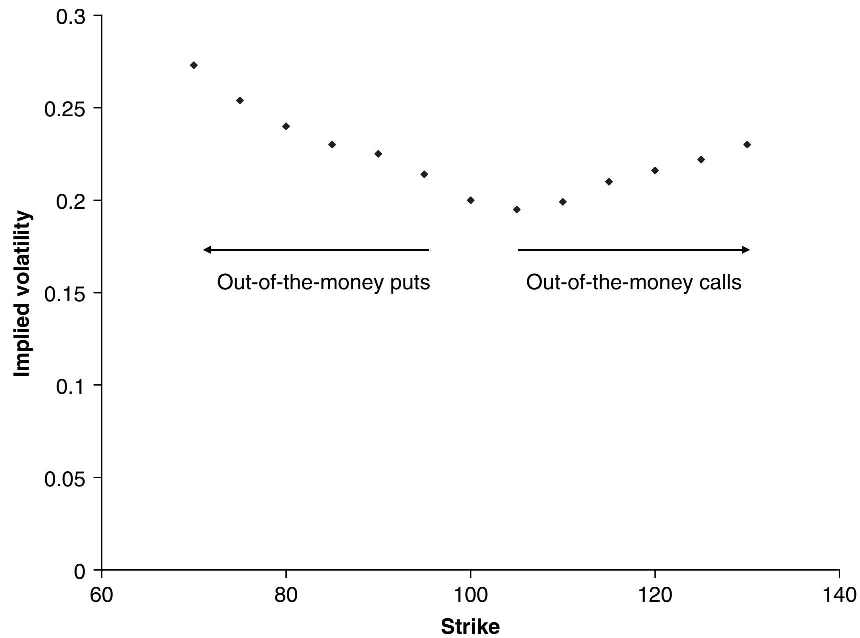


Figure 7.17 Implied volatility versus strike.

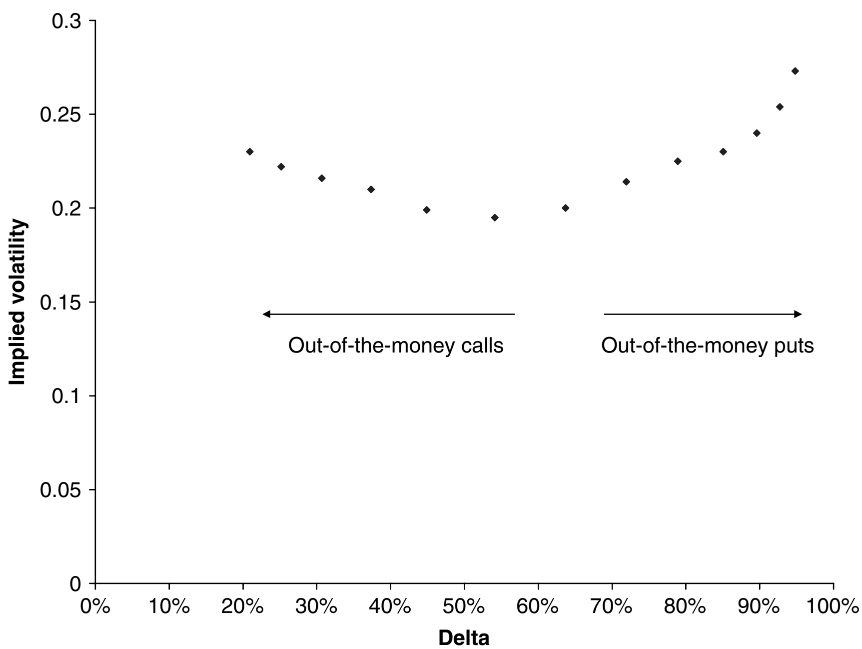
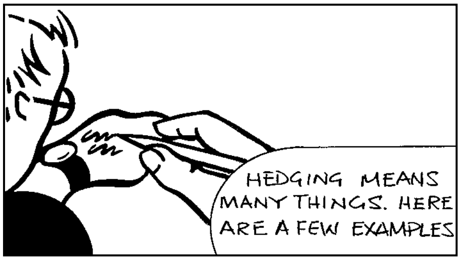


Figure 7.18 Implied volatility versus delta.



7.10 A CLASSIFICATION OF HEDGING TYPES

7.10.1 Why Hedge?

'Hedging' in its broadest sense means the reduction of risk by exploiting relationships or correlation between various risky investments (or bets). The concept is used widely in horse racing, other sports betting and, of course, high finance. The reason for hedging is that it can lead to an improved risk/return. In the classical Modern Portfolio Theory framework (Chapter 18), for example, it is usually possible to construct many portfolios having the same expected return but with different variance of returns ('risk'). Clearly, if you have two portfolios with the same expected return the one with the lower risk is the better investment.

7.10.2 The Two Main Classifications

Probably the most important distinction between types of hedging is between model-independent and model-dependent hedging strategies.

Model-independent hedging: An example of such hedging is Put-call Parity. There is a simple relationship between calls and puts on an asset (when they are both European and with the same strikes and expiries), the underlying stock and a zero-coupon bond with the same maturity. This relationship is completely independent of how the underlying asset changes in value. Another example is Spot-forward Parity. In neither case do we have to specify the dynamics of the asset, not even its volatility, to find a possible hedge. Such model-independent hedges are few and far between.

Model-dependent hedging: Most sophisticated finance hedging strategies depend on a model for the underlying asset. The obvious example is the hedging used in the Black–Scholes analysis that leads to a whole theory for the value of derivatives. In pricing derivatives we typically need to know at least the volatility of the underlying asset. If the model is wrong then the option value and any hedging strategy will also be wrong.

7.10.3 Delta Hedging

One of the building blocks of derivatives theory is **delta hedging.** This is the theoretically perfect elimination of all risk by using a very clever hedge between the option and its underlying. Delta hedging exploits the perfect correlation between the changes in the option value and the changes in the stock price. This is an example of 'dynamic' hedging; the hedge must be continually monitored and frequently adjusted by the sale or purchase of the underlying asset. Because of the frequent rehedging, any dynamic hedging strategy is going to result in losses due to transaction costs. In some markets this can be very important.

7.10.4 Gamma Hedging

To reduce the size of each rehedge and/or to increase the time between rehedges, and thus reduce costs, the technique of **gamma hedging** is often employed. A portfolio that is delta hedged is insensitive to movements in the underlying as long as those movements are quite small. There is a small error in this due to the convexity of the portfolio with respect to the underlying. Gamma hedging is a more accurate form of hedging that theoretically eliminates these second-order effects. Typically, one hedges one, exotic, say, contract with a vanilla contract and the underlying. The quantities of the vanilla and the underlying are chosen so as to make both the portfolio delta and the portfolio gamma instantaneously zero.

7.10.5 Vega Hedging

As I said above, the prices and hedging strategies are only as good as the model for the underlying. The key parameter that determines the value of a contract is the volatility of the underlying asset. Unfortunately, this is a very difficult parameter to measure or even estimate. Nor is it usually a constant as assumed in the simple theories. Obviously, the value of a contract depends on this parameter, and so to ensure that our portfolio value is insensitive to this parameter we can **vega hedge**. This means that we hedge one option with both the underlying and another option in such a way that both the delta and the vega, the sensitivity of the portfolio value to volatility, are zero. This is often quite satisfactory in practice but is usually theoretically inconsistent; we should not use a constant volatility (basic Black–Scholes) model to calculate sensitivities to parameters that are assumed not to vary. The distinction between variables (underlying asset price and time) and parameters (volatility, dividend yield, interest rate) is extremely important here. It is justifiable to rely on sensitivities of prices to variables, but usually not sensitivity to parameters. To get around this problem it is possible to model independently volatility etc. as variables themselves. In such a way it is possible to build up a consistent theory.

7.10.6 Static Hedging

There are quite a few problems with delta hedging, on both the practical and the theoretical side. In practice, hedging must be done at discrete times and is costly. Sometimes one has to buy or sell a prohibitively large number of the underlying in order to follow the theory. This is a problem with barrier options and options with discontinuous payoff. On the theoretical side, the model for the underlying is not perfect, at the very least we do not know parameter values accurately. Delta hedging alone leaves us very exposed to the model; this is model risk. Many of these problems can be reduced or eliminated if we follow a strategy of **static hedging** as well as delta hedging: buy or sell more liquid traded contracts to reduce the cashflows in the original contract. The static hedge is put into place now, and left until expiry. In the extreme case where an exotic contract has all of its cashflows matched by cashflows from traded options then its value is given by the cost of setting up the static hedge; a model is not needed. (But then the option wasn't exotic in the first place.)

7.10.7 Margin Hedging

Often what causes banks, and other institutions, to suffer during volatile markets is not the change in the paper value of their assets but the requirement to come up suddenly with a large amount of cash to cover an unexpected margin call. Recent examples where margin has caused significant damage are Metallgesellschaft and Long Term Capital Management. Writing options is very risky. The downside of buying an option is just the initial premium, while the upside may be unlimited. The upside of writing an option is limited, but the downside could be huge. For this reason, to cover the risk of default in the event of an unfavorable outcome, the clearing houses that register and settle options insist on the deposit of a margin by the writers of options. Margin comes in two forms, the initial margin and the maintenance margin. The initial margin is the amount deposited at the initiation of the contract. The total amount held as margin must stay above a prescribed maintenance margin. If it ever falls below this level then more money (or equivalent in bonds, stocks etc.) must be deposited. The amount of margin that must be deposited depends on the particular contract. A dramatic market move could result in a sudden large margin call that may be difficult to meet. To prevent this situation it is possible to **margin hedge**. That is, set up a portfolio such that margin calls on one part of the portfolio are balanced by refunds from other parts. Usually over-the-counter contracts have no associated margin requirements and so won't appear in the calculation.

7.10.8 Crash (Platinum) Hedging

The final variety of hedging that we discuss is specific to extreme markets. Market crashes have at least two obvious effects on our hedging. First of all, the moves are so large and rapid that they cannot be traditionally delta hedged. The convexity effect is not small. Second, normal market correlations become meaningless. Typically all correlations become one (or minus one). **Crash** or **Platinum hedging** exploits the latter effect in such a way as to minimize the worst possible outcome for the portfolio. The method, called CrashMetrics (Chapter 43), does not rely on difficult to measure parameters such as volatilities and so is a very robust hedge. Platinum hedging comes in two types: hedging the paper value of the portfolio and hedging the margin calls.

7.11 SUMMARY

In this chapter we went through the derivation of some of the most important formulae. We also saw the definitions and descriptions of the hedge ratios. Trading in derivatives would be no more than gambling if you took away the ability to hedge. Hedging is all about managing risk and reducing uncertainty.

FURTHER READING

- See Taleb (1997) for a lot of detailed analysis of vega.
- See Press *et al*. (1992) for more routines for finding roots, i.e. for finding implied volatilities.
- There are many 'virtual' option pricers on the internet. See, for example, `www.cboe.com`.
- I'm not going to spend much time on deriving or even presenting formulae. There are 1001 books that contain option formulae, there is even one book with 1001 formulae (Haug, 1997).
- See the series of articles by Thorp (2002) on how he derived the correct option pricing formulae in the late 1960s.
- Haug (2003) discusses the sophisticated trader use of the simple equations.

	Call	Put	Binary Call	Binary Put
Value V Black–Scholes value	$Se^{-D(T-t)}N(d_1)$ $-Ee^{-r(T-t)}N(d_2)$	$-Se^{-D(T-t)}N(-d_1)$ $+Ee^{-r(T-t)}N(-d_2)$	$e^{-r(T-t)}N(d_2)$	$e^{-r(T-t)}(1-N(d_2))$
Delta $\frac{\partial V}{\partial S}$ Sensitivity to underlying	$e^{-D(T-t)}N(d_1)$	$e^{-D(T-t)}(N(d_1)-1)$	$\frac{e^{-r(T-t)}N'(d_2)}{\sigma S\sqrt{T-t}}$	$-\frac{e^{-r(T-t)}N'(d_2)}{\sigma S\sqrt{T-t}}$
Gamma $\frac{\partial^2 V}{\partial S^2}$ Sensitivity of delta to underlying	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$	$-\frac{e^{-r(T-t)}d_1N'(d_2)}{\sigma^2S^2(T-t)}$	$\frac{e^{-r(T-t)}d_1N'(d_2)}{\sigma^2S^2(T-t)}$
Theta $\frac{\partial V}{\partial t}$ Sensitivity to time	$-\frac{\sigma Se^{-D(T-t)}N'(d_1)}{2\sqrt{T-t}}$ $+DSN(d_1)e^{-D(T-t)}$ $-rEe^{-r(T-t)}N(d_2)$	$-\frac{\sigma Se^{-D(T-t)}N'(-d_1)}{2\sqrt{T-t}}$ $-DSN(-d_1)e^{-D(T-t)}$ $+rEe^{-r(T-t)}N(-d_2)$	$re^{-r(T-t)}N(d_2)$ $+e^{-r(T-t)}N'(d_2)\times$ $\left(\frac{d_1}{2(T-t)}-\frac{r-D}{\sigma\sqrt{T-t}}\right)$	$re^{-r(T-t)}(1-N(d_2))$ $-e^{-r(T-t)}N'(d_2)\times$ $\left(\frac{d_1}{2(T-t)}-\frac{r-D}{\sigma\sqrt{T-t}}\right)$
Speed $\frac{\partial^3 V}{\partial S^3}$ Sensitivity of gamma to underlying	$-\frac{e^{-D(T-t)}N'(d_1)}{\sigma^2S^2(T-t)}\times$ $\left(d_1+\sigma\sqrt{T-t}\right)$	$-\frac{e^{-D(T-t)}N'(d_1)}{\sigma^2S^2(T-t)}\times$ $\left(d_1+\sigma\sqrt{T-t}\right)$	$-\frac{e^{-r(T-t)}N'(d_2)}{\sigma^2S^3(T-t)}\times$ $\left(-2d_1+\frac{1-d_1d_2}{\sigma\sqrt{T-t}}\right)$	$\frac{e^{-r(T-t)}N'(d_2)}{\sigma^2S^3(T-t)}\times$ $\left(-2d_1+\frac{1-d_1d_2}{\sigma\sqrt{T-t}}\right)$
Vega $\frac{\partial V}{\partial \sigma}$ Sensitivity to volatility	$S\sqrt{T-t}\,e^{-D(T-t)}N'(d_1)$	$S\sqrt{T-t}\,e^{-D(T-t)}N'(d_1)$	$-e^{-r(T-t)}N'(d_2)\times$ $\left(\sqrt{T-t}+\frac{d_2}{\sigma}\right)$	$e^{-r(T-t)}N'(d_2)\times$ $\left(\sqrt{T-t}+\frac{d_2}{\sigma}\right)$
Rho (r) $\frac{\partial V}{\partial r}$ Sensitivity to interest rate	$E(T-t)e^{-r(T-t)}N(d_2)$	$-E(T-t)e^{-r(T-t)}N(-d_2)$	$-(T-t)e^{-r(T-t)}N(d_2)$ $+\frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$	$-(T-t)e^{-r(T-t)}(1-N(d_2))$ $-\frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$
Rho (D) $\frac{\partial V}{\partial D}$ Sensitivity to dividend yield	$-(T-t)Se^{-D(T-t)}N(d_1)$	$(T-t)Se^{-D(T-t)}N(-d_1)$	$-\frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$	$\frac{\sqrt{T-t}}{\sigma}e^{-r(T-t)}N'(d_2)$

$$d_1=\frac{\log(S/E)+(r-D+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},\quad d_2=\frac{\log(S/E)+(r-D-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}=d_1-\sigma\sqrt{T-t},\quad N(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-(\frac{1}{2})\xi^2}d\xi\quad\text{and}\quad N'(x)=\frac{1}{\sqrt{2\pi}}e^{-(\frac{1}{2})x^2}$$

CHAPTER 8
simple generalizations of the Black–Scholes world

In this Chapter...

- complex dividend structures
- jump conditions
- time-dependent volatility, interest rate and dividend yield

8.1 INTRODUCTION

This chapter is an introduction to some of the possible generalizations of the 'Black–Scholes world.' In particular, I will discuss the effect of dividend payments on the underlying asset and how to incorporate time-dependent parameters into the framework. These subjects lead to some interesting and important mathematical and financial conclusions.

The generalizations are very straightforward. However, later, in Part Five, I describe other models of the financial world that take us a long way from Black–Scholes.

8.2 DIVIDENDS, FOREIGN INTEREST AND COST OF CARRY

In Chapter 5 I showed how to incorporate certain types of dividend structures into the Black–Scholes option pricing framework, and then in Chapter 7 I gave some formulae for the values of some common vanilla contracts, again with dividends on the underlying. The dividend structure that I dealt with was the very simplest from a mathematical point of view. I assumed that an amount was paid to the holder of the asset that was proportional to the value of the asset and that it was paid continuously. In other words, the owner of one asset received a dividend of $DS\,dt$ in a time step dt. This dividend structure is realistic if the underlying is an index on a large number of individual assets each receiving a lump sum dividend but with all these dividends spread out through the year. It is also a good model if the underlying is a currency in which case we simply take the 'dividend yield' to be the foreign interest rate. Similarly, if the underlying is a commodity with a cost of carry that is proportional to its value, then the 'dividend yield' is just the cost of carry (with a minus sign, we benefit from dividends but must pay out the cost of carry).

To recap, if the underlying receives a dividend of $DS\,dt$ in a time step dt when the asset price is S then

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D)S\frac{\partial V}{\partial S} - rV = 0.$$

However, if the underlying is a stock, then the assumption of constant and continuously-paid dividend yield is not a good one.

8.3 DIVIDEND STRUCTURES

Typically, dividends are paid out quarterly in the US and semi-annually or quarterly in the UK. The dividend is set by the board of directors of the company some time before it is paid out and the amount of the payment is made public. The amount is often chosen to be similar to previous payments, but will obviously reflect the success or otherwise of the company. The amount specified is a dollar amount, it is *not* a percentage of the stock price on the day that the payment is made. So reality differs from the above simple model in three respects:

- the amount of the dividend is not known until shortly before it is paid
- the payment is a given dollar amount, independent of the stock price
- the dividend is paid discretely, and not continuously throughout the year.

In what follows I am going to make some assumptions about the dividend. I will assume that

- the amount of the dividend is a known amount, possibly with some functional dependence on the asset value *at the payment date*
- the dividend is paid discretely on a known date.

Other assumptions that I could, but won't, make because of the subsequent complexity of the modeling are that the dividend amount and/or date are random, that the dividend amount is a function of the stock price on the day that the dividend is set, that the dividend depends on how well the stock has done in the previous quarter ...

8.4 DIVIDEND PAYMENTS AND NO ARBITRAGE

How does the stock react to the payment of a dividend? To put the question another way, if you have a choice whether to buy a stock just before or just after it goes ex-dividend, which should you choose?

Let me introduce some notation. The dates of dividends will be t_i and the amount of the dividend paid on that day will be D_i. This may be a function of the underlying asset, but it then must be a deterministic function. The moment just before the stock goes ex-dividend will be denoted by t_i^- and the moment just after will be t_i^+.

The person who buys the stock on or before t_i^- will also get the rights to the dividend. The person who buys it at t_i^+ or later will not receive the dividend. It looks like there is an advantage

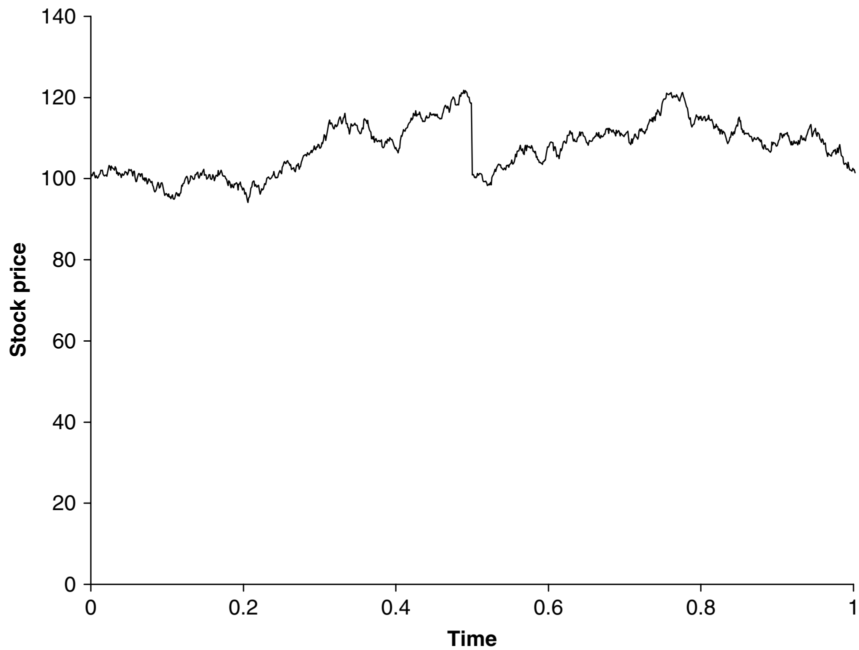


Figure 8.1 A stock price path across a dividend date.

in buying the stock just before the dividend date. Of course, this advantage is balanced by *a fall in the stock price as it goes ex-dividend*. Across a dividend date the stock falls by the amount of the dividend. If it did not, then there would be arbitrage opportunities. We can write

$$S(t_i^+) = S(t_i^-) - D_i. \tag{8.1}$$

In Figure 8.1 is shown an asset price path showing the fall in the asset price as it goes ex-dividend; the drop has been exaggerated.

This jump in the stock price will presumably have some effect on the value of an option. We will discuss this next.

8.5 THE BEHAVIOR OF AN OPTION VALUE ACROSS A DIVIDEND DATE

We have just seen how the underlying asset jumps in value, in a completely predictable way, across a dividend date.

Jump conditions tell us about the value of a dependent variable, an option price, when there is a discontinuous change in one of the independent variables. In the present case, there is a discontinuous change in the asset price due to the payment of a dividend but how does this affect the option price? Does the option price also jump? The jump condition relates the values of the option across the jump, from times t_i^- to t_i^+. The jump condition will be derived by a simple no-arbitrage argument.

To see what the jump condition should be, ask the question: 'By how much do I profit or lose when the stock price jumps?' If you hold the option then you do not see any of the dividend,

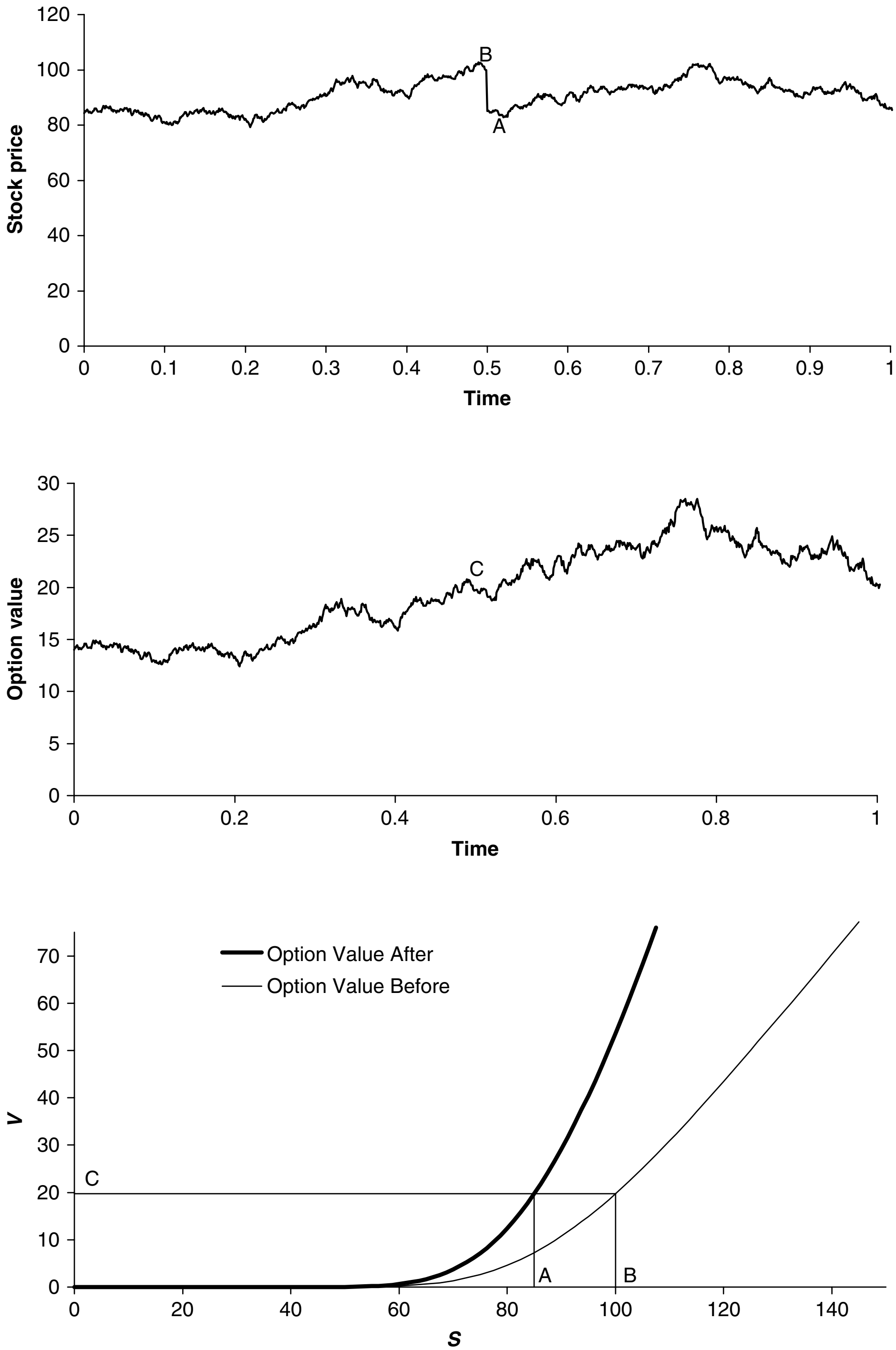


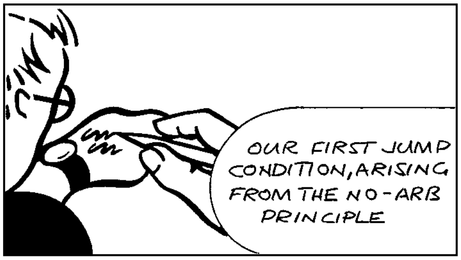
Figure 8.2 Top picture, a realization of the stock price showing a fall across the dividend date. Middle picture, the corresponding realization of the option price (in this example a call). Bottom picture, the option value as a function of the stock price just before and just after the dividend date.

that goes to the holder of the stock not you, the holder of the option. If the dividend amount and date are known in advance then there is no surprise in the fall in the stock price. The conclusion must be that the option does not change in value across the dividend date, its path is continuous. Continuity of the option value across a dividend date can be written as

$$V\left(S(t_i^-), t_i^-\right) = V\left(S(t_i^+), t_i^+\right) \tag{8.2}$$

or, in terms of the amount of the dividend,

$$\boxed{V\left(S, t_i^-\right) = V\left(S - D_i, t_i^+\right).} \tag{8.3}$$



The jump condition and its effect on the option value can be explained by reference to Figure 8.2. In this figure, the top picture shows a realization of the stock price with a fall across the dividend date. The middle picture shows the corresponding realization of a call option price. The bottom picture shows the option value as a function of the stock price just before and just after the dividend date. Observe the points 'A' and 'B' on these pictures. 'A' is the stock price after the dividend has been paid and 'B' is the price before. On the bottom picture we see the values of the option associated with these before and after asset prices. *These option values are the same* and are denoted by 'C.' Even though there is a fall in the asset value, the option value is unchanged because the whole V versus S plot changes. The relationship between the before and after values of the option are related by (8.3). I will give two examples.

Suppose that the dividend paid out is proportional to the asset value, $D_i = DS$. In this case

$$S(t_i^+) = (1 - D)S(t_i^-).$$

Equation (8.3) is then just

$$V\left(S, t_i^-\right) = V\left((1 - D)S, t_i^+\right).$$

The two option price curves are identical if one stretches the after curve by a factor of $(1 - D)^{-1}$ in the horizontal direction. Thus, even though the option value is continuous across a dividend date, the delta changes discontinuously.

If the dividend is independent of the stock price then

$$S(t_i^+) = S(t_i^-) - D_i,$$

where D_i is independent of the asset value. The before curve is then identical to the after curve, but shifted by an amount D_i.

8.6 COMMODITIES

Convenience yield is the benefit or premium associated with holding an underlying product or physical good, rather than a future position in that product. For example, there is an obvious benefit to the actual holding of barrels of oil. This naturally leads us to thinking of commodities as being of two types.

- **Investment commodities:** Commodities held for investment, such as gold.
- **Consumption commodities:** Commodities held for consumption, such as oil or wheat. Since they are held for consumption, and have a value associated with this, they may not be sold when the price rises. Crucially, this can make arbitrage arguments one sided.

8.6.1 Futures Prices and Arbitrage

The standard argument for the relationship between spot and futures prices does not hold for consumption commodities. Only an upper bound can be found. This is because there may be constraints on the selling of consumption commodities; the oil is needed for heating, transport etc. and cannot be reasonably sold.

The no-arbitrage argument can still be applied to investment commodity futures.

8.6.2 Storage Costs

No storage costs

If there are no storage costs then the relationship between spot and forward prices is

$$F = S\, e^{r(T-t)}.$$

Storage costs proportional to spot price, for investment commodities

If there are storage costs, and they are proportional to the price of the commodity then

$$F = S\, e^{(r+u)(T-t)}.$$

This is equivalent to there being a negative dividend yield.

Storage costs proportional to spot price, for consumption commodities

For consumption commodities all we can say is

$$F \leq S\, e^{(r+u)(T-t)}.$$

This is because holders of commodities will be reluctant to sell the commodity (at the spot price) which they are keeping for consumption. They may buy more oil if the spot price is low, but they won't sell it if the spot price is high since then they would have no fuel.

8.6.3 Convenience Yield

Since

$$F \leq S\, e^{(r+u)(T-t)}$$

we introduce the **convenience yield** y such that

$$F = S\, e^{(r+u-y)(T-t)},$$

with $y \geq 0$. For investment commodities $y = 0$, of course.

8.6.4 Cost of Carry

Cost of carry is the sum of storage cost, interest paid to finance the asset, less any income from the asset.

Examples: Usually the drift coefficient in the Black–Scholes equation is just r, but this will change if there is a cost of carry as follows.

1. No dividends etc.: r
2. Dividend yield: $r - D$
3. Foreign exchange: $r - r_f$
4. Commodity: $r + u$

The convenience yield is *not* included in this.

8.6.5 Effect on Options

If S is the spot price then options on the spot satisfy

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r + u)S\frac{\partial V}{\partial S} - rV = 0.$$

Change variables to

$$F = S\, e^{(r+u-y)(T-t)}, \quad H(F, t) = V(S, t)$$

to get

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2 F^2 \frac{\partial^2 H}{\partial F^2} + yF\frac{\partial H}{\partial F} - rH = 0.$$

This is the relevant equation incorporating cost of carry and the convenience yield.

8.7 STOCK BORROWING AND REPO

I have many times referred to selling stock for hedging purposes, going short the stock. But in practice how can one sell a quantity that one does not own and which is naturally something you would buy, in 'positive' quantities as opposed to negative quantities? To keep it real, let's imagine you want to go short a lawnmower. You don't own a lawnmower to sell, so what can you do? Easy, just borrow one from your neighbor and sell that! When your neighbor wants his lawnmower back you have to go out and buy another one to give him. If lawnmower prices have meanwhile fallen you will make a profit.

In the world of stocks and shares the same idea applies. If you want to go short a stock you must first borrow it. But this is not going to be costless, usually there is some payment to be made, like an interest charge on the amount you borrow. This should really be factored into any option pricing model. To quantify this, let's suppose that you have to pay interest at a rate of R on the value of the stock that you have borrowed. Now go through the Black–Scholes argument and include this cost in the analysis.

We begin with the lognormal random walk for the stock,

$$dS = \mu S\,dt + \sigma S\,dX.$$

And we set up a portfolio, long the option and short the stock,

$$\Pi = V(S, t) - \Delta S.$$

This then changes by

$$d\Pi = dV - \Delta\,dS,$$

which can be written, using Itô's lemma, as

$$d\Pi = \frac{\partial V}{\partial t}\,dt + \frac{\partial V}{\partial S}\,dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\,dt - \Delta\,dS.$$

Except that this is not quite right. The value of our portfolio changes because of the interest rate, the repo rate; we must pay for the borrowed stock. So scratch that equation, it should be

$$d\Pi = \frac{\partial V}{\partial t}\,dt + \frac{\partial V}{\partial S}\,dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\,dt - \Delta\,dS - R\max(\Delta, 0)S\,dt.$$

Let me explain the new term. First, the interest payment is on the *value* of the short position, that is ΔS. Second, because it is an interest *rate* the actual payment is proportional to the time step, dt. Finally, and this is the interesting part, the payment is only when Δ is positive, so that we are short the stock (remember the minus sign in front of the Δ in the portfolio). Hence the maximum function above.

To eliminate the random terms we still choose

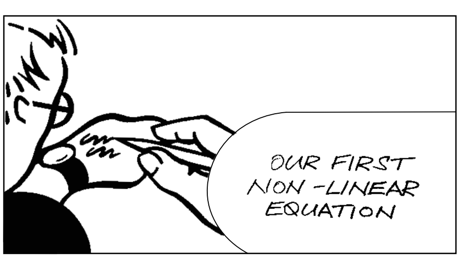
$$\Delta = \frac{\partial V}{\partial S},$$

leaving us with

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt - R\max(\Delta, 0)S\,dt.$$

Because this is deterministic we can set the return on the portfolio equal to the risk-free rate

$$d\Pi = r\Pi\,dt.$$



The end result is the partial differential equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV - RS\max\left(\frac{\partial V}{\partial S}, 0\right) = 0.$$

This is the equation for pricing derivatives in the presence of interest payments for shorting stocks. Note that it is actually a non-linear equation. The consequences of non-linearity will discussed in depth later on in the book.

8.8 TIME-DEPENDENT PARAMETERS

The next generalization concerns the term structure of parameters. In this section I show how to derive formulae for options when the interest rate, volatility and dividend yield/foreign interest rate are time dependent.

The Black–Scholes partial differential equation is valid as long as the parameters r, D and σ are known functions of time; in practice one often has a view on the future behavior of these parameters. For instance, you may want to incorporate the market's view on the direction of interest rates. Assume that you want to price options knowing $r(t)$, $D(t)$ and $\sigma(t)$. Note that when I write '$D(t)$' I am specifically assuming a time-dependent dividend yield, that is, the amount of the dividend is $D(t)S\,dt$ in a time step dt.

The equation that we must solve is now

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2(t)S^2\frac{\partial^2 V}{\partial S^2} + (r(t) - D(t))S\frac{\partial V}{\partial S} - r(t)V = 0, \tag{8.4}$$

where the dependence on t is shown explicitly.

Introduce new variables as follows:

$$\overline{S} = Se^{\alpha(t)}, \quad \overline{V} = Ve^{\beta(t)}, \quad \overline{t} = \gamma(t).$$

We are free to choose the functions α, β and γ and so we will choose them so as to eliminate all time-dependent coefficients from (8.4). After changing variables (8.4) becomes

$$\dot{\gamma}(t)\frac{\partial \overline{V}}{\partial \overline{t}} + \tfrac{1}{2}\sigma(t)^2\overline{S}^2\frac{\partial^2 \overline{V}}{\partial \overline{S}^2} + \big(r(t) - D(t) + \dot{\alpha}(t)\big)\overline{S}\frac{\partial \overline{V}}{\partial \overline{S}} - \big(r(t) + \dot{\beta}(t)\big)\overline{V} = 0, \tag{8.5}$$

where $\dot{} = d/dt$. By choosing

$$\beta(t) = \int_t^T r(\tau)\,d\tau$$

we make the coefficient of $\overline{V}$ zero and then by choosing

$$\alpha(t) = \int_t^T (r(\tau) - D(\tau))\,d\tau,$$

we make the coefficient of $\partial\overline{V}/\partial\overline{S}$ also zero. Finally, the remaining time dependence, in the volatility term, can be eliminated by choosing

$$\gamma(t) = \int_t^T \sigma^2(\tau)\,d\tau.$$

Now (8.5) becomes the much simpler equation

$$\frac{\partial \overline{V}}{\partial \overline{t}} = \tfrac{1}{2}\overline{S}^2\frac{\partial^2 \overline{V}}{\partial \overline{S}^2}. \tag{8.6}$$

The important point about this equation is that it has coefficients which are *independent of time*, and there is no mention of r, D or σ. If we use $\overline{V}(\overline{S}, \overline{t})$ to denote any solution of (8.6), then the corresponding solution of (8.5), in the original variables, is

$$V = e^{-\beta(t)}\overline{V}\left(Se^{\alpha(t)}, \gamma(t)\right). \tag{8.7}$$

Now use V_{BS} to mean any solution of the Black–Scholes equation for *constant* interest rate r_c, dividend yield D_c and volatility σ_c. This solution can be written in the form

$$V_{BS} = e^{-r_c(T-t)}\overline{V}_{BS}\left(Se^{-(r_c-D_c)(T-t)}, \sigma_c^2(T-t)\right) \tag{8.8}$$

for some function $\overline{V}_{BS}$. By comparing (8.7) and (8.8) it follows that the solution of the time-dependent parameter problem is the same as the solution of the constant parameter problem if we use the following substitutions:

$$\boxed{\begin{aligned} r_c &= \frac{1}{T-t}\int_t^T r(\tau)\,d\tau \\ D_c &= \frac{1}{T-t}\int_t^T D(\tau)\,d\tau \\ \sigma_c^2 &= \frac{1}{T-t}\int_t^T \sigma^2(\tau)\,d\tau \end{aligned}}$$

These formulae give the average, over the remaining lifetime of the option, of the interest rate, the dividend yield and the squared volatility.

Just to make things absolutely clear, here is the formula for a European call option with time-dependent parameters:

$$Se^{-\int_t^T D(\tau)\,d\tau}N(d_1) - Ee^{-\int_t^T r(\tau)\,d\tau}N(d_2)$$

where

$$d_1 = \frac{\log(S/E) + \int_t^T (r(\tau) - D(\tau))\,d\tau + \frac{1}{2}\int_t^T \sigma^2(\tau)\,d\tau}{\sqrt{\int_t^T \sigma^2(\tau)\,d\tau}}$$

and

$$d_2 = \frac{\log(S/E) + \int_t^T (r(\tau) - D(\tau))\,d\tau - \frac{1}{2}\int_t^T \sigma^2(\tau)\,d\tau}{\sqrt{\int_t^T \sigma^2(\tau)\,d\tau}}.$$

There are some conditions that I must attach to the use of these formulae. They are generally not correct if there is early exercise, or for certain types of exotic option. The question to ask to decide whether they are correct is: ‘Are all the conditions, final and boundary, preserved by the transformations?’

8.9 FORMULAE FOR POWER OPTIONS

An option with a payoff that depends on the asset price at expiry raised to some power is called a **power option**. Suppose that it has a payoff

$$\text{Payoff}(S^\alpha)$$

we can find a simple formula for the value of the option if we have a simple formula for an option with payoff given by

$$\text{Payoff}(S). \tag{8.9}$$

This is because of the lognormality of the underlying asset.

Writing

$$\mathcal{S} = S^\alpha$$

the Black–Scholes equation becomes, in the new variable $\mathcal{S}$,

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\alpha^2\sigma^2\mathcal{S}^2\frac{\partial^2 V}{\partial \mathcal{S}^2} + \alpha\left(\tfrac{1}{2}\sigma^2(\alpha-1)+r\right)\mathcal{S}\frac{\partial V}{\partial \mathcal{S}} - rV = 0.$$

Thus whatever the formula for the option value with simple payoff (8.9), the formula for the power version has S^α instead of S and adjustment made to σ, r and D.

8.10 THE log CONTRACT

The **log contract** has the payoff

$$\log(S/E).$$

The theoretical fair value for this contract is of the form

$$a(t) + b(t)\log(S/E).$$

Substituting this expression into the Black–Scholes equation results in

$$\dot{a} + \dot{b}\log(S/E) - \tfrac{1}{2}\sigma^2 b + (r-D)b - ra - rb\log(S/E) = 0,$$

where $\dot{}$ denotes d/dt. Equating terms in $\log(S/E)$ and those independent of S results in

$$b(t) = e^{-r(T-t)} \quad \text{and} \quad a(t) = \left(r - D - \tfrac{1}{2}\sigma^2\right)(T-t)e^{-r(T-t)}.$$

The two arbitrary constants of integration have been chosen to match the solution with the payoff at expiry.

This value is rather special in that the dependence of the option price on the underlying asset, S, and the volatility, σ, uncouples. One term contains S and no σ and the other contains σ and no S. We briefly saw in Chapter 7 the concept of vega hedging to eliminate volatility risk. It is conceivable, even though not entirely justifiably, that the simplicity of the log contract value makes it a useful weapon for hedging other contracts against fluctuations in volatility. Having said that, it's not exactly a highly liquid contract.

The log contract payoff can be positive or negative depending on whether $S > E$ or $S < E$. If we modify the payoff to be

$$\max(\log(S/E), 0)$$

then we have a genuine 'option' which may or may not be exercised. The value of this option is

$$e^{-r(T-t)}\sigma\sqrt{T-t}N'(d_2) + e^{-r(T-t)}\left(\log(S/E) + \left(r - D - \tfrac{1}{2}\sigma^2\right)(T-t)\right)N(d_2).$$

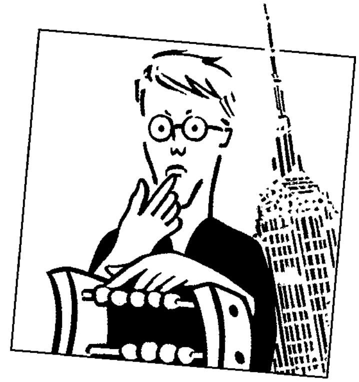
8.11 SUMMARY

In this chapter I made some very simple generalizations to the Black–Scholes world. I showed the effect of discretely paid dividends on the value of an option, deriving a jump condition by a no-arbitrage argument. Generally, this condition would be applied numerically and its implementation is discussed in Chapter 78. I also showed how time-dependent parameters can be incorporated into the pricing of simple vanilla options.

FURTHER READING

- See Merton (1973) for the original derivation of the Black–Scholes formulae with time-dependent parameters.
- For a model with stochastic dividends, see Geske (1978).
- The practical implications of discrete dividend payments are discussed by Gemmill (1992).
- See Neuberger (1994) for further info on the log contract.

CHAPTER 9
early exercise and American options



In this Chapter...

- the meaning of 'early exercise'
- the difference between European, American and Bermudan options
- how to value American options in the partial differential equation framework
- how to decide when to exercise early
- early exercise and dividends

9.1 INTRODUCTION

American options are contracts that may be exercised early, *prior* to expiry. For example, if the option is a call, we may hand over the exercise price and receive the asset whenever we wish. These options must be contrasted with European options for which exercise is only permitted *at* expiry. Most traded stock and futures options are American style, but most index options are European.

The right to exercise at any time at will is clearly valuable. The value of an American option cannot be less than an equivalent European option. But as well as giving the holder more rights, they also give him more headaches; when should he exercise? Part of the valuation problem is deciding when is the best time to exercise. This is what makes American options much more interesting than their European cousins. Moreover, the issues I am about to raise have repercussions in many other financial problems.

9.2 THE PERPETUAL AMERICAN PUT

There is a very simple example of an American option that we can examine for the insight that it gives us in the general case. This simple example is the **perpetual American put**. This contract can be exercised for a put payoff at *any* time. There is no expiry, that's why it is called a 'perpetual' option. So we can, at any time of *our* choosing, sell the underlying and receive an amount E. That is, the payoff is

$$\max(E - S, 0).$$

We want to find the value of this option before exercise.

The first point to note is that the solution is independent of time, $V(S)$. It depends only on the level of the underlying. This is a property of perpetual options when the contract details are time-homogeneous, provided that there is a finite solution. When we come to the general, non-perpetual, American option, we unfortunately lose this property. ('Unfortunately,' since it makes it easy for us to find the solution in this special case.)

The second point to note, which is important for all American options, is that the option value can never go below the early-exercise payoff. In the case under consideration

$$V \geq \max(E - S, 0). \tag{9.1}$$

Consider what would happen if this 'constraint' were violated. Suppose that the option value is less than $\max(E - S, 0)$, I could buy the option for V, immediately exercise it by handing over the asset (worth S) and receive an amount E. I thus make

$$-\text{cost of put} - \text{cost of asset} + \text{strike price} \;= -V - S + E > 0.$$

This is a riskless profit. If we believe that there are no arbitrage opportunities then we must believe (9.1).

While the option value is strictly greater than the payoff, it must satisfy the Black–Scholes equation; I return to this point in the next section. Recalling that the option is perpetual and therefore that the value is independent of t, it must satisfy

$$\tfrac{1}{2}\sigma^2 S^2 \frac{d^2V}{dS^2} + rS\frac{dV}{dS} - rV = 0.$$

This is the ordinary differential equation you get when the option value is a function of S only. The general solution of this second-order ordinary differential equation is

$$V(S) = AS + BS^{-2r/\sigma^2},$$

where A and B are arbitrary constants.

The first part of this solution (that with coefficient A) is simply the asset: the asset itself satisfies the Black–Scholes equation. If we can find A and B we have found the solution for the perpetual American put.

Clearly, for the perpetual American put the coefficient A must be zero; as $S \to \infty$ the value of the option must tend to zero. What about B?

Let us postulate that while the asset value is 'high' we won't exercise the option. But if it falls too low we immediately exercise the option, receiving $E - S$. (Common sense tells us we don't exercise when $S > E$.) Suppose that we decide that $S = S^*$ is the value at which we exercise, i.e. as soon as S reaches this value from above we exercise. How do we choose S^*?

When $S = S^*$ the option value must be the same as the exercise payoff:

$$V(S^*) = E - S^*.$$

It cannot be less, that would result in an arbitrage opportunity, and it cannot be more or we wouldn't exercise. Continuity of the option value with the payoff gives us one equation:

$$V(S^*) = B(S^*)^{-2r/\sigma^2} = E - S^*.$$

But since both B and S^* are unknown, we need one more equation. Let's look at the value of the option as a function of S^*, eliminating B using the above. We find that for $S > S^*$

$$V(S) = (E - S^*)\left(\frac{S}{S^*}\right)^{-2r/\sigma^2}. \tag{9.2}$$

We are going to choose S^* to *maximize the option's value at any time before exercise*. In other words, what choice of S^* makes V given by (9.2) as large as possible? The reason for this is obvious, if we can exercise whenever we like then we do so in such a way to maximize our worth. We find this value by differentiating (9.2) with respect to S^* and setting the resulting expression equal to zero:

$$\frac{\partial}{\partial S^*}(E - S^*)\left(\frac{S}{S^*}\right)^{-2r/\sigma^2} = \frac{1}{S^*}\left(\frac{S}{S^*}\right)^{-2r/\sigma^2}\left(-S^* + \frac{2r}{\sigma^2}(E - S^*)\right) = 0.$$

We find that

$$S^* = \frac{E}{1 + \sigma^2/2r}.$$

This choice maximizes $V(S)$ for *all* $S \geq S^*$. The solution with this choice for S^* and with the corresponding B given by

$$\frac{\sigma^2}{2r}\left(\frac{E}{1 + \sigma^2/2r}\right)^{1+2r/\sigma^2}$$

is shown in Figure 9.1.

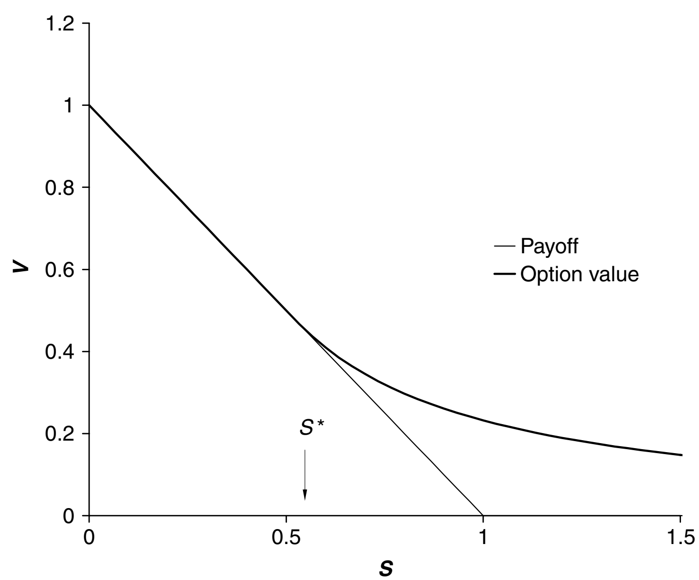


Figure 9.1 The solution for the perpetual American put.

The observant reader will notice something special about this function: The slope of the option value and the slope of the payoff function are the same at $S = S^*$. To see that this follows from the choice of S^* let us examine the difference between the option value and the payoff function:

$$(E - S^*)\left(\frac{S}{S^*}\right)^{-2r/\sigma^2} - (E - S).$$

Differentiate this with respect to S and you will find that the expression is zero at $S = S^*$.

This demonstrates, in a completely non-rigorous way, that if we want to maximize our option's value by a careful choice of exercise strategy, then this is equivalent to solving the Black–Scholes equation with continuity of option *value* and option *delta*, the slope. This is called the **high-contact** or **smooth-pasting condition**.

> The American option value is maximized by
> an exercise strategy that makes the
> option value and option delta continuous

We exercise the option as soon as the asset price reaches the level at which the option price and the payoff meet. This position, S^*, is called the **optimal exercise point**.

Another way of looking at the condition of continuity of delta is to consider what happens if the delta is not continuous at the exercise point. The two possibilities are shown in Figure 9.2. In this figure the curve (a) corresponds to exercise that is not optimal because it is premature; the option value is lower than it could be. In case (b) there is clearly an arbitrage opportunity.

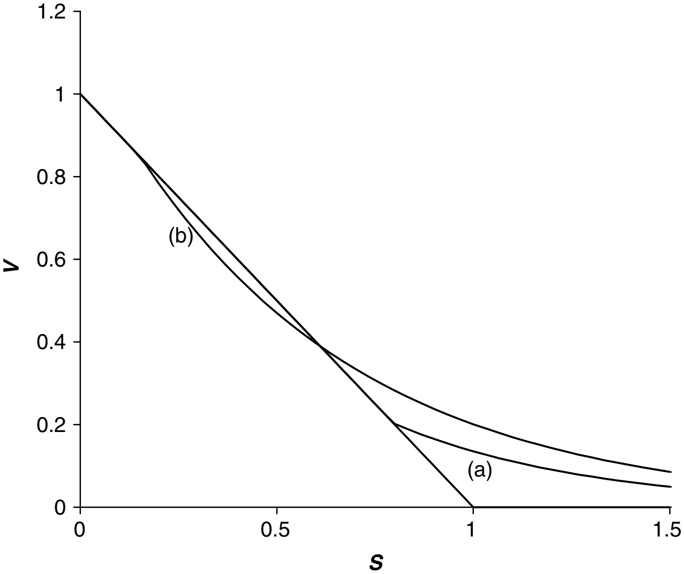


Figure 9.2 Option price when exercise is (a) too soon or (b) too late.

If we take case (a) but progressively delay exercise by lowering the exercise point, we will maximize the option value everywhere when the delta is continuous.

When there is a continuously paid and constant dividend yield on the asset, or the asset is a foreign currency, the relevant ordinary differential equation for the perpetual option is

$$\tfrac{1}{2}\sigma^2 S^2 \frac{d^2V}{dS^2} + (r - D)S\frac{dV}{dS} - rV = 0.$$

The general solution is now

$$AS^{\alpha^+} + BS^{\alpha^-},$$

where

$$\alpha^{\pm} = \frac{1}{\sigma^2}\left(-\left(r - D - \tfrac{1}{2}\sigma^2\right) \pm \sqrt{\left(r - D - \tfrac{1}{2}\sigma^2\right)^2 + 2r\sigma^2}\right),$$

with $\alpha^- < 0 < \alpha^+$. The perpetual American put now has value

$$BS^{\alpha^-},$$

where

$$B = -\frac{1}{\alpha^-}\left(\frac{E}{1 - 1/\alpha^-}\right)^{1-\alpha^-}.$$

It is optimal to exercise when S reaches the value

$$\frac{E}{1 - 1/\alpha^-}.$$

Before considering the formulation of the general American option problem, we consider one more special case.

9.3 PERPETUAL AMERICAN CALL WITH DIVIDENDS

The solution for the American perpetual call is

$$AS^{\alpha^+},$$

where

$$A = \frac{1}{\alpha^+}\left(\frac{E}{1 - 1/\alpha^+}\right)^{1-\alpha^+}$$

and it is optimal to exercise as soon as S reaches

$$S^* = \frac{E}{1 - 1/\alpha^+}$$

from below.

An interesting special case is when $D = 0$. Then the solution is $V = S$ and S^* becomes infinite. Thus it is never optimal to exercise the American perpetual call when there are no dividends on the underlying; its value is the same as the underlying. As we see in a moment, it is also never optimal to exercise the ordinary non-perpetual American call in the absence of dividends.

9.4 MATHEMATICAL FORMULATION FOR GENERAL PAYOFF

Now I build up the theory for American-style contracts with arbitrary payoff and expiry, following the standard Black–Scholes argument with minor modifications. The contract will no longer be perpetual and the option value will now be a function of both S and t.

Construct a portfolio of one American option with value $V(S,t)$ and short a number Δ of the underlying:

$$\Pi = V - \Delta S.$$

The change in value of this portfolio in excess of the risk-free rate is given by

$$d\Pi - r\Pi\, dt = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - r(V - \Delta S)\right) dt + \left(\frac{\partial V}{\partial S} - \Delta\right) dS.$$

With the choice

$$\Delta = \frac{\partial V}{\partial S}$$

this becomes

$$\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV\right) dt.$$

In the Black–Scholes argument for European options we set this expression equal to zero, since this precludes arbitrage. But it precludes arbitrage whether we are buying or selling the contract. When the contract is American the long/short relationship is asymmetrical, it is the holder of the exercise rights who controls the early-exercise feature. The writer of the option can do no more than sit back and enjoy the view. If V is the value of a long position in an American option then all we can say is that we can earn *no more* than the risk-free rate on our portfolio. Thus we arrive at the *inequality*

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \le 0. \tag{9.3}$$

The writer of the American option *can* make more than the risk-free rate if the holder does not exercise *optimally*. He also makes more profit if the holder of the option has a poor estimate of the volatility of the underlying and exercises in accordance with that estimate.

Equation (9.3) can easily be modified to accommodate dividends on the underlying.

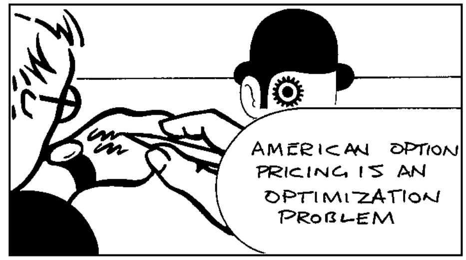
If the payoff for early exercise is $P(S,t)$, possibly time-dependent, then the no-arbitrage constraint

$$V(S,t) \ge P(S,t), \tag{9.4}$$

must apply everywhere. At expiry we have the final condition

$$V(S,T) = P(S,T). \tag{9.5}$$

> The option value is maximized if the owner of the option exercises such that
> $$\Delta = \frac{\partial V}{\partial S} \text{ is continuous} \qquad (9.6)$$



The American option valuation problem consists of (9.3), (9.4), (9.5) and (9.6).

If we substitute the Black–Scholes European call solution, in the absence of dividends, into the inequality (9.3) then it is clearly satisfied; it actually satisfies the *equality*. If we substitute the expression into the constraint (9.4) with $P(S,t) = \max(S-E, 0)$ then this too is satisfied. The conclusion is that the value of an American call option is the same as the value of a European call option when the underlying pays no dividends. Compare this with our above result that the perpetual call option should not be exercised—the American call option with a finite time to expiry should also not be exercised before expiry. To exercise before expiry would be 'sub-optimal.'

None of this is true if there are dividends on the underlying. Again, to see this simply substitute the expressions from Chapter 8 into the constraint. Since the call option has a value which approaches $Se^{-D(T-t)}$ as $S \to \infty$ there is clearly a point at which the European value fails to satisfy the constraint (9.4). *If the constraint is not satisfied* somewhere *then the problem has not been solved* anywhere. This is very important, our 'solution' must satisfy the inequalities everywhere or the 'solution' is invalid. This is due to the diffusive nature of the differential equation; an error in the solution at any point is immediately propagated *everywhere*.

The problem for the American option is what is known as a **free boundary problem**. In the European option problem we know that we must solve for all values of S from zero to infinity. When the option is American we do not know *a priori* where the Black–Scholes equation is to be satisfied; this must be found as part of the solution. This means that we do not know the position of the early exercise boundary. Moreover, except in special and trivial cases, this position is time-dependent. For example, we should exercise the American put if the asset value falls below $S^*(t)$, but how do we find $S^*(t)$?

Not only is this problem much harder than the fixed boundary problem (for example, where we know that we solve for S between zero and infinity), but this also makes the problem non linear. That is, if we have two solutions of the problem we do not get another solution if we add them together. This is easily shown by considering the perpetual American straddle on a dividend-paying stock. If this is defined as a *single* contract that may at any time be exercised for an amount $\max(S-E, 0) + \max(E-S, 0) = |S-E|$ then its value is not the same as the sum of a perpetual American put and a perpetual American call. Its solution is again of the form

$$V(S) = AS^{\alpha^+} + BS^{\alpha^-},$$

and I suggest that the reader find the solution for himself. The reason that this contract is not the sum of two other American options is that there is only one exercise opportunity; the two-option contract has one exercise opportunity per contract. If the contracts were both European then the sum of the two separate solutions would give the correct answer; the European valuation problem is linear. This contract can also be used to demonstrate that there can easily be more than one optimal exercise boundary. With the perpetual American straddle, as defined here, one should exercise either if the asset gets too low or too high. The exact positions of the boundaries can be determined by making the option and its delta everywhere continuous. One can imagine that if a contract has a really strange payoff, that there could be any number of free boundaries. Since we don't know *a priori* how many free boundaries there are going to

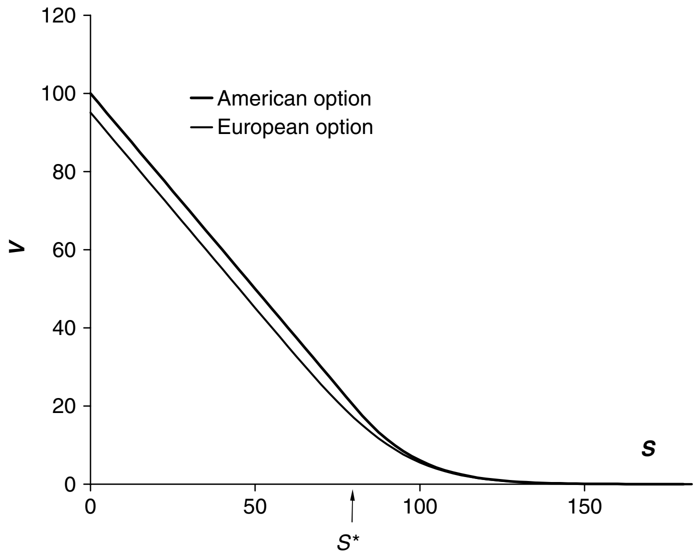


Figure 9.3 Values of a European and an American put, see text for parameter values. The optimal exercise point is marked.

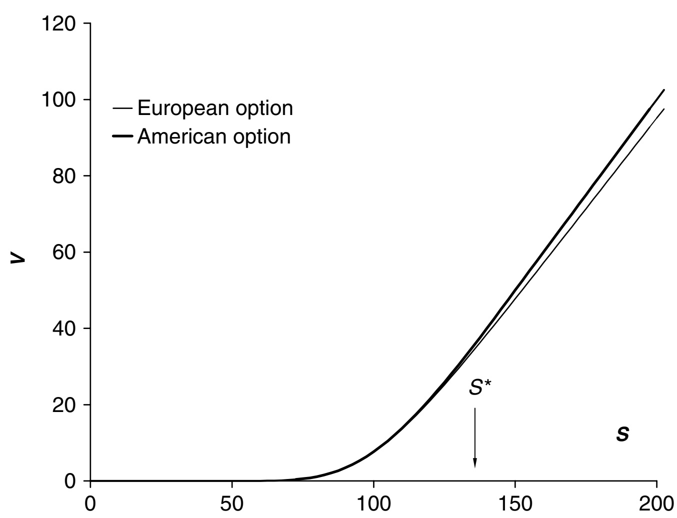
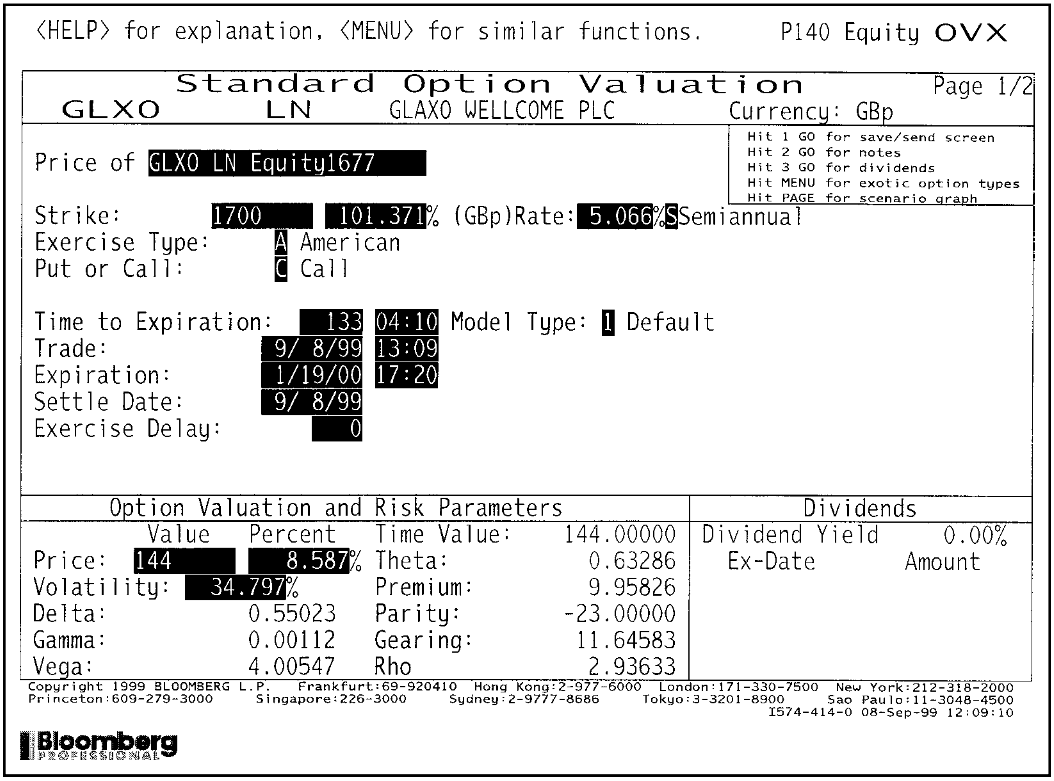


Figure 9.4 Values of a European and an American call, see text for parameter values. The optimal exercise point is marked.



<HELP> for explanation, <MENU> for similar functions. P140 Equity OVX

Standard Option Valuation Page 1/2
GLXO LN GLAXO WELLCOME PLC Currency: GBp

Hit 1 GO for save/send screen
Hit 2 GO for notes
Hit 3 GO for dividends
Hit MENU for exotic option types
Hit PAGE for scenario graph

Price of GLXO LN Equity1677

Strike: 1700 101.371% (GBp)Rate: 5.066%S Semiannual
Exercise Type: A American
Put or Call: C Call

Time to Expiration: 133 04:10 Model Type: 1 Default
Trade: 9/ 8/99 13:09
Expiration: 1/19/00 17:20
Settle Date: 9/ 8/99
Exercise Delay: 0

Option Valuation and Risk Parameters

	Value	Percent		
Price:	144	8.587%	Time Value:	144.00000
			Theta:	0.63286
Volatility:	34.797%		Premium:	9.95826
Delta:		0.55023	Parity:	-23.00000
Gamma:		0.00112	Gearing:	11.64583
Vega:		4.00547	Rho	2.93633

Dividends
Dividend Yield 0.00%
Ex-Date Amount

Copyright 1999 BLOOMBERG L.P. Frankfurt:69-920410 Hong Kong:2-977-6000 London:171-330-7500 New York:212-318-2000
Princeton:609-279-3000 Singapore:226-3000 Sydney:2-9777-8686 Tokyo:3-3201-8900 Sao Paulo:11-3048-4500
I574-414-0 08-Sep-99 12:09:10

Bloomberg PROFESSIONAL

Figure 9.5 Bloomberg option valuation. Source: Bloomberg L.P.

be (although common sense gives us a clue) it is useful to have a numerical method that can find these boundaries without having to be told how many to look for. I discuss these issues in Chapter 78.

In Figure 9.3 are shown the values of a European and American put with strike 100, volatility 20%, interest rate 5% and with one year to expiry. The position of the free boundary, the optimal exercise point, is marked. Remember that this point moves in time.

In Figure 9.4 are shown the values of a European and American call with strike 100, volatility 20%, interest rate 5% and with one year to expiry. There is a constant dividend yield of 5% on the underlying. (If there were no dividend payment then the two curves would be identical.) The position of the free boundary, the optimal exercise point, is marked.

Figure 9.5 shows the Bloomberg option valuation calculator, applied to an American option.

9.5 LOCAL SOLUTION FOR CALL WITH CONSTANT DIVIDEND YIELD

If we cannot find full solutions to non-trivial problems, we can at least find local solutions, solutions that are good approximations for some values of the asset at some times. We have seen the solution for the American call with dividends when there is a long time to expiry, what about close to expiry? I will state the results without any proof. The proofs are simple but tedious; the relevant literature is cited at the end of the chapter.

First let's consider the case $r > D$. This is usually true for options on equities, for which the dividend is small. Close to expiry the optimal exercise boundary is

$$S^*(t) \sim \frac{rE}{D}\left(1 + 0.9034\ldots\sigma\sqrt{\tfrac{1}{2}(T-t)} + \cdots\right).$$

The call should be exercised if the asset rises above this value. Note that as $T - t \to \infty$ we have from the perpetual call analysis that the free boundary tends to

$$\frac{E}{1 - 1/\alpha^+}.$$

If the asset value rises above the free boundary it is better to exercise the option to receive the dividends than to continue holding it.

Near the point $t = T$, $S = Er/D$ the option price is approximately

$$V \sim S - E + E(T-t)^{3/2} f\left(\frac{\log(SD/Er)}{\sqrt{T-t}}\right),$$

where

$$f(x) = -\frac{2r}{\sigma^2}x + 0.075\ldots\left((x^2+4)e^{-\frac{1}{4}x^2} + \tfrac{1}{2}(x^2+6x)\int_{-\infty}^{x} e^{-\frac{1}{4}s^2}\,ds\right).$$

When $D = 0$ there is no free boundary; it is never optimal to exercise early.

When $r < D$ the free boundary 'starts' from $S = E$ at time $t = T$. The local analysis is more subtle. (There is a nasty $\sqrt{(T-t)\log(T-t)}$ term.)

9.6 OTHER DIVIDEND STRUCTURES

Other dividend structures present no difficulties. The discretely paid dividend does have an interesting effect on the option value, and on the jump condition across the dividend date.

In Chapter 8 I showed that if there is a discretely paid dividend then the asset falls by the amount of the dividend:

$$S(t_d^+) = S(t_d^-) - D.$$

If the dividend takes the asset value from $S(t_d^-)$ to $S(t_d^+)$ then we must apply the jump condition

$$V\left(S(t_d^-), t_d^-\right) = V\left(S(t_d^+), t_d^+\right),$$

across the dividend date T_d. Thus

$$V\left(S, t_d^-\right) = V\left(S - D, t_d^+\right).$$

This ensures that the realized option value is continuous. When the option is American it is possible that such a jump condition takes the option value below the payoff just before the dividend date. This is not allowed. If we find that this happens, we must impose the no-arbitrage constraint that the option value is at least the payoff function. Thus the jump condition becomes

$$\boxed{V\left(S, t_d^-\right) = \max\left(V\left(S - D, t_d^+\right), P(S, t_d^-)\right)}$$

But this means that the realized option value may no longer be continuous, is this correct? Yes, this does not matter because continuity is only lost *if one should have already optimally exercised before the dividend is paid*.

9.7 ONE-TOUCH OPTIONS

We saw the European binary option in Chapters 2 and 7. The payoff for that option is \$1 if the asset is above, for a binary call, a specified level at expiry. The **one-touch option** is an American version of this. This contract can be exercised at any time for a fixed amount, \$1 if the asset is above some specified level. There is no benefit in holding the option once the level has been reached therefore it should be exercised immediately the level is reached for the first time, hence the name 'one touch.' These contracts fall into the class of 'once exotic now vanilla,' due to their popularity. They are particularly useful for hedging other contracts that also have a payoff that depends on whether or not the specified level is reached.

Since they are American-style options we must decide as part of the solution when to exercise optimally. As I have said, they would clearly be exercised as soon as the level is reached. This makes an otherwise complicated free boundary problem into a rather simple *fixed* boundary problem. For a one-touch call, we must solve the Black–Scholes equation with $V(S_u, t) = 1$, where S_u is the strike price of the contract, and $V(S, T) = 0$. We only need solve for S less than S_u. The solution of this problem is

$$V(S, t) = \left(\frac{S_u}{S}\right)^{2r/\sigma^2} N(d_6) + \frac{S}{S_u} N(d_1),$$

with the usual d_1, and

$$d_6 = \frac{\log(S/S_u) - (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}.$$

(The subscript '6' is to make the notation consistent with that in a later chapter.) The option value is shown in Figure 9.6 and the delta in Figure 9.7.

The problem for the one-touch put is obvious and the solution is

$$V(S, t) = \left(\frac{S_l}{S}\right)^{2r/\sigma^2} N(-d_6) + \frac{S}{S_l} N(-d_1),$$

with

$$d_6 = \frac{\log(S/S_l) - (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}.$$

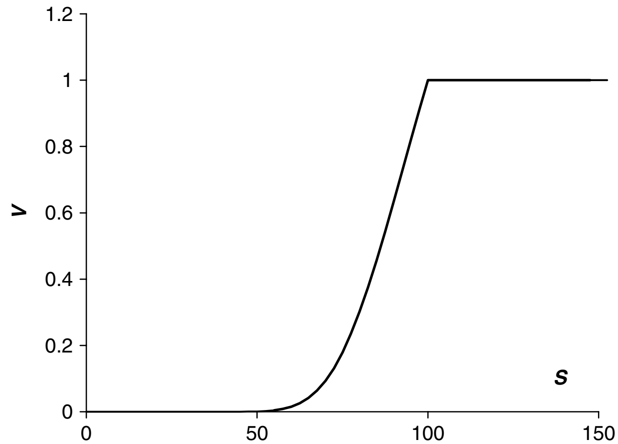


Figure 9.6 The value of a one-touch call option.

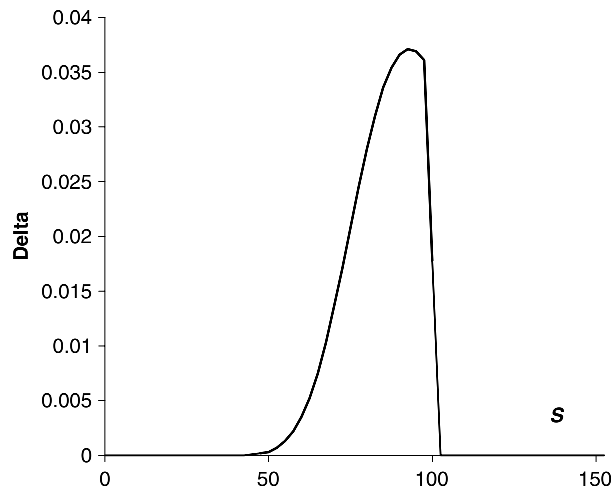


Figure 9.7 The delta of a one-touch call option.

The double one-touch option has both an upper and a lower level on which the payoff of \$1 is received. Thus $V(S_l, t) = V(S_u, t) = 1$. The solution, shown in Figure 9.8, can be found by Fourier series (see Chapter 6). Note that the value is not the sum of a one-touch call and a one-touch put.

9.8 OTHER FEATURES IN AMERICAN-STYLE CONTRACTS

The American option can be made even more interesting in many ways. Two possibilities are described in this section. These are the intermittent exercise opportunities of the Bermudan option, and the make-your-mind-up feature where the decision to exercise and the actual exercise occur at different times. Another rather simple extension is the **Israeli option** in which the writer can cancel the option early, on payment of a penalty charge.

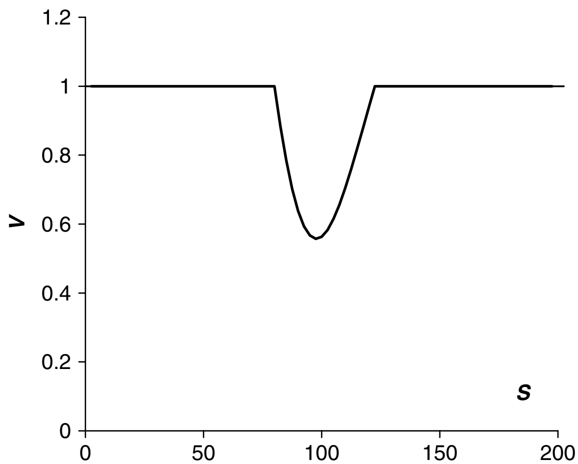


Figure 9.8 The value of a double one-touch option.

9.8.1 Bermudan Options

It is common for contracts that allow early exercise to permit the exercise only at certain specified times, and not at *all* times before expiry. For example, exercise may only be allowed on Thursdays between certain times. An option with such intermittent exercise opportunities is called a **Bermudan option**. All that this means mathematically is that the constraint (9.4) is only 'switched on' at these early exercise dates. The pricing of a such a contract numerically is, as we shall see, no harder than the pricing of American options when exercise is permitted at all times.

This situation can be made more complicated by the dependence of the exercise dates on a second asset. For example, early exercise is permitted only when a second asset is above a certain level. This makes the contract a multi-asset contract, see Chapter 11.

9.8.2 Make Your Mind Up

In some contracts the decision to exercise must be made before exercise takes place. For example, we must give two weeks' warning before we exercise, and we cannot change our mind. This contract is not hard to value theoretically. Suppose that we must give a warning of time τ. If at time t we decide to exercise at time $t+\tau$ then on exercise we receive a certain deterministic amount. To make the analysis easier to explain, assume that there is no time dependence in this payoff, so that on exercise we receive $P(S)$. The value of this payoff at a time τ earlier is $V^\tau(S,\tau)$ where $V^\tau(S,t)$ is the solution of

$$\frac{\partial V^\tau}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2V^\tau}{\partial S^2}+rS\frac{\partial V^\tau}{\partial S}-rV^\tau=0$$

with

$$V^\tau(S,0)=P(S).$$

This would have to be modified if the problem were time-inhomogeneous.

Obviously, we have

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \le 0.$$

Because $V^\tau(S,\tau)$ is the value of the contract at decision time if we have decided to exercise then our early-exercise constraint becomes

$$V(S,t) \ge V^\tau(S,\tau).$$

As an example, suppose that we get a payoff of $S - E$, this is $P(S)$. Note that there is no $\max(\cdot)$ function in this; we have said we will exercise and exercise we must, even if the asset is out of the money. The function $V^\tau(S,\tau)$ is clearly $S - Ee^{-r\tau}$ so that our **make-your-mind-up option** satisfies the constraint

$$V(S,t) \ge S - Ee^{-r\tau}.$$

A further complication is to allow one change of mind. That is, we say we will exercise in two weeks' time, but when that date comes we change our mind, and do not exercise. But the next time we say we will exercise, we must. This is also not too difficult to price theoretically.

The trick is to introduce two functions for the option value, $V_0(S,t)$ and $V_1(S,t)$. The former is the value before making the first decision to exercise, the latter is the value having made that decision but having changed your mind. We also need $V_0^\tau(S,t',t)$ and $V_1^\tau(S,t)$. The latter is simply the earlier V^τ. The former is slightly more complicated. In $V_0^\tau(S,t',t)$ the t' represents the time at which the option will be exercised or exercise is declined. The t represents the time before that date.

The problem for V_1 is exactly the same as for the basic make-your-mind-up option i.e.

$$\frac{\partial V_1}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + rS\frac{\partial V_1}{\partial S} - rV_1 \le 0$$

with

$$V_1(S,t) \ge V_1^\tau(S,\tau)$$

where

$$\frac{\partial V_0^\tau}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_0^\tau}{\partial S^2} + rS\frac{\partial V_0^\tau}{\partial S} - rV_0^\tau = 0$$

with

$$V_1^\tau(S,0) = P(S).$$

The function $V_0^\tau(S,t',t)$ satisfies

$$\frac{\partial V_0^\tau}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_0^\tau}{\partial S^2} + rS\frac{\partial V_0^\tau}{\partial S} - rV_0^\tau = 0$$

(with time derivatives with respect to t and *not* t') with

$$V_0^\tau(S,t',0) = \max(P(S), V_1(S,t')).$$

Then we have

$$\frac{\partial V_0}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_0}{\partial S^2} + rS\frac{\partial V_0}{\partial S} - rV_0 \leq 0$$

with the optimality constraint

$$V_0(S,t) \geq V_0^{\tau}(S, t+\tau, \tau).$$

Obviously, we can introduce more levels if we are permitted to change our minds a specified number of times.

In Part Two we will see many problems where we must introduce more than one function to value a single contract.

9.9 OTHER ISSUES

The pricing of American options and all the issues that this raises are important for many reasons. Some of these we describe here, but we will come back to the ideas again and again.

9.9.1 Non-linearity

The pricing of American options is a non-linear problem because of the free boundary. There are other non-linear problems in finance, some are non-linear because of the free boundary and some because the governing differential equation is itself non-linear. Non-linearity can be important for several reasons. Most obviously, non-linear problems are harder to solve than linear problems, usually requiring numerical solution.

Non-linear governing equations are found in Chapter 48 for models of pricing with transaction costs, Chapter 52 for uncertain parameter models, Chapter 58 for models of market crashes, and Chapter 59 for models of options used for speculative purposes.

9.9.2 Free-boundary Problems

Free-boundary problems, in other contexts, will be found scattered throughout the book. Again, the solution must almost always be found numerically. As an example of a free-boundary problem that is not quite an American option (but is similar), consider the **instalment option**. In this contract the owner must keep paying a premium, on prescribed dates, to keep the contract alive. If the premium is not paid then the contract lapses. Consider two cases, the first is when the premium is paid out continuously day by day, and the second, more realistic case, is when the premium is paid at discrete intervals. Part of the valuation is to decide whether or not it is worth paying the premium, or whether the contract should be allowed to lapse.

First, consider the case of continuous payment of a premium. If we pay out a constant rate $L\,dt$ in a time step dt to keep the contract alive then we must solve

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV - L \leq 0.$$

The term L represents the continual input of cash. But we would only pay the premium if it is, in some sense, 'worth it.' As long as the contract value is positive, we should maintain the payments. If the contract value ever goes negative, we should let the contract lapse. However, we can do better than this. If we impose the constraint

$$V(S,t) \geq 0,$$

with continuity of the delta, and let the contract lapse if ever $V = 0$ then we give our contract the *highest value possible*. This is very much like the American option problem, but now we must optimally cease to pay the premium (instead of optimally exercising).

Now let us consider the more realistic discrete payment case. Suppose that payments of L (not $L\,dt$) are made discretely at time t_i. The value of the contract must increase in value from before the premium is paid to just after it is paid. The reason for this is clear. Once we have paid the premium on date t_i we do not have to worry about handing over any more money until time t_{i+1}. The rise in value exactly balances the premium, L:

$$V(S,t_i^-) = V(S,t_i^+) - L,$$

where the superscripts $+$ and $-$ refer to times just after and just before the premium is paid. But we would only hand over L if the contract would be worth more than L at time t_i^+. Thus we arrive at the jump condition

$$V(S,t_i^-) = \max(V(S,t_i^+) - L, 0).$$

If $V(S,t_i^+) \leq L$ then it is optimal to discontinue payment of the premiums.

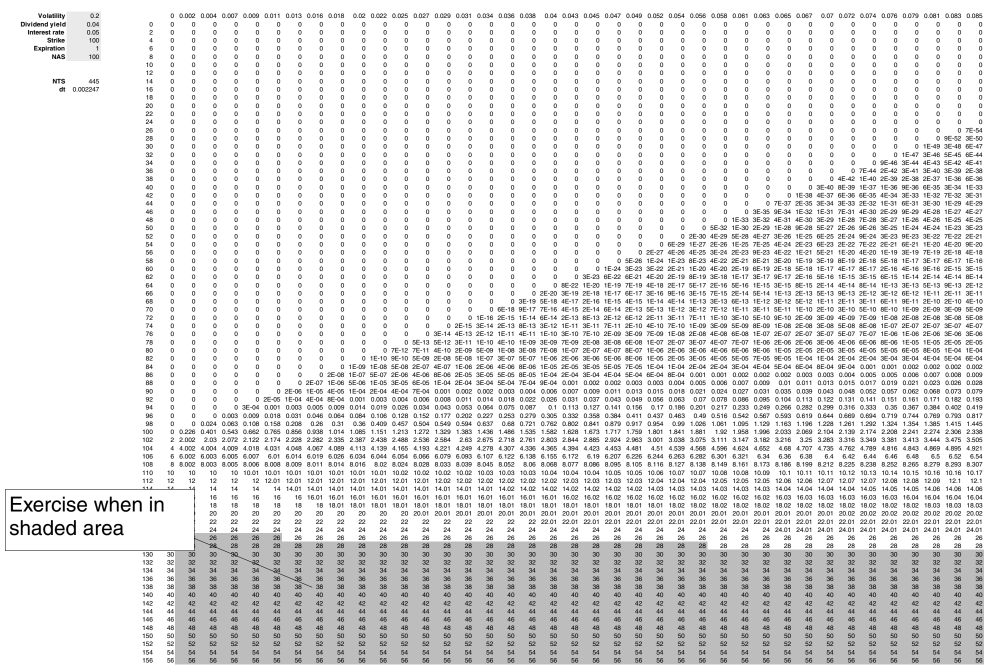


Figure 9.9 Spreadsheet showing the value of an American call option on a stock paying dividends. Shaded area is where you should exercise.

In practice, the premium L is chosen so that the value of the contract at initiation is exactly equal to L. This means that the start date is just like any other payment date.

9.9.3 Numerical Solution

Although free-boundary problems must usually be solved numerically, this is not difficult as we shall see in later chapters. We solve the relevant equation by either a finite-difference method or the binomial method.

The other numerical method that I describe is the Monte Carlo simulation. If there is any early-exercise feature in a contract this makes solution by Monte Carlo more complicated. I discuss this issue in Chapter 80.

The next three figures, Figures 9.9, 9.10, and 9.11, show the output of an explicit finite-difference method for the value of American call and put and Bermudan put respectively. You can't read the numbers but you can see that Excel's conditional formatting has been used to show the regions where the option value and the payoff are the same. Here you should exercise the option. In these figures time is in the top row. The first long column on the left is the stock price. Time goes from right to left, so that the second long column is the payoff. The second row down represents option value when stock price is zero, so the further down the row the higher the stock price.

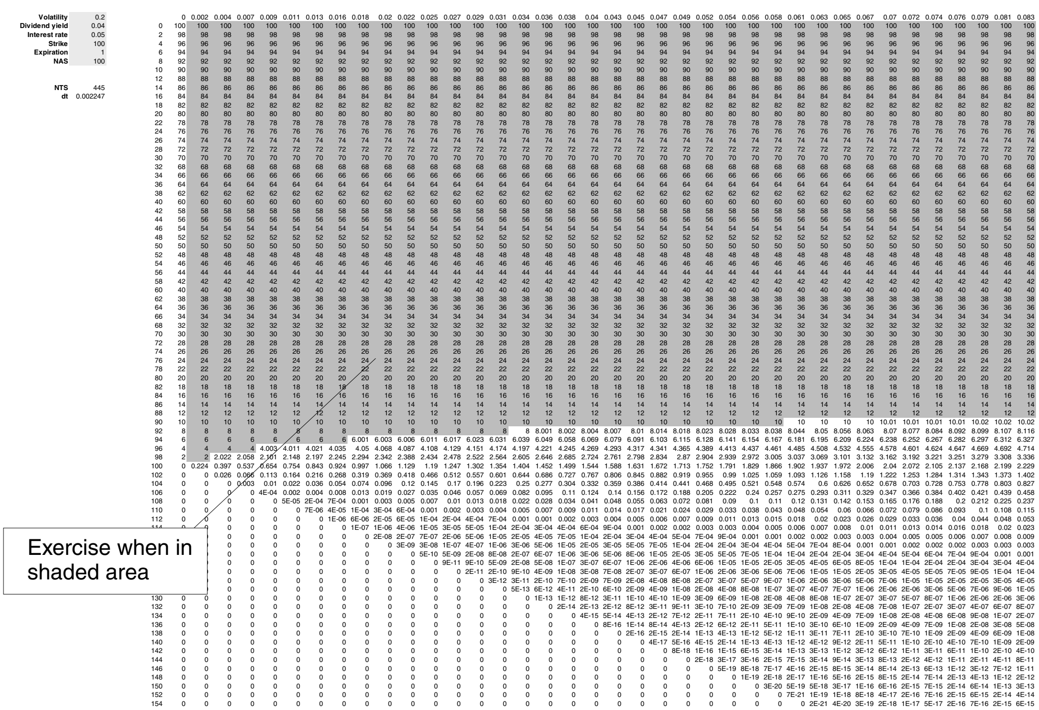


Figure 9.10 Spreadsheet showing the value of an American put option. Shaded area is where you should exercise.

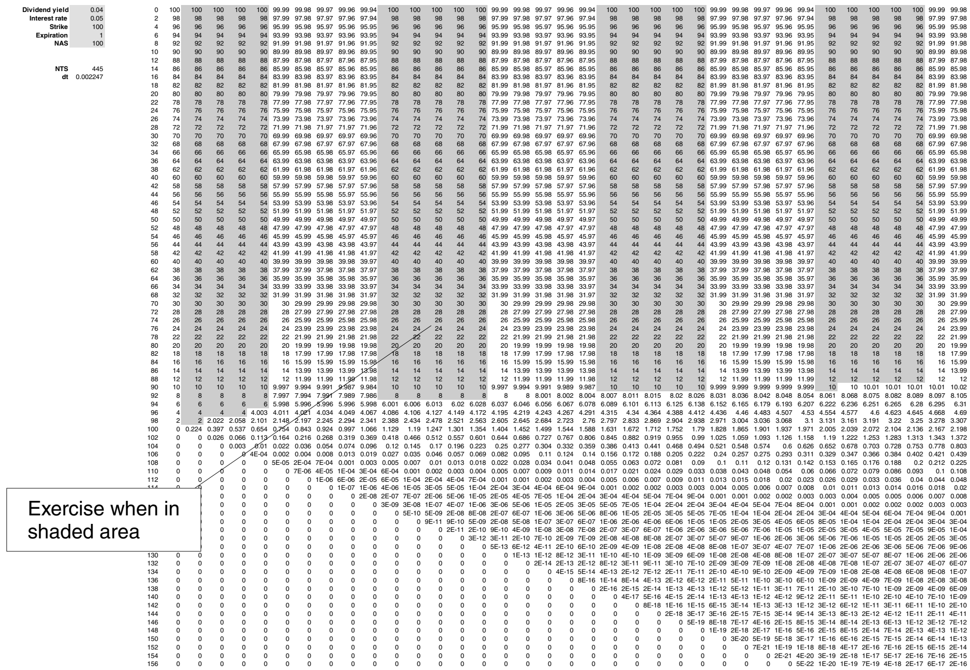


Figure 9.11 Spreadsheet showing the value of a Bermudan put option on a stock paying dividends. Shaded area is where you should exercise.

9.10 SUMMARY

This chapter raised a lot of issues that will be important for much of the rest of the book. The reader should familiarize himself with these concepts. Most importantly, free boundary problems and optimal strategies occur in many guises. So, even if a contract is not explicitly called 'American,' these modeling issues could well be present.

The exercise of American options at times other than 'optimal' is discussed at length in Chapter 63.

FURTHER READING

- See Merton (1992) and Duffie (1992) for further discussion of the 'high contact' condition.
- Solutions for American option problems can be found in Roll (1977), Whaley (1981), Johnson (1983) and Barone-Adesi & Whaley (1987).
- See Rupf, Dewynne, Howison & Wilmott (1993) for the local solution of the American call problem and Barles, Burdeau, Romano & Samsen (1995) for the put. Kruske & Keller (1998) also study the local solution of the put problem and go to a higher order of accuracy.
- The exercise strategy of the holder of an American option and its effect on the profit of both the holder and the writer are discussed in Ahn & Wilmott (1998).
- For the numerical solution of the early exercise boundary see Abboud & Zhang (2004).

CHAPTER 10
probability density functions and first-exit times

In this Chapter...

- the transition probability density function
- how to derive the forward and backward equations for the transition probability density function
- how to use the transition probability density function to solve a variety of problems
- first-exit times and their relevance to American options

10.1 INTRODUCTION

Modern finance theory, especially derivatives theory, is based on the random movement of financial quantities. In the main, the building block is the Wiener process and Normal distributions. I have shown how to derive deterministic equations for the values of options in this random world, but I have said little about the way that the future may actually evolve, which direction a stock is expected to move, or what the probability is of the option expiring in the money. This may seem perverse, but the majority of derivative theory uses ideas of hedging and no arbitrage so as to avoid dealing with the issue of randomness; uncertainty is bad. Nevertheless, it is important to acknowledge the underlying randomness, to study it, to determine properties about possible future outcomes, if one is to have a thorough understanding of financial markets.

10.2 THE TRANSITION PROBABILITY DENSITY FUNCTION

The results of this chapter will be useful for equities, currencies, interest rates or anything that evolves according to a stochastic differential equation. For that reason, I will describe the theories in terms of the general stochastic differential equation

$$dy = A(y, t)\, dt + B(y, t)\, dX \tag{10.1}$$

for the variable y. In our lognormal equity world we would have $A = \mu y$ and $B = \sigma y$, and then we would write S in place of y.

To analyze the probabilistic properties of the random walk, I will introduce the **transition probability density function** $p(y, t; y', t')$ defined by

$$\text{Prob}(a < y < b \text{ at time } t' | y \text{ at time } t) = \int_a^b p(y, t; y', t')\, dy'.$$

In words this is 'the probability that the random variable y lies between a and b at time t' in the future, given that it started out with value y at time t.'

Think of y and t as being current values with y' and t' being future values. The transition probability density function can be used to answer the question, 'What is the probability of the variable y being in a certain range at time t' given that it started out with value y at time t?'

The transition probability density function $p(y, t; y', t')$ satisfies two equations, one involving derivatives with respect to the future state and time (y' and t') and called the forward equation, and the other involving derivatives with respect to the current state and time (y and t) and called the backward equation. These two equations are parabolic partial differential equations not dissimilar to the Black–Scholes equation.[1]

I derive these two equations in the next few sections, using a simple trinomial approximation to the random walk for y.

10.3 A TRINOMIAL MODEL FOR THE RANDOM WALK

By far the easiest and most straightforward way to derive the forward and backward equations is via a trinomial approximation to the continuous-time random walk. This approximation is shown in Figure 10.1.

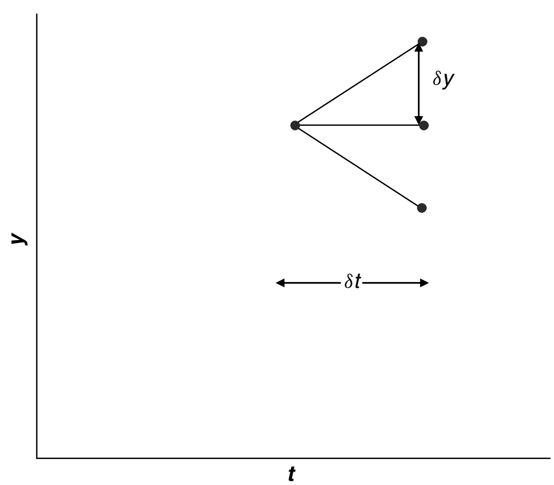


Figure 10.1 The trinomial approximation to the random walk for y.

[1] One of them, the backward equation, is *very* similar to the Black–Scholes equation.

The variable y can either rise, fall or take the same value after a time step δt. These movements have certain probabilities associated with them. I am going to choose the size of the rise and the fall to be the same, with probabilities such that the mean and standard deviation of the discrete-time approximation are the same as the mean and standard deviation of the continuous-time model over the same time step. I have three quantities to play with here, the jump size δy, the probability of a rise and the probability of a fall, but only two quantities to fix, the mean and standard deviation. The probability of not moving is such that the three probabilities sum to one. I will thus carry around the quantity δy which will drop out from the final equation.

I will use $\phi^+(y,t)$ and $\phi^-(y,t)$ to be the probabilities of a rise and fall respectively. The *mean* of the change in y after the time step is thus

$$\phi^+\delta y + (1-\phi^+-\phi^-).0 + \phi^-(-\delta y) = (\phi^+-\phi^-)\,\delta y.$$

Since $\mathrm{Var}[\cdot] = E[(\cdot)^2] - E[\cdot]^2$ the variance is

$$\phi^+\delta y^2 + (1-\phi^+-\phi^-).0^2 + \phi^-(-\delta y)^2 - \left((\phi^+-\phi^-)\,\delta y\right)^2 = \delta y^2\left(\phi^+ + \phi^- - (\phi^+-\phi^-)^2\right).$$

In both of these the arguments of ϕ^+ and ϕ^- are y and t.

The mean of the change in the continuous-time version of the random walk is, from Equation (10.1),

$$A(y,t)\,\delta t$$

and the variance is

$$B(y,t)^2\,\delta t.$$

(These are correct only to leading order, the discrete versions are exact.)

To match the mean and standard deviation we choose

$$\phi^+(y,t) = \tfrac{1}{2}\frac{\delta t}{\delta y^2}(B(y,t)^2 + A(y,t)\,\delta y)$$

and

$$\phi^-(y,t) = \tfrac{1}{2}\frac{\delta t}{\delta y^2}(B(y,t)^2 - A(y,t)\,\delta y).$$

Although I said that we have two equations for three unknowns, we must have

$$\delta y = O(\sqrt{\delta t}),$$

otherwise the diffusive properties of the problem are lost.

Now we are set to find the equations for the transition probability density function.

10.4 THE FORWARD EQUATION

In Figure 10.2 is shown a trinomial representation of the random walk. The variable y takes the value y' at time t', but how did it get there?

In our trinomial walk we can only get to the point y' from the three values $y' + \delta y$, y' and $y' - \delta y$. The probability of being at y' at time t' is related to the probabilities of being at the previous three values and *moving in the right direction*: The probability of being at y' at time

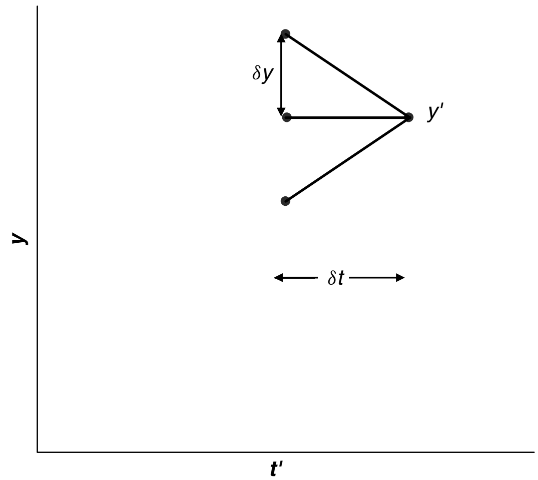


Figure 10.2 The trinomial approximation to the random walk used in finding the forward equation.

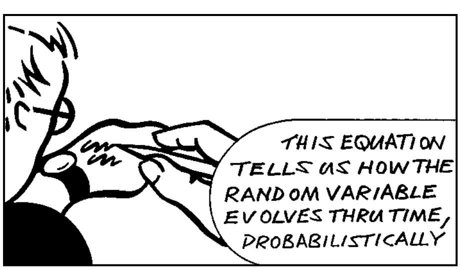
t' is equal to the probability of being at $y' + \delta y$ at the previous time $t' - \delta t$ and moving down plus the probability of being at y' at the previous time $t' - \delta t$ and moving sideways plus the probability of being at $y' - \delta y$ at the previous time $t' - \delta t$ and moving up. In symbols this is

$$\begin{aligned} p(y, t; y', t') = {} & \phi^-(y' + \delta y, t' - \delta t)p(y, t; y' + \delta y, t' - \delta t) \\ & + (1 - \phi^-(y', t' - \delta t) - \phi^+(y', t' - \delta t))p(y, t; y', t' - \delta t) \\ & + \phi^+(y' - \delta y, t' - \delta t)p(y, t; y' - \delta y, t' - \delta t). \end{aligned}$$

We can easily expand each of the terms in Taylor series about the point y', t'. For example,

$$p(y, t; y' + \delta y, t' - \delta t) \approx p(y, t; y', t') + \delta y \frac{\partial p}{\partial y'} + \tfrac{1}{2}\delta y^2 \frac{\partial^2 p}{\partial y'^2} - \delta t \frac{\partial p}{\partial t'} + \cdots.$$

I will omit the rest of the details, but the result is



$$\boxed{\begin{aligned} \frac{\partial p}{\partial t'} = {} & \tfrac{1}{2} \frac{\partial^2}{\partial y'^2}(B(y', t')^2 p) \\ & - \frac{\partial}{\partial y'}(A(y', t')p). \end{aligned}} \tag{10.2}$$

This is the **Fokker–Planck** or **forward Kolmogorov equation**. It is a forward parabolic partial differential equation, requiring initial conditions at time t and to be solved for $t' > t$.

This equation is to be used if there is some special state now and you want to know what could happen later. For example, you know the current value of y and want to know the distribution of values at some later date.

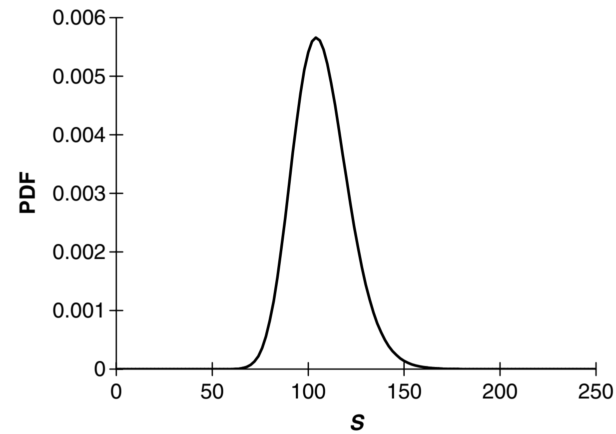


Figure 10.3 The probability density function for the lognormal random walk.

Example

The most important example to us is that of the distribution of equity prices in the future. If we have the random walk

$$dS = \mu S\,dt + \sigma S\,dX$$

then the forward equation becomes

$$\frac{\partial p}{\partial t'} = \frac{1}{2}\frac{\partial^2}{\partial S'^2}(\sigma^2 S'^2 p) - \frac{\partial}{\partial S'}(\mu S' p).$$

A special solution of this is the one having a delta function initial condition

$$p(S, t; S', t) = \delta(S' - S),$$

representing a variable that begins with certainty with value S at time t. The solution of this problem is

$$\boxed{p(S, t; S', t') = \frac{1}{\sigma S'\sqrt{2\pi(t'-t)}} e^{-(\log(S/S') + (\mu - 1/2\sigma^2)(t'-t))^2/2\sigma^2(t'-t)}.} \qquad (10.3)$$

This is plotted as a function of S' in Figure 10.3 and as a function of both S' and t' in Figure 10.4.

10.5 THE STEADY-STATE DISTRIBUTION

Some random walks have a steady-state distribution. That is, in the long run as $t' \to \infty$ the distribution $p(y, t; y', t')$ as a function of y' settles down to be independent of the starting state y and time t. Loosely speaking, this requires at least that the random walk is time homogeneous, i.e. that A and B are independent of t, asymptotically. Some random walks have no such steady state even though they have a time-independent equation; the lognormal random walk either grows without bound or decays to zero.

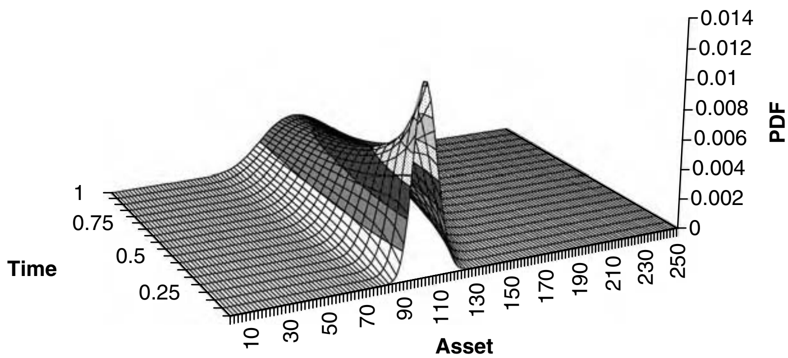
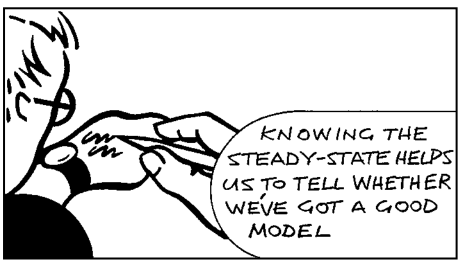


Figure 10.4 The probability density function for the lognormal random walk evolving through time.



If there is a steady-state distribution $p_\infty(y')$ then it satisfies

$$\tfrac{1}{2}\frac{d^2}{dy'^2}(B_\infty^2 p_\infty) - \frac{d}{dy'}(A_\infty p_\infty) = 0.$$

In this equation A_∞ and B_∞ are the functions in the limit $t \to \infty$. We'll see this equation used several times in later chapters, sometimes to calculate $p_\infty(y')$ knowing A and B and sometimes to calculate A knowing $p_\infty(y')$ and B.

10.6 THE BACKWARD EQUATION

Now we come to find the backward equation. This will be useful if we want to calculate probabilities of reaching a specified final state from various initial states. It will be a backward parabolic partial differential equation requiring conditions imposed in the future, and solved backwards in time. The equation I am about to derive is very similar to the Black–Scholes equation for the fair value of an option, indeed, the value of an option can be interpreted as an expectation over possible future states; much more of this later. Here is the derivation.

The maths may be easier for the backward equation, however, the justification for the relationship between the probabilities of being at the four 'nodes' is more subtle than in the derivation of the forward equation.

So, let's look at a concrete example (see Figure 10.5).

At 5pm you are in the office. (This is the point (y, t).) At 6pm you will be at one of three places: The Pub; Still at the office, working late; Madame Jojo's. (These are the points $(y + \delta y, t + \delta t)$, $(y, t + \delta t)$ and $(y - \delta y, t + \delta t)$.)

We are going to look at the probability that at midnight you are tucked up in bed. (This is the point (y', t').)

Remember that $p(y, t; y', t')$ represents the probability of being at the future point (y', t'), bed at midnight, given that you started at (y, t), the office at 5pm. You can only get to the bed at midnight via either the pub, the office or Madame Jojo's at 6pm. What happens after 6pm doesn't matter (you may not even remember!), we are only concerned with the probability that you are in bed at midnight, not how you got there.

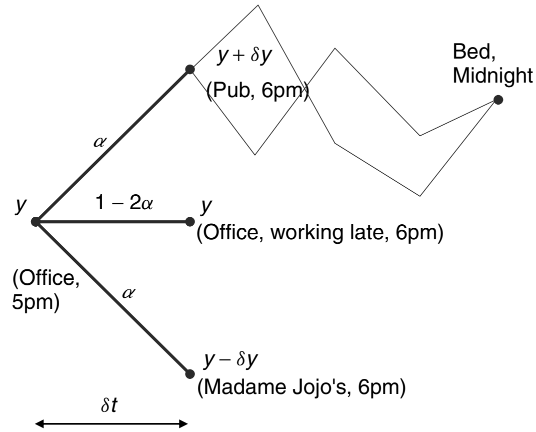


Figure 10.5 The pub crawl example used in deriving the backward equation.

In words: The probability of being in bed at midnight (given that you were in the office at 5pm) is the probability of going to the pub and being there at 6pm (given that you were in the office at 5pm) plus the probability of working late at the office and being there at 6pm (given that you were in the office at 5pm) plus the probability of going to Madame Jojo's and being there at 6pm (given that you were in the office at 5pm).

The 'given that you were in the office at 5pm' is mathematically the ' $; y, t)$ ' bit of the transition probability density function.

In symbols we can write this as

$$p(y, t; y', t') = \phi^{+}(y, t)p(y + \delta y, t + \delta t; y', t') + (1 - \phi^{+}(y, t) - \phi^{-}(y, t))p(y, t + \delta t; y', t') + \phi^{-}(y, t)p(y - \delta y, t + \delta t; y', t').$$

The Taylor series expansion leads to the **backward Kolmogorov equation**

$$\boxed{\frac{\partial p}{\partial t} + \tfrac{1}{2}B(y, t)^2 \frac{\partial^2 p}{\partial y^2} + A(y, t)\frac{\partial p}{\partial y} = 0.} \qquad (10.4)$$

Example

The transition probability density function (10.3) for the lognormal random walk satisfies this equation, but note the different independent variables.

10.7 FIRST-EXIT TIMES

The **first-exit time** is the time at which the random variable reaches a given boundary. Perhaps we want to know how long before a certain level is reached or perhaps we want to know how long before an American option should be optimally exercised. An example of a first-exit time is given in Figure 10.6.

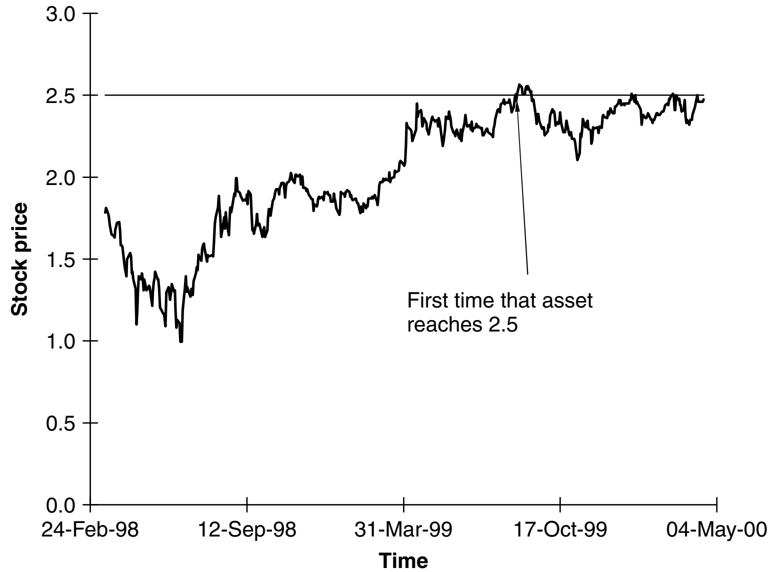


Figure 10.6 An example of a first-exit time.

Questions to ask about first-exit times are 'What is the probability of an asset level being reached before a certain time?', 'How long do we expect it to take for an interest rate to fall to a given level?' I will address these problems now.

10.8 CUMULATIVE DISTRIBUTION FUNCTIONS FOR FIRST-EXIT TIMES

What is the probability of your favorite asset doubling or halving in value in the next year? This is a question that can be answered by the solution of a simple diffusion equation. It is an example of the more general question, 'What is the probability of a random variable leaving a given range before a given time?' This question is illustrated in Figure 10.7.

Let me introduce the function $C(y, t; t')$ as the probability of the variable y leaving the region Ω before time t'. This function can be thought of as a cumulative distribution function. This function also satisfies the backward equation

$$\frac{\partial C}{\partial t} + \tfrac{1}{2}B(y, t)^2\frac{\partial^2 C}{\partial y^2} + A(y, t)\frac{\partial C}{\partial y} = 0.$$

What makes the problem different from that for the transition probability density function are the boundary and final conditions. If the variable y is actually *on* the boundary of the region Ω then clearly the probability of exiting is one:

$$C(y, t, t') = 1 \quad \text{on the edge of } \Omega.$$

On the other hand if we are inside the region Ω at time t' then there is no time left for the variable to leave the region and so the probability is zero. Thus we have

$$C(y, t', t') = 0.$$

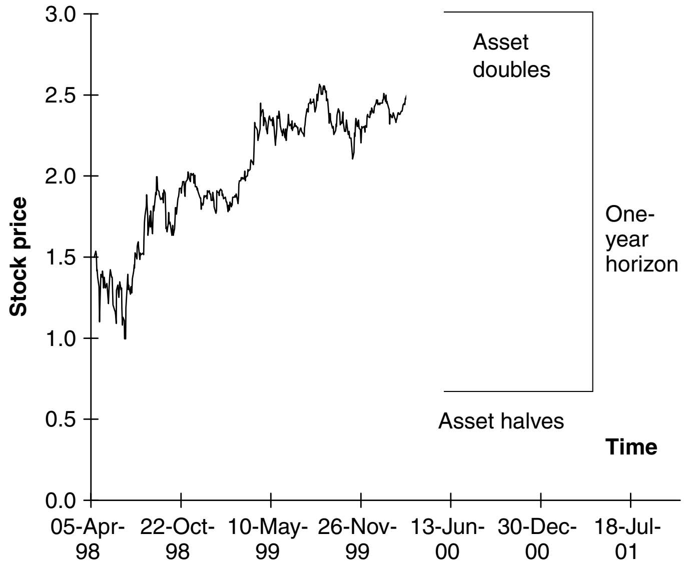


Figure 10.7 What is the probability of the asset leaving the region before the given time?

10.9 EXPECTED FIRST-EXIT TIMES

In the previous section I showed how to calculate the probability of leaving a given region. We can use this function to find the *expected* time to exit. Once we have found C then it is simple to find the **expected first-exit time**. Let me call the expected first-exit time $u(y, t)$. It is a function of where we start out, y and t.

Since C is a cumulative distribution function the expected first-exit time can be written as

$$u(y, t) = \int_t^{\infty} (t' - t) \frac{\partial C}{\partial t'} \, dt'.$$

After an integration by parts we get

$$u(y, t) = \int_t^{\infty} 1 - C(y, t; t') \, dt'.$$

The function C satisfies the backward equation in y and t so that, after differentiating under the integral sign, we find that u satisfies the equation

$$\frac{\partial u}{\partial t} + \tfrac{1}{2} B(y, t)^2 \frac{\partial^2 u}{\partial y^2} + A(y, t) \frac{\partial u}{\partial y} = -1. \tag{10.5}$$

Since C is one on the boundary of Ω, u must be zero around the boundary of the region. What about the final condition? Typically one solves over a region Ω that is bounded in time, for example as shown in Figure 10.8.

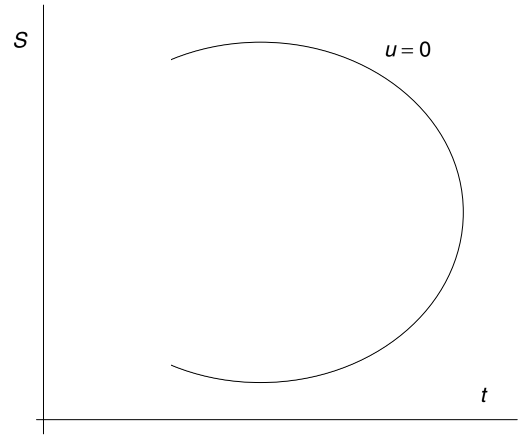


Figure 10.8 The first-exit time problem.

Example

When the stochastic differential equation is independent of time, that is, both A and B are functions of y only, and the region Ω is also time homogeneous, then there may be a steady-state solution of (10.5). Returning to the logarithmic asset problem, what is the expected time for the asset to leave the range (S_0, S_1)? The answer to this question is the solution of

$$\tfrac{1}{2}\sigma^2 S^2 \frac{d^2u}{dS^2} + \mu S \frac{du}{dS} = -1,$$

with

$$u(S_0) = u(S_1) = 0,$$

and is

$$u(S) = \frac{1}{\frac{1}{2}\sigma^2 - \mu}\left(\log(S/S_0) - \frac{1 - (S/S_0)^{1-2\mu/\sigma^2}}{1 - (S_1/S_0)^{1-2\mu/\sigma^2}}\log(S_1/S_0)\right).$$

10.10 ANOTHER EXAMPLE OF OPTIMAL STOPPING

You hold some investment, it goes up in value, it goes down in value. Generally speaking, it's going nowhere fast. When should you sell it? Let's make two big assumptions.

First of all, let's say that you know the statistical/stochastic properties of your investment's value so that you can write

$$dS = \mu(S)\,dt + \sigma(S)\,dX$$

for some known $\mu(S)$ and $\sigma(S)$ (independent of time to keep it simple).

Second, let's assume that you want to sell at the time which *maximizes the expected value of your investment*, with suitable allowance being taken for the time value of money. We'll use V to denote this maximum.

Since we want to calculate an expectation, the relevant equation to be solved is the backward Kolmogorov equation. And because the governing stochastic differential equation is time homogeneous and we have no finite time horizon we must solve for $V(S)$ with no t dependence:

$$\tfrac{1}{2}\sigma(S)^2\frac{d^2V}{dS^2}+\mu(S)\frac{dV}{dS}-rV=0.$$

The last term on the left is the usual time-value-of-money term.

This must be solved subject to

$$V\geq S$$

with continuity of V and dV/dS. This constraint ensures that we maximize our expected value.

Example

When we have a logarithmic asset so that $\mu(S)=\mu S$ and $\sigma(S)=\sigma S$ we get some interesting results.

The general solution of the second order ordinary differential equation is then

$$AS\alpha+ + BS^{\alpha-}$$

where A and B are arbitrary constants and

$$\alpha^{\pm}=\frac{1}{\sigma^2}\left(-\mu+\tfrac{1}{2}\sigma^2\pm\sqrt{\left(\mu-\tfrac{1}{2}\sigma^2\right)^2+2r\sigma^2}\right).$$

If $\mu>r$ then there is no finite solution for V (with a finite time horizon) and one should never sell the asset. If $\mu<r$ then $V=S$ and one should immediately sell the asset. (Remember that these simple results only apply under the second assumption, the risk inherent in the position has not been accounted for.)

How does all this change when there are dividends on the asset? That's an exercise for the reader.

10.11 EXPECTATIONS AND BLACK–SCHOLES

The transition probability density $p(S,t;S',t')$ for S following the random walk

$$dS=\mu S\,dt+\sigma S\,dX$$

satisfies

$$\frac{\partial p}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 p}{\partial S^2}+\mu S\frac{\partial p}{\partial S}=0. \qquad (10.6)$$

This is the backward Kolmogorov equation.

To calculate the expected value of some function $F(S)$ at time T we must solve (10.6) for the function $p_F(S,t)$ with

$$p_F(S,T)=F(S).$$

If the function $F(S)$ represents an amount of money received at time T then it is natural to ask what is the present value of the expected amount received. In other words, what is the expected amount of an option's payoff? To calculate this present value we simply multiply by the discount factor, this gives

$$e^{-r(T-t)}p_F(S,t), \tag{10.7}$$

when interest rates are constant. In this we have calculated the present value today of the expected payoff received at time T, given that today, time t, the asset value is S. Call the function (10.7) $V(S,t)$, what equation does it satisfy? Substituting

$$p_F(S,t) = e^{r(T-t)}V(S,t)$$

into (10.6) we find that $V(S,t)$ satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} - rV = 0.$$

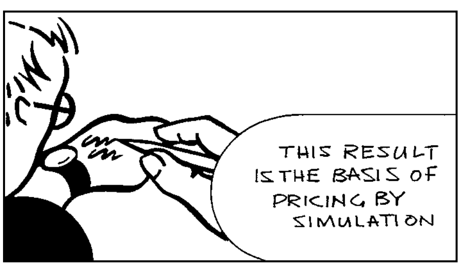
This looks very like the Black–Scholes equation, with one small difference. In the Black–Scholes equation there is no μ.

Now forget that there is such a quantity as μ and replace it in this equation with r. The resulting equation is *exactly* the Black–Scholes equation. There must be something special about the random walk in which μ is replaced by r:

$$dS = rS\,dt + \sigma S\,dX. \tag{10.8}$$

This is called the **risk-neutral random walk**. It has the same drift as money in the bank.

Our conclusion is that



> The fair value of an option is the present value of the expected payoff at expiry under a risk-neutral random walk for the underlying

We can write

$$\boxed{\text{option value} = e^{-r(T-t)}E\left[\text{payoff}(S)\right]}$$

provided that the expectation is with respect to the risk-neutral random walk, not the *real* one.

Think of the phrase 'risk neutral' as meaning 'not caring about risk' or, more technically, 'not requiring any extra return for taking risk.'

This result is the main contribution to finance theory of what is known as the martingale approach to pricing. It is very elegant and powerful. But also a bit of a dead end. You don't need thousands of books on option pricing if the answer is simply that an option price is the present value of an expected payoff. The most exciting areas of finance are those where the model leads to more interesting results. In this regard, the martingale approach is a bit of a one-trick pony. However, I would suggest you read the book by Joshi (2003) which compares and contrasts the martingale approach with others.

In this expression we see the short-term interest rate playing two distinct roles. First, it is used for discounting the payoff to the present. This is the term $e^{-r(T-t)}$ outside the expectation. Second, the return on the asset in the risk-neutral world is expected to be $rS\,dt$ in a time step dt.

10.12 A COMMON MISCONCEPTION

Ever heard that the delta is the probability of an option ending up in the money? It's a commonly held belief and one that is totally wrong. Let's look at a call option and see how this confusion has arisen, and try to correct it.

If the underlying follows the lognormal random walk

$$dS = \mu S\,dt + \sigma S\,dX$$

then it is a relatively straightforward task to find the formula for the probability of the stock ending above some level E at time T given that it starts at S at time t, see Chapter 10 for ways in which this can be done. The formula is just

$$N(d_1')$$

where

$$d_1' = \frac{\log(S/E) + (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

Now superficially this looks like the formula for a call option's delta. On closer inspection you will see that the d_1' and the d_1 are similar but not identical. There are two differences, and therefore two reasons why the option's delta is *not* the probability of an option ending up in the money. These two reasons are

1. There is a μ in the formula for the probability. There is no μ in the option's delta. We want to know what the real probability of ending up in the money is, and option prices have nothing to do with real probabilities.
2. There is a sign difference.

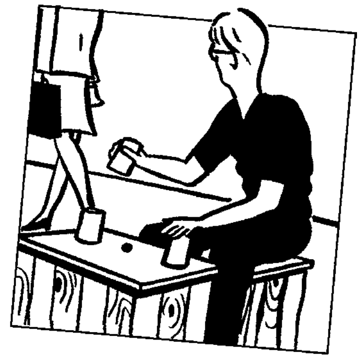
Given the important distinction between real and risk neutral and the even more important difference between plus and minus, how on earth did this misconception arise? And who is to blame? String him up!

10.13 SUMMARY

Although probability theory underpins most finance theory, it is possible to go a long way without even knowing what a transition probability density function is. But it is important in many circumstances to remember the foundation of uncertainty and to examine the future in a probabilistic sense. A couple of obvious examples spring to mind. First, if you own an American option when do you expect to exercise it? The value depends theoretically on the parameter σ in the asset price random walk but the expected time to exercise also depends on μ; the payoff may be certain because of hedging but you cannot be certain whether you will still hold the option at expiry. The second example concerns speculation with options. What if you *don't* hedge? In that case your final payoff is at the mercy of the markets, it is uncertain and can only be described probabilistically. Both of these problems can be addressed via transition probability density functions. In Chapter 59 we explore the random behavior of unhedged option positions.

FURTHER READING

- A very good book on probability theory that carefully explains the derivation, meaning and use of transition probability density functions is Cox & Miller (1965).
- The two key references that rigorously relate option values and results from probability theory are by Harrison & Kreps (1979) and Harrison & Pliska (1981).
- Atkinson & Wilmott (1993) discuss transition densities for moving averages of asset prices.



CHAPTER 11
multi-asset options

In this Chapter...

- how to model the behavior of many assets simultaneously
- estimating correlation between asset price movements
- how to value and hedge options on many underlying assets in the Black–Scholes framework
- the pricing formula for European non-path-dependent options on dividend-paying assets
- how to price and hedge quantos and the role of correlation

11.1 INTRODUCTION

In this chapter I introduce the idea of higher dimensionality by describing the Black–Scholes theory for options on more than one underlying asset. This theory is perfectly straightforward; the only new idea is that of correlated random walks and the corresponding multifactor version of Itô's lemma.

Although the modeling and mathematics is easy, the final step of the pricing and hedging, the 'solution,' can be extremely hard indeed. I explain what makes a problem easy, and what makes it hard, from the numerical analysis point of view.

11.2 MULTI-DIMENSIONAL LOGNORMAL RANDOM WALKS

The basic building block for option pricing with one underlying is the lognormal random walk

$$dS = \mu S\,dt + \sigma S\,dX.$$

This is readily extended to a world containing many, say d, assets via models for each underlying

$$dS_i = \mu_i S_i\,dt + \sigma_i S_i\,dX_i.$$

Here S_i is the price of the ith asset, $i = 1, \ldots, d$, and μ_i and σ_i are the drift and volatility of that asset respectively and dX_i is the increment of a Wiener process. We can still continue to think of dX_i as a random number drawn from a Normal distribution with mean zero and standard deviation $dt^{1/2}$ so that

$$E[dX_i] = 0 \quad \text{and} \quad E[dX_i^2] = dt$$

but the random numbers dX_i and dX_j are **correlated**:

$$E[dX_i\, dX_j] = \rho_{ij}\, dt.$$

Here ρ_{ij} is the correlation coefficient between the ith and jth random walks. The symmetric matrix with ρ_{ij} as the entry in the ith row and jth column is called the **correlation matrix**. For example, if we have seven underlyings $d = 7$ the correlation matrix will look like this:

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} & \rho_{16} & \rho_{17} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} & \rho_{25} & \rho_{26} & \rho_{27} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} & \rho_{35} & \rho_{36} & \rho_{37} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 & \rho_{45} & \rho_{46} & \rho_{47} \\ \rho_{51} & \rho_{52} & \rho_{53} & \rho_{54} & 1 & \rho_{56} & \rho_{57} \\ \rho_{61} & \rho_{62} & \rho_{63} & \rho_{64} & \rho_{65} & 1 & \rho_{67} \\ \rho_{71} & \rho_{72} & \rho_{73} & \rho_{74} & \rho_{75} & \rho_{76} & 1 \end{pmatrix}$$

Note that $\rho_{ii} = 1$ and $\rho_{ij} = \rho_{ji}$. The correlation matrix is positive definite, so that $\mathrm{y}^T \mathbf{\Sigma} \mathrm{y} \geq 0$. The **covariance matrix** is simply

$$\mathrm{M}\mathbf{\Sigma}\mathrm{M},$$

where M is the matrix with the σ_i along the diagonal and zeros everywhere else.

To be able to manipulate functions of many random variables we need a multidimensional version of Itô's lemma. If we have a function of the variables $S_1, \ldots, S_d$ and t, $V(S_1, \ldots, S_d, t)$, then

$$dV = \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} \right) dt + \sum_{i=1}^{d} \frac{\partial V}{\partial S_i} dS_i.$$

We can get to this same result by using Taylor series and the rules of thumb:

$$dX_i^2 = dt \quad \text{and} \quad dX_i dX_j = \rho_{ij}\, dt.$$

11.3 MEASURING CORRELATIONS

If you have time series data at intervals of δt for all d assets you can calculate the correlation between the returns as follows. First, take the price series for each asset and calculate the return over each period. The return on the ith asset at the kth data point in the time series is simply

$$R_i(t_k) = \frac{S_i(t_k + \delta t) - S_i(t_k)}{S_i(t_k)}.$$

The historical volatility of the ith asset is

$$\sigma_i = \sqrt{\frac{1}{\delta t(M-1)}\sum_{k=1}^{M}(R_i(t_k) - \overline{R}_i)^2}$$

where M is the number of data points in the return series and $\overline{R}_i$ is the mean of all the returns in the series.

The covariance between the returns on assets i and j is given by

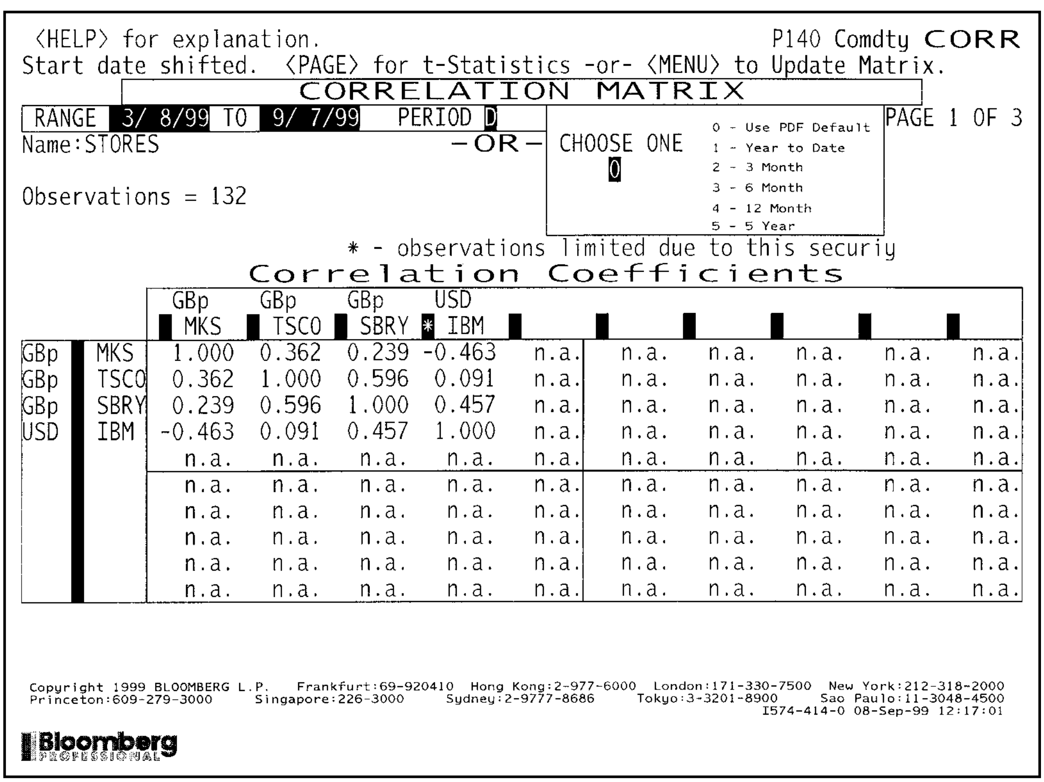
$$\frac{1}{\delta t(M-1)}\sum_{k=1}^{M}(R_i(t_k) - \overline{R}_i)(R_j(t_k) - \overline{R}_j).$$

The correlation is then

$$\frac{1}{\delta t(M-1)\sigma_i\sigma_j}\sum_{k=1}^{M}(R_i(t_k) - \overline{R}_i)(R_j(t_k) - \overline{R}_j).$$

In Excel correlation between two time series can be found using the CORREL worksheet function, or Tools | Data Analysis | Correlation.

Figure 11.1 shows the correlation matrix for Marks & Spencer, Tesco, Sainsbury and IBM.



<HELP> for explanation. P140 Comdty CORR
Start date shifted. <PAGE> for t-Statistics -or- <MENU> to Update Matrix.

CORRELATION MATRIX

RANGE 3/ 8/99 TO 9/ 7/99 PERIOD D
Name:STORES

Observations = 132

-OR- CHOOSE ONE 0

0 - Use PDF Default
1 - Year to Date
2 - 3 Month
3 - 6 Month
4 - 12 Month
5 - 5 Year

PAGE 1 OF 3

* - observations limited due to this securiy

Correlation Coefficients

		GBp MKS	GBp TSCO	GBp SBRY	USD * IBM						
GBp	MKS	1.000	0.362	0.239	-0.463	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
GBp	TSCO	0.362	1.000	0.596	0.091	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
GBp	SBRY	0.239	0.596	1.000	0.457	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
USD	IBM	-0.463	0.091	0.457	1.000	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Copyright 1999 BLOOMBERG L.P. Frankfurt:69-920410 Hong Kong:2-977-6000 London:171-330-7500 New York:212-318-2000
Princeton:609-279-3000 Singapore:226-3000 Sydney:2-9777-8686 Tokyo:3-3201-8900 Sao Paulo:11-3048-4500
I574-414-0 08-Sep-99 12:17:01

Bloomberg PROFESSIONAL

Figure 11.1 Some correlations. Source: Bloomberg L.P.

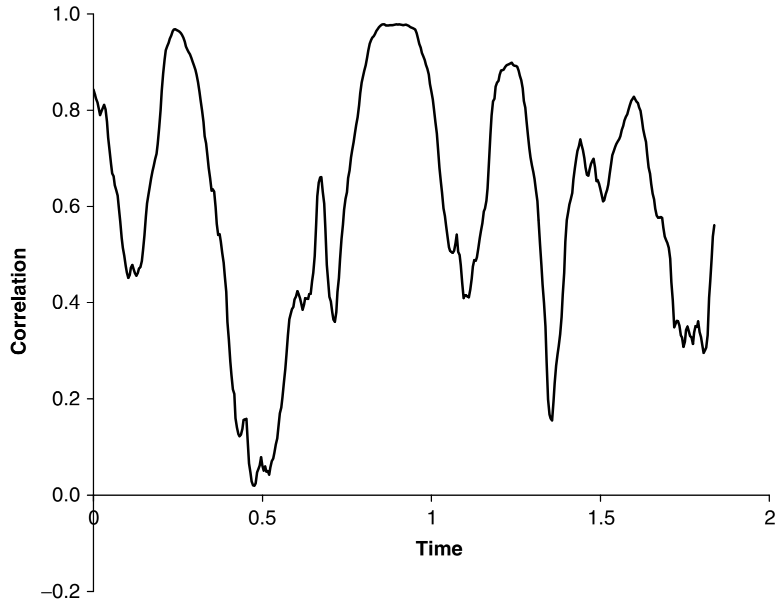


Figure 11.2 A correlation time series.

Correlations measured from financial time series data are notoriously unstable. If you split your data into two equal groups, up to one date and beyond that date, and calculate the correlations for each group you may find that they differ quite markedly. You could calculate a 60-day correlation, say, from several years' data and the result would look something like Figure 11.2. You might want to use a historical 60-day correlation if you have a contract of that maturity. But, as can be seen from the figure, such a historical correlation should be used with care; correlations are even more unstable than volatilities.

The other possibility is to back out an **implied correlation** from the quoted price of an instrument. The idea behind that approach is the same as with implied volatility, it gives an estimate of the market's perception of correlation.

11.4 OPTIONS ON MANY UNDERLYINGS

Options with many underlyings are called **basket options**, **options on baskets** or **rainbow options**. The theoretical side of pricing and hedging is straightforward, following the Black–Scholes arguments but now in higher dimensions.

Set up a portfolio consisting of one basket option and short a number Δ_i of each of the assets S_i:

$$\Pi = V(S_1, \dots, S_d, t) - \sum_{i=1}^{d} \Delta_i S_i.$$

The change in this portfolio is given by

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} \right) dt + \sum_{i=1}^{d} \left(\frac{\partial V}{\partial S_i} - \Delta_i \right) dS_i.$$

If we choose

$$\Delta_i = \frac{\partial V}{\partial S_i}$$

for each i, then the portfolio is hedged, and risk-free. Setting the return equal to the risk-free rate we arrive at

$$\frac{\partial V}{\partial t} + \tfrac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + r \sum_{i=1}^{d} S_i \frac{\partial V}{\partial S_i} - rV = 0. \tag{11.1}$$

This is the multidimensional version of the Black–Scholes equation. The modifications that need to be made for dividends are obvious. When there is a dividend yield of D_i on the ith asset we have

$$\boxed{\frac{\partial V}{\partial t} + \tfrac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + \sum_{i=1}^{d} (r - D_i) S_i \frac{\partial V}{\partial S_i} - rV = 0}$$

11.5 THE PRICING FORMULA FOR EUROPEAN NON-PATH-DEPENDENT OPTIONS ON DIVIDEND-PAYING ASSETS

Because there is a Green's function for this problem (see Chapter 6) we can write down the value of a European non-path-dependent option with payoff of Payoff$(S_1, \ldots, S_d)$ at time T:

$$\boxed{\begin{aligned} V &= e^{-r(T-t)} \left(2\pi (T-t)\right)^{-d/2} (\mathrm{Det}\boldsymbol{\Sigma})^{-1/2} (\sigma_1 \cdots \sigma_d)^{-1} \\ &\int_0^\infty \cdots \int_0^\infty \frac{\mathrm{Payoff}(S_1' \cdots S_d')}{S_1' \cdots S_d'} \exp\left(-\frac{1}{2} \boldsymbol{\alpha}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha}\right) dS_1' \cdots dS_d'. \\ \alpha_i &= \frac{1}{\sigma_i (T-t)^{1/2}} \left(\log\left(\frac{S_i}{S_i'}\right) + \left(r - D_i - \frac{\sigma_i^2}{2}\right)(T-t) \right) \end{aligned}} \tag{11.2}$$

This has included a constant continuous dividend yield of D_i on each asset.

11.6 EXCHANGING ONE ASSET FOR ANOTHER: A SIMILARITY SOLUTION

An **exchange option** gives the holder the right to exchange one asset for another, in some ratio. The payoff for this contract at expiry is

$$\max(q_1 S_1 - q_2 S_2, 0),$$

where q_1 and q_2 are constants.

The partial differential equation satisfied by this option in a Black–Scholes world is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i\partial S_j} + \sum_{i=1}^{2}(r - D_i)S_i\frac{\partial V}{\partial S_i} - rV = 0.$$

A dividend yield has been included for both assets. Since there are only two underlyings the summations in these only go up to two.

This contract is special in that there is a similarity reduction. Let's postulate that the solution takes the form

$$V(S_1, S_2, t) = q_1 S_2 H(\xi, t),$$

where the new variable is

$$\xi = \frac{S_1}{S_2}.$$

If this is the case, then instead of finding a function V of three variables, we only need find a function H of two variables, a much easier task. The final condition for H would then be

$$H(\xi, T) = \max\left(\xi - \frac{q_2}{q_1}, 0\right),$$

so that the ratio q_2/q_1 plays the role of the strike.

Changing variables from S_1, S_2 to ξ we must use the following for the derivatives.

$$\frac{\partial}{\partial S_1} = \frac{1}{S_2}\frac{\partial}{\partial \xi}, \quad \frac{\partial}{\partial S_2} = -\frac{\xi}{S_2}\frac{\partial}{\partial \xi},$$

$$\frac{\partial^2}{\partial S_1^2} = \frac{1}{S_2^2}\frac{\partial^2}{\partial \xi^2}, \quad \frac{\partial^2}{\partial S_2^2} = \frac{\xi^2}{S_2^2}\frac{\partial^2}{\partial \xi^2} + \frac{2\xi}{S_2^2}\frac{\partial}{\partial \xi}, \quad \frac{\partial^2}{\partial S_1\partial S_2} = -\frac{\xi}{S_2^2}\frac{\partial^2}{\partial \xi^2} - \frac{1}{S_2^2}\frac{\partial}{\partial \xi}.$$

The time derivative is unchanged. The partial differential equation now becomes

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma'^2\xi^2\frac{\partial^2 H}{\partial \xi^2} + (D_2 - D_1)\xi\frac{\partial H}{\partial \xi} - D_2 H = 0.$$

where

$$\sigma' = \sqrt{\sigma_1^2 - 2\rho_{12}\sigma_1\sigma_2 + \sigma_2^2}.$$

You will recognize this equation as being the Black–Scholes equation for a single stock with D_2 in place of r, D_1 in place of the dividend yield on the single stock and with a volatility of σ'.

From this it follows, retracing our steps and writing the result in the original variables, that

$$V(S_1, S_2, t) = q_1 S_1 e^{-D_1(T-t)} N(d_1') - q_2 S_2 e^{-D_2(T-t)} N(d_2')$$

where

$$d_1' = \frac{\log(q_1 S_1/q_2 S_2) + (D_2 - D_1 + \frac{1}{2}\sigma'^2)(T-t)}{\sigma'\sqrt{T-t}} \quad \text{and} \quad d_2' = d_1' - \sigma'\sqrt{T-t}.$$

One fascinating thing about the exchange option is the value of the hedged portfolio. Notice that the quantity of S_1 you hold is

$$\frac{\partial V}{\partial S_1} = q_1 \frac{\partial H}{\partial \xi}$$

and the quantity of S_2 you hold is

$$\frac{\partial V}{\partial S_2} = q_1 H - q_1 \xi \frac{\partial H}{\partial \xi}.$$

If you use these expressions in the portfolio value equation

$$\Pi = V - \Delta_1 S_1 - \Delta_2 S_2$$

you will find that the portfolio has zero value.

11.7 QUANTOS

There is one special, and very important type of multi-asset option. This is the cross-currency contract called a **quanto**. The quanto has a payoff defined with respect to an asset or an index (or an interest rate) in one country, but then the payoff is converted to another currency for payment. An example of such a contract would be a call on the Nikkei Dow index but paid in US dollars. This contract is exposed to the dollar-yen exchange rate and the Nikkei Dow index. We could write down the differential equation directly assuming that the underlyings satisfy lognormal random walks with correlation ρ. But we will build up the problem from first principles to demonstrate what hedging must take place.

Define $S_\$$ to be the yen-dollar exchange rate (number of dollars per yen[1]) and S_N is the level of the Nikkei Dow index. We assume that they satisfy

$$dS_\$ = \mu_\$ S_\$ \, dt + \sigma_\$ S_\$ \, dX_\$ \quad \text{and} \quad dS_N = \mu_N S_N \, dt + \sigma_N S_N \, dX_N,$$

with a correlation coefficient ρ between them.

Construct a portfolio consisting of the quanto in question, hedged with yen and the Nikkei Dow index:

$$\Pi = V(S_\$, S_N, t) - \Delta_\$ S_\$ - \Delta_N S_N S_\$.$$

Note that every term in this equation is measured in dollars. $\Delta_\$$ is the number of yen we hold short, so $-\Delta_\$ S_\$$ is the dollar value of that yen. Similarly, with the term $-\Delta_N S_N S_\$$ we have converted the yen-denominated index S_N into dollars; Δ_N is the amount of the index held short.

[1] This is the opposite of market convention; currencies are usually quoted against the dollar with pound sterling being an exception. I use this way around for this problem just to simplify the algebra.

The change in the value of the portfolio is due to the change in the value of its components and the interest received on the yen:

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma_\$^2 S_\$^2 \frac{\partial^2 V}{\partial S_\$^2} + \rho\sigma_\$\sigma_N S_\$ S_N \frac{\partial^2 V}{\partial S_\$ \partial S_N} + \tfrac{1}{2}\sigma_N^2 S_N^2 \frac{\partial^2 V}{\partial S_N^2}\right.$$
$$\left. - \rho\sigma_\$\sigma_N \Delta_N S_\$ S_N - r_f \Delta_\$ S_\$\right) dt + \left(\frac{\partial V}{\partial S_\$} - \Delta_\$ - \Delta_N S_N\right) dS_\$$$
$$+ \left(\frac{\partial V}{\partial S_N} - \Delta_N S_\$\right) dS_N.$$

There is a term in the above that we have not seen before, the $-\rho\sigma_\$\sigma_N S_\$ S_N$. This is due to the increment of the product $-\Delta_N S_N S_\$$. There is also the interest received by the yen holding, we *have* seen such a term before. We now choose

$$\Delta_\$ = \frac{\partial V}{\partial S_\$} - \frac{S_N}{S_\$}\frac{\partial V}{\partial S_N} \quad \text{and} \quad \Delta_N = \frac{1}{S_\$}\frac{\partial V}{\partial S_N}$$

to eliminate the risk in the portfolio. Setting the return on this riskless portfolio equal to the *US risk-free rate of interest* $r_\$$, since Π is measured entirely in dollars, yields

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma_\$^2 S_\$^2 \frac{\partial^2 V}{\partial S_\$^2} + \rho\sigma_\$\sigma_N S_\$ S_N \frac{\partial^2 V}{\partial S_\$ \partial S_N} + \tfrac{1}{2}\sigma_N^2 S_N^2 \frac{\partial^2 V}{\partial S_N^2}$$
$$+ S_\$ \frac{\partial V}{\partial S_\$}\left(r_\$ - r_f\right) + S_N \frac{\partial V}{\partial S_N}\left(r_f - \rho\sigma_\$\sigma_N\right) - r_\$ V = 0.$$

This completes the formulation of the pricing equation. The equation is valid for any contract with underlying measured in one currency but paid in another. To fully specify our particular quanto we must give the final conditions on $t = T$:

$$V(S_\$, S_N, T) = \max(S_N - E, 0).$$

Note that as far as the payoff is concerned we don't much care what $S_\$$ is (only that we can hedge with it somehow). We do *not* multiply this by the exchange rate. Because of the simple form of the payoff we can look for a solution that is independent of the exchange rate. Trying

$$V(S_\$, S_N, t) = W(S_N, t)$$

we find that

$$\frac{\partial W}{\partial t} + \tfrac{1}{2}\sigma_N^2 S_N^2 \frac{\partial^2 W}{\partial S_N^2} + S_N \frac{\partial W}{\partial S_N}\left(r_f - \rho\sigma_\$\sigma_N\right) - r_\$ V = 0.$$

This is the simple one-factor Black–Scholes equation. If we compare this equation with the Black–Scholes equation with a constant dividend yield we see that pricing the quanto is equivalent to using a dividend yield of

$$r_\$ - r_f + \rho\sigma_\$\sigma_N.$$

The only noticeable effect of the cross-currency feature on the option value is an adjustment to a dividend yield. This yield depends on the volatility of the exchange rate and the correlation between the underlying and the exchange rate. This is a common result for the simpler quantos.

11.8 TWO EXAMPLES

In Figure 11.3 is shown the term sheet for 'La Tricolore' Capital-guaranteed Note. This contract pays off the *second* best performing of three currencies against the French Franc, but only if the second-best performing has appreciated against the Franc, otherwise it pays off at par. This contract does not have any unusual features, and has a value that can be written as a three-dimensional integral, of the form (11.2). But what would the payoff function be? You wouldn't use a partial differential equation to price this contract. Instead you would estimate the multiple integral directly by the methods of Chapter 81.

The next example, whose term sheet is shown in Figure 11.4, is of basket equity swap. This rather complex, high-dimensional contract, is for a swap of interest payment based on three-month LIBOR and the level of an index. The index is made up of the weighted average of 20 pharmaceutical stocks. To make matters even more complex, the index uses a time averaging of the stock prices.

11.9 OTHER FEATURES

Basket options can have many of the other features that we have seen or will see. This includes early exercise and, discussed later, path dependency. Sometimes the payoff is in one asset

'La Tricolore' Capital-guaranteed Note

Issuer	XXXX
Principal Amount	FRF 100,000,000
Issue Price	98.75%
Maturity Date	Twelve months after Issue Date
Coupon	Zero
Redemption Amount	If at least two of the following three appreciation indices, namely: $\frac{\text{USD/FRF} - 6.0750}{6.0750};\ \frac{\text{GBP/FRF} - 10.2000}{10.2000};\ \frac{\text{JPY/FRF} - 0.05120}{0.05120}$ are positive at Maturity, the Note will redeem in that currency whose appreciation index is the second highest of the three; in all other circumstances the Note will redeem at Par in FRF. If the Note redeems in a currency other than FRF, the amount of that currency shall be calculated by dividing the FRF Principal Amount by the spot Currency/FRF exchange rate prevailing on the Issue Date.

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 11.3 Term sheet for 'La Tricolore' Capital-guaranteed Note.

International Pharmaceutical Basket Equity Swap

Indicative terms	
Trade Date	[ ]
Initial Valuation Date	[ ]
Effective Date	[ ]
Final Valuation Date	26th September 2002
Averaging Dates	The monthly anniversaries of the Initial Valuation Date commencing 26th March 2002 and up to and including the Expiration Date
Notional Amount	US$25,000,000
Counterparty floating amounts (US$ LIBOR)	
Floating Rate Payer	[ ]
Floating Rate Index	USD-LIBOR
Designated Maturity	Three months
Spread	Minus 0.25%
Day Count Fraction	Actual/360
Floating Rate Payment Dates	Each quarterly anniversary of the Effective Date
Initial Floating Rate Index	[ ]
The Bank Fixed and Floating Amounts (Fee, Equity Option)	
Fixed Amount Payer	XXXX
Fixed Amount	1.30% of Notional Amount
Fixed Amount Payment Date	Effective Date
Basket	A basket comprising 20 stocks and constructed as described in attached Appendix
Initial Basket Level	Will be set at 100 on the Initial Valuation Date
Floating Equity Amount Payer	XXXX
Floating Equity Amount	Will be calculated according to the performance of the basket of stocks in the following way:

$$\text{Notional Amount} * \max\left[0, \left(\frac{\text{BASKET}_{\text{average}} - 100}{100}\right)\right]$$

where

$$\text{BASKET}_{\text{average}} = 100 * \sum_{20\text{ stocks}} \left(\text{Weight} * \frac{P_{\text{average}}}{P_{\text{initial}}}\right)$$

Figure 11.4 Term sheet for a basket equity swap.

And for each stock the Weight is given in the Appendix
P_initial is the local currency price of each stock on the Initial Valuation Date
P_average is the arithmetic average of the local currency price of each stock on each of the Averaging Dates

Floating Equity Amount Payment Date	Termination Date

Appendix
Each of the following stocks are equally weighted (5%):
Astra (Sweden), Glaxo Wellcome (UK), Smithkline Beecham (UK), Zeneca Group (UK), Novartis (Switzerland), Roche Holding Genus (Switzerland), Sanofi (France), Synthelabo (France), Bayer (Germany), Abbott Labs (US), Bristol Myers Squibb (US), American Home Products (US), Amgen (US), Eli Lilly (US), Medtronic (US), Merck (US), Pfizer (US), Schering-Plough (US), Sankyo (Japan), Takeda Chemical (Japan).

This indicative termsheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 11.4 (*continued*)

with a feature such as early exercise being dependent on another asset. This would still be a multifactor problem; in this particular case there are two sources of randomness.

Continuing with this example, suppose that the payoff is $P(S_1)$ at expiry. Also suppose that the option can be exercised early receiving this payoff, but only when $S_2 > E_2$. To price this contract we must find $V(S_1, S_2, t)$ where

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_i\sigma_j\rho_{ij}S_iS_j\frac{\partial^2 V}{\partial S_i\,\partial S_j} + \sum_{i=1}^{2} rS_i\frac{\partial V}{\partial S_i} - rV = 0,$$

with

$$V(S_1, S_2, T) = P(S_1)$$

subject to

$$V(S_1, S_2, t) \geq P(S_1) \quad \text{for} \quad S_2 > E_2,$$

and continuity of V and its first derivatives.

If we can price and hedge an option on a single asset with a set of characteristics, then theoretically we can price and hedge a multi-asset version as well. The practice of pricing and hedging may be much harder as I mention below.

11.10 REALITIES OF PRICING BASKET OPTIONS

The factors that determine the ease or difficulty of pricing and hedging multi-asset options are

- existence of a closed-form solution
- number of underlying assets, the dimensionality
- path dependency
- early exercise

We have seen all of these except path dependency, which is one of the subjects of Part Two.
The solution technique that we use will generally be one of

- finite-difference solution of a partial differential equation
- numerical integration
- Monte Carlo simulation

These methods are the subjects of Part Six.

11.10.1 Easy Problems

If we have a closed-form solution then our work is done; we can easily find values and hedge ratios. This is provided that the solution is in terms of sufficiently simple functions for which there are spreadsheet functions or other libraries. If the contract is European with no path-dependency then the solution may be of the form (11.2). If this is the case, then we often have to do the integration numerically. This is not difficult. Several methods are described in Chapter 81, including Monte Carlo integration and the use of low-discrepancy sequences.

11.10.2 Medium Problems

If we have low dimensionality, less than three or four, say, the finite-difference methods are the obvious choice. They cope well with early exercise and many path-dependent features can be incorporated, though usually at the cost of an extra dimension.

For higher dimensions, Monte Carlo simulations are good. They cope with all path-dependent features. Unfortunately, they are not very efficient for American-style early exercise.

11.10.3 Hard Problems

The hardest problems to solve are those with both high dimensionality, for which we would like to use Monte Carlo simulation, and with early exercise, for which we would like to use finite-difference methods. There is currently no numerical method that copes well with such a problem.

11.11 REALITIES OF HEDGING BASKET OPTIONS

Even if we can find option values and the greeks, they are often very sensitive to the level of the correlation. But as I have said, the correlation is a very difficult quantity to measure. So the hedge ratios are very likely to be inaccurate. If we are delta hedging then we need accurate estimates of the deltas. This makes basket options very difficult to delta hedge successfully.

When we have a contract that is difficult to delta hedge we can try to reduce sensitivity to parameters, and the model, by hedging with other derivatives. This was the basis of vega hedging, mentioned in Chapter 7. We could try to use the same idea to reduce sensitivity to the correlation. Unfortunately, that is also difficult because there just aren't enough contracts traded that depend on the right correlations.

11.12 CORRELATION VERSUS COINTEGRATION

The correlations between financial quantities are notoriously unstable. One could easily argue that a theory should not be built up using parameters that are so unpredictable. I would tend to agree with this point of view. One could propose a stochastic correlation model, but that approach has its own problems.

An alternative statistical measure to correlation is **cointegration**. Very loosely speaking, two time series are cointegrated if a linear combination has constant mean and standard deviation. In other words, the two series never stray too far from one another. This is probably a more robust measure of the linkage between two financial quantities but as yet there is little derivatives theory based on the concept.

11.13 SUMMARY

The new ideas in this chapter were the multifactor, correlated random walks for assets, and Itô's lemma in higher dimensions. These are both simple concepts, and we will use them often, especially in interest-rate-related topics.

FURTHER READING

- See Hamilton (1994) for further details of the measurement of correlation and cointegration.
- The first solution of the exchange option problem was by Margrabe (1978).
- For analytical results, formulae or numerical algorithms for the pricing of some other multifactor options see Stulz (1982), Johnson (1987), Boyle, Evnine & Gibbs (1989), Boyle & Tse (1990), Rubinstein (1991) and Rich & Chance (1993).
- See Emanuel Derman's autobiography for discussion of quantos (Derman, 2004).
- For details of cointegration, what it means and how it works see the papers by Alexander & Johnson (1992, 1994).
- Krekel *et al.* (2004) compare different pricing methods for basket options.

CHAPTER 12
how to delta hedge

In this Chapter. . .

- how to make money if your volatility forecast is more accurate than the market's
- different ways of delta hedging
- how much profit should you expect to make

12.1 INTRODUCTION

In this chapter we are going to assert boldly that there is money to be made from options, because options may be mispriced by the market. In simple terms, *there are arbitrage opportunities*. Shock, horror! I know that the whole of Chicago University has just thrown down this book in disgust. However, I hope the rest of you will enjoy the contents of this chapter for it puts into concrete mathematics some ideas that are most important, and frighteningly under-explained in the literature. This is the subject of how to delta hedge when your estimate of future actual volatility differs from that of the market as measured by the implied volatility.

As I hinted above, to some people, saying that actual volatility and implied volatility can be inconsistent with each other is a heresy, for it implies arbitrage and hence free money. Well, it's only as free as the model is accurate, that is, not at all. But even so, if there is such a difference (and vol arb hedge funds certainly think there is) then we can only get at that money by hedging, and if we have two estimates of volatility which one goes into the famous Black–Scholes delta formula?

We'll see how you can hedge using a delta based on *either* actual volatility or on implied volatility. Neither is wrong, they just have different risk/return profiles.

If you do doubt that implied volatility and actual volatility can be in disagreement then take a look at Figure 12.1. This is simply a plot of the distributions of the logarithms of the VIX and of the rolling realized SPX volatility. The VIX is an implied volatility measure based on the SPX index and so you would expect it and the realized SPX volatility to bear some resemblance. Not quite, as can be seen in the figure. The implied volatility VIX seems to be higher than the realized volatility. Both of these volatilities are approximately lognormally distributed (since their logarithms appear to be Normal), especially the realized volatility. The VIX distribution is somewhat truncated on the left. The mean of the realized volatility, about 15%, is significantly less than the mean of the VIX, about 20%, but its standard deviation is greater.

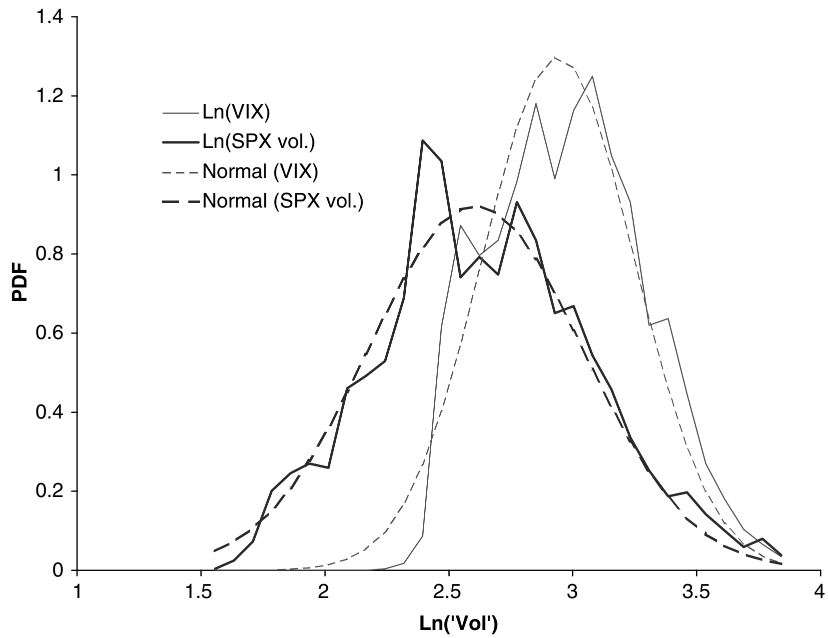


Figure 12.1 Distributions of the logarithms of the VIX and the rolling realized SPX volatility, and the Normal distributions for comparison.

12.2 WHAT IF IMPLIED AND ACTUAL VOLATILITIES ARE DIFFERENT?

Actual volatility is the amount of 'noise' in the stock price; it is the coefficient of the Wiener process in the stock returns model; it is the amount of randomness that 'actually' transpires. Implied volatility is how the market is pricing the option currently. Since the market does not have perfect knowledge about the future these two numbers can and will be different.

Actual volatility being different from implied volatility is the heart of this chapter. Let's look at the simple case of exploiting such a difference by buying or selling options, but not delta hedging them.

Imagine that we have a forecast for volatility over the remaining life of an option, this volatility is forecast to be constant, and, crucially, our forecast turns out to be correct.

If you believe that actual volatility is higher than implied you might want to buy a straddle because there is then a good chance that the stock will move so far before expiry that you will get a payoff of more than the premium you paid, even after allowing for the time value of money. This is a very simple strategy, requiring no maintenance. There is one big problem with this however. It is risky. Sometimes you'll win, sometimes you'll lose. *Unless you can do this strategy many, many times you could end up losing a great deal.* Even if you are right about the actual volatility being large the stock might end up at the money, and you lose out. The relationship between actual volatility and the range over which an asset moves is a probabilistic one, there are no guarantees that a high volatility results in large moves.

If you buy an at-the-money straddle close to expiry, the profit you expect to make from this strategy is approximately

$$\sqrt{\frac{2(T-t)}{\pi}}\,(\sigma-\tilde{\sigma})\,S.$$

The expression uses the close to expiry and ATM approximation we saw in Chapter 7. The notation is obvious: σ is the actual volatility, assumed constant, and $\tilde{\sigma}$ is the implied volatility. Note that this is an *expectation*. It is also a *real* expectation, however the real drift doesn't appear in this expression because it is an approximation valid only when close to expiration.

The standard deviation of the profit (the risk) is approximately

$$\sqrt{1-\frac{2}{\pi}}\quad\sigma S\sqrt{T-t}.$$

Observe how this depends on the actual volatility and not on the implied volatility. This standard deviation is of the same order of magnitude as the expected profit. That is a lot of risk. We can improve the risk-reward profile by delta hedging as we shall see next. The main purpose of writing down the above expressions is to show how they are linear in the two volatilities.

12.3 IMPLIED VERSUS ACTUAL; DELTA HEDGING BUT USING WHICH VOLATILITY?

Let's buy an underpriced option, or portfolio as above, but now, to improve risk and reward, we will delta hedge to expiry. This is a less risky strategy.

But which delta do you choose? Delta based on actual or implied volatility? This is one of those questions that people always ask, and one that no one seems to know the full answer to.

Scenario: Implied volatility for an option is 20%, but we believe that actual volatility is 30%. Question: How can we make money if our forecast is correct? Answer: Buy the option and delta hedge. But which delta do we use? We know that

$$\Delta = N(d_1)$$

where

$$N(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{s^2}{2}}\,ds$$

and

$$d_1=\frac{\ln(S/E)+\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

We can all agree on S, E, $T-t$ and r (almost), but not on σ, so should we use $\sigma=0.2$ or 0.3, implied volatility or actual? In this example,

$$\sigma = \text{ actual volatility, 30\%}$$

and

$$\tilde{\sigma} = \text{ implied volatility, 20\%.}$$

which of these goes into the d_1?

12.4 CASE 1: HEDGE WITH ACTUAL VOLATILITY, σ

By hedging with actual volatility we are replicating a short position in a *correctly priced* option. The payoffs for our long option and our short replicated option will exactly cancel. The profit we make will be exactly the difference in the Black–Scholes prices of an option with 30% volatility and one with 20% volatility. (Assuming that the Black–Scholes assumptions hold.) If $V(S,t;\sigma)$ is the Black–Scholes formula then the guaranteed profit is

$$V(S,t;\sigma) - V(S,t;\tilde{\sigma}).$$

But how is this guaranteed profit realized? Let's do the math on a mark-to-market basis.

In the following, superscript 'a' means actual and 'i' means implied; these can be applied to deltas and option values. For example, Δ^a is the delta using the actual volatility in the formula. V^i is the theoretical option value using the implied volatility in the formula. Note also that V, Δ, Γ and Θ are all simple, known, Black–Scholes formulae.

The model is as usual

$$dS = \mu S\,dt + \sigma S\,dX.$$

Now, set up a portfolio by buying the option for V^i and hedge with Δ^a of the stock. The values of each of the components of our portfolio are shown in Table 12.1.

Leave this hedged portfolio overnight, and come back to it the next day. The new values are shown in Table 12.2. (I have included a continuous dividend yield in this.)

Therefore we have made, mark to market,

$$dV^i - \Delta^a\,dS - r(V^i - \Delta^a\;S)\,dt - \Delta^a DS\,dt.$$

Since the option would be correctly valued at V^a, we have

$$dV^a - \Delta^a\,dS - r(V^a - \Delta^a\;S)\,dt - \Delta^a DS\,dt = 0.$$

Table 12.1 Portfolio composition and values, today.

Component	Value
Option	V^i
Stock	$-\Delta^a\;S$
Cash	$-V^i + \Delta^a\;S$

Table 12.2 Portfolio composition and values, tomorrow.

Component	Value
Option	$V^i + dV^i$
Stock	$-\Delta^a\;S - \Delta^a\,dS$
Cash	$(-V^i + \Delta^a\;S)(1 + r\,dt) - \Delta^a DS\,dt$

So we can write the mark-to-market profit over one time step as

$$dV^i - dV^a + r(V^a - \Delta^a\ S)\,dt - r(V^i - \Delta^a\ S)\,dt$$
$$= dV^i - dV^a - r(V^i - V^a)\,dt = e^{rt}\,d(e^{-rt}(V^i - V^a)).$$

That is the profit from time t to $t + dt$. The present value of this profit at time t_0 is

$$e^{-r(t-t_0)}e^{rt}d(e^{-rt}(V^i - V^a)) = e^{rt_0}d(e^{-rt}(V^i - V^a)).$$

So the total profit from t_0 to expiration is

$$e^{rt_0}\int_{t_0}^{T} d(e^{-rt}(V^i - V^a)) = V^a - V^i.$$

This confirms what I said earlier about the guaranteed total profit by expiration.

We can also write that one time step mark-to-market profit (using Itô's lemma) as

$$\Theta^i\,dt + \Delta^i\,dS + \tfrac{1}{2}\sigma^2 S^2\Gamma^i\,dt - \Delta^a\,dS - r(V^i - \Delta^a S)\,dt - \Delta^a DS\,dt$$

$$= \Theta^i\,dt + \mu S(\Delta^i - \Delta^a)\,dt + \tfrac{1}{2}\sigma^2 S^2\Gamma^i\,dt - r(V^i - V^a)\,dt + (\Delta^i - \Delta^a)\sigma S\,dX - \Delta^a DS\,dt$$

$$= (\Delta^i - \Delta^a)\sigma S\,dX + (\mu + D)S(\Delta^i - \Delta^a)\,dt + \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)S^2\Gamma^i\,dt$$

(using Black–Scholes with $\sigma = \tilde{\sigma}$)

$$= \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)S^2\Gamma^i\,dt + (\Delta^i - \Delta^a)\left((\mu - r + D)S\,dt + \sigma S\,dX\right).$$

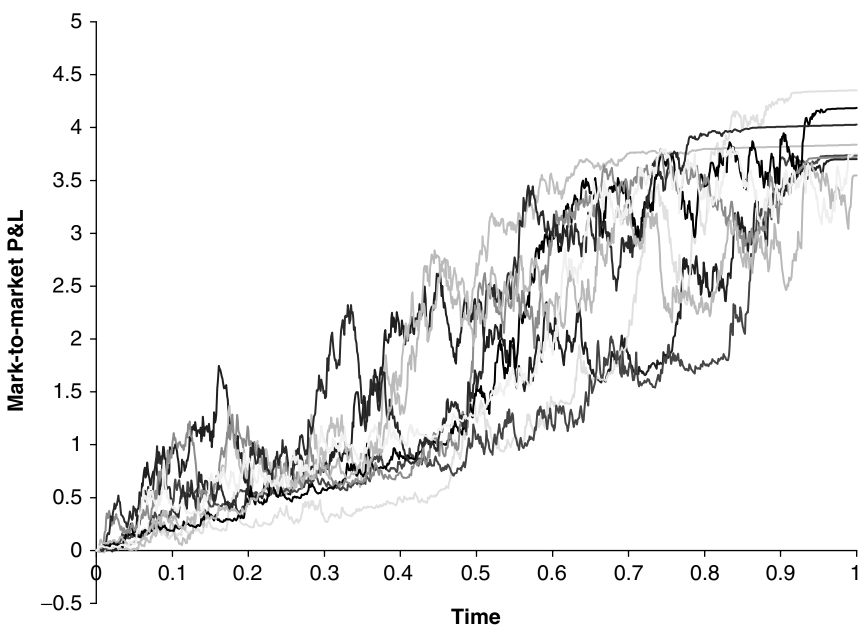


Figure 12.2 P&L for a delta-hedged option on a mark-to-market basis, hedged using actual volatility.

The conclusion is that the final profit is guaranteed (the difference between the theoretical option values with the two volatilities) but how that is achieved is random, because of the dX term in the above. On a mark-to-market basis you could lose before you gain. Moreover, the mark-to-market profit depends on the real drift of the stock, μ. This is illustrated in Figure 12.2, which shows several realizations of the same delta-hedged position. Note that the final P&L is not *exactly* the same in each case because of the effect of hedging discretely, a topic discussed in Chapter 47.

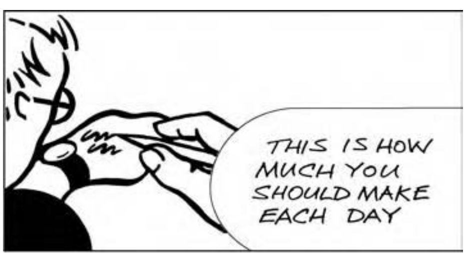
When S changes, so will V. But these changes do not cancel each other out. From a risk management point of view this is not ideal.

There is a simple analogy for this behavior. It is similar to owning a bond. For a bond there is a guaranteed outcome, but we may lose on a mark-to-market basis in the meantime.

12.5 CASE 2: HEDGE WITH IMPLIED VOLATILITY, $\tilde{\sigma}$

Compare and contrast now with the case of hedging using a delta based on implied volatility. By hedging with implied volatility we are balancing the random fluctuations in the mark-to-market option value with the fluctuations in the stock price. The evolution of the portfolio value is then 'deterministic' as we shall see.

Buy the option today, hedge using the implied delta, and put any cash in the bank earning r. The mark-to-market profit from today to tomorrow is



$$\begin{aligned} &dV^i - \Delta^i\,dS - r(V^i - \Delta^i S)\,dt - \Delta^i DS\,dt \\ &\quad = \Theta^i\,dt + \tfrac{1}{2}\sigma^2 S^2\Gamma^i\,dt - r(V^i - \Delta^i S)\,dt - \Delta^i DS\,dt \\ &\quad = \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)S^2\Gamma^i\,dt. \end{aligned} \tag{12.1}$$

This is a far nicer way to make money. Observe how the daily profit is deterministic, there aren't any dX terms. From a risk management perspective this is much better behaved. There is another, rather wonderful, advantage of hedging using implied volatility ... we don't even need to know what actual volatility is. And to make a profit all we need to know is that actual is always going to be greater than implied (if we are buying) or always less (if we are selling). This takes some of the pressure off forecasting volatility accurately in the first place.

Add up the present value of all of these profits to get a total profit of

$$\tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)\int_{t_0}^{T} e^{-r(t-t_0)} S^2\Gamma^i\,dt.$$

This is always positive, but highly path dependent. Being path dependent it will depend on the drift μ. If we start off at the money and the drift is very large (positive or negative) we will find ourselves quickly moving into territory where gamma and hence (12.1) is small, so that there will be not much profit to be made. The best that could happen would be for the stock to end up close to the strike at expiration, this would maximize the total profit. This path dependency is shown in Figure 12.3. The figure shows several realizations of the same delta-hedged position. Note that the lines are not perfectly smooth, again because of the effect of hedging discretely.

The simple analogy is now just putting money in the bank. The P&L is always increasing in value but the end result is random.

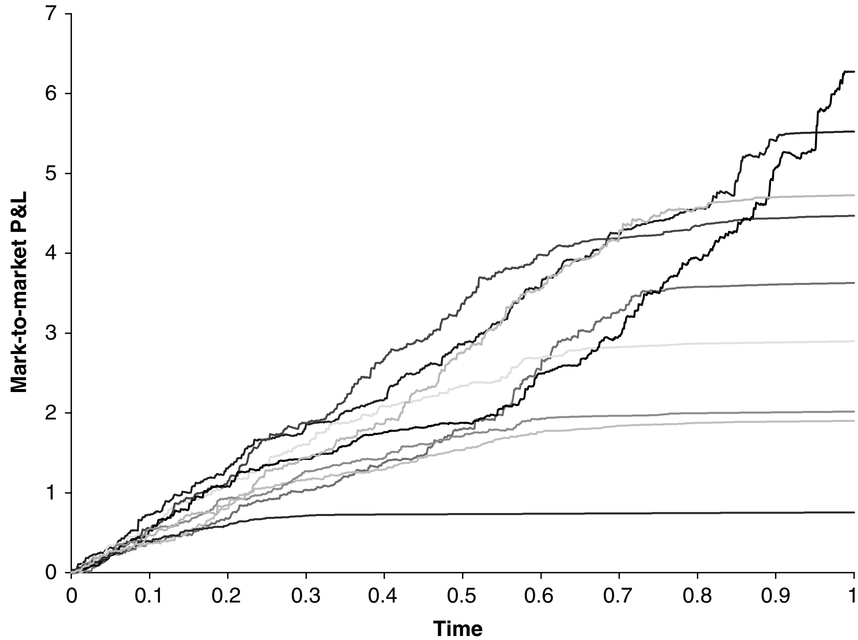


Figure 12.3 P&L for a delta-hedged option on a mark-to-market basis, hedged using implied volatility.

Peter Carr (2005) and Henrard (2001) show that if you hedge using a delta based on a volatility σ_h then the PV of the total profit is given by

$$V(S, t; \sigma_h) - V(S, t; \tilde{\sigma}) + \tfrac{1}{2}\left(\sigma^2 - \sigma_h^2\right) \int_{t_0}^{T} e^{-r(t-t_0)} S^2 \Gamma^h \, dt, \qquad (12.2)$$

where the superscript on the gamma means that it uses the Black–Scholes formula with a volatility of σ_h.

12.5.1 The Expected Profit after Hedging using Implied Volatility

When you hedge using delta based on implied volatility the profit each 'day' is deterministic but the present value of total profit by expiration is path dependent, and given by

$$\tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2) \int_{t_0}^{T} e^{-r(s-t_0)} S^2 \Gamma^i \, ds.$$

Introduce

$$I = \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2) \int_{t_0}^{t} e^{-r(s-t_0)} S^2 \Gamma^i \, ds.$$

Since therefore

$$dI = \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2) e^{-r(t-t_0)} S^2 \Gamma^i \, dt$$

we can write down the following partial differential equation for the *real* expected value, $P(S, I, t)$, of I:

$$\frac{\partial P}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + \mu S \frac{\partial P}{\partial S} + \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2) e^{-r(t-t_0)} S^2 \Gamma^i \, \frac{\partial P}{\partial I} = 0,$$

with

$$P(S, I, T) = I.$$

Look for a solution of this equation of the form

$$P(S, I, t) = I + H(S, t)$$

so that

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 H}{\partial S^2} + \mu S \frac{\partial H}{\partial S} + \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)\, e^{-r(t-t_0)} S^2 \Gamma^i = 0.$$

The source term can be simplified to

$$\frac{E(\sigma^2 - \tilde{\sigma}^2) e^{-r(T-t_0)} e^{-d_2^2/2}}{2\tilde{\sigma}\sqrt{2\pi(T-t)}}.$$

Change variables to

$$x = \log(S/E) + \left(\mu - \tfrac{1}{2}\sigma^2\right)\tau \quad \text{and} \quad \tau = T - t$$

and write

$$H = w(x, \tau).$$

The resulting partial differential equation is then a bit nicer. Details can be found in the appendix to this chapter

After some manipulations we end up with the expected profit initially ($t = t_0$, $I = 0$) being the single integral

$$\frac{E e^{-r(T-t_0)}(\sigma^2 - \tilde{\sigma}^2)}{2\sqrt{2\pi}} \int_{t_0}^{T} \frac{1}{\sqrt{\sigma^2(s-t_0) + \tilde{\sigma}^2(T-s)}}$$
$$\times \exp\left(-\frac{\left(\log(S/E) + \left(\mu - \tfrac{1}{2}\sigma^2\right)(s-t_0) + \left(r - D - \tfrac{1}{2}\tilde{\sigma}^2\right)(T-s)\right)^2}{2(\sigma^2(s-t_0) + \tilde{\sigma}^2(T-s))}\right) ds.$$

Results are shown in the following figures.

In Figure 12.4 is shown the expected profit versus the growth rate μ. Parameters are $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $E = 110$, $T = 1$, $\tilde{\sigma} = 0.2$. Observe that the expected profit has a maximum. This will be at the growth rate that ensures, roughly speaking, that the stock ends up close to at the money at expiration, where gamma is largest. In the figure is also shown the profit to be made when hedging with actual volatility. For most realistic parameters regimes the maximum expected profit hedging with implied is similar to the guaranteed profit hedging with actual.

In Figure 12.5 is shown expected profit versus E and μ. You can see how the higher the growth rate the larger the strike price at the maximum. The contour map is shown in Figure 12.6.

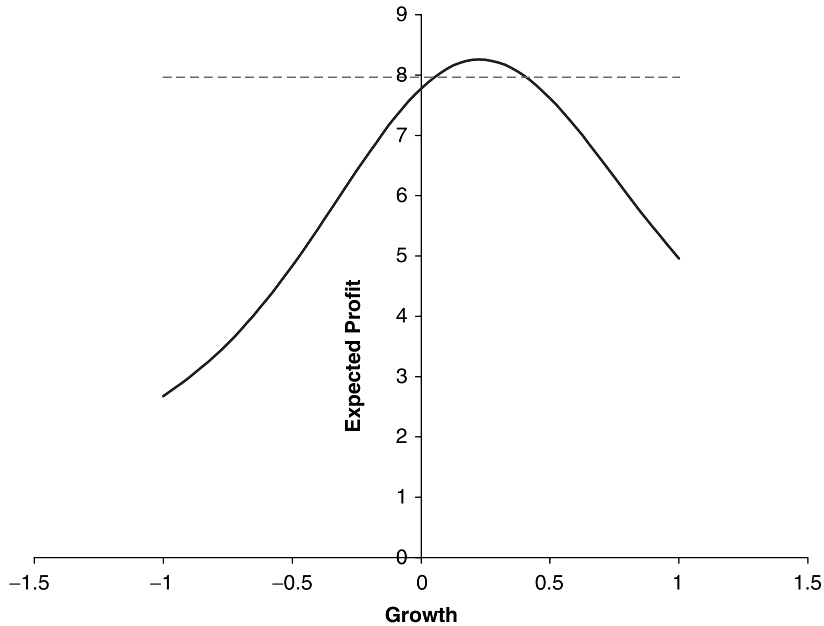


Figure 12.4 Expected profit, hedging using implied volatility, versus growth rate μ; $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $E = 110$, $T = 1$, $\tilde{\sigma} = 0.2$. The dashed line is the profit to be made when hedging with actual volatility.

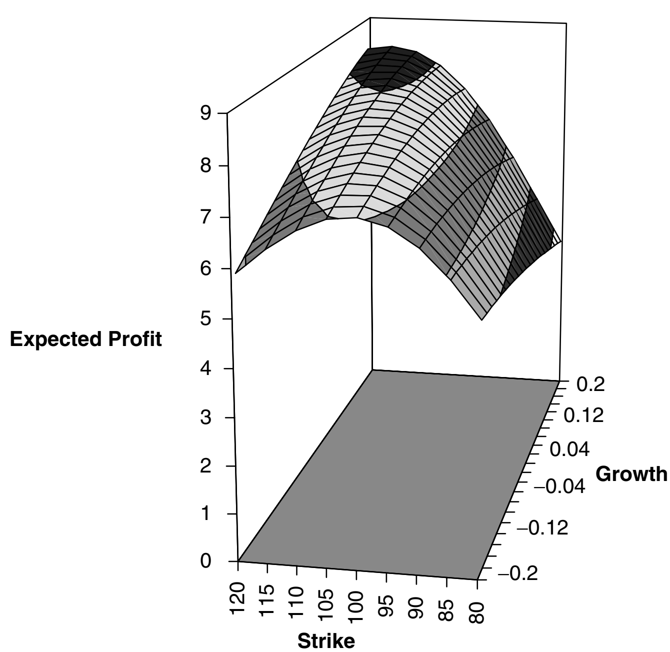


Figure 12.5 Expected profit, hedging using implied volatility, versus growth rate μ and strike E; $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $T = 1$, $\tilde{\sigma} = 0.2$.

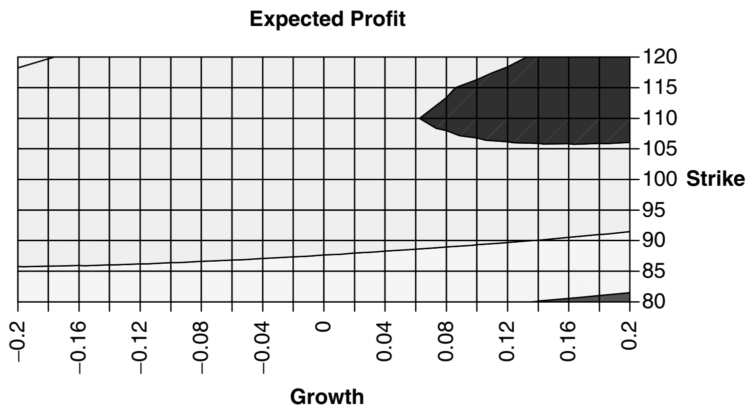


Figure 12.6 Contour map of expected profit, hedging using implied volatility, versus growth rate μ and strike E; $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $T = 1$, $\tilde{\sigma} = 0.2$.

The effect of skew is shown in Figure 12.7. Here I have used a linear negative skew, from 22.5% at a strike of 75, falling to 17.5% at the 125 strike. The at-the-money implied volatility is 20% which in this case is the actual volatility. This picture changes when you divide the expected profit by the price of the option (puts for lower strikes, call for higher), see Figure 12.8. There is no maximum, profitability increases with distance away from the money. Of course, this doesn't take into account the risk, the standard deviation associated with such trades.

12.5.2 The Variance of Profit after Hedging using Implied Volatility

Once we have calculated the expected profit from hedging using implied volatility we can calculate the variance in the final profit. Using the above notation, the variance will be the expected value of I^2 less the square of the average of I. So we will need to calculate $v(S, I, t)$ where

$$\frac{\partial v}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 v}{\partial S^2} + \mu S \frac{\partial v}{\partial S} + \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)\, e^{-r(t-t_0)} S^2 \Gamma^i \, \frac{\partial v}{\partial I} = 0,$$

with

$$v(S, I, T) = I^2.$$

The details of finding this function v are rather messy, but a solution can be found of the form

$$v(S, I, t) = I^2 + 2I\ H(S, t) + G(S, t).$$

The initial variance is $G(S_0, t_0) - F(S_0, t_0)^2$, where

$$G(S_0, t_0) = \frac{E^2(\sigma^2 - \tilde{\sigma}^2)^2 e^{-2r(T-t_0)}}{4\pi\sigma\tilde{\sigma}} \int_{t_0}^{T} \int_{s}^{T}$$

$$\times \frac{e^{p(u,s;S_0,t_0)}}{\sqrt{s-t_0}\sqrt{T-s}\sqrt{\sigma^2(u-s) + \tilde{\sigma}^2(T-u)}\sqrt{\frac{1}{\sigma^2(s-t_0)} + \frac{1}{\tilde{\sigma}^2(T-s)} + \frac{1}{\sigma^2(u-s)+\tilde{\sigma}^2(T-u)}}}\, du\, ds \tag{12.3}$$

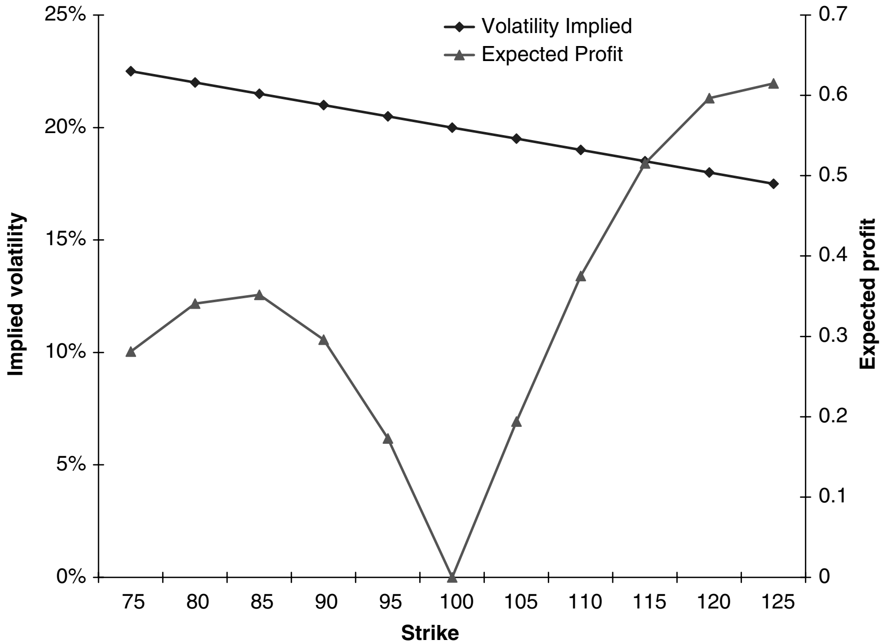


Figure 12.7 Effect of skew, expected profit, hedging using implied volatility, versus strike E; $S = 100$, $\mu = 0$, $\sigma = 0.2$, $r = 0.05$, $D = 0$, $T = 1$.

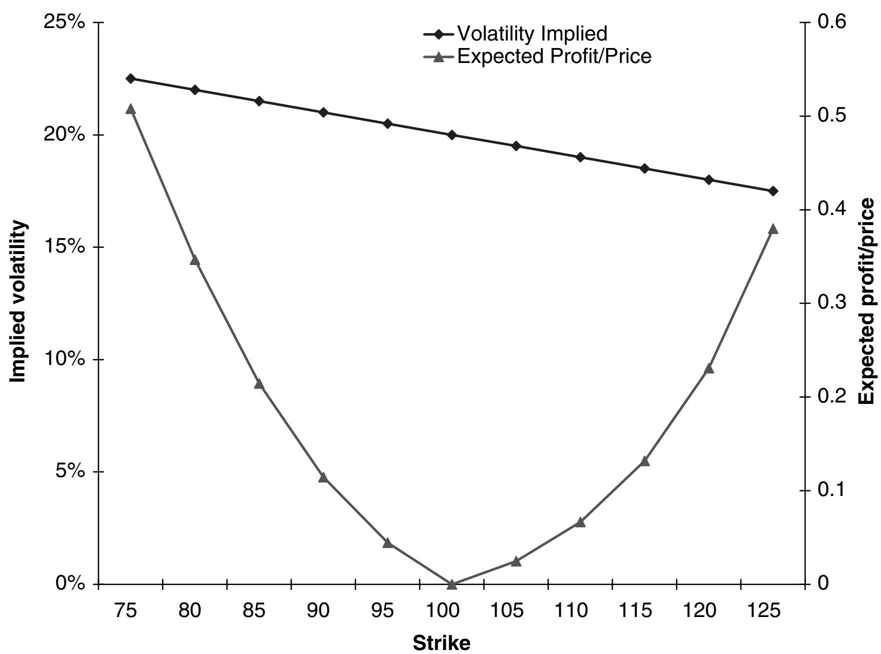


Figure 12.8 Effect of skew, ratio of expected profit to price, hedging using implied volatility, versus strike E; $S = 100$, $\mu = 0$, $\sigma = 0.2$, $r = 0.05$, $D = 0$, $T = 1$.

where

$$p(u,s;S_0,t_0) = -\tfrac{1}{2}\frac{(x+\alpha(T-s))^2}{\tilde{\sigma}^2(T-s)} - \tfrac{1}{2}\frac{(x+\alpha(T-u))^2}{\sigma^2(u-s)+\tilde{\sigma}^2(T-u)}$$
$$+\tfrac{1}{2}\frac{\left(\dfrac{x+\alpha(T-s)}{\tilde{\sigma}^2(T-s)}+\dfrac{x+\alpha(T-u)}{\sigma^2(u-s)+\tilde{\sigma}^2(T-u)}\right)^2}{\dfrac{1}{\sigma^2(s-t_0)}+\dfrac{1}{\tilde{\sigma}^2(T-s)}+\dfrac{1}{\sigma^2(u-s)+\tilde{\sigma}^2(T-u)}}$$

and

$$x = \ln(S_0/E) + \left(\mu - \tfrac{1}{2}\sigma^2\right)(T-t_0), \quad \text{and} \quad \alpha = \mu - \tfrac{1}{2}\sigma^2 - r + D + \tfrac{1}{2}\tilde{\sigma}^2.$$

The derivation of this can be found in the appendix to this chapter.

In Figure 12.9 is shown the standard deviation of profit versus growth rate, $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $E = 110$, $T = 1$, $\tilde{\sigma} = 0.2$. Figure 12.10 shows the standard deviation of profit versus strike, $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $\mu = 0.1$, $T = 1$, $\tilde{\sigma} = 0.2$.

Note that in these plots the expectations and standard deviations have not been scaled with the cost of the options.

In Figure 12.11 is shown expected profit divided by cost versus standard deviation divided by cost, as both strike and expiration vary. In these plots $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $\mu = 0.1$, $\tilde{\sigma} = 0.2$. To some extent, although we emphasize only *some*, these diagrams can be interpreted in a classical mean-variance manner, see Chapter 18. The main criticism is, of course, that we are not working with Normal distributions, and, furthermore, there is no downside, no possibility of any losses.

Figure 12.12 completes the earlier picture for the skew, since it now contains the standard deviation.

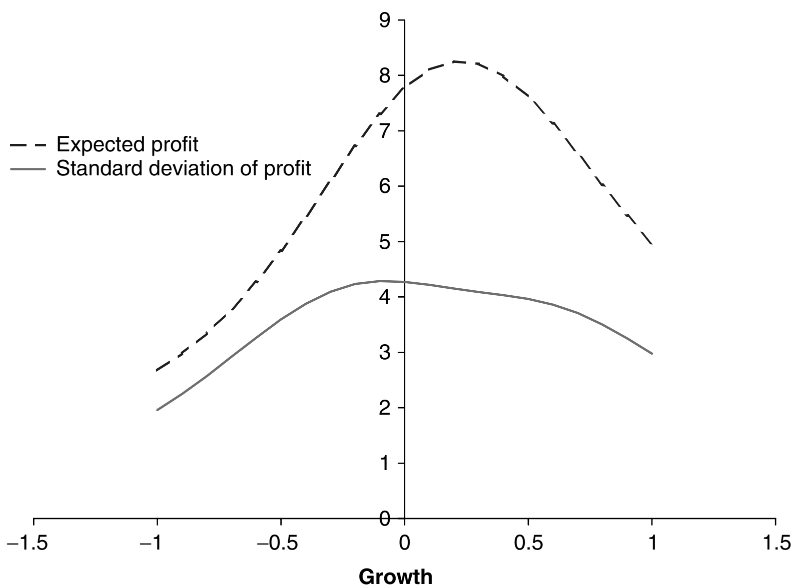


Figure 12.9 Standard deviation of profit, hedging using implied volatility, versus growth rate μ; $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $E = 110$, $T = 1$, $\tilde{\sigma} = 0.2$. (The expected profit is also shown.)

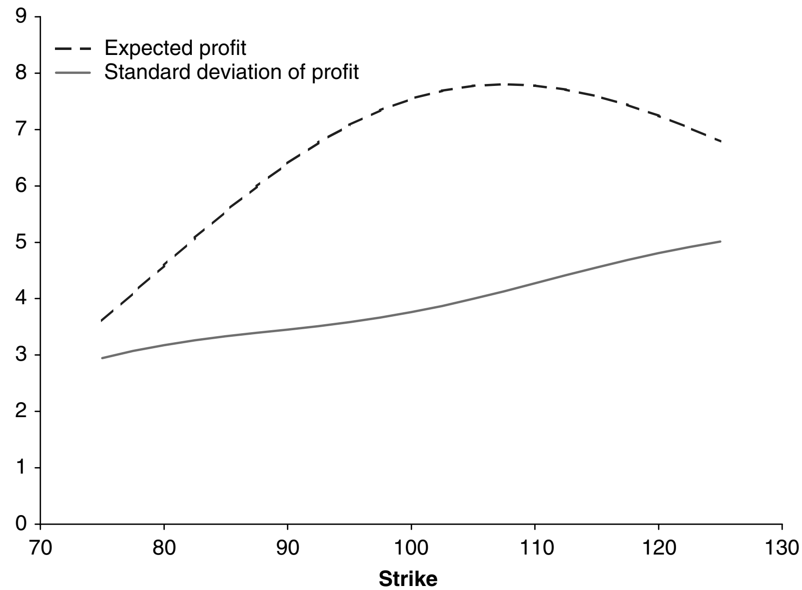


Figure 12.10 Standard deviation of profit, hedging using implied volatility, versus strike E; $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $\mu = 0$, $T = 1$, $\tilde{\sigma} = 0.2$. (The expected profit is also shown.)

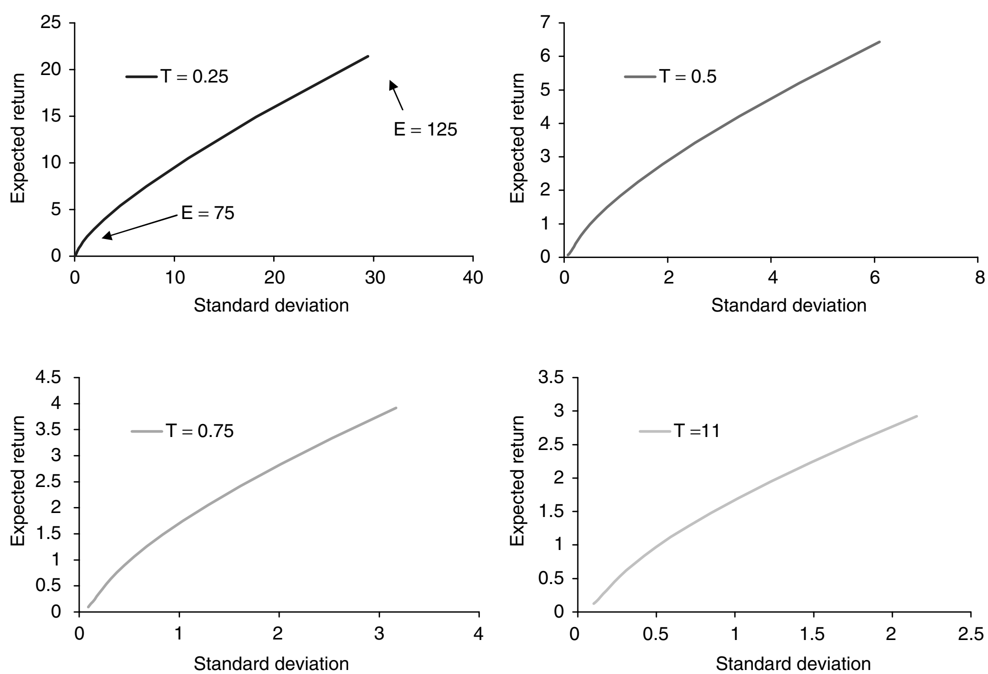


Figure 12.11 Scaled expected profit versus scaled standard deviation; $S = 100$, $\sigma = 0.4$, $r = 0.05$, $D = 0$, $\mu = 0.1$, $\tilde{\sigma} = 0.2$. Four different expirations, varying strike.

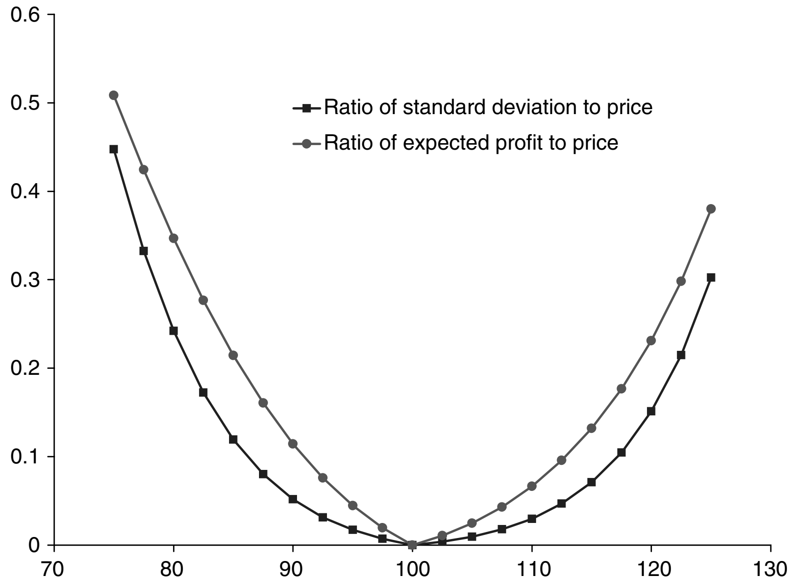


Figure 12.12 Effect of skew, ratio of expected profit to price, and ratio of standard deviation to price, versus strike E; $S = 100$, $\mu = 0$, $\sigma = 0.2$, $r = 0.05$, $D = 0$, $T = 1$.

12.5.3 Hedging with Different Volatilities

We will briefly examine hedging using volatilities other than actual or implied, using the general expression for profit given by (12.2).

The expressions for the expected profit and standard deviations now must allow for the $V(S, t; \sigma_h) - V(S, t; \tilde{\sigma})$, since the integral of gamma term can be treated as before if one replaces $\tilde{\sigma}$ with σ_h in this term. Results are presented in the next two figures.

In Figure 12.13 is shown the expected profit and standard deviation of profit when hedging with various volatilities. The thin, dotted lines, continuing on from the bold lines, represent hedging with volatilities outside the implied-actual range. The chart also shows standard deviation of profit, and minimum and maximum. Parameters are $E = 90$, $S = 100$, $\mu = -0.1$, $\sigma = 0.4$, $r = 0.1$, $D = 0$, $T = 1$, and $\tilde{\sigma} = 0.2$. Note that it is possible to lose money if you hedge at below implied, but hedging with a higher volatility you will not be able to lose until hedging with a volatility of approximately 70%. In this example, the expected profit decreases with increasing hedging volatility.

Figure 12.14 shows the same quantities but now for an option with a strike price of 110. The upper hedging volatility, beyond which it is possible to make a loss, is now slightly higher. The expected profit now increases with increasing hedging volatility.

In practice which volatility one uses is often determined by whether one is constrained to mark to market or mark to model. If one is able to mark to model then one is not necessarily concerned with the day-to-day fluctuations in the mark-to-market profit and loss and so it is natural to hedge using actual volatility. This is usually not far from optimal in the sense of possible expected total profit, and it has no standard deviation of final profit. However, it is common to have to report profit and loss based on market values. This constraint may be imposed by a risk management department, by prime brokers, or by investors who may monitor

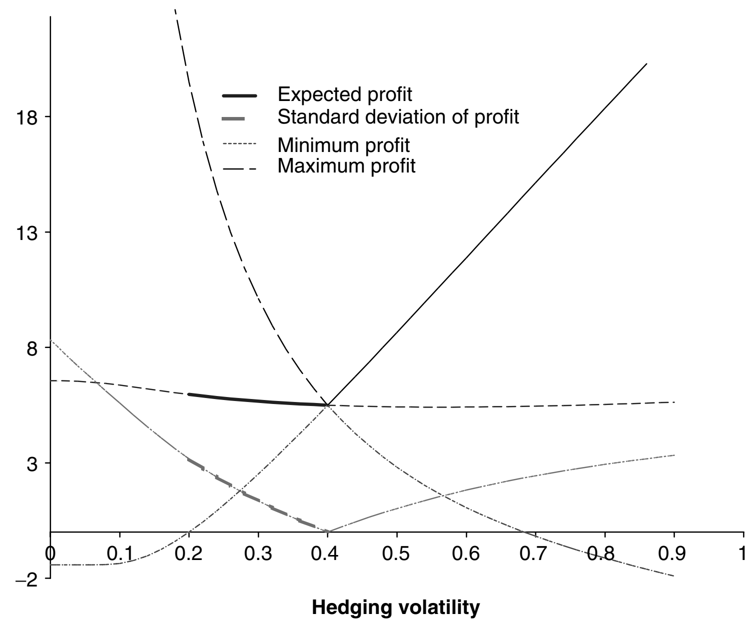


Figure 12.13 Expected profit and standard deviation of profit hedging with various volatilities. $E = 90$, $S = 100$, $\mu = -0.1$, $\sigma = 0.4$, $r = 0.1$, $D = 0$, $T = 1$, $\tilde{\sigma} = 0.2$.

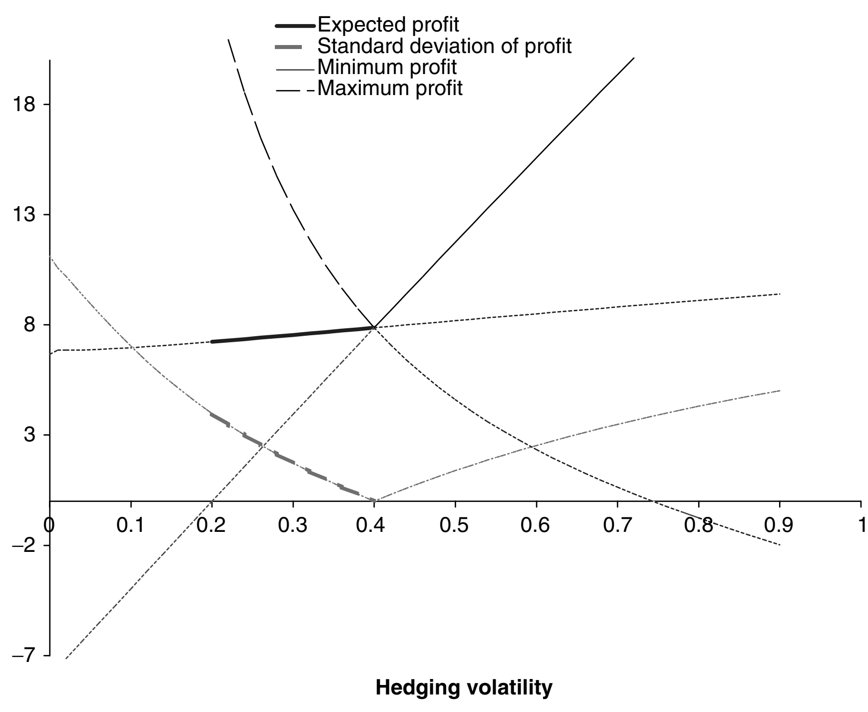


Figure 12.14 Expected profit and standard deviation of profit hedging with various volatilities. $E = 110$, $S = 100$, $\mu = -0.1$, $\sigma = 0.4$, $r = 0.1$, $D = 0$, $T = 1$, $\tilde{\sigma} = 0.2$.

the mark-to-market profit on a regular basis. In this case it is more usual to hedge based on implied volatility to avoid the daily fluctuations in the profit and loss.

For the remainder of this chapter we will only consider the case of hedging using a delta based on implied volatility.

12.6 PORTFOLIOS WHEN HEDGING WITH IMPLIED VOLATILITY

A natural extension to the above analysis is to look at portfolios of options, options with different strikes and expirations. Since only an option's gamma matters when we are hedging using implied volatility, calls and puts are effectively the same since they have the same gamma.

The profit from a portfolio is now

$$\frac{1}{2}\sum_k q_k(\sigma^2-\tilde{\sigma}_k^2)\int_{t_0}^{T_k} e^{-r(s-t_0)}S^2\Gamma_k^i\,ds,$$

where k is the index for an option, and q_k is the quantity of that option. Introduce

$$I=\frac{1}{2}\sum_k q_k(\sigma^2-\tilde{\sigma}_k^2)\int_{t_0}^{t} e^{-r(s-t_0)}S^2\Gamma_k^i\,ds, \tag{12.4}$$

as a new state variable, and the analysis can proceed as before. Note that since there may be more than one expiration date since we have several different options, it must be understood in Equation (12.4) that Γ_k^i is zero for times beyond the expiration of the option.

The governing differential operator for expectation, variance, etc. is then

$$\frac{\partial}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2}{\partial S^2}+\mu S\frac{\partial}{\partial S}+\frac{1}{2}\sum_k(\sigma^2-\tilde{\sigma}_k^2)\,e^{-r(t-t_0)}S^2\Gamma_k^i\,\frac{\partial}{\partial I}=0,$$

with final condition representing expectation, variance, etc.

12.6.1 Expectation

The solution for the present value of the expected profit ($t=t_0$, $S=S_0$, $I=0$) is simply the sum of individual profits for each option,

$$F(S_0,t_0)=\sum_k q_k\frac{E_ke^{-r(T_k-t_0)}(\sigma^2-\tilde{\sigma}_k^2)}{2\sqrt{2\pi}}\int_{t_0}^{T_k}\frac{1}{\sqrt{\sigma^2(s-t_0)+\tilde{\sigma}_k^2(T_k-s)}}$$
$$\times\exp\left(-\frac{\left(\ln(S_0/E_k)+\left(\mu-\frac{1}{2}\sigma^2\right)(s-t_0)+\left(r-D-\frac{1}{2}\tilde{\sigma}_k^2\right)(T_k-s)\right)^2}{2(\sigma^2(s-t_0)+\tilde{\sigma}_k^2(T_k-s))}\right)ds.$$

The derivation can be found in this chapter's appendix.

12.6.2 Variance

The variance is more complicated, obviously, because of the correlation between all of the options in the portfolio. Nevertheless, we can find an expression for the initial variance as $G(S_0, t_0) - F(S_0, t_0)^2$ where

$$G(S_0, t_0) = \sum_j \sum_k q_j q_k G_{jk}(S_0, t_0)$$

where

$$G_{jk}(S_0, t_0) = \frac{E_j E_k (\sigma^2 - \tilde{\sigma}_j^2)(\sigma^2 - \tilde{\sigma}_k^2) e^{-r(T_j - t_0) - r(T_k - t_0)}}{4\pi\sigma\tilde{\sigma}_k} \int_{t_0}^{\min(T_j, T_k)} \int_s^{T_j}$$

$$\times \frac{e^{p(u,s;S_0,t_0)}}{\sqrt{s-t_0}\sqrt{T_k - s}\sqrt{\sigma^2(u-s) + \tilde{\sigma}_j^2(T_j - u)}\sqrt{\frac{1}{\sigma^2(s-t_0)} + \frac{1}{\tilde{\sigma}_k^2(T_k - s)} + \frac{1}{\sigma^2(u-s) + \tilde{\sigma}_j^2(T_j - u)}}} du\, ds \tag{12.5}$$

where

$$p(u, s; S_0, t_0) = -\frac{1}{2}\frac{(\ln(S_0/E_k) + \overline{\mu}(s - t_0) + \overline{r}_k(T_k - s))^2}{\tilde{\sigma}_k^2(T_k - s)}$$

$$- \frac{1}{2}\frac{(\ln(S_0/E_j) + \overline{\mu}(u - t_0) + \overline{r}_j(T_j - u))^2}{\sigma^2(u - s) + \tilde{\sigma}_j^2(T_j - u)}$$

$$+ \frac{1}{2}\frac{\left(\frac{\ln(S_0/E_k) + \overline{\mu}(s-t_0) + \overline{r}_k(T_k - s)}{\tilde{\sigma}_k^2(T_k - s)} + \frac{\ln(S_0/E_j) + \overline{\mu}(u - t_0) + \overline{r}_j(T_j - u)}{\sigma^2(u-s) + \tilde{\sigma}_j^2(T_j - u)}\right)^2}{\frac{1}{\sigma^2(s-t_0)} + \frac{1}{\tilde{\sigma}_k^2(T_k - s)} + \frac{1}{\sigma^2(u-s) + \tilde{\sigma}_j^2(T_j - u)}}$$

and

$$\overline{\mu} = \mu - \tfrac{1}{2}\sigma^2, \quad \overline{r}_j = r - D - \tfrac{1}{2}\tilde{\sigma}_j^2 \quad \text{and} \quad \overline{r}_k = r - D - \tfrac{1}{2}\tilde{\sigma}_k^2.$$

The derivation can be found in this chapter's appendix.

12.6.3 Portfolio Optimization Possibilities

There is clearly plenty of scope for using the above formulae in portfolio optimization problems. Here I give one example.

The stock is currently at 100. The growth rate is zero, actual volatility is 20%, zero dividend yield and the interest rate is 5%. Table 12.3 shows the available options, and associated parameters. Observe the negative skew. The out-of-the-money puts are overvalued and the out-of-the-money calls are undervalued. (The 'Profit Total Expected' row assumes that we buy a single one of that option.)

Using the above formulae we can find the portfolio that maximizes or minimizes target quantities (expected profit, standard deviation, ratio of profit to standard deviation). Let us consider the simple case of maximizing the expected return, while constraining the standard

Table 12.3 Available options.

	A	B	C	D	E
Type	Put	Put	Call	Call	Call
Strike	80	90	100	110	120
Expiration	1	1	1	1	1
Volatility, Implied	0.250	0.225	0.200	0.175	0.150
Option Price, Market	1.511	3.012	10.451	5.054	1.660
Option Value, Theory	0.687	2.310	10.451	6.040	3.247
Profit Total Expected	−0.933	−0.752	0.000	0.936	1.410

Table 12.4 An optimal portfolio.

	A	B	C	D	E
Type	Put	Put	Call	Call	Call
Strike	80	90	100	110	120
Quantity	−2.10	−2.25	0	1.46	1.28

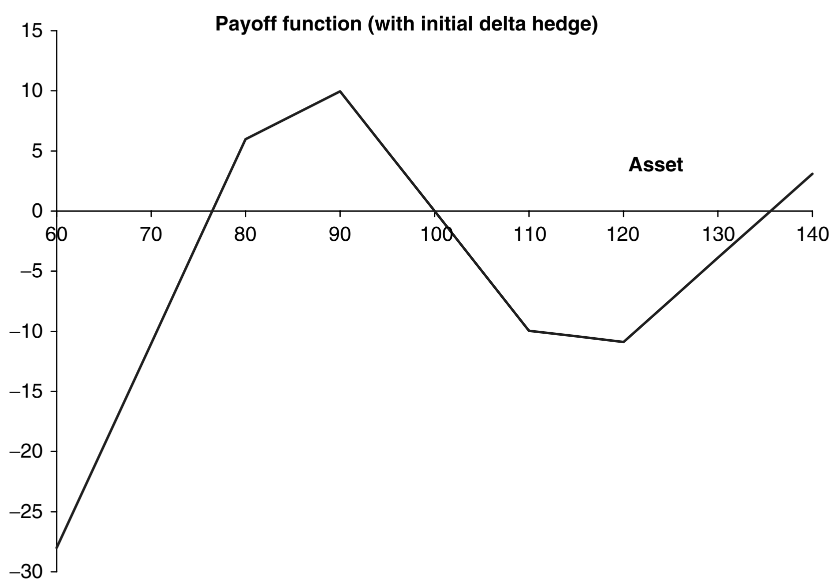


Figure 12.15 Payoff with initial delta hedge for optimal portfolio; $S = 100$, $\mu = 0$, $\sigma = 0.2$, $r = 0.05$, $D = 0$, $T = 1$. See text for additional parameters and information.

deviation to be one. This is a very natural strategy when trying to make a profit from volatility arbitrage while meeting constraints imposed by regulators, brokers, investors etc. The result is given in Table 12.4.

The payoff function (with its initial delta hedge) is shown in Figure 12.15. This optimization has effectively found an ideal risk reversal trade. This portfolio would cost −\$0.46 to set up, i.e. it would bring in premium. The expected profit is \$6.83.

12.7 HEDGING WHEN IMPLIED VOLATILITY IS STOCHASTIC

It is demonstrably true that implied volatility tends to vary, whereas it requires quite advanced statistics to make the same observation about actual volatility. So, let's repeat some of the analyses above assuming that

$$d\tilde{\sigma} = a\,dt + b\,dX_2$$

with the stochastic differential equation for S being the usual, with random component dX_1 and actual volatility constant. There will be a correlation coefficient between dX_1 and dX_2 of ρ. See also Carr & Verma (2005).

12.7.1 Case 1: Hedge with Actual Volatility, σ

The argument that the final profit is guaranteed is not affected by having implied volatility stochastic, except insofar as you may get the opportunity to close the position early if implied volatility reaches the level of actual.

The mark-to-market profit now becomes

$$\begin{aligned}&\tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)S^2\Gamma^i\,dt + (\Delta^i - \Delta^a)\,((\mu - r + D)S\,dt + \sigma S\,dX_1)\\&\quad + \frac{\partial V^i}{\partial \tilde{\sigma}}d\tilde{\sigma} + \tfrac{1}{2}b^2\frac{\partial^2 V^i}{\partial \tilde{\sigma}^2}dt + \rho\sigma bS\frac{\partial^2 V^i}{\partial S\,\partial \tilde{\sigma}}\,dt.\end{aligned}$$

Remember that V^i is the Black–Scholes formula for the option.

12.7.2 Case 2: Hedge with Implied Volatility, $\tilde{\sigma}$?

There is a question mark in the title of this section because it is not clear that using the delta based on implied volatility is meaningful when implied volatility is stochastic. So for the moment we will just hedge using 'Δ,' to be chosen.

The mark-to-market profit is

$$\begin{aligned}&dV^i - \Delta\,dS - r(V^i - \Delta\;S)dt - \Delta DS\,dt\\&\quad = \Theta^i\,dt + \Delta^i\,dS + \tfrac{1}{2}\sigma^2S^2\Gamma^i\,dt + \frac{\partial V^i}{\partial \tilde{\sigma}}d\tilde{\sigma} + \tfrac{1}{2}b^2\frac{\partial^2 V^i}{\partial \tilde{\sigma}^2}\,dt\\&\qquad + \rho\sigma bS\frac{\partial^2 V^i}{\partial S\,\partial \tilde{\sigma}}\,dt - \Delta\,dS - r(V^i - \Delta\;S)\,dt - \Delta DS\,dt.\end{aligned}$$

The variance of this is

$$b^2\left(\frac{\partial V^i}{\partial \tilde{\sigma}}\right)^2 + \sigma^2S^2(\Delta^i - \Delta)^2 + 2\rho\sigma bS(\Delta^i - \Delta)\frac{\partial V^i}{\partial \tilde{\sigma}} + \cdots.$$

(multiplied by dt). This is minimized by the choice

$$\Delta = \Delta^i + \frac{\rho b}{\sigma S}\frac{\partial V^i}{\partial \sigma}.$$

This confirms that to minimize risk on a mark-to-market basis you must adjust the delta by the vega.

With this choice of delta we find that the mark-to-market profit contains a deterministic and a random term. The random term is

$$\left(-\frac{\rho b}{\sigma S} dS + d\tilde{\sigma}\right) \frac{\partial V^i}{\partial \tilde{\sigma}}.$$

This has standard deviation

$$b\sqrt{1-\rho^2}\,\frac{\partial V^i}{\partial \tilde{\sigma}}\,dt^{1/2}.$$

Using the Black–Scholes formulae the deterministic part becomes

$$e^{-D(T-t)}\sqrt{T-t}\,SN'(d_1)$$
$$\times\left(\frac{1}{2}\frac{\sigma^2-\tilde{\sigma}^2}{\tilde{\sigma}(T-t)} - \frac{1}{2}b^2 d_1\frac{\partial d_1}{\partial \tilde{\sigma}} + \rho b\sigma - \frac{\rho b\sigma d_1}{\tilde{\sigma}\sqrt{T-t}} + a + (r-\mu-D)\frac{\rho b}{\sigma}\right) \quad (12.6)$$

(multiplied by dt).

12.8 HOW DOES IMPLIED VOLATILITY BEHAVE?

Now is the natural time to talk a little bit about how implied volatility behaves in practice.

As the stock price goes up and down randomly we often see that the implied volatility of each option will also vary. This may or may not be consistent with certain models, and may or may not be consistent with no arbitrage. But more importantly, what does it mean for making money if we think that the market is wrong? Below are a couple of 'models' for how implied volatility might change as the market moves.

12.8.1 Sticky Strike

In this model implied volatility remains constant for each option (i.e. each strike and expiration). Effectively each option inhabits its own little Black–Scholes world of constant volatility. This behavior seems to be most common in the equity markets. As far as making a profit if the implied volatility is different from actual volatility then the first half of this chapter is clearly very relevant.

12.8.2 Sticky Delta

Since the delta of an option is a function of its moneyness, S/E, the sticky delta behavior could also be called the sticky moneyness rule. This behavior is commonly seen in the FX markets, possibly because there it is usual to quote option prices/volatilities for options with specific deltas rather than specific strike. (There is, of course, a one-to-one correspondence for vanillas.)

In this model we have

$$\tilde{\sigma} = g(S/E, t).$$

Therefore

$$d\tilde{\sigma} = \left(\frac{\partial g}{\partial t} + \mu \frac{S}{E} \frac{\partial g}{\partial \xi} + \tfrac{1}{2}\sigma^2 \frac{S^2}{E^2} \frac{\partial^2 g}{\partial \xi^2} \right) dt + \sigma \frac{S}{E} \frac{\partial g}{\partial \xi}\, dX_1,$$

where $\xi = S/E$. The most important point about this expression is that it is perfectly correlated with dS, $\rho = 1$, so that perfect hedging (in the mark-to-market sense) is possible. We can substitute for a and b into Equation (12.6) to get the 'daily' mark-to-market profit.

A variation on this theme is to have implied volatility at different strikes being proportional to the ATM volatility and a function of the moneyness, such as

$$\tilde{\sigma} = \sigma_{ATM}\; g\left(\frac{\log(S/E)}{\sqrt{T-t}} \right).$$

See Natenberg (1994) for details. Of course, this then requires a model for the behavior of the ATM volatility.

12.8.3 Time-periodic Behavior

Just to make matters even more interesting, there appears to be a day-of-the-week effect in implied volatility. Figures 12.16–12.19 show how the VIX volatility index (a measure of the implied volatility of the ATM SPX adjusted to always have an expiration of 30 days) changes

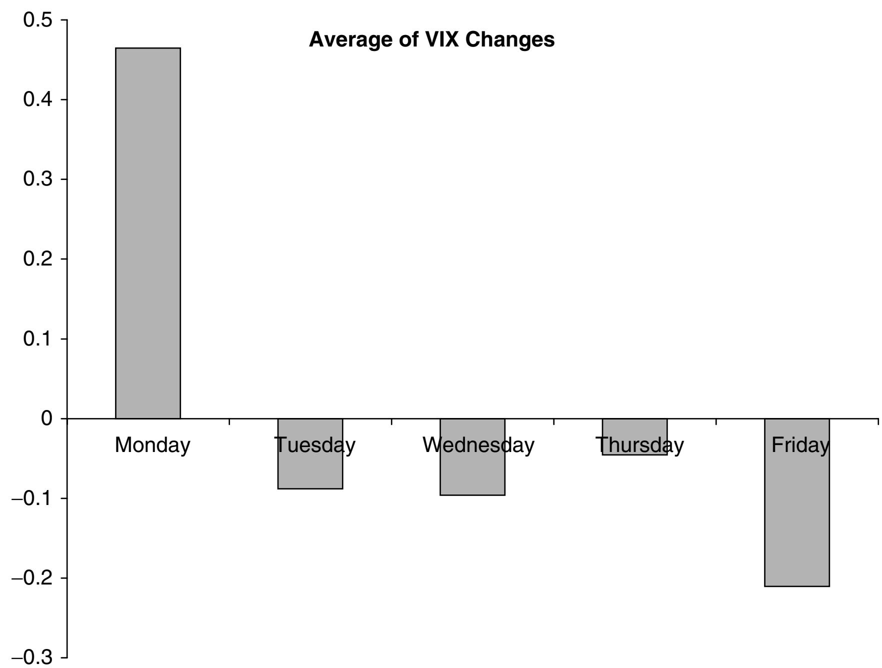


Figure 12.16 Average change in level of VIX versus day of week.

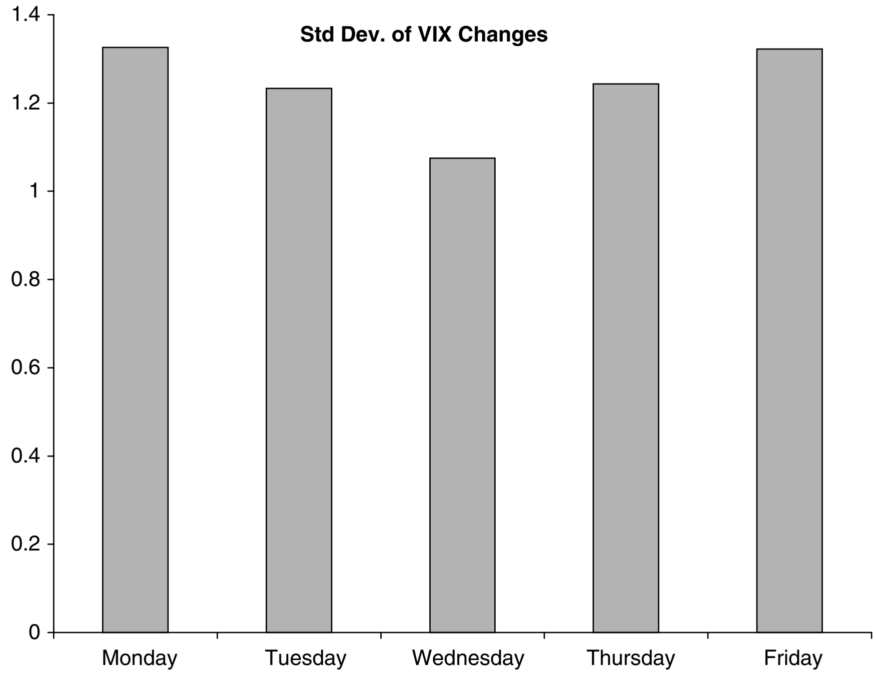


Figure 12.17 Standard deviation of change in level of VIX versus day of week.

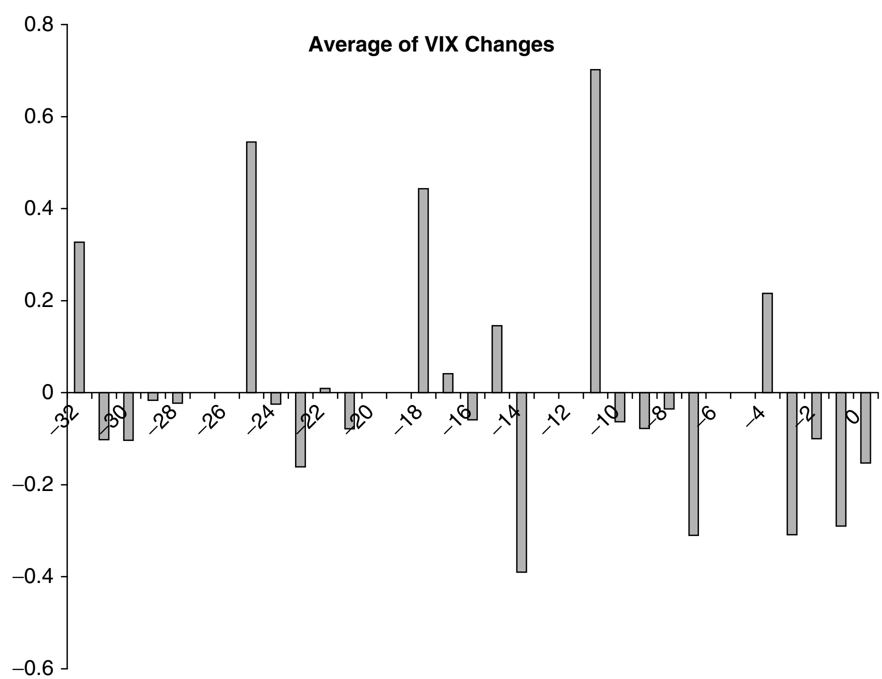


Figure 12.18 Average change in level of VIX versus days before next expiration.

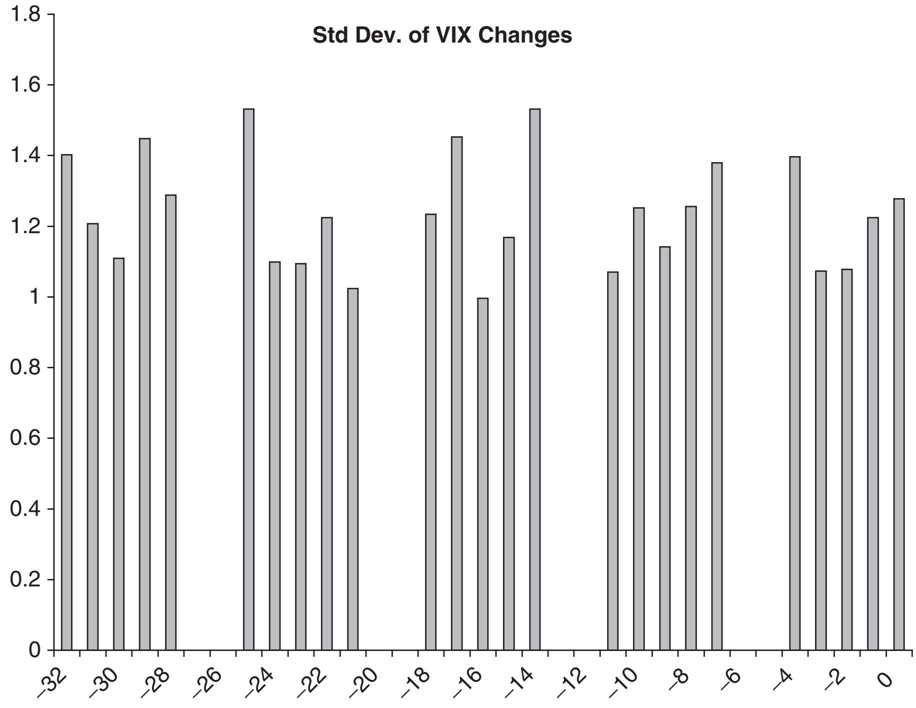


Figure 12.19 Standard deviation of change in level of VIX versus days before next expiration.

with day of the week and number of days to next expiration. Both average changes and standard deviation are shown.[1]

12.9 SUMMARY

In this chapter we have seen some hints of how we can start to move away from the Black–Scholes world, and perhaps even profit from options.

FURTHER READING

- See Derman (1999) for a description of sticky strike and delta, and other volatility regimes.
- Natenberg's book (Natenberg, 1994) is still the classic reference for volatility trading.
- See Carr's (Carr, 2005) excellent FAQs paper for further insight into which volatility to use for hedging. Also Henrard (2001), who examined the role of the real drift rate.
- Ahmad & Wilmott (2005) delve even deeper into the subject of hedging with different volatilities.

[1] Of course, some of this is no doubt related to the role of weekends in the calculation of volatility. There is a lot of 'potential' for volatility over weekends in the sense that there is plenty of news that comes out that will impact on market prices when markets open on the Monday.

APPENDIX: DERIVATION OF RESULTS

Preliminary Results

In the following derivations we often require the following simple results.
First,

$$\int_{-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}. \tag{12.7}$$

Second, the solution of

$$\frac{\partial w}{\partial \tau} = \frac{\partial^2 w}{\partial x^2} + f(x,\tau)$$

that is initially zero and is zero at plus and minus infinity is

$$\frac{1}{2\sqrt{\pi}}\int_{-\infty}^{\infty}\int_0^{\tau}\frac{f(x',\tau')}{\sqrt{\tau-\tau'}}e^{-(x-x')^2/4(\tau-\tau')}\,d\tau'\,dx'. \tag{12.8}$$

Finally, the transformations

$$x = \ln(S/E) + \frac{2}{\sigma^2}\left(\mu - \tfrac{1}{2}\sigma^2\right)\tau \quad \text{and} \quad \tau = \frac{\sigma^2}{2}(T-t),$$

turn the operator

$$\frac{\partial}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2}{\partial S^2} + \mu S\frac{\partial}{\partial S}$$

into

$$\tfrac{1}{2}\sigma^2\left(-\frac{\partial}{\partial \tau} + \frac{\partial^2}{\partial x^2}\right). \tag{12.9}$$

Expectation, Single Option

The equation to be solved for $F(S,t)$ is

$$\frac{\partial F}{\partial t} + \tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 F}{\partial S^2} + \mu S\frac{\partial F}{\partial S} + \tfrac{1}{2}(\sigma^2 - \tilde{\sigma}^2)\,e^{-r(t-t_0)}S^2\Gamma^i = 0,$$

with zero final and boundary conditions. Using the above changes of variables this becomes $F(S,t) = w(x,\tau)$ where

$$\frac{\partial w}{\partial \tau} = \frac{\partial^2 w}{\partial x^2} + \frac{E(\sigma^2 - \tilde{\sigma}^2)e^{-r(T-t_0)}e^{-d_2^2/2}}{\sigma\tilde{\sigma}\sqrt{\pi\tau}}$$

where

$$d_2 = \frac{\sigma}{\tilde{\sigma}}\frac{x - \frac{2}{\sigma^2}\left(\mu - \frac{1}{2}\sigma^2\right)\tau + \frac{2}{\sigma^2}\left(r - D - \frac{1}{2}\tilde{\sigma}^2\right)\tau}{\sqrt{2\tau}}.$$

The solution of this problem is, using (12.8),

$$\frac{1}{2\pi}\frac{E(\sigma^2-\tilde{\sigma}^2)e^{-r(T-t_0)}}{\sigma\tilde{\sigma}}\int_{-\infty}^{\infty}\int_0^{\tau}\frac{1}{\sqrt{\tau'}}\frac{1}{\sqrt{\tau-\tau'}}$$

$$\times\exp\left(-\frac{(x-x')^2}{4(\tau-\tau')}-\frac{\sigma^2}{4\tilde{\sigma}^2\tau'}\left(x'-\frac{2}{\sigma^2}\left(\mu-\tfrac{1}{2}\sigma^2\right)\tau'+\frac{2}{\sigma^2}\left(r-D-\tfrac{1}{2}\tilde{\sigma}^2\right)\tau'\right)^2\right)\,d\tau'\,dx'.$$

If we write the argument of the exponential function as

$$-a(x'+b)^2+c$$

we have the solution

$$\frac{1}{2\pi}\frac{E(\sigma^2-\tilde{\sigma}^2)e^{-r(T-t_0)}}{\sigma\tilde{\sigma}}\int_0^{\tau}\frac{1}{\sqrt{\tau'}}\frac{1}{\sqrt{\tau-\tau'}}\int_{-\infty}^{\infty}\exp\left(-a(x'+b)^2+c\right)\,dx'\,d\tau'$$

$$=\frac{1}{2\sqrt{\pi}}\frac{E(\sigma^2-\tilde{\sigma}^2)e^{-r(T-t_0)}}{\sigma\tilde{\sigma}}\int_0^{\tau}\frac{1}{\sqrt{\tau'}}\frac{1}{\sqrt{\tau-\tau'}}\frac{1}{\sqrt{a}}\exp(c)\,d\tau'.$$

It is easy to show that

$$a=\frac{1}{4(\tau-\tau')}+\frac{\sigma^2}{4\tilde{\sigma}^2\tau'}$$

and

$$c=-\frac{\sigma^2}{4\tilde{\sigma}^2\tau'(\tau-\tau')}\frac{\left(x-\frac{2\tau'}{\sigma^2}\left(\mu-\tfrac{1}{2}\sigma^2-r+D+\tfrac{1}{2}\tilde{\sigma}^2\right)\right)^2}{\dfrac{1}{\tau-\tau'}+\dfrac{\sigma^2}{\tilde{\sigma}^2\tau'}}.$$

With

$$s-t=\frac{2}{\sigma^2}\tau'$$

we have

$$c=-\frac{\left(\ln(S/E)+\left(\mu-\tfrac{1}{2}\sigma^2\right)(s-t)+\left(r-D-\tfrac{1}{2}\tilde{\sigma}^2\right)(T-s)\right)^2}{2(\sigma^2(s-t)+\tilde{\sigma}^2(T-s))}.$$

From this follows, that the expected profit initially ($t=t_0$, $S=S_0$, $I=0$) is

$$\frac{Ee^{-r(T-t_0)}(\sigma^2-\tilde{\sigma}^2)}{2\sqrt{2\pi}}\int_{t_0}^{T}\frac{1}{\sqrt{\sigma^2(s-t_0)+\tilde{\sigma}^2(T-s)}}$$

$$\times\exp\left(-\frac{\left(\ln(S_0/E)+\left(\mu-\tfrac{1}{2}\sigma^2\right)(s-t_0)+\left(r-D-\tfrac{1}{2}\tilde{\sigma}^2\right)(T-s)\right)^2}{2(\sigma^2(s-t_0)+\tilde{\sigma}^2(T-s))}\right)ds.$$

Variance, Single Option

The problem for the expectation of the square of the profit is

$$\frac{\partial v}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 v}{\partial S^2}+\mu S\frac{\partial v}{\partial S}+\tfrac{1}{2}(\sigma^2-\tilde{\sigma}^2)\,e^{-r(t-t_0)}S^2\Gamma^i\,\frac{\partial v}{\partial I}=0, \tag{12.10}$$

with

$$v(S,I,T)=I^2.$$

A solution can be found of the form

$$v(S,I,t)=I^2+2I\;H(S,t)+G(S,t).$$

Substituting this into Equation (12.10) leads to the following equations for H and G (both to have zero final and boundary conditions):

$$\frac{\partial H}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 H}{\partial S^2}+\mu S\frac{\partial H}{\partial S}+\tfrac{1}{2}(\sigma^2-\tilde{\sigma}^2)\,e^{-r(t-t_0)}S^2\Gamma^i=0;$$

$$\frac{\partial G}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 G}{\partial S^2}+\mu S\frac{\partial G}{\partial S}+(\sigma^2-\tilde{\sigma}^2)\,e^{-r(t-t_0)}S^2\Gamma^iH=0.$$

Comparing the equations for H and the earlier F we can see that

$$H=F=\frac{Ee^{-r(T-t_0)}(\sigma^2-\tilde{\sigma}^2)}{2\sqrt{2\pi}}\int_t^T\frac{1}{\sqrt{\sigma^2(s-t)+\tilde{\sigma}^2(T-s)}}$$
$$\times\exp\left(-\frac{\left(\ln(S/E)+\left(\mu-\frac{1}{2}\sigma^2\right)(s-t)+\left(r-D-\frac{1}{2}\tilde{\sigma}^2\right)(T-s)\right)^2}{2(\sigma^2(s-t)+\tilde{\sigma}^2(T-s))}\right)ds.$$

Notice in this that the expression is a present value at time $t=t_0$, hence the $e^{-r(T-t_0)}$ term at the front. The rest of the terms in this must be kept as the running variables S and t.

Returning to variables x and τ, the governing equation for $G(S,t)=w(x,\tau)$ is

$$\frac{\partial w}{\partial \tau}=\frac{\partial^2 w}{\partial x^2}+\frac{2}{\sigma^2}\frac{E\sigma(\sigma^2-\tilde{\sigma}^2)e^{-r(T-t_0)}e^{-d_2^2/2}}{4\tilde{\sigma}\sqrt{\pi\tau}}$$
$$\times\frac{E(\sigma^2-\tilde{\sigma}^2)e^{-r(T-t_0)}}{2\sigma\tilde{\sigma}\sqrt{\pi}}\int_0^\tau\frac{1}{\sqrt{\tau'}}\frac{1}{\sqrt{\tau-\tau'}}\frac{1}{\sqrt{a}}\exp(c)\,d\tau' \tag{12.11}$$

where

$$d_2=\frac{\sigma}{\tilde{\sigma}}\frac{x-\frac{2}{\sigma^2}\left(\mu-\frac{1}{2}\sigma^2+r-D-\frac{1}{2}\tilde{\sigma}^2\right)\tau}{\sqrt{2\tau}},$$

and a and c are as above.

The solution is therefore

$$\frac{1}{2\sqrt{\pi}}\frac{2}{\sigma^2}\frac{E\sigma(\sigma^2-\tilde{\sigma}^2)e^{-r(T-t_0)}}{4\tilde{\sigma}\sqrt{\pi}}\frac{E(\sigma^2-\tilde{\sigma}^2)e^{-r(T-t_0)}}{2\sigma\tilde{\sigma}\sqrt{\pi}}\int_{-\infty}^{\infty}\int_0^{\tau}\frac{f(x',\tau')e^{-d_2^2/2}}{\sqrt{\tau-\tau'}}$$
$$\times\, e^{-(x-x')^2/4(\tau-\tau')}\,d\tau'\,dx'.$$

Where now

$$f(x',\tau')=\frac{1}{\sqrt{\tau'}}\int_0^{\tau'}\frac{1}{\sqrt{\tau''}}\frac{1}{\sqrt{\tau'-\tau''}}\frac{1}{\sqrt{a}}\exp\left(c\right)\,d\tau''$$

and in a and c all τs become τ's and all τ's become τ''s, and in d_2 all τs become τ's and all xs become x's.

The coefficient in front of the integral signs simplifies to

$$\frac{1}{8\pi^{3/2}}\frac{E^2(\sigma^2-\tilde{\sigma}^2)^2e^{-2r(T-t_0)}}{\sigma^2\tilde{\sigma}^2}.$$

The integral term is of the form

$$\int_{-\infty}^{\infty}\int_0^{\tau}\int_0^{\tau'}\cdots\,d\tau''\,d\tau'\,dx',$$

with the integrand being the product of an algebraic term

$$\frac{1}{\sqrt{\tau'}\sqrt{\tau''}\sqrt{\tau-\tau'}\sqrt{\tau'-\tau''}\sqrt{a}}$$

and an exponential term

$$\exp\left(-\tfrac{1}{2}d_2^2-\frac{(x-x')^2}{4(\tau-\tau')}+c\right).$$

This exponent is, in full,

$$-\frac{1}{4\tau'}\frac{\sigma^2}{\tilde{\sigma}^2}\left(x'-\frac{2}{\sigma^2}\left(\mu-\tfrac{1}{2}\sigma^2\right)\tau'+\frac{2}{\sigma^2}\left(r-D-\tfrac{1}{2}\tilde{\sigma}^2\right)\tau'\right)^2-\frac{(x-x')^2}{4(\tau-\tau')}$$
$$-\frac{\sigma^2}{4\tilde{\sigma}^2\tau''(\tau'-\tau'')}\frac{\left(x'-\dfrac{2\tau''}{\sigma^2}\left(\mu-\tfrac{1}{2}\sigma^2-r+D+\tfrac{1}{2}\tilde{\sigma}^2\right)\right)^2}{\dfrac{1}{\tau'-\tau''}+\dfrac{\sigma^2}{\tilde{\sigma}^2\tau''}}.$$

This can be written in the form

$$-d(x'+f)^2+g,$$

where

$$d=\frac{1}{4}\frac{\sigma^2}{\tilde{\sigma}^2}\frac{1}{\tau'}+\frac{1}{4}\frac{1}{\tau-\tau'}+\frac{1}{4}\frac{\sigma^2}{\sigma^2(\tau'-\tau'')+\tilde{\sigma}^2\tau''}$$

and

$$g = -\frac{\sigma^2}{4\tilde{\sigma}^2\tau'}\left(x - \frac{2\alpha\tau'}{\sigma^2}\right)^2 - \frac{\sigma^2}{4(\sigma^2(\tau'-\tau'') + \tilde{\sigma}^2\tau'')}\left(x - \frac{2\alpha\tau''}{\sigma^2}\right)^2$$
$$+ \frac{1}{4}\frac{\left(\frac{\sigma^2}{\tilde{\sigma}^2\tau'}\left(x - \frac{2\alpha\tau'}{\sigma^2}\right) + \frac{\sigma^2}{(\sigma^2(\tau'-\tau'') + \tilde{\sigma}^2\tau'')}\left(x - \frac{2\alpha\tau''}{\sigma^2}\right)\right)^2}{\frac{\sigma^2}{\tilde{\sigma}^2}\frac{1}{\tau'} + \frac{1}{\tau-\tau'} + \frac{\sigma^2}{\sigma^2(\tau'-\tau'') + \tilde{\sigma}^2\tau''}},$$

where

$$\alpha = \mu - \tfrac{1}{2}\sigma^2 - r + D + \tfrac{1}{2}\tilde{\sigma}^2.$$

Using Equation (12.7) we end up with

$$\frac{1}{4\pi^{3/2}}\frac{E^2(\sigma^2 - \tilde{\sigma}^2)^2 e^{-2r(T-t_0)}}{\sigma^2\tilde{\sigma}^2}$$
$$\int_0^{\tau}\int_0^{\tau'} \frac{1}{\sqrt{\tau'}\sqrt{\tau''}\sqrt{\tau-\tau'}\sqrt{\tau'-\tau''}\sqrt{a}}\sqrt{\frac{\pi}{d}}\exp(g)\,d\tau''\,d\tau'.$$

Changing variables to

$$\tau = \frac{\sigma^2}{2}(T-t), \quad \tau' = \frac{\sigma^2}{2}(T-s), \quad \text{and} \quad \tau'' = \frac{\sigma^2}{2}(T-u),$$

and evaluating at $S = S_0$, $t = t_0$, gives the required result for the variance.

Expectation, Portfolio of Options

This expression follows from the additivity of expectations.

Variance, Portfolio of Options

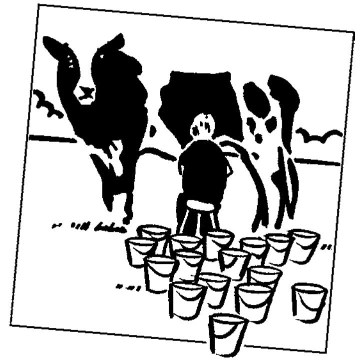
The manipulations and calculations required for the analysis of the portfolio variance are similar to that for a single contract. There is again a solution of the form

$$v(S, I, t) = I^2 + 2I\ H(S, t) + G(S, t).$$

The main differences are that we have to carry around two implied volatilities, $\tilde{\sigma}_j$ and $\tilde{\sigma}_k$, and two expirations, T_j and T_k. We will find that the solution for the variance is the sum of terms satisfying diffusion equations with source terms like in Equation (12.11). The subscript 'k' is then associated with the gamma term, and so appears outside the integral in the equivalent of (12.11), and the subscript 'j' is associated with the integral and so appears in the integrand.

There is one additional subtlety in the derivations and that concerns the expirations. We must consider the general case $T_j \neq T_k$. The integrations in (12.5) must only be taken over the intervals up until the options have expired. The easiest way to apply this is to use the convention that the gammas are zero after expiration. For this reason the s integral is over t_0 to $\min(T_j, T_k)$.

CHAPTER 13
fixed-income products and analysis: yield, duration and convexity



In this Chapter...

- the names and properties of the basic and most important fixed-income products
- the definitions of features commonly found in fixed-income products
- simple ways to analyze the market value of the instruments: yield, duration and convexity
- how to construct yield curves and forward rates

13.1 INTRODUCTION

This chapter is an introduction to some basic instruments and concepts in the world of fixed income, that is, the world of cashflows that are in the simplest cases independent of any stocks, commodities etc. I will describe the most elementary of fixed-income instruments, the coupon-bearing bond, and show how to determine various properties of such bonds to help in their analysis.

This chapter is self contained, and does not require any knowledge from earlier chapters. A lot of it is also not really necessary reading for anything that follows. The reason for this is that, although the concepts and techniques I describe here are used in practice and are *useful* in practice, it is difficult to present a completely coherent theory for more sophisticated products in this framework. Part Three is all about such coherent frameworks.

13.2 SIMPLE FIXED-INCOME CONTRACTS AND FEATURES

13.2.1 The Zero-coupon Bond

The **zero-coupon bond** is a contract paying a known fixed amount, the **principal**, at some given date in the future, the **maturity** date T. For example, the bond pays \$100 in 10 years' time, see Figure 13.1. We're going to scale this payoff, so that in future all principals will be \$1.

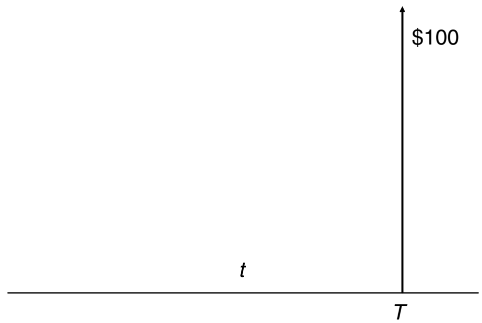


Figure 13.1 The zero-coupon bond.

This promise of future wealth is worth something now: it cannot have zero or negative value. Furthermore, except in extreme circumstances, the amount you pay initially will be smaller than the amount you receive at maturity.

We discussed the idea of time value of money in Chapter 1. This is clearly relevant here and we will return to this in a moment.

13.2.2 The Coupon-bearing Bond

A **coupon-bearing bond** is similar to the above except that as well as paying the principal at maturity, it pays smaller quantities, the coupons, at intervals up to and including the maturity date (see Figure 13.2).

These coupons are usually prespecified fractions of the principal. For example, the bond pays \$1 in 10 years and 2%, i.e. 2 cents, every six months. This would be called a 4% coupon. This bond is clearly more valuable than the bond in the previous example because of the coupon payments. We can think of the coupon-bearing bond as a portfolio of zero-coupon bearing bonds; one zero-coupon bearing bond for each coupon date with a principal being the same as the original bond's coupon, and then a final zero-coupon bond with the same maturity as the original.

Figure 13.3 is an excerpt from *The Wall Street Journal Europe* of 14th April 2005 showing US Treasury Bonds, Notes and Bills. Observe that there are many different 'rates' or coupons, and different maturities. The values of the different bonds will depend on the size of the coupon, the maturity and the market's view of the future behavior of interest rates.

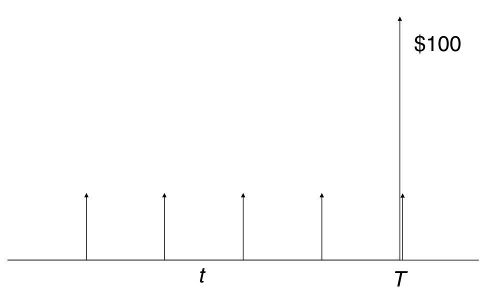


Figure 13.2 The coupon-bearing bond.

Figure 13.3 *The Wall Street Journal Europe* of 14th April 2005 Treasury Bonds, Notes and Bills. Reproduced by permission of Dow Jones & Company, Inc.

13.2.3 The Money Market Account

Everyone who has a bank account has a **money market account**. This is an account that accumulates interest compounded at a rate that varies from time to time. The rate at which interest accumulates is usually a short-term and unpredictable rate. In the sense that money held in a money market account will grow at an unpredictable rate, such an account is risky when compared with a one-year zero-coupon bond. On the other hand, the money market account can be closed at any time but if the bond is sold before maturity there is no guarantee how much it will be worth at the time of the sale.

13.2.4 Floating Rate Bonds

In its simplest form a **floating interest rate** is the amount that you get on your bank account (see Figure 13.4). This amount varies from time to time, reflecting the state of the economy and in response to pressure from other banks for your business. This uncertainty about the interest rate you receive is compensated by the flexibility of your deposit; it can be withdrawn at any time.

The most common measure of interest is **London Interbank Offer Rate** or **LIBOR**. LIBOR comes in various maturities, one month, three month, six month etc., and is the rate of interest offered between Eurocurrency banks for fixed-term deposits.

Sometimes the coupon payment on a bond is not a prescribed dollar amount but depends on the level of some 'index,' measured at the time of the payment or before. Typically, we cannot know at the start of the contract what level this index will be at when the payment is made. We will see examples of such contracts in later chapters.

13.2.5 Forward Rate Agreements

A **Forward Rate Agreement** (**FRA**) is an agreement between two parties that a prescribed interest rate will apply to a prescribed principal over some specified period in the future. The cashflows in this agreement are as follows: party A pays party B the principal at time T_1 and B pays A the principal plus agreed interest at time $T_2 > T_1$. The value of this exchange at the time the contract is entered into is generally not zero and so there will be a transfer of cash from one party to the other at the start date.

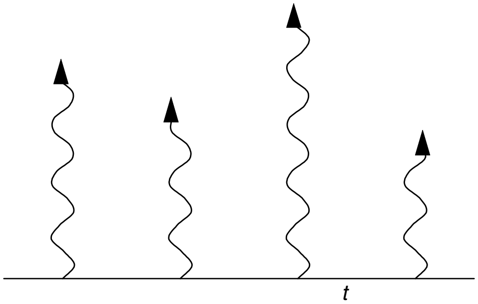


Figure 13.4 The floating-rate bond.

13.2.6 Repos

A **repo** is a repurchase agreement. It is an agreement to sell some security to another party and buy it back at a fixed date and for a fixed amount. The price at which the security is bought back is greater than the selling price and the difference implies an interest rate called the **repo rate**. The commonest repo is the overnight repo in which the agreement is renegotiated daily. If the repo agreement extends for 30 days it is called a **term repo**.

A **reverse repo** is the borrowing of a security for a short period at an agreed interest rate.

Repos can be used to lock in future interest rates. For example, buy a six-month Treasury bill today and repo it out for three months. There is no cash flow today since the bond has been paid for (money out) and then repoed (same amount in). In three months' time you will have to repurchase the bill at the agreed price; this is an outflow of cash. In six months you receive the principal. Money out in three months, money in six months; for there to be no arbitrage the equivalent interest rate should be that currently prevailing between three and six months' time.

13.2.7 STRIPS

STRIPS stands for 'Separate Trading of Registered Interest and Principal of Securities'. The coupons and principal of normal bonds are split up, creating artificial zero-coupon bonds of longer maturity than would otherwise be available.

13.2.8 Amortization

In all of the above products I have assumed that the principal remains fixed at its initial level. Sometimes this is not the case; the principal can **amortize** or decrease during the life of the contract. The principal is thus paid back gradually and interest is paid on the amount of the principal outstanding. Such amortization is arranged at the initiation of the contract and may be fixed, so that the rate of decrease of the principal is known beforehand, or can depend on the level of some index: if the index is high the principal amortizes faster for example. We see an example of a complex amortizing structure in Chapter 32.

13.2.9 Call Provision

Some bonds have a **call provision**. The issuer can call back the bond on certain dates or at certain periods for a prescribed, possibly time-dependent, amount. This lowers the value of the bond. The mathematical consequences of this are discussed in Chapter 30.

13.3 INTERNATIONAL BOND MARKETS

13.3.1 United States of America

In the US, bonds of maturity less than one year are called **bills** and are usually zero coupon. Bonds with maturity 2–10 years are called **notes**. They are coupon bearing with coupons every six months. Bonds with maturity greater than 10 years are called **bonds**. Again they are coupon bearing. In this book I tend to call all of these 'bonds,' merely specifying whether or not they have coupons.

Bonds traded in the United States foreign bond market but which are issued by non-US institutions are called **Yankee** bonds.

Since the beginning of 1997 the US government has also issued bonds linked to the rate of inflation.

13.3.2 United Kingdom

Bonds issued by the UK government are called **gilts**. Some of these bonds are callable, some are irredeemable, meaning that they are perpetual bonds having a coupon but no repayment of principal. The government also issues convertible bonds which may be converted into another bond issue, typically of longer maturity. Finally, there are index-linked bonds having the amount of the coupon and principal payments linked to a measure of inflation, the Retail Price Index (RPI).

13.3.3 Japan

Japanese Government Bonds (**JGB**s) come as short-term treasury bills, medium-term, long-term (10-year maturity) and super long-term (20-year maturity). The long- and super long-term bonds have coupons every six months. The short-term bonds have no coupons and the medium-term bonds can be either coupon-bearing or zero-coupon bonds.

Yen denominated bonds issued by non-Japanese institutions are called **Samurai** bonds.

13.4 ACCRUED INTEREST

The market price of bonds quoted in the newspapers are **clean prices**. That is, they are quoted without any **accrued interest**. The accrued interest is the amount of interest that has built up since the last coupon payment:

$$\text{accrued interest} = \text{interest due in full period} \times \frac{\text{number of days since last coupon date}}{\text{number of days in period between coupon payments}}.$$

The actual payment is called the **dirty price** and is the sum of the quoted clean price and the accrued interest.

13.5 DAY-COUNT CONVENTIONS

Because of such matters as the accrual of interest between coupon dates there naturally arises the question of how to accrue interest over shorter periods. Interest is accrued between two dates according to the formula

$$\frac{\text{number of days between the two dates}}{\text{number of days in period}} \times \text{interest earned in reference period.}$$

There are three main ways of calculating the 'number of days' in the above expression.

- **Actual/Actual** Simply count the number of calendar days
- **30/360** Assume there are 30 days in a month and 360 days in a year
- **Actual/360** Each month has the right number of days but there are only 360 days in a year

13.6 CONTINUOUSLY AND DISCRETELY COMPOUNDED INTEREST

To be able to compare fixed-income products we must decide on a convention for the measurement of interest rates. So far, we have used a continuously compounded rate, meaning that the present value of \$1 paid at time T in the future is

$$e^{-rT} \times \$1$$

for some r. We have seen how this follows from the cash-in-the-bank or money market account equation

$$dM = rM\,dt.$$

This is the convention used in the options world.

Another common convention is to use the formula

$$\frac{1}{(1 + r')^T} \times \$1,$$

for present value, where r' is some interest rate. This represents discretely compounded interest and assumes that interest is accumulated *annually* for T years. The formula is derived from calculating the present value from a single-period payment, and then compounding this for each year. This formula is commonly used for the simpler type of instruments such as coupon-bearing bonds. The two formulae are identical, of course, when

$$r = \log(1 + r').$$

This gives the relationship between the continuously compounded interest rate r and the discrete version r'. What would the formula be if interest was discretely compounded twice per year?

In this book we tend to use the continuous definition of interest rates.

13.7 MEASURES OF YIELD

There is such a variety of fixed-income products, with different coupon structures, amortization, fixed and/or floating rates, that it is necessary to be able to compare different products consistently. Suppose you have to choose between a ten-year zero-coupon bond and a 20-year coupon-bearing bond. One has no income for ten years but then gets a big lump sum, the other has a trickle of income but you have to wait much longer for the big amount.

One way to do this is through measures of how much each contract earns; there are several measures of this all coming under the name **yield**.

13.7.1 Current Yield

The simplest measurement of how much a contract earns is the **current yield**. This measure is defined by

$$\text{current yield} = \frac{\text{annual \$ coupon income}}{\text{bond price}}.$$

For example, consider the 10-year bond that pays 2 cents every six months and \$1 at maturity. This bond has a total income per annum of 4 cents. Suppose that the quoted market price of this bond is 88 cents. The current yield is simply

$$\frac{0.04}{0.88} = 4.5\%.$$

This measurement of the yield of the bond makes no allowance for the payment of the principal at maturity, nor for the time value of money if the coupon payment is reinvested, nor for any capital gain or loss that may be made if the bond is sold before maturity. It is a relatively unsophisticated measure, concentrating very much on short-term properties of the bond.

13.7.2 The Yield to Maturity (YTM) or Internal Rate of Return (IRR)

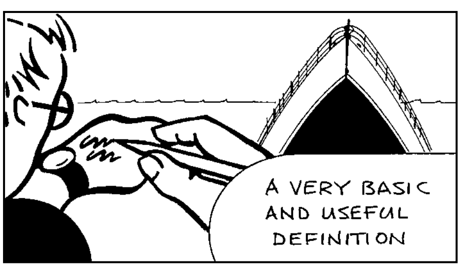
Suppose that we have a zero-coupon bond maturing at time T when it pays one dollar. At time t it has a value $Z(t;T)$. Applying a constant rate of return of y between t and T, then one dollar received at time T has a present value of $Z(t;T)$ at time t, where

$$Z(t;T) = e^{-y(T-t)}.$$

It follows that

$$y = -\frac{\log Z}{T-t}.$$

Let us generalize this. Suppose that we have a coupon-bearing bond. Discount all coupons and the principal to the present by using some interest rate y. The present value of the bond, at time t, is then



$$V = Pe^{-y(T-t)} + \sum_{i=1}^{N} C_i e^{-y(t_i-t)}, \tag{13.1}$$

where P is the principal, N the number of coupons, and C_i the coupon paid on date t_i.

If the bond is a traded security then we know the price at which the bond can be bought. If this is the case then we can calculate the **yield to maturity** or **internal rate of return** as the value y that we must put into Equation (13.1) to make V equal to the traded price of the bond. This calculation must be performed by some trial and error/iterative procedure. For example, in the bond in Table 13.1 we have a principal of \$1 paid in five years and coupons of three cents (three percent) paid every six months.

Suppose that the market value of this bond is 96 cents. We ask 'What is the internal rate of return we must use to give these cash flows a total present value of 96 cents?' This value is the yield to maturity. In the fourth column in this table is the present value (PV) of each of the cashflows using a rate of 6.8406%: since the sum of these present values is 96 cents the YTM or IRR is 6.8406%.

This yield to maturity is a valid measure of the return on a bond if we intend to hold it to maturity.

To calculate the yield to maturity of a portfolio of bonds simply treat all the cashflows as if they were from the one bond and calculate the value of the whole portfolio by adding up the market values of the component bonds.

Table 13.1 An example of a coupon-bearing bond.

Time	Coupon	Principal repayment	PV (discounting at 6.8406%)
0			0
0.5	.03		0.0290
1.0	.03		0.0280
1.5	.03		0.0270
2.0	.03		0.0262
2.5	.03		0.0253
3.0	.03		0.0244
3.5	.03		0.0236
4.0	.03		0.0228
4.5	.03		0.0220
5.0	.03	1.00	0.7316
		Total	0.9600

13.8 THE YIELD CURVE

The plot of yield to maturity against time to maturity is called the **yield curve**. For the moment assume that this has been calculated from zero-coupon bonds and that these bonds have been issued by a perfectly creditworthy source.

If the bonds have coupons then the calculation of the yield curve is more complicated and the 'forward curve,' described below, is a better measure of the interest rate pertaining at some time in the future. Figure 13.5 shows the yield curve for US Treasuries as it was on 9th September 1999.

13.9 PRICE/YIELD RELATIONSHIP

From Equation (13.1) we can easily see that the relationship between the price of a bond and its yield is of the form shown in Figure 13.6 (assuming that all cash flows are positive). On this figure is marked the current market price and the current yield to maturity.

Since we are often interested in the sensitivity of instruments to the movement of certain underlying factors it is natural to ask how does the price of a bond vary with the yield, or vice versa. To a first approximation this variation can be quantified by a measure called the duration.

Figure 13.7 shows the Price/Yield relationship for a specific five-year US Treasury.

13.10 DURATION

From Equation (13.1) we find that

$$\frac{dV}{dy} = -(T-t)Pe^{-y(T-t)} - \sum_{i=1}^{N} C_i(t_i - t)e^{-y(t_i-t)}.$$

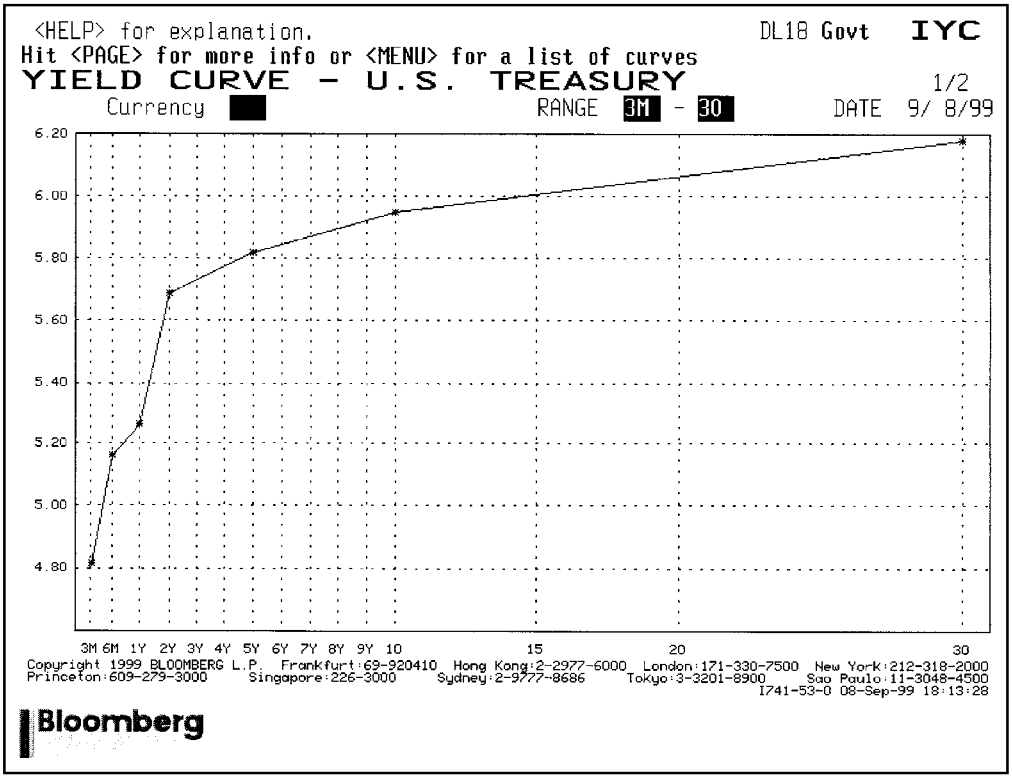


Figure 13.5 Yield curve for US Treasuries. Source: Bloomberg L.P.

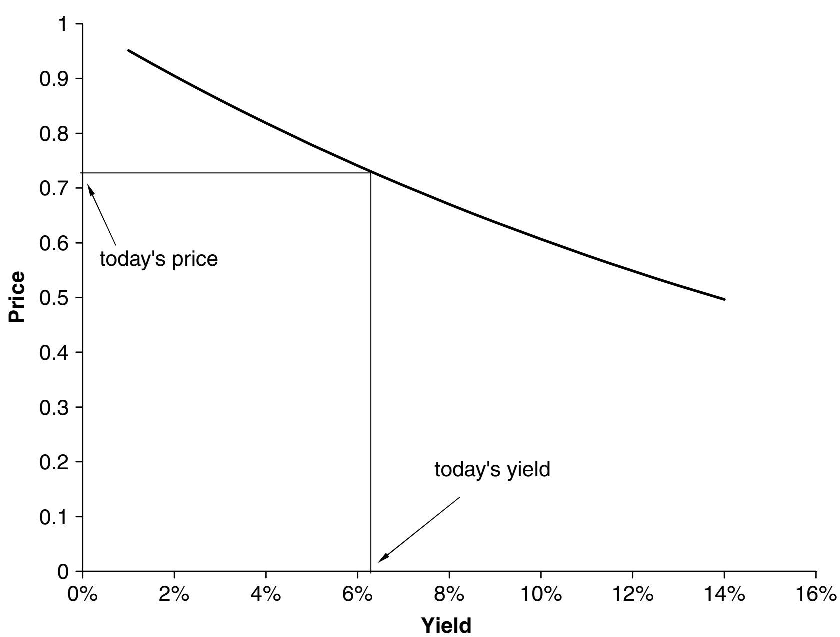


Figure 13.6 The Price/Yield relationship.

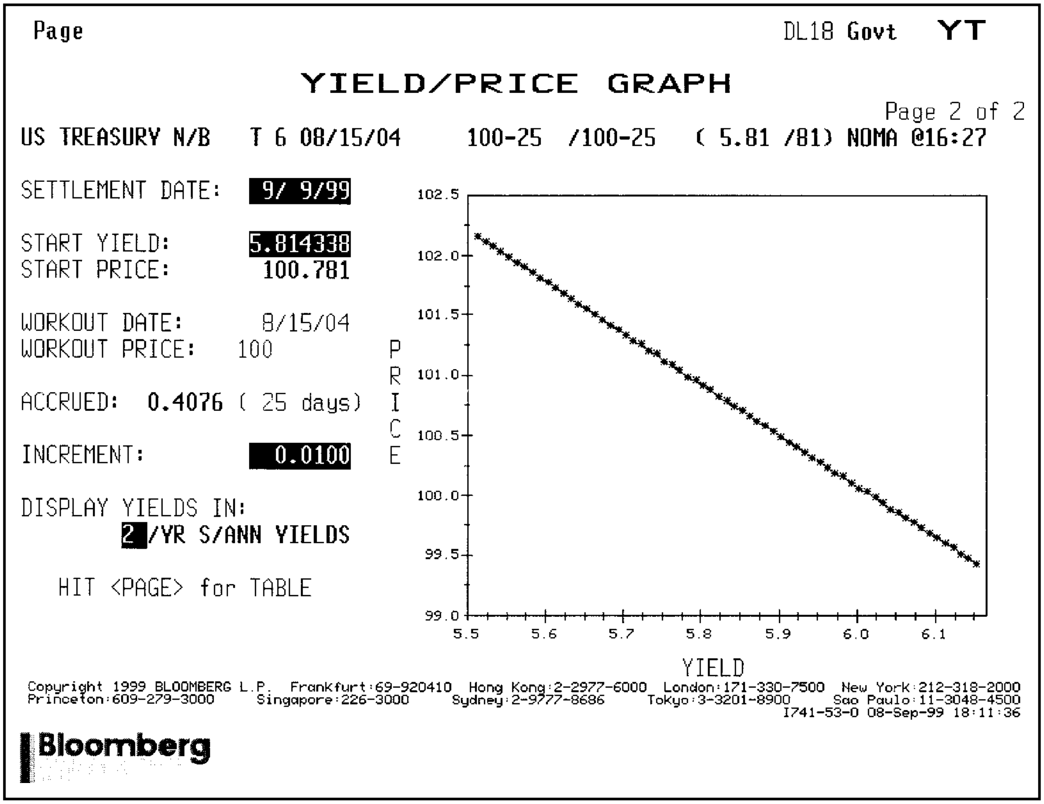
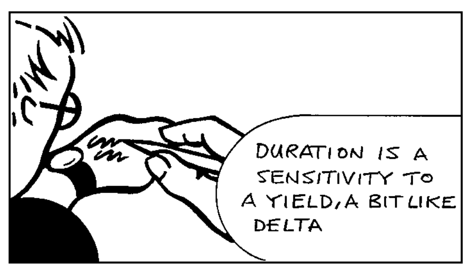


Figure 13.7 Bloomberg's Price/Yield graph. Source: Bloomberg L.P.

This is the slope of the Price/Yield curve. The quantity

$$-\frac{1}{V}\frac{dV}{dy}$$

is called the **Macaulay duration**. (The **modified duration** is similar but uses the discretely compounded rate.) In the expression for the duration the time of each coupon payment is weighted by its present value. The higher the value of the present value of the coupon the more it contributes to the duration. Also, since y is measured in units of inverse time, the units of the duration are time. The duration is a measure of the average life of the bond. It is easily shown that the Macaulay duration for a zero-coupon bond is the same as its maturity.



Let's take a look at the idea of average time. Suppose we asked to what zero-coupon bond is our (coupon-bearing) bond equivalent? That is, what maturity would an 'equivalent' bond have? Take the actual bond's value and equate it to a zero-coupon bond, having the same yield but an unknown maturity (and unknown quantity!):

$$V = Pe^{-y(T-t)} + \sum_{i=1}^{N} C_i e^{-y(t_i-t)} = X\, e^{-y(\overline{T}-t)}.$$

Differentiate both sides with respect to y:

$$\frac{dV}{dy} = -(T-t)Pe^{-y(T-t)} - \sum_{i=1}^{N} C_i(t_i - t)e^{-y(t_i-t)} = -X\ (\overline{T}-t)\ e^{-y(\overline{T}-t)}.$$

Finally, divide both sides by $-V$:

$$-\frac{1}{V}\frac{dV}{dy} = \cdots = \overline{T} - t.$$

Hence the statement about the bond's *average* life, or effective maturity.

For small movements in the yield, the duration gives a good measure of the change in value with a change in the yield. For larger movements we need to look at higher order terms in the Taylor series expansion of $V(y)$.

One of the most common uses of the duration is in plots of yield versus duration for a variety of instruments. An example is shown in Figure 13.8. Look at the bond marked 'CPU.' This bond has a coupon of 4.75% paid twice per year, callable from June 1998 and maturing in June 2000. We can use this plot to group together instruments with the same or similar durations and make comparisons between their yields. Two bonds having the same duration but with one bond having a higher yield might be suggestive of value for money in the higher-yielding bond, or of credit risk issues. However, such indicators of relative value must be used with care. It is possible for two bonds to have vastly different cashflow profiles yet have the same duration; one may have a maturity of 30 years but an average life and hence a duration of seven years, whereas another may be a seven-year zero-coupon bond. Clearly, the former has 23 years more risk than the latter.

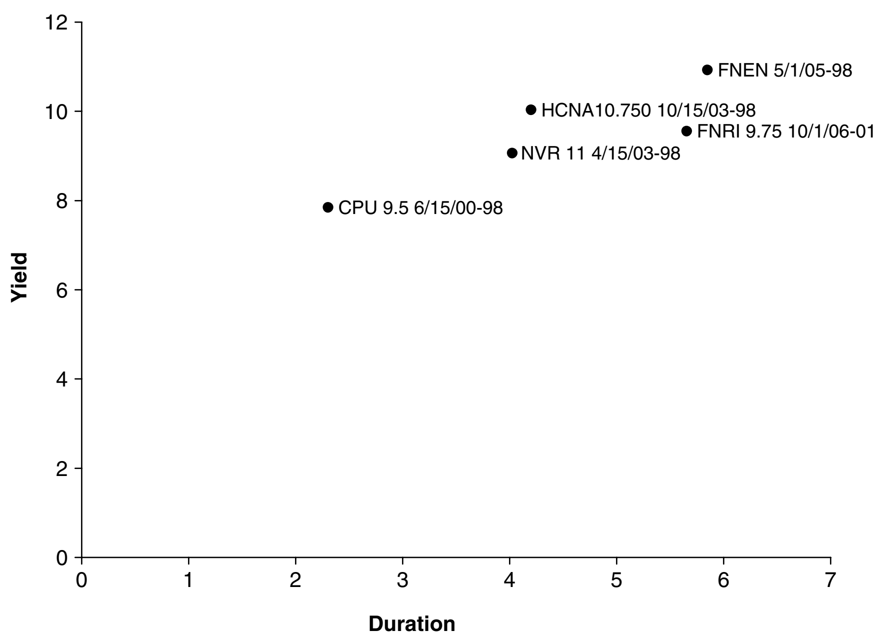


Figure 13.8 Yield versus duration; measuring the relative value of bonds.

13.11 CONVEXITY

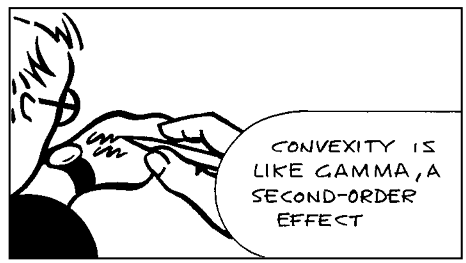
The Taylor series expansion of V gives

$$\frac{dV}{V} = \frac{1}{V}\frac{dV}{dy}\delta y + \frac{1}{2V}\frac{d^2V}{dy^2}(\delta y)^2 + \cdots,$$

where δy is a change in yield. For very small movements in the yield, the change in the price of a bond can be measured by the duration. For larger movements we must take account of the curvature in the Price/Yield relationship.

The **dollar convexity** is defined as

$$\frac{d^2V}{dy^2} = (T-t)^2 Pe^{-y(T-t)} + \sum_{i=1}^{N} C_i(t_i-t)^2 e^{-y(t_i-t)}$$



and the **convexity** is

$$\frac{1}{V}\frac{d^2V}{dy^2}.$$

To see how these can be used, examine Figure 13.9.

In this figure we see the Price/Yield relationship for two bonds having the same value and duration when the yield is around 8%, but then they have different convexities. Bond A has a greater convexity than bond B. This figure suggests that bond A is better value than B because a small change in the yields results in a higher value for A. When we develop a consistent theory for pricing bonds when interest rates are stochastic we will see how the absence of arbitrage will lead to relationships between such quantities as yield, duration and convexity, not unlike the Black–Scholes equation.

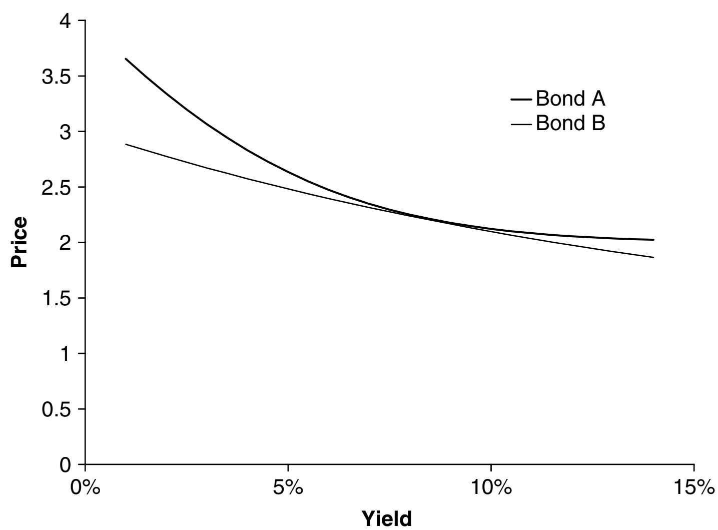
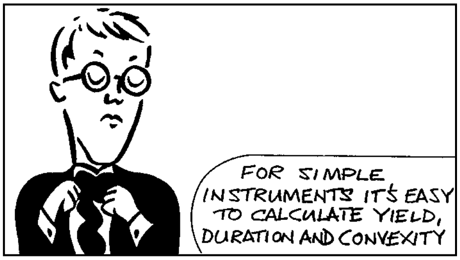


Figure 13.9 Two bonds with the same price and duration but different convexities.



	A	B	C	D	E	F	G	H	I	J	K
1					Date	Coupon	Principal	PVs	Time	Time^2	
2					0				wtd	wtd	
3		YTM	4.95%		0.5	2%		0.0195	0.0098	0.0049	
4		Mkt price	**0.921**		1	2%		0.0190	0.0190	0.0190	
5		Th. Price	0.921		1.5	2%		0.0186	0.0279	0.0418	
6		Error	1.4E-08		2	2%		0.0181	0.0362	0.0725	
7		Duration	8.2544		2.5	2%		0.0177	0.0442	0.1104	
8		Convexity	76.8728		3	2%		0.0172	0.0517	0.1551	
9					3.5	2%		0.0168	0.0589	0.2060	
10					4	2%		0.0164	0.0656	0.2625	
11					4.5	2%		0.0160	0.0720	0.3241	
12					5	2%		0.0156	0.0781	0.3903	
13					5.5	2%		0.0152	0.0838	0.4607	
14					6	2%		0.0149	0.0892	0.5349	
15					6.5	2%		0.0145	0.0942	0.6124	
16					7	2%		0.0141	0.0990	0.6929	
17					7.5	2%		0.0138	0.1035	0.7760	
18					8	2%		0.0135	0.1077	0.8613	
19					8.5	2%		0.0131	0.1116	0.9485	
20					9	2%		0.0128	0.1153	1.0374	
21					9.5	2%		0.0125	0.1187	1.1276	
22					10	2%	1	0.6216	6.2161	62.1614	
23											
24											
25											
26											
27											
28											
29											
30											
31											
32											
33											
34											
35											
36											
37											

= SUM(H3:H22)

= C4-C5

= SUM(I3:I22)/C5

= SUM(J3:J22)/C5

= F20*EXP(-E20*$C$3)

= (G22+F22)*EXP(-E22*$C$3)

= E20*H20

= E20*I20

Goal Seek

Set cell: $C$6

To value: 0

By changing cell: $C$3

Ok Cancel

Figure 13.10 A spreadsheet showing the calculation of yield, duration and convexity.

The calculation of yield to maturity, duration and convexity is shown in Figure 13.10. Inputs are in the grey boxes.

13.12 AN EXAMPLE

Figure 13.11 shows the yield analysis screen from Bloomberg. The yield, duration and convexity have been calculated for a specific US Treasury. Figures 13.12 and 13.13 show time series of the price and yield respectively.

13.13 HEDGING

In measuring and using yields to maturity, it must be remembered that the yield is the rate of discounting that makes the present value of a bond the same as its market value. A yield is thus identified with each individual instrument. It is perfectly possible for the yield on one instrument to rise while another falls, especially if they have significantly different maturities or durations. Nevertheless, one often wants to hedge movements in one bond with movements in another. This is commonly achieved by making one big assumption about the relative movements of yields on the two bonds. Bond A has a yield of y_A, bond B has a yield of y_B, they have different

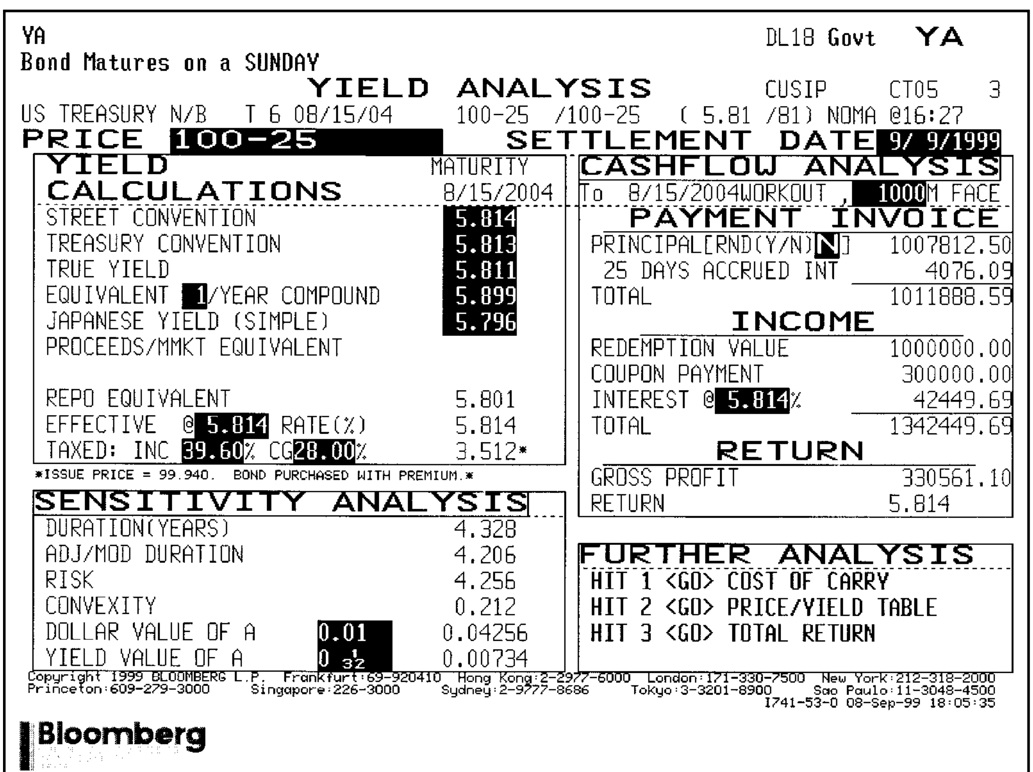


Figure 13.11 Yield analysis. Source: Bloomberg L.P.

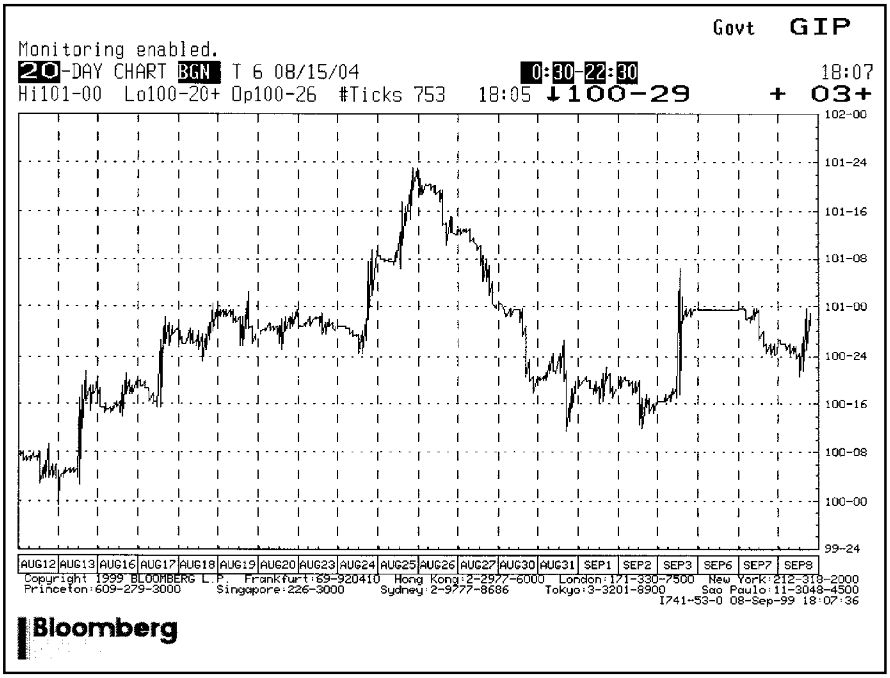


Figure 13.12 Price time series. Source: Bloomberg L.P.

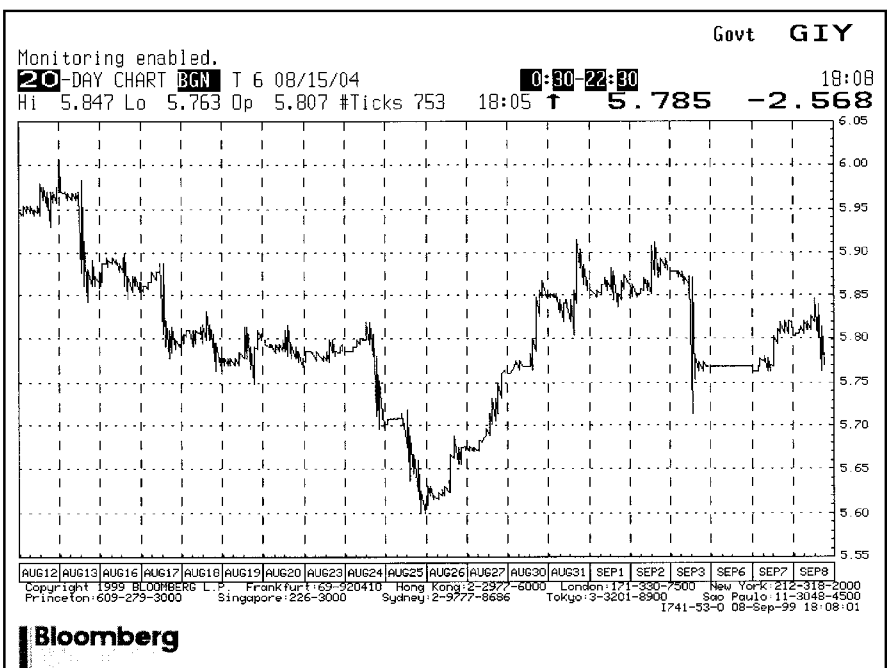


Figure 13.13 Yield time series. Source: Bloomberg L.P.

maturities and durations but we will assume that a move of $x\%$ in A's yield is accompanied by a move of $x\%$ in B's yield. This is the assumption of **parallel shifts** in the yield curve. If this is the case, then if we hold A bonds and B bonds in the inverse ratio of their durations (with one long position and one short) we will be leading-order hedged:

$$\Pi = V_A(y_A) - \Delta V_B(y_B),$$

with the obvious notation for the value and yield of the two bonds. The change in the value of this portfolio is

$$\delta\Pi = \frac{\partial V_A}{\partial y_A}x - \Delta\frac{\partial V_B}{\partial y_B}x + \text{higher-order terms.}$$

Choose

$$\Delta = \frac{\partial V_A}{\partial y_A} \bigg/ \frac{\partial V_B}{\partial y_B}$$

to eliminate the leading-order risk. The higher-order terms depend on the convexity of the two instruments.

Of course, this is a simplification of the real situation; there may be little relationship between the yields on the two instruments, especially if the cash flows are significantly different. In this case there may be twisting or arching of the yield curve.

13.14 TIME-DEPENDENT INTEREST RATE

In this section we examine bond pricing when we have an interest rate that is a known function of time. The interest rate we consider will be what is known as a **short-term interest rate** or **spot interest rate** $r(t)$. This means that the rate $r(t)$ is to apply at time t: interest is compounded at this rate at each moment in time but *this rate may change*; generally we assume it to be time dependent.

If the spot interest rate $r(t)$ is a known function of time, then the bond price is also a function of time only: $V = V(t)$. (The bond price is, of course, also a function of maturity date T, but I suppress that dependence except when it is important.) We begin with a zero-coupon bond example. Because we receive 1 at time $t = T$ we know that $V(T) = 1$. I now derive an equation for the value of the bond at a time before maturity, $t < T$.

Suppose we hold one bond. The change in the value of that bond in a time-step dt (from t to $t + dt$) is

$$\frac{dV}{dt}\,dt.$$

Arbitrage considerations again lead us to equate this with the return from a bank deposit receiving interest at a rate $r(t)$:

$$\frac{dV}{dt} = r(t)V.$$

The solution of this equation is

$$V(t;T) = e^{-\int_t^T r(\tau)\,d\tau}. \tag{13.2}$$

Now let's introduce coupon payments. If during the period t to $t + dt$ we have received a coupon payment of $K(t)\,dt$, which may be either in the form of continuous or discrete payments or a combination, our holdings including cash change by an amount

$$\left(\frac{dV}{dt} + K(t)\right) dt.$$

Again setting this equal to the risk-free rate $r(T)$ we conclude that

$$\frac{dV}{dt} + K(t) = r(t)V. \tag{13.3}$$

Dropping the parameter T, the solution of this ordinary differential equation is easily found to be

$$V(t) = e^{-\int_t^T r(\tau)\,d\tau} \left(1 + \int_t^T K(t')e^{\int_{t'}^T r(\tau)\,d\tau} dt'\right); \tag{13.4}$$

the arbitrary constant of integration has been chosen to ensure that $V(T) = 1$.

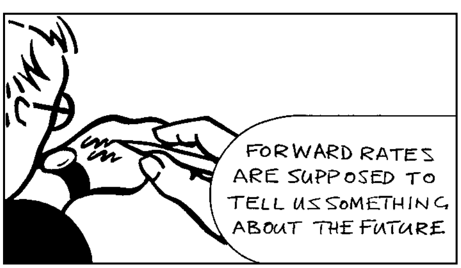
13.15 DISCRETELY PAID COUPONS

Equation (13.4) allows for the payment of a coupon. But what if the coupon is paid discretely, as it is in practice, for example, every six months, say? We can arrive at this result by a financial argument that will be useful later. Since the holder of the bond receives a coupon, call it K_c, at time t_c there must be a jump in the value of the bond across the coupon date. That is, the values before and after this date differ by K_c:

$$V(t_c^-) = V(t_c^+) + K_c.$$

This will be recognized as a jump condition. This time the realized bond price is *not* continuous. After all, there is a discrete payment at the coupon date. This jump condition will still apply when we come to consider stochastic interest rates.

Having built up a simple framework in which interest rates are time dependent I now show how to derive information about these rates from the market prices of bonds.



13.16 FORWARD RATES AND BOOTSTRAPPING

The main problem with the use of yield to maturity as a measure of interest rates is that it is not consistent across instruments. One five-year bond may have a different yield from another five-year bond if they have different coupon structures. It is therefore difficult to say that there is a single interest rate associated with a maturity.

One way of overcoming this problem is to use **forward rates**.

Forward rates are interest rates that are assumed to apply over given periods *in the future* for *all* instruments. This contrasts with yields which are assumed to apply up to maturity, with a different yield for each bond.

Let us suppose that we are in a perfect world in which we have a continuous distribution of zero-coupon bonds with all maturities T. Call the prices of these at time t, $Z(t;T)$. Note the use of Z for zero-coupon.

The **implied forward rate** is the curve of a time-dependent spot interest rate that is consistent with the market price of instruments. If this rate is $r(\tau)$ at time τ then it satisfies

$$Z(t;T) = e^{-\int_t^T r(\tau)\,d\tau}.$$

On rearranging and differentiating this gives

$$r(T) = -\frac{\partial}{\partial T}(\log Z(t;T)).$$

This is the forward rate for time T as it stands today, time t. Tomorrow the whole curve (the dependence of r on the future) may change. For that reason we usually denote the forward rate at time t applying at time T in the future as $F(t;T)$ where

$$F(t;T) = -\frac{\partial}{\partial T}(\log Z(t;T)).$$

Writing this in terms of yields $y(t;T)$ we have

$$Z(t;T) = e^{-y(t;T)(T-t)}$$

and so

$$F(t;T) = y(t;T) + \frac{\partial y}{\partial T}.$$

This is the relationship between yields and forward rates when everything is nicely differentiable (see Figure 13.14).

13.16.1 Discrete Data

In the less-than-perfect real world we must do with only a discrete set of data points. We continue to assume that we have zero-coupon bonds but now we will only have a discrete set of them. We can still find an implied forward rate curve as follows.

Rank the bonds according to maturity, with the shortest maturity first. The market prices of the bonds will be denoted by Z_i^M where i is the position of the bond in the ranking.

Using only the first bond, ask the question 'What interest rate is implied by the market price of the bond?' The answer is given by y_1, the solution of

$$Z_1^M = e^{-y_1(T_1-t)},$$

i.e.

$$y_1 = -\frac{\log(Z_1^M)}{T_1 - t}.$$

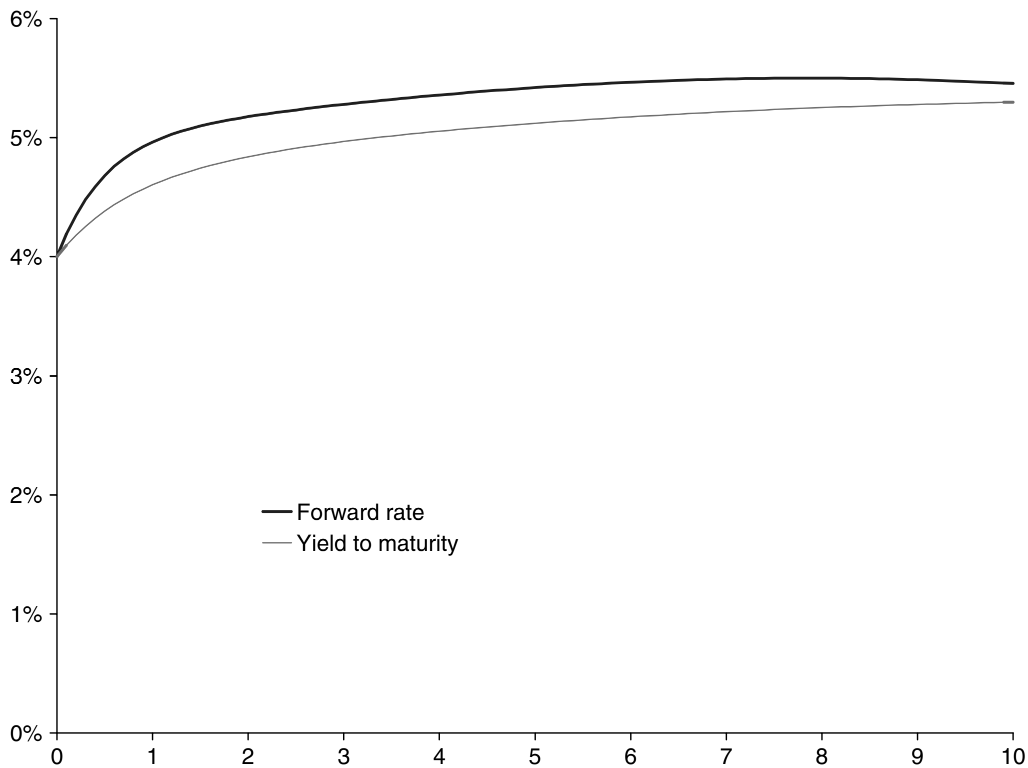


Figure 13.14 The yields and the forward rates.

This rate will be the rate that we use for discounting between the present and the maturity date T_1 of the first bond. And it will be applied to *all* instruments whenever we want to discount over this period.

Now move on to the second bond having maturity date T_2. We know the rate to apply between now and time T_1, but at what interest rate must we discount between dates T_1 and T_2 to match the theoretical and market prices of the second bond? The answer is y_2 which solves the equation

$$Z_2^M = e^{-y_1(T_1-t)}e^{-y_2(T_2-T_1)},$$

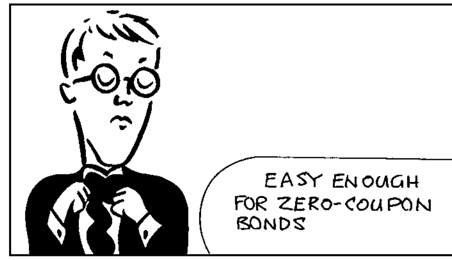
i.e.

$$y_2 = -\frac{\log\left(Z_2^M/Z_1^M\right)}{T_2 - T_1}.$$

By this method of **bootstrapping** we can build up the forward rate curve. Note how the forward rates are applied between two dates, for which period we have assumed they are constant. Figure 13.15, gives an example.

	A	B	C	D	E
1	**Time to**	**Market**	**Yield to**	**Forward**	
2	**maturity**	**price z-c b**	**maturity**	**rate**	
3	0.25	0.9809	7.71%	7.71%	
4	0.5	0.9612	7.91%	8.12%	
5	1	0.9194	8.40%	8.89%	
6	2	0.8436	8.50%	8.60%	
7	3	0.7772	8.40%	8.20%	
8	5	0.644	8.80%	9.40%	
9	7	0.5288	9.10%	9.85%	
10	10	0.3985	9.20%	9.43%	
11					
12	= -LN(B10)/A10				
13					
14					
15		= (C10*A10-C9*A9)/(A10-A9)			
16					

Figure 13.15 A spreadsheet showing the calculation of yields and forward rates from zero-coupon bonds.



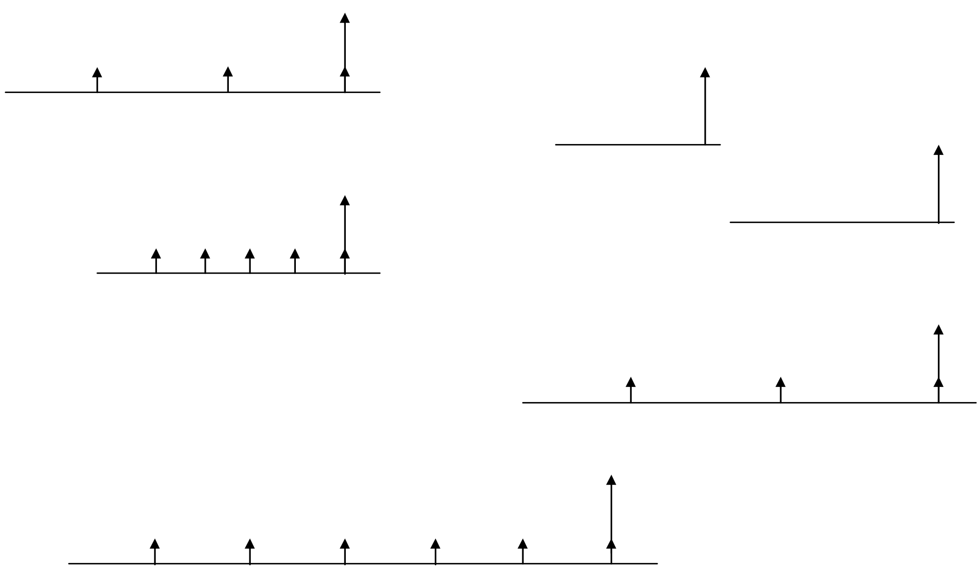


Figure 13.16 The universe of bonds.

This method can easily be extended to accommodate coupon-bearing bonds. Again rank the bonds by their maturities, but now we have the added complexity that we may only have one market value to represent the sum of several cashflows. Thus one often has to make some assumptions to get the right number of equations for the number of unknowns. See Figures 13.16–13.21.

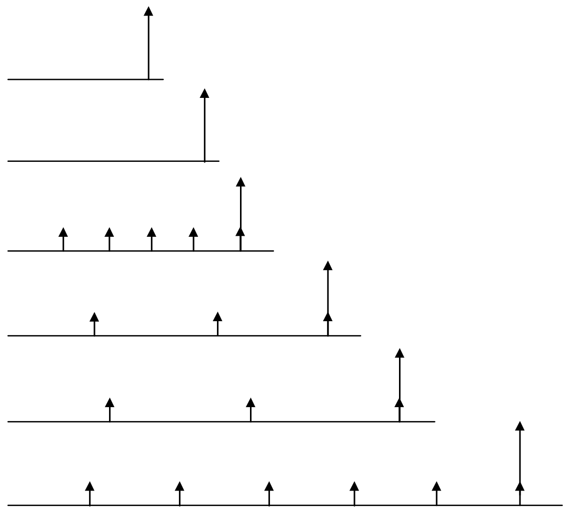


Figure 13.17 The universe of bonds, ranked in order of maturity.

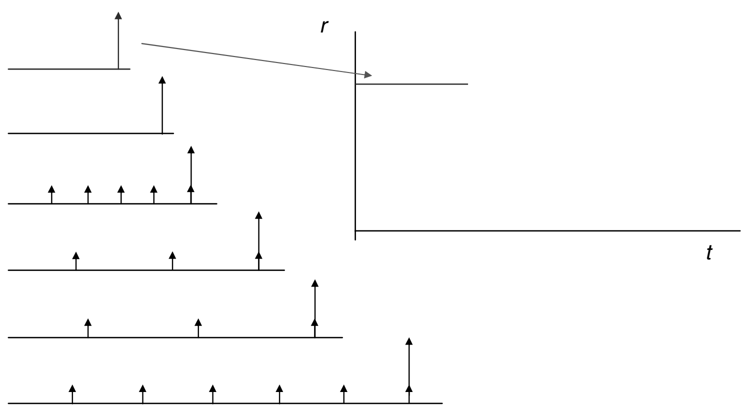


Figure 13.18 The first-maturing bond gives us a forward rate from now until its maturity.

13.16.2 On a Spreadsheet

Given the market price of zero-coupon bonds it is very easy to calculate yields and forward rates, as shown in the spreadsheet (Figure 13.15). Inputs are in the grey boxes.

The yields and forward rates for this data are shown in Figures 13.22 and 13.23. Note that in each case the yield begins at zero maturity and extends up to the maturity of each bond. The forward rates pick up where the last forward rates left off.

There are far more swaps of different maturities than there are bonds, so that in practice swaps are used to build up the forward rates by bootstrapping. Fortunately, there is a simple decomposition of swaps prices into the prices of zero-coupon bonds so that bootstrapping is still relatively straightforward. Swaps are discussed in more detail in Chapter 14.

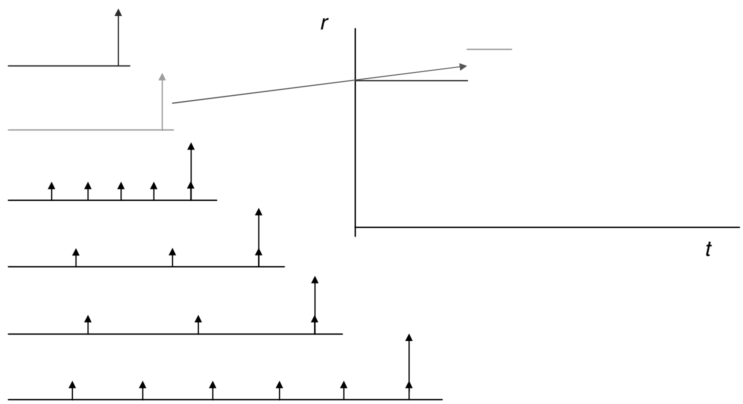


Figure 13.19 The second-maturing bond gives us a forward rate from maturity of the previous bond until its own maturity.

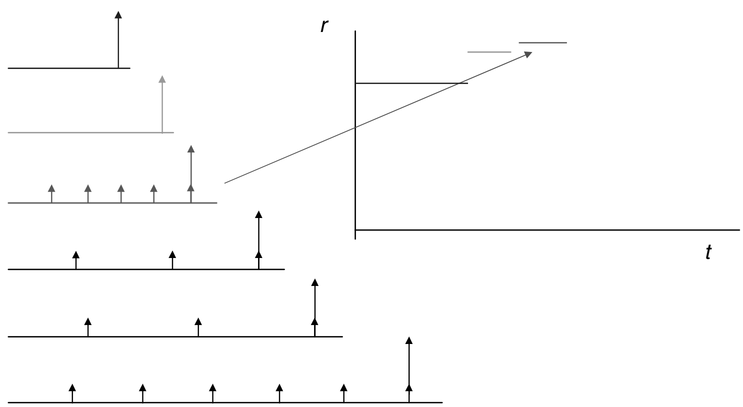


Figure 13.20 The third-maturing bond gives us a forward rate from maturity of the previous bond until its own maturity.

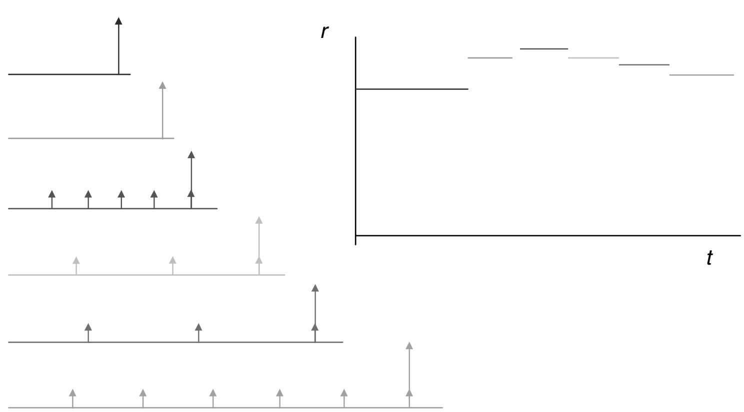


Figure 13.21 And so on.

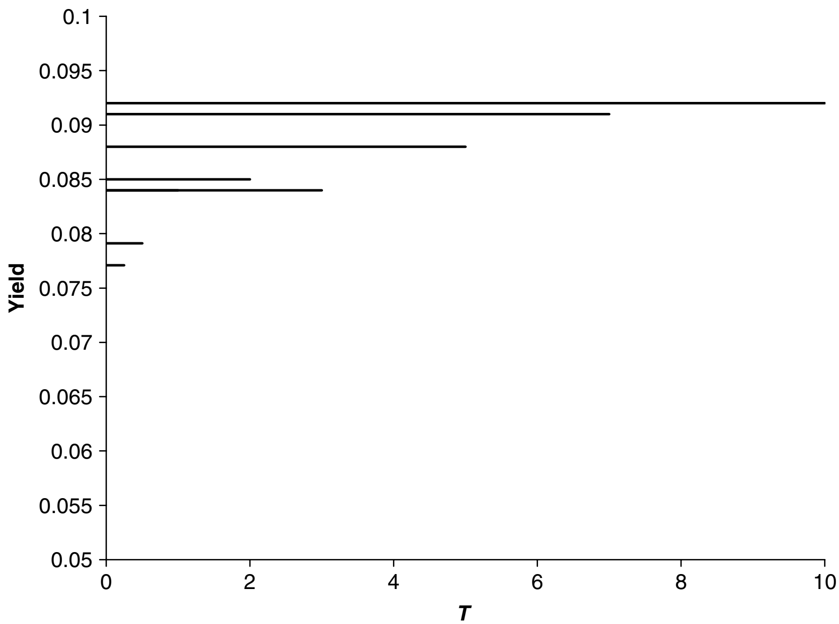


Figure 13.22 Yield to maturities.

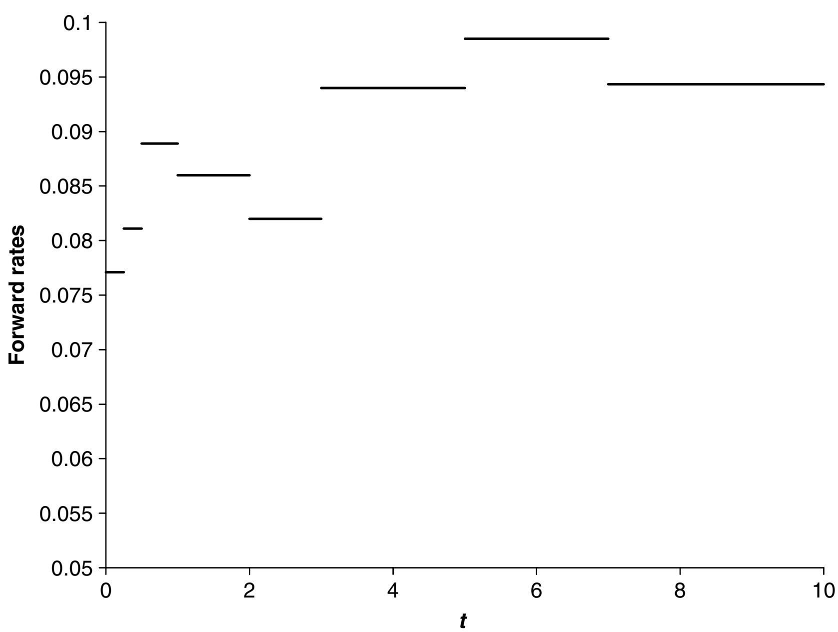


Figure 13.23 Forward rates.

13.17 INTERPOLATION

We have explicitly assumed in the previous section that the forward rates are piecewise constant, jumping from one value to the next across the maturity of each bond. Other methods of 'interpolation' are also possible. For example, the forward rate curve could be made continuous, with piecewise constant gradient. Some people like to use cubic splines. The correct way of 'joining the dots' (for there are only a finite number of market prices) has been the subject of much debate. If you want to know what rate to apply to a two-and-a-half-year cashflow and the nearest bonds are at two and three years then you will have to make some assumptions; there is no 'correct' value. Perhaps the best that can be done is to bound the rate.

13.18 SUMMARY

There are good and bad points about the interest rate model of this chapter. First, I mention the advantages.

Compare the simplicity of the mathematics in this chapter with that in previous chapters on option pricing. Clearly there is benefit in having models for which the analysis is so simple. Computation of many values and hedging can be performed virtually instantaneously on even slow computers. Moreover, it may be completely unnecessary to have a more complex model. For example, if we want to compare simple cashflows it may be possible to value one bond directly by summing other bonds, if their cashflows can be made to match. Such a situation, although uncommon, is market-independent modeling. Even if exact cashflow matches are not possible, there may be sufficiently close agreement for the differences to be estimated or at least bounded; large errors are easily avoided.

On the other hand, it is common experience that interest rates are unpredictable, random, and for complex products the movement of rates is the most important factor in their pricing. To assume that interest rates follow forward rates would be financial suicide in such cases. Think back to Jensen's inequality. There is therefore a need for models more closely related to the stochastic models we have seen in earlier chapters.

In this chapter we saw simple yet powerful ways to analyze simple fixed-income contracts. These methods are used very frequently in practice, far more frequently than the complex methods we later discuss for the pricing of interest rate derivatives. The assumptions underlying the techniques, such as deterministic forward rates, are only relevant to simple contracts. As we have seen in the options world, more complex products with non-linear payoffs, require a model that incorporates the stochastic nature of variables. Stochastic interest rates will be the subject of later chapters.

FURTHER READING

- The work of Macaulay (1938) on duration wasn't used much prior to the 1960s, but now it is considered fundamental to fixed income analysis.
- See Fabozzi (1996) for a discussion of yield, duration and convexity in greater detail. He explains how the ideas are extended to more complicated instruments.
- The argument about how to join the yield curve dots is as meaningless as the argument between the Little-Endians and Big-Endians of Swift (1726).
- See Walsh (2003) for issues concerning curve building.

CHAPTER 14
swaps

In this Chapter...

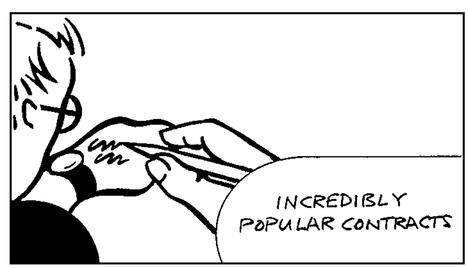
- the specifications of basic interest rate swap contracts
- the relationship between swaps and zero-coupon bonds
- exotic swaps

14.1 INTRODUCTION

A **swap** is an agreement between two parties to exchange, or swap, future cashflows. The size of these cashflows is determined by some formulae, decided upon at the initiation of the contract. The swaps may be in a single currency or involve the exchange of cashflows in different currencies.

The swaps market is big. The total notional principal amount is, in US dollars, currently comfortably in *14* figures. This market really began in 1981 although there were a small number of swap-like structures arranged in the 1970s. Initially the most popular contracts were currency swaps, discussed below, but very quickly they were overtaken by the interest rate swap.

14.2 THE VANILLA INTEREST RATE SWAP



In the **interest rate swap** the two parties exchange cashflows that are represented by the interest on a notional principal. Typically, one side agrees to pay the other a fixed interest rate and the cashflow in the opposite direction is a **floating rate**. The parties to a swap are shown schematically in Figure 14.1. One of the commonest floating rates used in a swap agreement is LIBOR, London Interbank Offer Rate.

Commonly in a swap, the exchange of the fixed and floating interest payments occur every six months. In this case the relevant LIBOR rate would be the six-month rate. At the maturity of the contract the principal is *not* exchanged.

Let me give an example of how such a contract works.

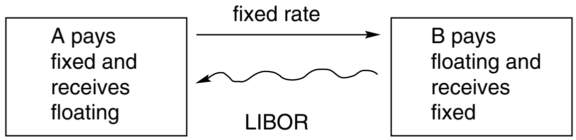


Figure 14.1 The parties to an interest rate swap.

Example

Suppose that we enter into a five-year swap on 4th August 2006, with semi-annual interest payments. We will pay to the other party a rate of interest fixed at 6% on a notional principal of \$100 million; the counterparty will pay us six-month LIBOR. The cashflows in this contract are shown in Figure 14.2. The straight lines denote a fixed rate of interest and thus a known amount; the curly lines are floating rate payments.

The first exchange of payments is made on 4th February 2007, six months after the deal is signed. How much money changes hands on that first date? We must pay $0.03 \times \$100{,}000{,}000 = \$3{,}000{,}000$. The cashflow in the opposite direction will be at six-month LIBOR, *as quoted six months previously* i.e. at the initiation of the contract. This is a very important point. The LIBOR rate is set six months before it is paid, so that in the first exchange of payments the floating side is known. This makes the first exchange special.

The second exchange takes place on 4th August 2007. Again we must pay \$3,000,000, but now we receive LIBOR, as quoted on 4th February 2007. Every six months there is an exchange of such payments, with the fixed leg always being known and the floating leg being known six months before it is paid. This continues until the last date, 4th August 2011.

Why is the floating leg set six months before it is paid? This 'minor' detail makes a large difference to the pricing of swaps, believe it or not. It is no coincidence that the time between payments is the same as the tenor of LIBOR that is used, six months in this example. This convention has grown up because of the meaning of LIBOR; it is the rate of interest on a fixed-term maturity, set now and paid at the end of the term. Each floating leg of the swap is like a single investment of the notional principal six months prior to the payment of the interest. Hold that thought, we return to this point in a couple of sections, to show the simple relationship between a swap and bonds.

There is also the **LIBOR in arrears swap** in which the LIBOR rate paid on the swap date is the six-month rate set that day, not the rate set six months before.

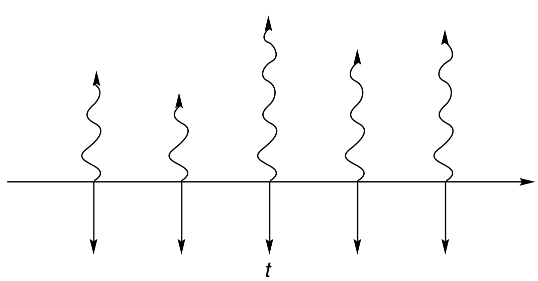


Figure 14.2 A schematic diagram of the cashflows in an interest rate swap.

14.3 COMPARATIVE ADVANTAGE

Swaps were first created to exploit **comparative advantage**. This is when two companies who want to borrow money are quoted fixed and floating rates such that by exchanging payments between themselves they benefit, at the same time benefitting the intermediary who puts the deal together. Here's an example.

Two companies A and B want to borrow \$50 MM, to be paid back in two years. They are quoted the interest rates for borrowing at fixed and floating rates shown in Table 14.1.

Note that both must pay a premium over LIBOR to cover risk of default, which is perceived to be greater for company B.

Ideally, company A wants to borrow at floating and B at fixed. If they each borrow directly then they pay the following in total (Table 14.2):

$$\text{six-month LIBOR} + 30\,\text{bps} + 8.2\% = \text{six-month LIBOR} + 8.5\%.$$

However, if A borrowed at fixed and B at floating they'd only be paying

$$\text{six-month LIBOR} + 100\,\text{bps} + 7\% = \text{six-month LIBOR} + 8\%.$$

That's a saving of 0.5%.

Let's suppose that A borrows fixed and B floating, even though that's not what they want. Their total interest payments are six-month LIBOR plus 8%. Now let's see what happens if we throw a swap into the pot.

* A is currently paying 7% and B six-month LIBOR plus 1%. They enter into a swap in which A pays LIBOR to B and B pays 6.95% to A. They have swapped interest payments.

Looked at from A's perspective they are paying 7% and LIBOR while receiving 6.95%, a net floating payment of LIBOR plus 5 bps. Not only is this floating, as A originally wanted, but it is 25 bps better than if they had borrowed directly at the floating rate. There's still another 25 bps missing, and, of course, B gets this. B pays LIBOR plus 100 bps and also 6.95% to A while receiving LIBOR from A. This nets out at 7.95%, which is fixed, as required, and 25 bps less than the original deal.

Where did I get the 6.95% from? Let's do the same calculation with 'x' instead of 6.95.

Go back to *. A is currently paying 7% and B six-month LIBOR plus 1%. They enter into a swap in which A pays LIBOR to B and B pays x% to A. They have swapped interest payments.

Looked at from A's perspective they are paying 7% and LIBOR while receiving 'x'%, a net floating payment of LIBOR plus $7 - x$%. Now we want A to benefit by 25 bps over the original

Table 14.1 Borrowing rates for companies A and B.

	Fixed	Floating
A	7%	six-month LIBOR + 30 bps
B	8.2%	six-month LIBOR + 100 bps

Table 14.2 Borrowing rates with no swap involved.

A	six-month LIBOR + 30 bps (floating)
B	8.2% (fixed)

deal, this is half the 50 bps advantage. (I've just unilaterally decided to divide the advantage equally, 25 bps each.) So...

$$\text{LIBOR} + 7 - x + 0.25 = \text{LIBOR} + 0.3,$$

i.e.

$$x = 6.95\%.$$

In practice the two counterparties would deal through an intermediary who would take a piece of the action.

Although comparative advantage was the original reason for the growth of the swaps market, it is no longer the reason for the popularity of swaps. Swaps are now very vanilla products existing in many maturities and more liquid than simple bonds.

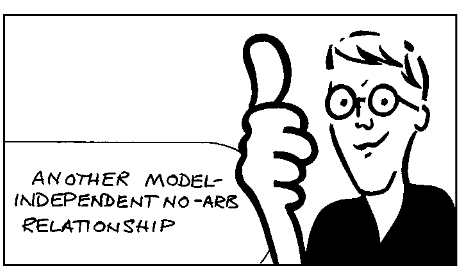
14.4 THE SWAP CURVE

When the swap is first entered into it is usual for the deal to have no value to either party. This is done by a careful choice of the fixed rate of interest. In other words, the 'present value,' let us say, of the fixed side and the floating side both have the same value, netting out to zero. Consider the two extreme scenarios, very high fixed leg and very low fixed leg. If the fixed leg is very high the receiver of fixed has a contract with a high value. If the fixed leg is low the receiver has a contract that is worth a negative amount. Somewhere in between is a value that makes the deal valueless. The fixed leg of the swap is chosen for this to be the case.

Such a statement throws up many questions: How is the fixed leg decided upon? Why should both parties agree that the deal is valueless?

There are two ways to look at this. One way is to observe that a swap can be decomposed into a portfolio of bonds (as we see shortly) and so its value is not open to question if we are given the yield curve. However, in practice the calculation goes the other way. The swaps market is so liquid, at so many maturities, that it is the prices of swaps that drive the prices of bonds. The fixed leg of a **par swap** (having no value) is determined by the market.

The rates of interest in the fixed leg of a swap are quoted at various maturities. These rates make up the **swap curve**, see Figure 14.3.



14.5 RELATIONSHIP BETWEEN SWAPS AND BONDS

There are two sides to a swap, the fixed-rate side and the floating-rate side. The fixed interest payments, since they are all known in terms of actual dollar amount, can be seen as the sum of zero-coupon bonds. If the fixed rate of interest is r_s and the time between payments is τ then the fixed payments add up to

$$r_s \tau \sum_{i=1}^{N} Z(t; T_i).$$

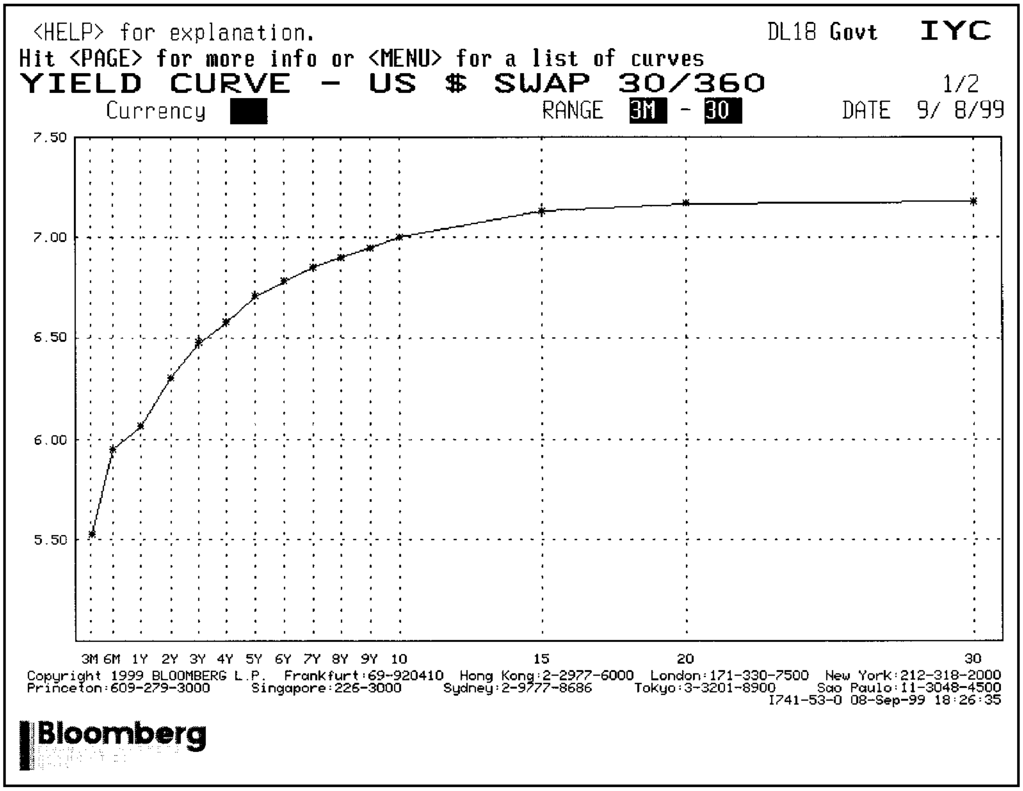


Figure 14.3 The swap curve. Source: Bloomberg L.P.

This is the value today, time t, of all the fixed-rate payments. Here there are N payments, one at each T_i, and τ is the time between the T_is. Of course, this is multiplied by the notional principal, but assume that we have scaled this to one.

To see the simple relationship between the floating leg and zero-coupon bonds I draw some schematic diagrams and compare the cashflows. A single floating leg payment is shown in Figure 14.4. At time T_i there is payment of $r_\tau \tau$ of the notional principal, where r_τ is the period τ rate of LIBOR, set at time $T_i - \tau$. I add and subtract \$1 at time T_i to get the second diagram. The first and the second diagrams obviously have the same present value. Now recall the precise definition of LIBOR. It is the interest rate paid on a fixed-term deposit. Thus the $\$1 + r_\tau \tau$ at time T_i is the same as \$1 at time $T_i - \tau$. This gives the third diagram. It follows that the single floating rate payment is equivalent to two zero-coupon bonds. A single floating leg of a swap at time T_i is *exactly* equal to a deposit of \$1 at time $T_i - \tau$ and a withdrawal of \$1 at time T_i.

Now add up all the floating legs as shown in Figure 14.5, note the cancellation of all \$1 (dashed) cashflows except for the first and last. This shows that the floating side of the swap has value

$$1 - Z(t; T_N).$$

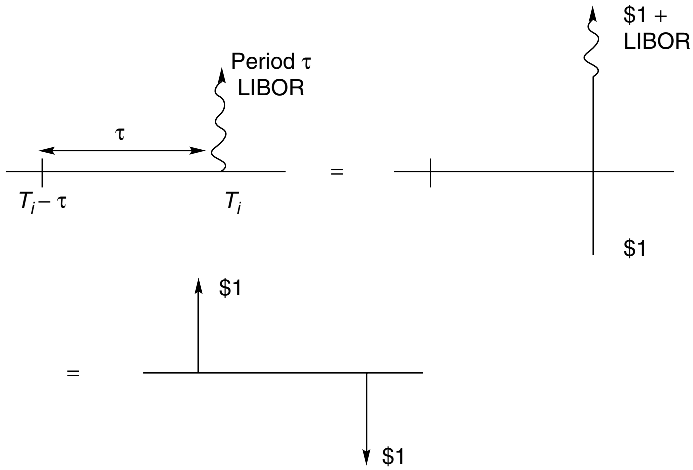


Figure 14.4 A schematic diagram of a single floating leg in an interest rate swap and equivalent portfolios.

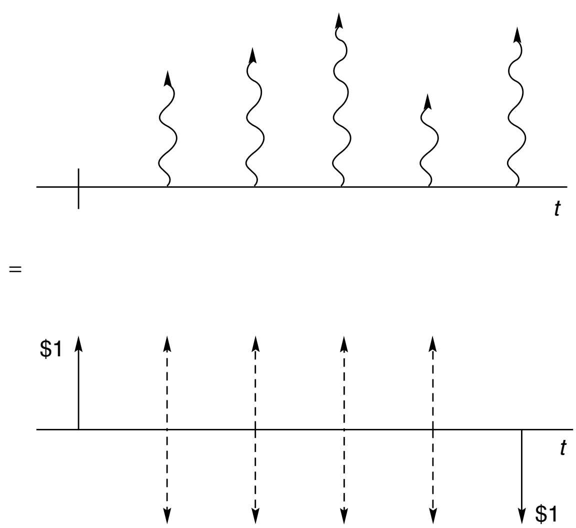


Figure 14.5 A schematic diagram of all the floating legs in a swap.

Bring the fixed and floating sides together to find that the value of the swap, to the receiver of the fixed side, is

$$r_s\tau\sum_{i=1}^{N}Z(t;T_i)-1+Z(t;T_N).$$

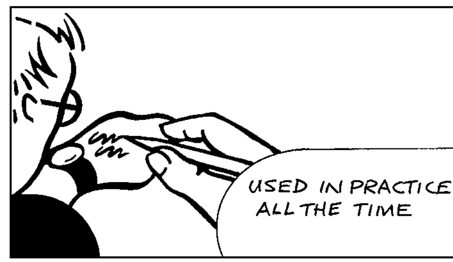
This result is *model independent*. This relationship is independent of any mathematical model for bonds or swaps.

At the start of the swap contract the rate r_s is usually chosen to give the contract par value, i.e. zero value initially. Thus

$$\boxed{r_s = \frac{1 - Z(t;T_N)}{\tau \sum_{i=1}^{N} Z(t;T_i)}.} \tag{14.1}$$

This is the quoted swap rate.

14.6 BOOTSTRAPPING



Swaps are now so liquid and exist for an enormous range of maturities that their prices determine the yield curve and not vice versa. In practice one is given $r_s(T_i)$ for many maturities T_i and one uses (14.1) to calculate the prices of zero-coupon bonds and thus the yield curve. For the first point on the discount-factor curve we must solve

$$r_s(T_1) = \frac{1 - Z(t;T_1)}{\tau Z(t;T_1)},$$

i.e.

$$Z(t;T_1) = \frac{1}{1 + r_s(T_1)\tau}.$$

After finding the first j discount factors the $j+1$th is found from

$$\boxed{Z(t;T_{j+1}) = \frac{1 - r_s(T_{j+1})\tau \sum_{i=1}^{j} Z(t;T_i)}{1 + r_s(T_{j+1})\tau}}$$

Figure 14.6 shows the forward curve derived from the data in Figure 14.3 by bootstrapping.

14.7 OTHER FEATURES OF SWAPS CONTRACTS

The above is a description of the vanilla interest rate swap. There any many features that can be added to the contract that make it more complicated, and most importantly, model dependent. A few of these features are mentioned here.

Callable and puttable swaps

A **callable** or **puttable swap** allows one side or the other to close out the swap at some time before its natural maturity. If you are receiving fixed and the floating rate rises more than you had expected you would want to close the position. Mathematically we are in the early exercise world of American-style options. The problem is model dependent and is discussed in Chapter 32.

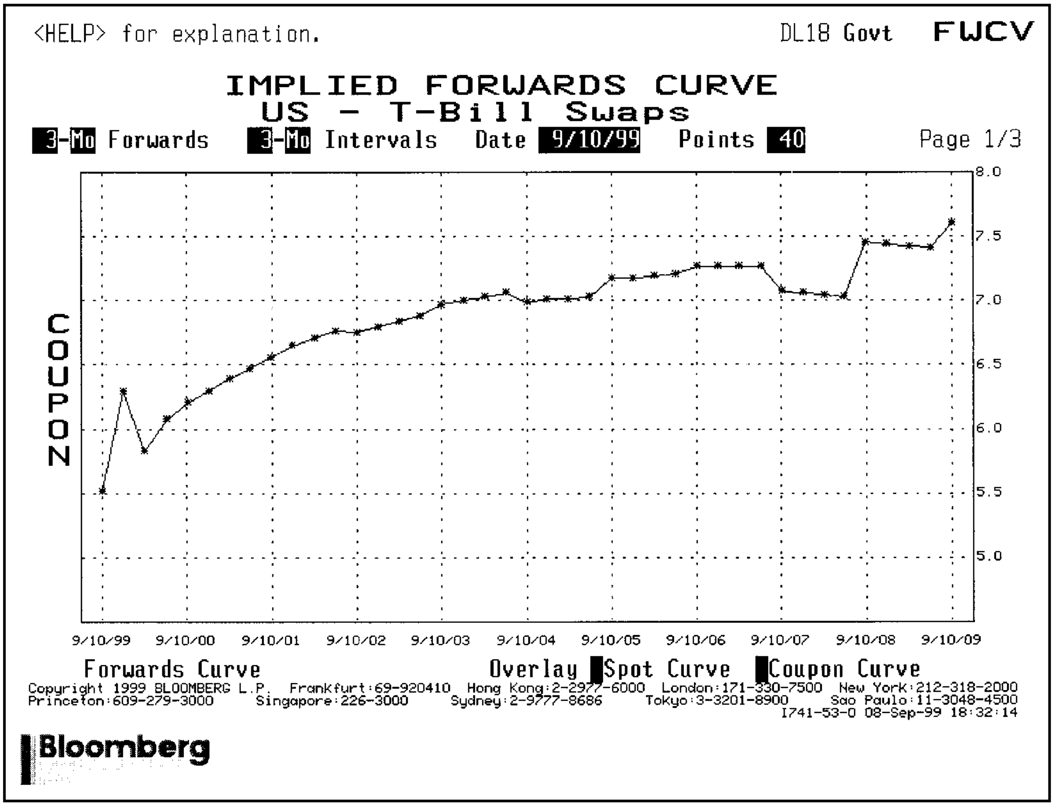


Figure 14.6 Forward rates derived from the swap curve by bootstrapping. Source: Bloomberg L.P.

Extendible swaps

The holder of an **extendible swap** can extend the maturity of a vanilla swap at the original swap rate.

Index amortizing rate swaps

The principal in the vanilla swap is constant. In some swaps the principal declines with time according to a prescribed schedule. The index amortizing rate swap is more complicated still with the amortization depending on the level of some index, say LIBOR, at the time of the exchange of payments. We will see this contract in detail in Chapter 32.

14.8 OTHER TYPES OF SWAP

14.8.1 Basis Rate Swap

In the **basis rate swap** the floating legs of the swap are defined in terms of two distinct interest rates. For example, the prime rate versus LIBOR. A bank may have outstanding loans based on this prime rate but itself may have to borrow at LIBOR. It is thus exposed to **basis risk** and this can be reduced with a suitable basis rate swap.

14.8.2 Equity Swaps

The basic **equity swap** is an agreement to exchange two payments, one being an agreed interest rate (either fixed or floating) and the other depending on an equity index. This equity component is usually measured by the total return on an index, both capital gains and dividend are included. The principal is not exchanged.

The **equity basis swap** is an exchange of payments based on *two* different indices.

14.8.3 Currency Swaps

A **currency swap** is an exchange of interest payments in one currency for payments in another currency. The interest rates can both be fixed, both floating or one of each. As well as the exchange of interest payments there is also an exchange of the principals (in two different currencies) at the beginning of the contract and at the end.

To value the fixed-to-fixed currency swap we need to calculate the present values of the cashflows in each currency. This is easily done, requiring the discount factors for the two currencies. Once this is done we can convert one present value to the other currency using the current *spot* exchange rate. If floating interest payments are involved we first decompose them into a portfolio of bonds (if possible) and value similarly.

14.9 SUMMARY

The need and ability to be able to exchange one type of interest payment for another is fundamental to the running of many businesses. This has put swaps among the most liquid of financial contracts. This enormous liquidity makes swaps such an important product that one has to be very careful in their pricing. In fact, swaps are so liquid that you do not price them in any theoretical way; to do so would be highly dangerous. Instead they are almost treated like an 'underlying' asset. From the market's view of the value we can back out, for example, the yield curve. We are helped in this by the fine detail of the swaps structure; the cashflows are precisely defined in a way that makes them exactly decomposable into zero-coupon bonds. And this can be done in a completely model-independent way. To finish this chapter I want to stress the importance of not using a model when a set of cashflows can be perfectly, statically and model-independently, hedged by other cashflows. Any mispricing, via a model, no matter how small could expose you to large and risk-free losses.

FURTHER READING

- Two good technical books on swaps are by Das (1994) and Miron & Swannell (1991).
- The pocketbook by Ungar (1996) describes the purpose of the swaps market, how it works and the different types of swaps, with no mathematics.

CHAPTER 15
the binomial model

In this Chapter...

- a simple model for an asset price random walk
- delta hedging
- no arbitrage
- the basics of the binomial method for valuing options
- risk neutrality

15.1 INTRODUCTION

We have seen in Chapter 3 a model for equities and other assets that is based on the mathematical theory of stochastic calculus. There is another, equally popular, approach that leads to the same partial differential equation, the Black–Scholes equation, in a way that some people find more 'accessible,' which can be made equally 'rigorous.' This approach, via the **binomial model** for equities, is the subject of this chapter.

Undoubtedly, one of the reasons for the popularity of this model is that it can be implemented without any higher mathematics (such as differential calculus) and there is actually no need to derive a partial differential equation before this implementation. This is a positive point, however the downside is that it is harder to attain greater levels of sophistication or numerical analysis in this setting.

Before I describe this model I want to stress that the binomial model may be thought of as being either a genuine *model* for the behavior of equities, or, alternatively, as a numerical method for the solution of the Black–Scholes equation.[1] Most importantly, we see the ideas of delta hedging, risk elimination and risk-neutral valuation occurring in another setting.

The binomial model is very important because it shows how to get away from a reliance on closed-form solutions. Indeed, it is extremely important to have a way of valuing options that only relies on a simple model and fast, accurate numerical methods. Often in real life, a contract may contain features that make analytic solution very hard or impossible. Some of these features may be just a minor modification to some other, easily-priced, contract but even minor changes to a contract can have important effects on the value and especially on the method of solution. The classic example is the American put. Early exercise may seem to be a small change to a contract but the difference between the values of a European and an American

[1] In this case, it is very similar to an explicit finite-difference method, of which more later.

put can be large and certainly there is no simple closed-form solution for the American option and its value must be found numerically.

Having said what is good about the binomial I would like to say what, to my mind, is bad about it.

First, as a model of stock price behavior it is poor. The binomial model says that the stock can either go up by a known amount or down by a known amount; there are but two possible stock prices 'tomorrow.' This is clearly unrealistic. This is important because from this model follow certain results that hinge entirely on there only being two prices for the stock tomorrow. Introduce a third state and the results collapse.

Second, as a numerical scheme it is prehistoric compared with modern numerical methods. We go into these numerical methods in some detail here but several volumes could be written on sophisticated numerical methods alone. I would advise the reader to study the binomial model for the intuition it gives, but do not rely on it for numerical calculations.

The intuition that one gets from the binomial method *is* useful. Indeed, it is said that the binomial model even helps MBA students understand options. I don't believe in such dumbing down. I don't think that quantitative finance should be dumbed down, just like I don't believe that brain surgery should be dumbed down.

My advice is that once you have become comfortable with the ideas that come out of this chapter you should relegate the binomial method to the back of your mind.

15.2 EQUITIES CAN GO DOWN AS WELL AS UP

The most 'accessible' approach to option pricing is the **binomial model**. This requires only basic arithmetic and no complicated stochastic calculus. In this model we will see the ideas of hedging and no arbitrage used. The end result is a simple algorithm for determining the correct value for an option.

We are going to examine a very simple model for the behavior of a stock, and based on this model see how to value options.

- We will have a stock, and a call option on that stock expiring tomorrow.
- The stock can either rise or fall by a known amount between today and tomorrow.
- Interest rates are zero.

Figure 15.1 gives an example. The stock is currently worth \$100 and can rise to \$101 or fall to \$99 between today and tomorrow.

Which of the two prices is realized tomorrow is completely random. There is a certain probability of the stock rising and one minus that probability of the stock falling. In this example the probability of a rise to 101 is 0.6, so that the probability of falling to 99 is 0.4 (see Figure 15.2).

Now let's introduce the call option on the stock. This call option has a strike of \$100 and expires tomorrow.

If the stock price rises to 101, what will then be the option's payoff? See Figure 15.3. It is just $101 - 100 = 1$.

And if the stock falls to 99 tomorrow, what is then the payoff? See Figure 15.4. The answer is zero, the option has expired out of the money.

If the stock rises the option is worth \$1, and if it falls it is worth \$0 (Figure 15.5). There is a 0.6 probability of getting \$1 and 0.4 probability of getting zero. Interest rates are zero ...

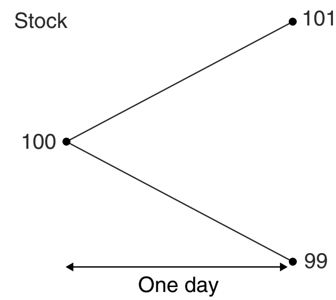


Figure 15.1 The stock can rise or fall over the next day; only two future prices are possible.

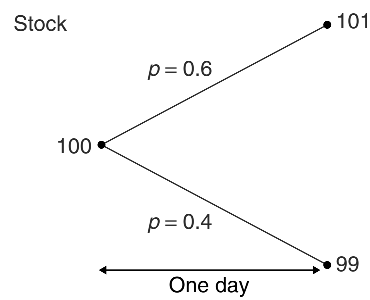


Figure 15.2 Probabilities associated with the future stock prices.

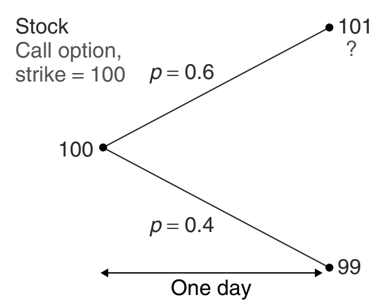


Figure 15.3 What is the option payoff if the stock rises?

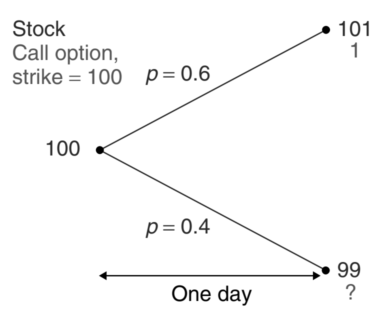


Figure 15.4 What is the option payoff if the stock falls?

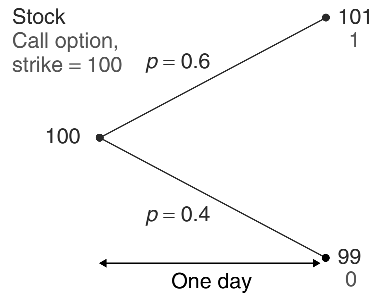


Figure 15.5 Now we know the option values in both 'states of the world.'

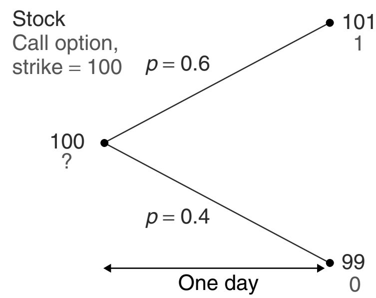


Figure 15.6 What is the option worth today?

What is the option worth today (Figure 15.6)?

No, the answer is *not* 0.6. If that is what you thought, based on calculating simple expectations then I successfully 'led you up the garden path' to the wrong answer.

15.3 THE OPTION VALUE

The correct answer is...

$$\frac{1}{2}.$$

Why?

To see how this can be the only correct answer we must first construct a *portfolio* consisting of one option and short $\frac{1}{2}$ of the underlying stock. This portfolio is shown in Figure 15.7.

If the stock rises to 101 then this portfolio is worth

$$1 - \tfrac{1}{2} \times 101;$$

the one being from the option payoff and the $-\frac{1}{2} \times 101$ being from a short ($-$) position ($\frac{1}{2}$) in the stock (now worth 101).

If the stock falls to 99 then this portfolio is worth

$$0 - \tfrac{1}{2} \times 99;$$

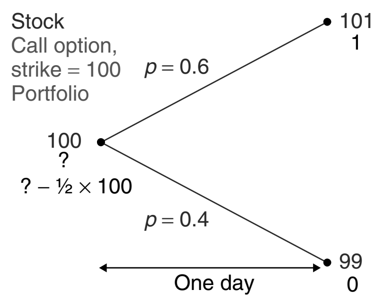


Figure 15.7 Long one option, short half of the stock.

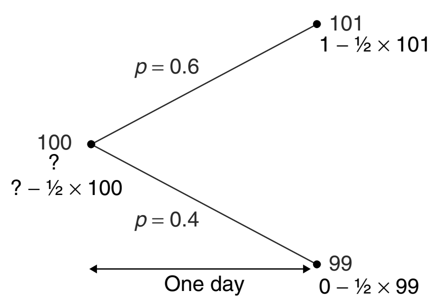


Figure 15.8 The portfolio values at expiration.

the zero being from the option payoff and the $-\frac{1}{2} \times 99$ being from a short ($-$) position ($\frac{1}{2}$) in the stock (now worth 99) (see Figure 15.8).

In either case, tomorrow, at expiration, the portfolio takes the value

$$-\frac{99}{2}$$

and that is regardless of whether the stock rises or falls.

We have constructed a perfectly risk-free portfolio.

If the portfolio is worth $-99/2$ tomorrow, and interest rates are zero, how much is this portfolio worth today?

It must also be worth $-99/2$ today.

This is an example of **no arbitrage**: There are two ways to ensure that we have $-99/2$ tomorrow.

1. Buy one option and sell one half of the stock.
2. Put the money under the mattress.

Both of these 'portfolios' must be worth the same today. Therefore, using '?' as in the figures to represent the unknown option value

$$? - \tfrac{1}{2} \times 100 = \text{the option value} - \tfrac{1}{2} \times 100 = -\tfrac{1}{2} \times 99$$

and so

$$? = \text{the option value} = \frac{1}{2}.$$

15.4 WHICH PART OF OUR 'MODEL' DIDN'T WE NEED?

The value of an option does not depend on the probability of the stock rising or falling. This is equivalent to saying that the stock growth rate is irrelevant for option pricing. This is because we have **hedged** the option with the stock, see Figure 15.9. We do not care whether the stock rises or falls, see Figure 15.10.

We *do* care about the stock price range, however. The stock volatility is very important in the valuation of options.

Three questions follow from the above simple argument:

- Why should this 'theoretical price' be the 'market price'?
- How did I know to sell $\frac{1}{2}$ of the stock for hedging?
- How does this change if interest rates are non-zero?

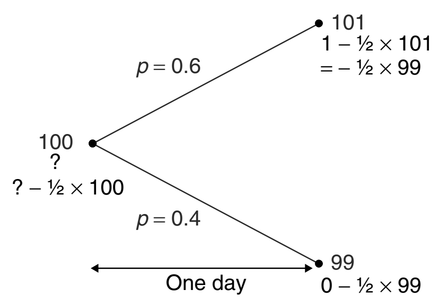


Figure 15.9 The portfolio is 'hedged.'

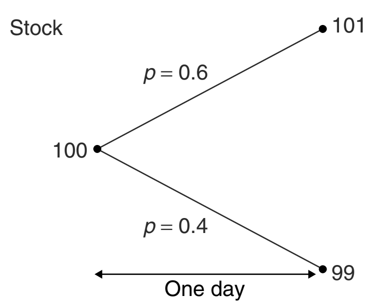


Figure 15.10 Which parameter(s) didn't we need?

15.5 WHY SHOULD THIS 'THEORETICAL PRICE' BE THE 'MARKET PRICE'?

This one is simple, because if it's not, then there is risk-free money to be made. If the option costs less than 0.5 simply buy it and hedge to make a profit. If it is worth more than 0.5 in the market then sell it and hedge, and make a guaranteed profit.

It's not quite this simple because it is possible for arbitrage opportunities to exist, and to exist for a long time. We also really need for there to be some practical mechanism for arbitrage to be 'removed.' In practice this means we really need there to be a couple of agents perhaps undercutting each other in such a way that the arb opportunity disappears: A sells the option for 0.55, gets all the business and makes a guaranteed 0.05 profit. Along comes B who sells the option for just 0.53; now he takes away all the business from A, who responds by dropping his price to 0.52 etc. So really, supply and demand should act to make the option price converge to the 0.5.

15.5.1 The Role of Expectations

The expected payoff is definitely 0.6 for this option. It's just that this has nothing to do with the option's value. Let's take a quick look at the role of this expectation.

Would anyone pay 0.6 or more for the option? No, unless they were **risk seeking**.

Would anyone pay 0.55? Perhaps, if they liked the idea of an expected return of

$$\frac{0.6 - 0.55}{0.55} \approx 9\%.$$

(But they would be better off replicating the option payoff with the stock and cash.) The person writing the option would be very pleased with the guaranteed profit of 0.05.

15.6 HOW DID I KNOW TO SELL $\frac{1}{2}$ OF THE STOCK FOR HEDGING?

Introduce a symbol! Use Δ to denote the quantity of stock that must be sold for hedging. We start off with one option, $-\Delta$ of the stock, giving a portfolio value of

$$? - \Delta \times 100.$$

Tomorrow the portfolio is worth

$$1 - \Delta \times 101$$

if the stock rises, or

$$0 - \Delta \times 99$$

if it falls.

The key step is the next one; make these two equal to each other:

$$1 - \Delta \times 101 = 0 - \Delta \times 99.$$

Therefore

$$\Delta(101 - 99) = 1$$
$$\Delta = 0.5.$$

Another example

Stock price is 100, and can rise to 103 or fall to 98. Value a call option with a strike price of 100. Interest rates are zero.

Again use Δ to denote the quantity of stock that must be sold for hedging.

The portfolio value is

$$? - \Delta \times 100.$$

Tomorrow the portfolio is worth either

$$3 - \Delta 103$$

or

$$0 - \Delta 98.$$

So we must make

$$3 - \Delta 103 = 0 - \Delta 98.$$

That is,

$$\Delta = \frac{3 - 0}{103 - 98} = \frac{3}{5} = 0.6.$$

The portfolio value tomorrow is then

$$-0.6 \times 98.$$

With zero interest rate, the portfolio value today must equal the risk-free portfolio value tomorrow:

$$? - 0.6 \times 100 = -0.6 \times 98.$$

Therefore the option value is 1.2.

15.6.1 The General Formula for Δ

Delta hedging means choosing Δ such that the portfolio value does not depend on the direction of the stock.

When we generalize this (using symbols instead of numbers later on) we will find that

$$\Delta = \frac{\text{Range of option payoffs}}{\text{Range of stock prices}}.$$

We can think of Δ as the sensitivity of the option to changes in the stock.

15.7 HOW DOES THIS CHANGE IF INTEREST RATES ARE NON-ZERO?

Simple. We delta hedge as before to construct a risk-free portfolio. (Exactly the same delta.) Then we present value that back in time, by multiplying by a discount factor.

Example

Same as the first example, but now $r = 0.1$.

The discount factor for going back one day is

$$\frac{1}{1 + 0.1/252} = 0.9996.$$

The portfolio value today must be the *present value* of the portfolio value tomorrow

$$? - 0.5 \times 100 = -0.5 \times 99 \times 0.9996.$$

So that

$$? = 0.51963.$$

15.8 IS THE STOCK ITSELF CORRECTLY PRICED?

Earlier, I tried to trick you into pricing the option by looking at the expected payoff. Suppose, for the sake of argument that I had been successful in this. I would then have asked you what was the expected stock price tomorrow, forgetting the option (see Figure 15.11).

The expected stock value tomorrow is

$$0.6 \times 101 + 0.4 \times 99 = 100.2.$$

In an expectation's sense, the stock itself seems incorrectly priced. Shouldn't it be valued at 100.2 today? Well, we already kind of know that expectations aren't the way to price options. But we can go further than that, and make some positive statements.

We pay less than the future expected value because the stock is risky. We want a positive expected return to compensate for the risk. This is an idea we will be seeing in detail later on, in Chapter 18 on portfolio management.

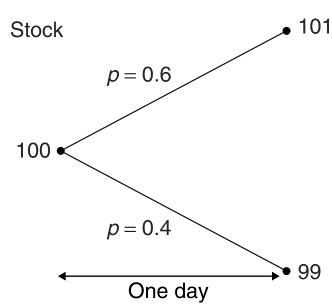


Figure 15.11 What is the expected stock price?

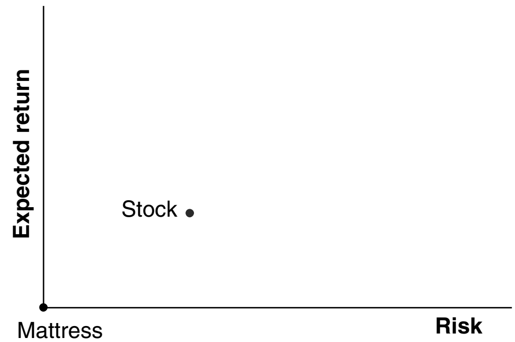


Figure 15.12 Risk and return for the stock and the risk-free investment (putting the money under the mattress).

We can plot the stock (and all investments) on our risk/return diagram (see Figure 15.12). Risk is measured by standard deviation and return is the expected return. The figure shows two investments, the stock and, at the origin, the bank investment. The bank investment has zero risk, and in our above examples, has zero expected return (that's why in the figure I've referred to it as putting money under the mattress).

We will see in Chapter 18 that we can get to other places in the risk/return space by dividing our money between several investments. In the present case, if we put half our money under the mattress and half in the stock we will find ourselves with an investment that is exactly half way between the two dots in the figure. We can get to any point on the straight line between the risk-free dot and the stock dot by splitting our money between these two, we can even get to any place on the extrapolated straight line by borrowing money at the risk-free rate to invest in the stock.

15.9 COMPLETE MARKETS

The option also has an expected return and a risk. In our example the expected return and the risk for the option are both much, much greater than for the stock. We can plot the option on the same risk/return diagram. Where do you think it might be? Above the extrapolated line, on it, or below it?

It turns out that the option lies *on* the straight line (see Figure 15.13). This means that we can 'replicate' an option's risk and return characteristic with stock and the risk-free investment. Option payoffs can be **replicated** by stocks and cash. Any two points on the straight lines can be used to get us to any other point. So, we can get a risk-free investment using the option and the stock, and this is hedging. And, the stock can be replicated by cash and the option.

A conclusion of this analysis is that options are redundant in this 'world,' i.e. in this model. We say that **markets are complete**. The practical implication of complete markets is that options are hedgeable and therefore can be priced without any need to know probabilities. We can hedge an option with stock to 'replicate' a risk-free investment (Figure 15.14) and we can replicate an option using stock and a risk-free investment (Figures 15.15 and 15.16). We could also, of course, replicate stock with an option and risk-free.

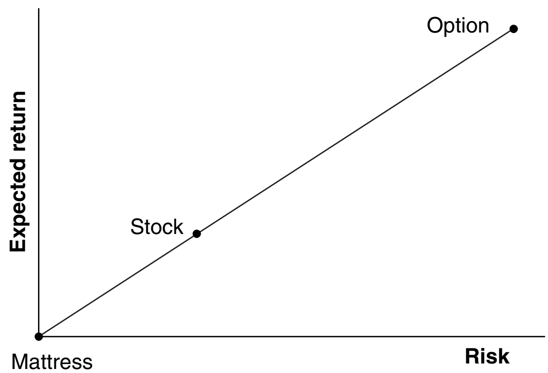


Figure 15.13 Now we have three investments, including the option.

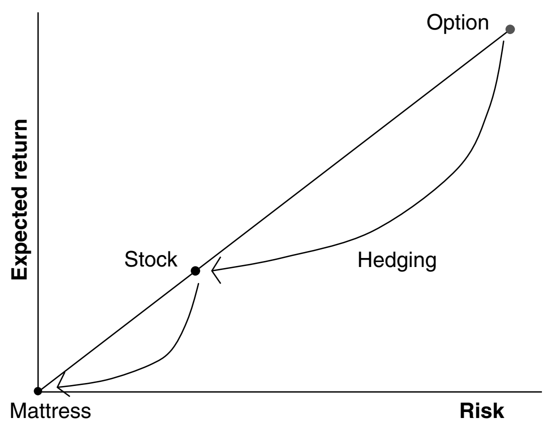


Figure 15.14 Hedging.

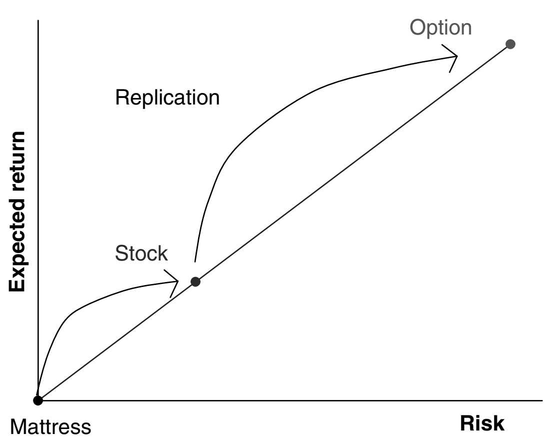


Figure 15.15 Replication.

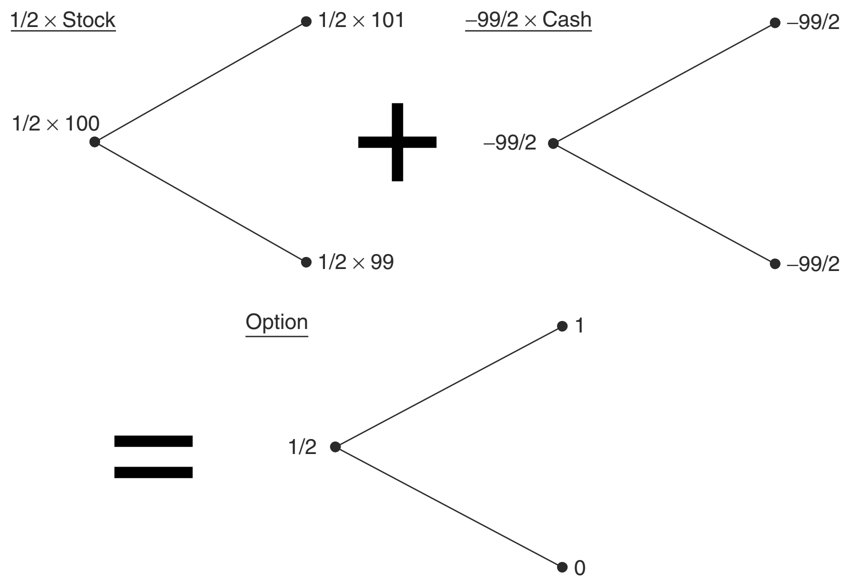
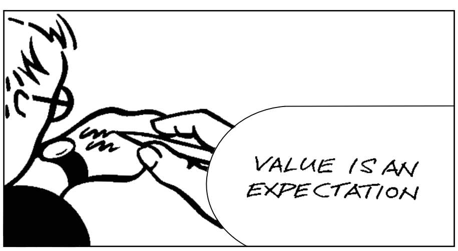


Figure 15.16 $1/2 \times$ Stock $-$ $99/2 \times$ Cash $=$ Option.



15.10 THE REAL AND RISK-NEUTRAL WORLDS

In our world, the **real world**, we have used our statistical skills to estimate the future possible stock prices (99 and 101) and the probabilities of reaching them (0.4 and 0.6).

Some properties of the real world are listed below.

- We know all about delta hedging and risk elimination.
- We are very sensitive to risk, and expect greater return for taking risk.
- It turns out that only the two stock prices matter for option pricing, not the probabilities.

People often refer to the **risk-neutral world** in which people don't care about risk. The risk-neutral world has the following characteristics:

- We don't care about risk, and don't expect any extra return for taking unnecessary risk.
- We don't ever need statistics for estimating probabilities of events happening.
- We believe that everything is priced using simple expectations.

Imagine yourself in the risk-neutral world, looking at the stock price model. Suppose all you know is that the stock is currently worth \$100 and could rise to \$101 or fall to \$99.

If the stock is correctly priced today, using simple expectations, what would you deduce to be the probabilities of the stock price rising or falling (see Figure 15.17)?

The symmetry makes the answer to this rather obvious. If the stock is correctly priced using real expectations then the probabilities ought to be 50% chance of a rise and 50% chance of a fall. The calculation we have just performed goes as follows ...

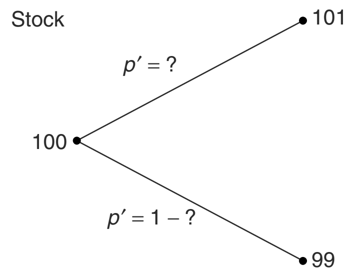


Figure 15.17 What is the probability of the stock rising?

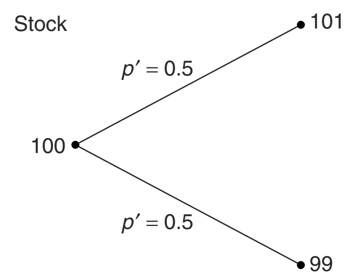


Figure 15.18 Risk-neutral probabilities.

On the risk-neutral planet they calculate **risk-neutral probabilities** p' (see Figure 15.18) from the equation

$$p' \times 101 + (1 - p') \times 99 = 100.$$

From which $p' = 0.5$.

Do not think that this p' is in any sense real. No, the real probabilities are still 60% and 40%. This calculation assumes something that is fundamentally wrong, that simple expectations are used for pricing.

Never mind, let's stay with this risk-neutral world and see what they think the option value is. We won't tell them yet that the calculation they have just done is 'wrong.'

How would they then value the call option? Since they reckon the probabilities to be 50–50 and they use simple expectations to calculate values with no regard to risk then they would price the option using the expected payoff with their probabilities i.e.

$$0.5 \times 1 + 0.5 \times 0 = 0.5.$$

See Figure 15.19. This is called the **risk-neutral expectation**.

Damn and blast! They have found the correct answer for the wrong reasons! To put it in a nutshell, they have twice used their basic assumption of pricing via simple expectations to get to the correct answer. Two wrongs in this case do make a right.

And this technique will always work.

In the risk-neutral world they have exactly the same price for the option (but for different reasons).

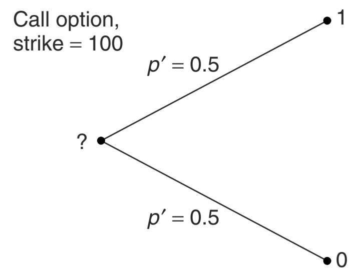


Figure 15.19 Pricing the option.

15.10.1 Non-zero Interest Rates

When interest rates are non-zero we must perform exactly the same operations, but whenever we equate values at different times we must allow for present valuing.

With $r = 0.1$ we calculate the risk-neutral probabilities from

$$0.9996 \times \left(p' \times 101 + (1 - p') \times 99\right) = 100.$$

So

$$p' = 0.51984.$$

The expected option payoff is now

$$0.51984 \times 1 + (1 - 0.51984) \times 0 = 0.51984.$$

And the present value of this is

$$0.9996 \times 0.51984 = 0.51963.$$

And this must be the option value. (It is the same as we derived the 'other' way.)

Risk-neutral pricing is a very powerful technique, and we will be seeing a lot more of it. Just remember one thing for the moment, that the risk-neutral probability p' that we have just calculated (the 0.5 in the first example) is not real, it does not exist, it is a mathematical construct. The real probability of the stock price was always in our example 0.6, it's just that this never was used in our calculations.

15.11 AND NOW USING SYMBOLS

In the binomial model we assume that the asset, which initially has the value S, can, during a time step δt, either

- rise to a value $u \times S$ or
- fall to a value $v \times S$,

with $0 < v < 1 < u$ (see Figure 15.20).

- The probability of a rise is p and so the probability of a fall is $1 - p$.

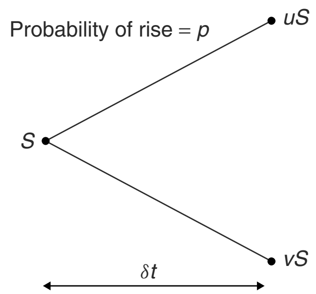


Figure 15.20 The model, using symbols.

Note: By *multiplying* the asset price by constants rather than *adding* constants, we will later be able to build up a whole tree of prices. This will be a discrete-time version of a lognormal random walk again.

- The three constants u, v and p are chosen to give the binomial walk the same drift and standard deviation as the asset we are trying to model.

This choice is far from unique. We have three parameters to choose, u, v and p, but only two statistical quantities to fit, μ and σ. This can be done in an infinite number of ways. The way I describe here is the best for teaching purposes.

For example,

$$u = 1 + \sigma\sqrt{\delta t},$$
$$v = 1 - \sigma\sqrt{\delta t}$$

and

$$p = \frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma}.$$

Let's check that these work.

15.11.1 Average Asset Change

The expected asset price after one time step is

$$puS + (1-p)vS = \left(\frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(1 + \sigma\sqrt{\delta t}\right)S + \left(\frac{1}{2} - \frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(1 - \sigma\sqrt{\delta t}\right)S$$
$$= (1 + \mu\,\delta t)S.$$

So the expected change in the asset is $\mu S\,\delta t$.

- The expected **return** is $\mu\,\delta t$

Correct.

15.11.2 Standard Deviation of Asset Price Change

The variance of change in asset price is

$$\begin{aligned} &S^2\left(p(u-1-\mu\,\delta t)^2+(1-p)(v-1-\mu\,\delta t)^2\right)\\ &= S^2\left(\left(\frac{1}{2}+\frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(\sigma\sqrt{\delta t}-\mu\,\delta t\right)^2+\left(\frac{1}{2}-\frac{\mu\sqrt{\delta t}}{2\sigma}\right)\left(\sigma\sqrt{\delta t}+\mu\,\delta t\right)^2\right)\\ &= S^2(\sigma^2\delta t-\mu^2\delta t^2). \end{aligned}$$

The standard deviation of asset changes is (approximately) $S\sigma\sqrt{\delta t}$.

- The standard deviation of returns is (approximately) $\sigma\sqrt{\delta t}$

Correct (ish).

15.12 AN EQUATION FOR THE VALUE OF AN OPTION

Suppose that we know the value of the option at the time $t+\delta t$. For example, this time may be the expiration of the option, say.

Now construct a portfolio at time t consisting of one option and a short position in a quantity Δ of the underlying. At time t this portfolio has value

$$\Pi = V - \Delta S,$$

where the option value V is for the moment unknown. You'll recognize this as exactly what we did before, but now we're using symbols instead of numbers.

At time $t+\delta t$ the option takes one of two values, depending on whether the asset rises or falls

$$V^+ \quad \text{or} \quad V^-.$$

At the same time the portfolio becomes either

$$V^+ - \Delta u S \quad \text{or} \quad V^- - \Delta v S.$$

Since we know V^+, V^-, u, v and S the values of both of these expressions are just linear functions of Δ.

15.12.1 Hedging

Having the freedom to choose Δ, we can make the value of this portfolio the same whether the asset rises or falls. This is ensured if we make

$$V^+ - \Delta u S = V^- - \Delta v S.$$

This means that we should choose

$$\Delta = \frac{V^+ - V^-}{(u-v)S} \tag{15.1}$$

for hedging.

The portfolio value is then

$$V^+ - \Delta u S = V^+ - \frac{u(V^+ - V^-)}{(u-v)}$$

if the stock rises or

$$V^- - \Delta v S = V^- - \frac{v(V^+ - V^-)}{(u-v)}$$

if it falls.

And, of course these two expressions are the same.

Let's denote this portfolio value by

$$\Pi + \delta\Pi.$$

This just means the original portfolio value plus the change in value.

15.12.2 No Arbitrage

Since the value of the portfolio has been guaranteed, we can say that its value must coincide with the value of the original portfolio plus any interest earned at the risk-free rate; this is the no-arbitrage argument.

Thus

$$\delta\Pi = r\Pi\,\delta t.$$

Putting everything together we get

$$\Pi + \delta\Pi = \Pi + r\Pi\,\delta t = \Pi(1 + r\,\delta t) = V^+ - \frac{u(V^+ - V^-)}{(u-v)}$$

with

$$\Pi = V - \Delta S = V - \frac{V^+ - V^-}{(u-v)S}S = V - \frac{V^+ - V^-}{(u-v)}$$

And the end result is

$$(1 + r\,\delta t)\left(V - \frac{V^+ - V^-}{(u-v)}\right) = V^- - \frac{v(V^+ - V^-)}{(u-v)}.$$

Rearranging as an equation for V we get

$$(1 + r\,\delta t)V = (1 + r\,\delta t)\frac{V^+ - V^-}{u-v} + \frac{uV^- - vV^+}{(u-v)}.$$

This is an equation for V given V^+, and V^-, the option values at the next time step, and the parameters u and v describing the random walk of the asset.

But it can be written more elegantly.

This equation can also be written as

$$(1 + r\delta t)V = p'V^+ + (1 - p')V^-, \tag{15.2}$$

where

$$p' = \frac{1}{2} + \frac{r\sqrt{\delta t}}{2\sigma}. \tag{15.3}$$

The left-hand side of Equation (15.2) is the future value of today's option value.

The right-hand side of Equation (15.2) is just like an expectation; it's the sum of probabilities multiplied by events.

If only the expression contained p, the real probability of a stock rise, then this expression would be the expected value at the next time step.

We see that the probability of a rise or fall is irrelevant as far as option pricing is concerned since p did not appear in Equation (15.2). But what if we interpret p' as a probability? Then we could 'say' that the option price is the present value of an expectation. But not the real expectation.

We are back with risk-neutral expectations again.

Let's compare the expression for p' with the expression for the actual probability p:

$$p' = \frac{1}{2} + \frac{r\sqrt{\delta t}}{2\sigma}$$

but

$$p = \frac{1}{2} + \frac{\mu\sqrt{\delta t}}{2\sigma}.$$

The two expressions differ in that where one has the interest rate r the other has the drift μ, but otherwise they are the same. Strange.

- We call p' the **risk-neutral probability**. It's like the real probability, but the real probability if the drift rate were r instead of μ.

Observe that the risk-free interest rate plays two roles in option valuation. It's used once for discounting to give present value, and it's used as the drift rate in the asset price random walk.

15.13 WHERE DID THE PROBABILITY *p* GO?

What happened to the probability p and the drift rate μ?

Interpreting p' as a probability, (15.2) is the statement that

- the option value at any time is the present value of the risk-neutral expected value at any later time.

In reading books or research papers on mathematical finance you will often encounter the expression 'risk-neutral' this or that, including the expression risk-neutral probability. You can think of an option value as being the present value of an expectation, only it's not the real expectation.

Don't worry, we'll come back to this several more times until you get the hang of it.

15.14 COUNTERINTUITIVE?

- Two stocks A and B.
- Both have same value, same volatility and are denominated in the same currency.
- Both have call options with the same strike and expiration.
- Stock A is doubling in value every year, stock B is halving.

Therefore both call options have the same value. But which would you buy? It can be difficult to accept initially that option prices don't depend on the direction that the stock is going.

15.15 THE BINOMIAL TREE

The binomial model, just introduced, allows the stock to move up or down a prescribed amount over the next time step. If the stock starts out with value S then it will take either the value uS or vS after the next time step. We can extend the random walk to the next time step. After two time steps the asset will be at either u^2S, if there were two up moves, uvS, if an up was followed by a down or vice versa, or v^2S, if there were two consecutive down moves. After three time steps the asset can be at u^3S, u^2vS, etc. One can imagine extending this random walk out all the way until expiry. The resulting structure looks like Figure 15.21 where the nodes represent the values taken by the asset. This structure is called the **binomial tree**. Observe how the tree bends due to the geometric nature of the asset growth. Often this tree is drawn as in Figure 15.22 because it is easier to draw, but this doesn't quite capture the correct structure.

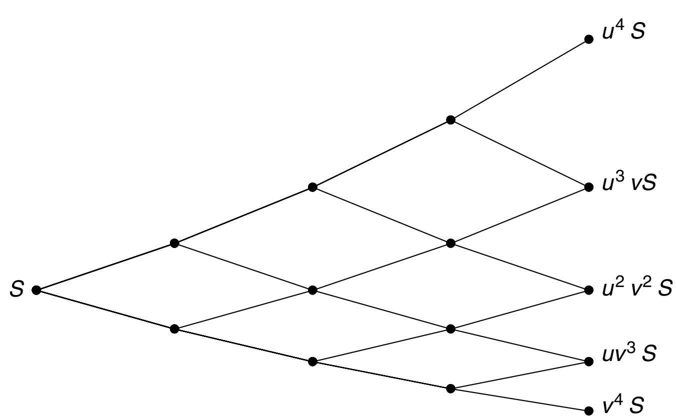


Figure 15.21 The binomial tree.

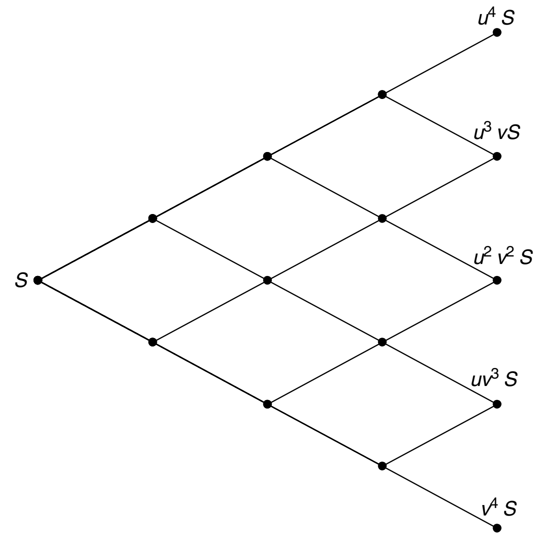


Figure 15.22 The binomial tree: a schematic version.

The top and bottom branches of the tree at expiry can only be reached by one path each, either all up or all down moves, whereas there will be several paths possible for each of the intermediate values at expiry. Therefore the intermediate values are more likely to be reached than the end values if one were doing a simulation. The binomial tree therefore contains within it an approximation to the probability density function for the lognormal random walk.

15.16 THE ASSET PRICE DISTRIBUTION

The probability of reaching a particular node in the binomial tree depends on the number of distinct paths to that node and the probabilities of the up and down moves. Since up and down moves are approximately equally likely and since there are more paths to the interior prices than to the two extremes, we will find that the probability distribution of future prices is roughly bell shaped. In Figure 15.23 is shown the number of paths to each node after four time steps and the probability of getting to each. In Figure 15.24 this is interpreted as probability density functions at a sequence of times.

15.17 VALUING BACK DOWN THE TREE

We certainly know V^+ and V^- at expiry, time T, because we know the option value as a function of the asset then; this is the payoff function.

If we know the value of the option at expiry we can find the option value at the time $T - \delta t$ for all values of S on the tree. But knowing these values means that we can find the option values one step further back in time.

- Thus we work our way back down the tree until we get to the root.

This root is the current time and asset value, and thus we find the option value today.

This algorithm is shown schematically in Figures 15.25 to 15.32.

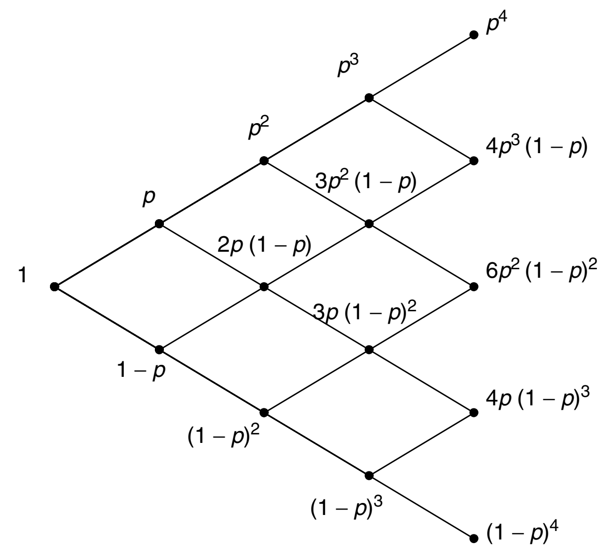


Figure 15.23 Counting paths.

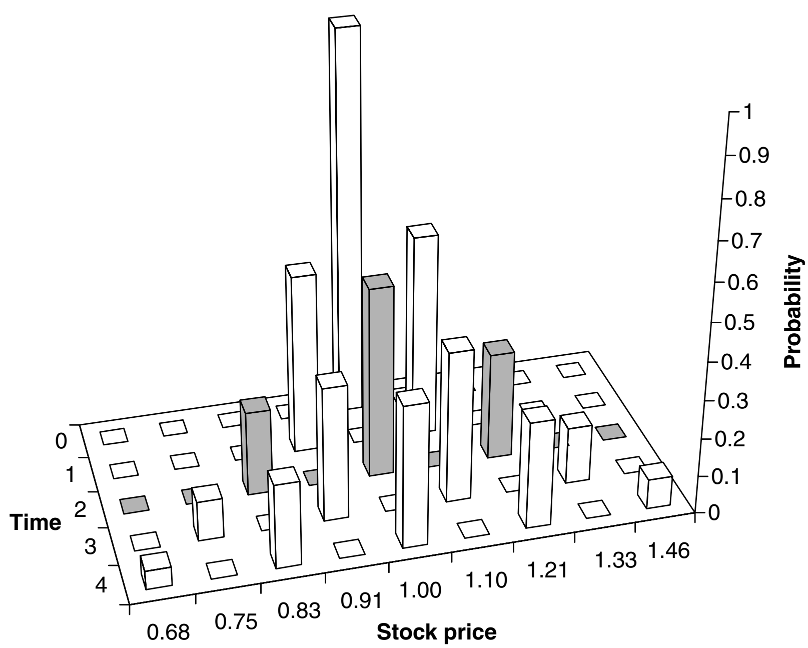


Figure 15.24 The probability distribution of future asset prices.

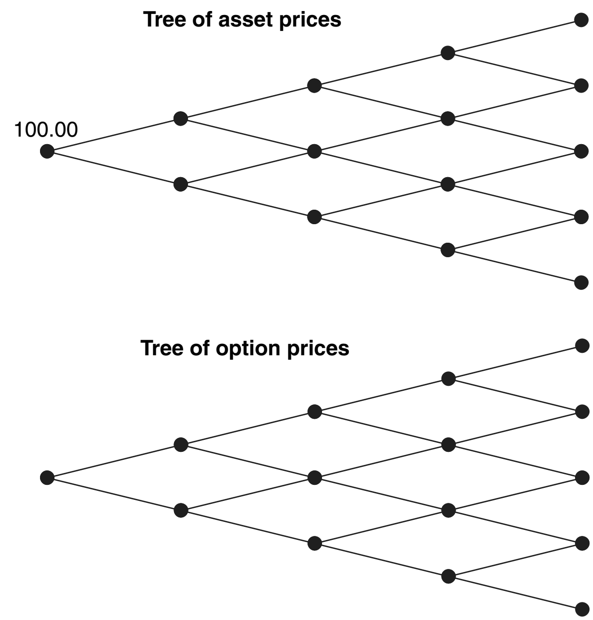


Figure 15.25 The two trees, asset and option.

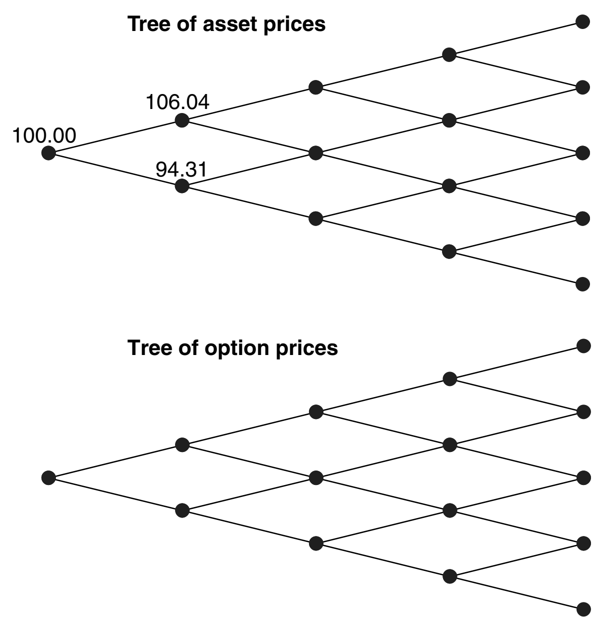


Figure 15.26 Start building up the stock-price tree.

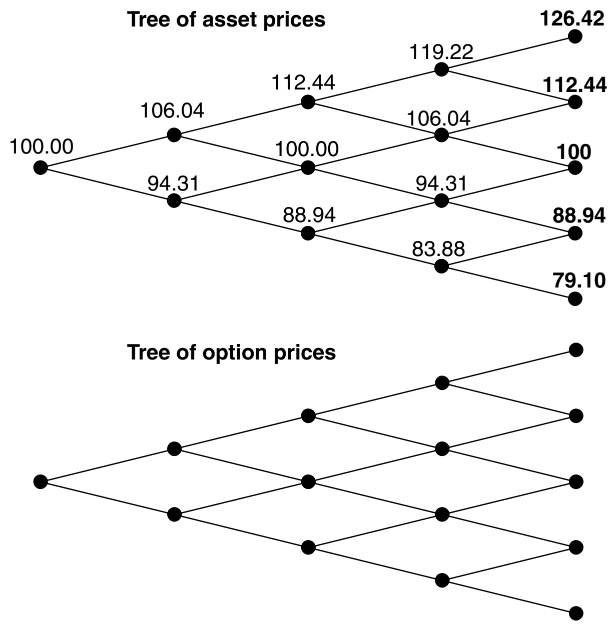


Figure 15.27 The finished stock tree.

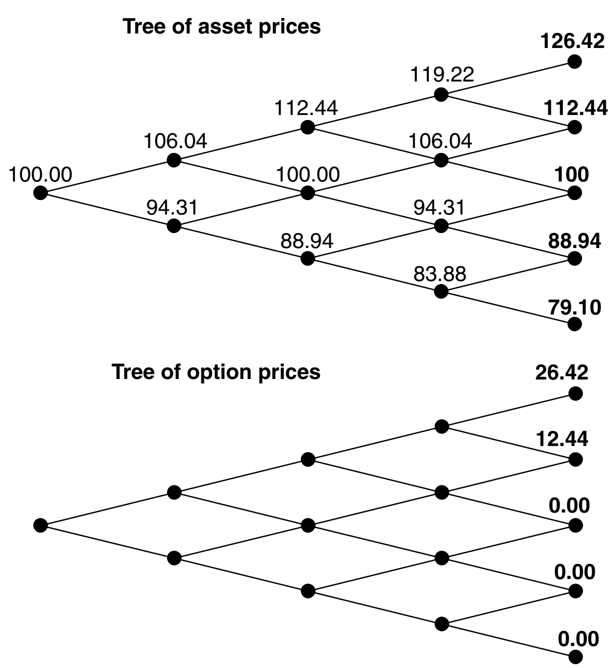


Figure 15.28 The option payoff.

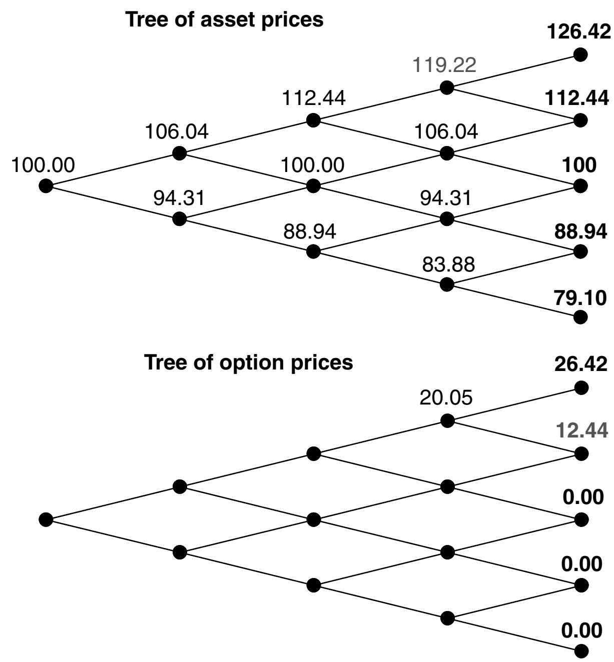


Figure 15.29 Work backwards one 'node' at a time.

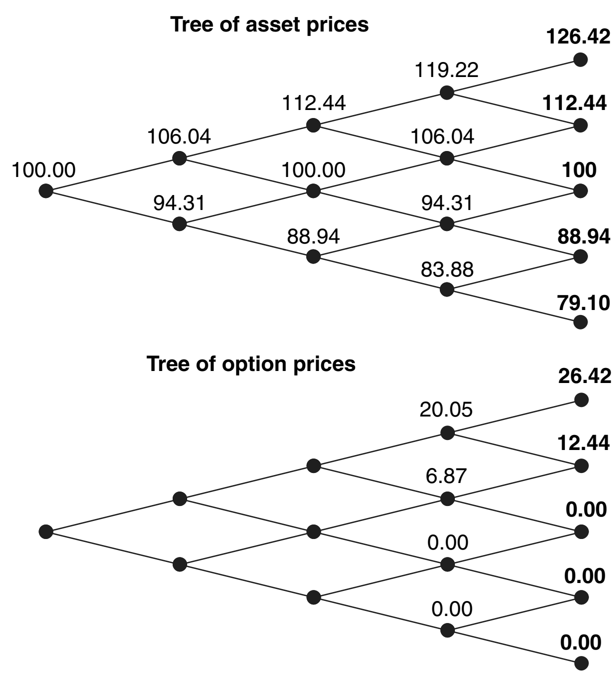


Figure 15.30 First time step completed.

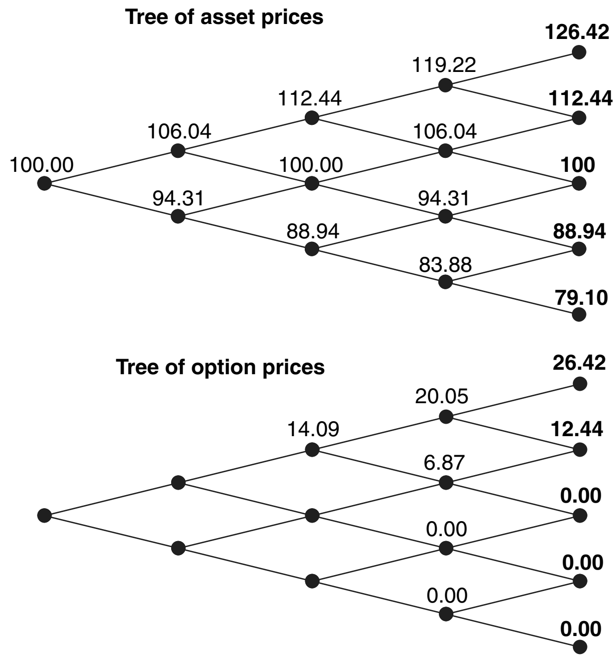


Figure 15.31 Starting on next time step.

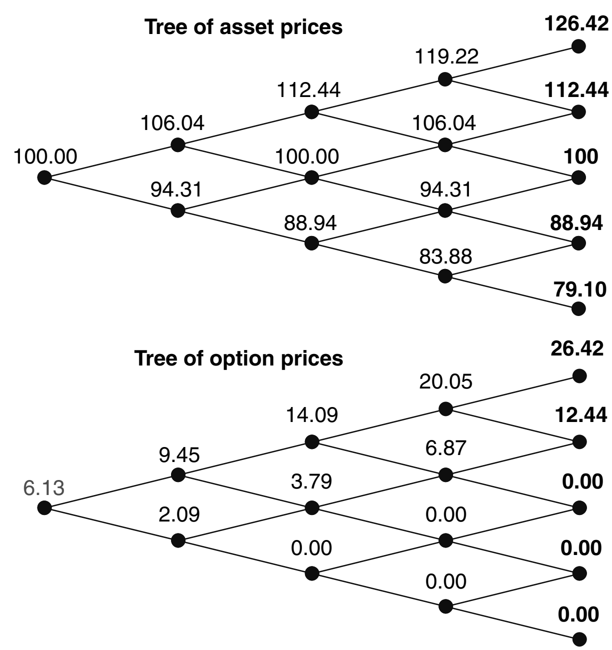
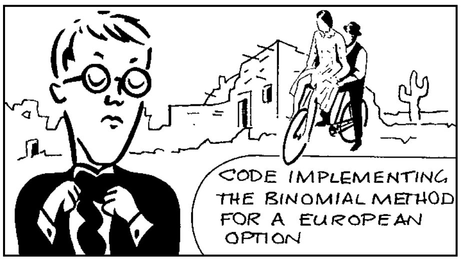


Figure 15.32 The finished option-price tree. Today's option price is therefore 6.13.

15.18 PROGRAMMING THE BINOMIAL METHOD

In practice, the binomial method is programmed rather than done on a spreadsheet. Here is a function that takes inputs for the underlying and the option, using an externally-defined payoff function.[2] Key points to note about this program concern the building up of the arrays for the asset `S()` and the option `V()`. First of all, the asset array is built up only in order to find the final values of the asset at each node at the final time step, expiry. The asset values on other nodes are never used. Second, the argument `j` refers to how far up the asset is from the lowest node *at that time step*.




```
Function Price(Asset As Double, Volatility As Double, _
                       IntRate As Double, Strike As _
                       Double, Expiry As Double, _
                       NoSteps As Integer)
ReDim S(0 To NoSteps)
ReDim V(0 To NoSteps)
time step = Expiry / NoSteps
DiscountFactor = Exp(-IntRate * time step)
temp1 = Exp((IntRate + Volatility * Volatility) _
                                          * time step)
temp2 = 0.5 * (DiscountFactor + temp1)
u = temp2 + Sqr(temp2 * temp2 - 1)
d = 1 / u
p = (Exp(IntRate * time step) - d) / (u - d)

S(0) = Asset
For n = 1 To NoSteps
    For j = n To 1 Step -1
        S(j) = u * S(j - 1)
    Next j
        S(0) = d * S(0)
Next n

For j = 0 To NoSteps
    V(j) = Payoff(S(j), Strike)
Next j

For n = NoSteps To 1 Step -1
    For j = 0 To n - 1
        V(j) = (p * V(j + 1) + (1 - p) * V(j)) _
            * DiscountFactor
    Next j
Next n
Price = V(0)
End Function
```

Here is the externally-defined payoff function `Payoff(S, Strike)` for a call.

```
Function Payoff(S, K)
Payoff = 0
If S > K Then Payoff = S - K
End Function
```

[2] This parameterization of the binomial method is the one explained in the appendix of this chapter.

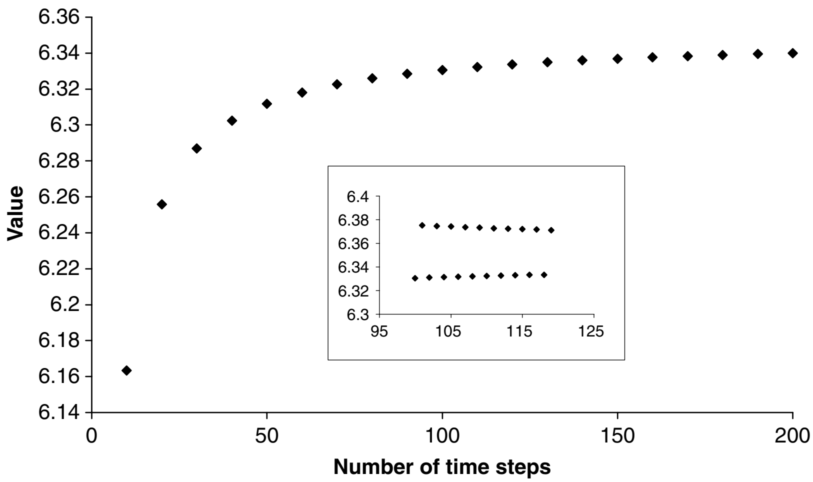


Figure 15.33 Option price as a function of number of time steps.

Since I never use the asset nodes other than at expiry I could have used only the one array in the above, with the same array being used for both S and V. I have kept them separate to make the program more transparent. Also, I could have saved the values of V at all of the nodes, whereas in the above I have only saved the node at the present time. Saving all the values will be important if you want to see how the option value changes with the asset price and time, if you want to calculate greeks for example.

In Figure 15.33 I show a plot of the calculated option price against the number of time steps using this algorithm. The inset figure is a close up. Observe the oscillation. In this example, an odd number of time steps gives an answer that is too high and an even an answer that is too low.

15.19 THE GREEKS

The greeks are defined as derivatives of the option value with respect to various variables and parameters. It is important to distinguish whether the differentiation is with respect to a variable or a parameter (it could, of course, be with respect to both). If the differentiation is only with respect to the asset price and/or time then there is sufficient information in our binomial tree to estimate the derivative. It may not be an accurate estimate, but it will be *an* estimate. The option's delta, gamma and theta can all be estimated from the tree.

On the other hand, if you want to examine the sensitivity of the option with respect to one of the parameters, then you must perform another binomial calculation. This applies to the option's vega and rho for example.

Let me take these two cases in turn.

From the binomial model the option's delta is defined by

$$\frac{V^+ - V^-}{(u - v)S}.$$

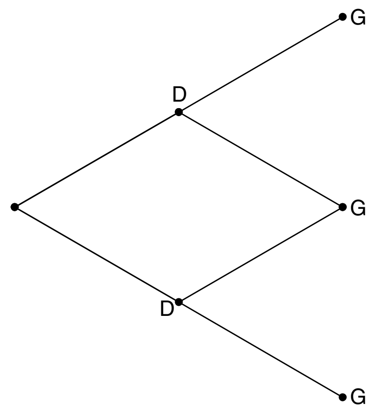


Figure 15.34 Calculating the delta and gamma.

We can calculate this quantity directly from the tree. Referring to Figure 15.34, the delta uses the option value at the two points marked 'D,' together with today's asset price and the parameters u and v. This is a simple calculation.

In the limit as the time step approaches zero, the delta becomes

$$\frac{\partial V}{\partial S}.$$

The gamma of the option is also defined as a derivative of the option with respect to the underlying:

$$\frac{\partial^2 V}{\partial S^2}.$$

To estimate this quantity using our tree is not so clear. It will be much easier when we use a finite-difference grid. However, gamma is a measure of how much we must rehedge at the next time step. But we can calculate the delta at points marked with a D in Figure 15.34 from the option value one time step further in the future. The gamma is then just the change in the delta from one of these to the other divided by the distance between them. This calculation uses the points marked 'G' in Figure 15.34.

The theta of the option is the sensitivity of the option price to time, assuming that the asset price does not change. Again, this is easier to calculate from a finite-difference grid. An obvious choice for the discrete-time definition of theta is to interpolate between V^+ and V^- to find a theoretical option value *had the asset not changed* and use this to estimate

$$\frac{\partial V}{\partial t}.$$

This results in

$$\frac{\frac{1}{2}(V^+ + V^-) - V}{\delta t}.$$

As the time step gets smaller and smaller these greeks approach the Black–Scholes continuous-time values.

Estimating the other type of greeks, the ones involving differentiation with respect to parameters, is slightly harder. They are harder to calculate in the sense that you must perform a second binomial calculation. I will illustrate this with the calculation of the option's vega.

The vega is the sensitivity of the option value to the volatility

$$\frac{\partial V}{\partial \sigma}.$$

Suppose we want to find the option value and vega when the volatility is 20%. The most efficient way to do this is to calculate the option price twice, using a binomial tree, with two different values of σ. Calculate the option value using a volatility of $\sigma \pm \epsilon$, for a small number ϵ; call the values you find $V_{\pm}$. The option value is approximated by the average value

$$V = \tfrac{1}{2}(V_+ + V_-)$$

and the vega is approximated by

$$\frac{V_+ - V_-}{2\epsilon}.$$

The idea can be applied to other greeks.

15.20 EARLY EXERCISE

American-style exercise is easy to implement in a binomial setting. The algorithm is identical to that for European exercise with one exception. We use the same binomial tree, with the same u, v and p, but there is a slight difference in the formula for V. We must ensure that there are no arbitrage opportunities at any of the nodes.

For reasons which will become apparent, I'm going to change my notation now, making it more complex but more informative. Introduce the notation S_j^n to mean the asset price at the nth time step, at the node j from the bottom, $0 \le j \le n$. This notation is consistent with the code above. In our lognormal world we have

$$S_j^n = S u^j v^{n-j},$$

where S is the current asset price. Also introduce V_j^n as the option value at the same node. Our ultimate goal is to find V_0^0 knowing the payoff, i.e. knowing V_j^M for all $0 \le j \le M$ where M is the number of time steps.

Returning to the American option problem, arbitrage is possible if the option value goes below the payoff at any time. If our theoretical value falls below the payoff then it is time to exercise. If we do then exercise the option, its value and the payoff must be the same. If we find that

$$\frac{V_{j+1}^{n+1} - V_j^{n+1}}{u - v} + \frac{1}{1 + r\,\delta t}\frac{uV_j^{n+1} - vV_{j+1}^{n+1}}{u - v} \ge \text{Payoff}(S_j^n)$$

then we use this as our new value. But if

$$\frac{V_{j+1}^{n+1} - V_j^{n+1}}{u - v} + \frac{1}{1 + r\,\delta t}\frac{uV_j^{n+1} - vV_{j+1}^{n+1}}{u - v} < \text{Payoff}(S_j^n)$$

we should exercise, giving us a better value of

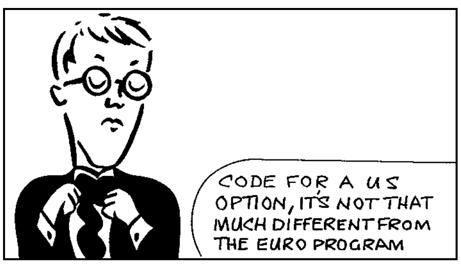
$$V_j^n = \text{Payoff}(S_j^n).$$

We can put these two together to get

$$V_j^n = \max\left(\frac{V_{j+1}^{n+1} - V_j^{n+1}}{u - v} + \frac{1}{1 + r\,\delta t}\frac{uV_j^{n+1} - vV_{j+1}^{n+1}}{u - v}, \text{Payoff}(S_j^n)\right)$$

instead of (15.2). This ensures that there are no arbitrage opportunities. This modification is easy to code, but note that the payoff is a function of the asset price at the node in question. This is new, not seen in the European problem for which we did not have to keep track of the asset values on each of the nodes.

Below is a function for calculating the value of an American-style option. Note the differences between this program and the one for European-style exercise. The code is the same except that we keep track of more information and the line that updates the option value incorporates the no-arbitrage condition.




```
Function USPrice(Asset As Double, Volatility As _
                      Double, IntRate As Double, _
                      Strike As Double, Expiry As _
                      Double, NoSteps As Integer)
ReDim S(0 To NoSteps, 0 To NoSteps)
ReDim V(0 To NoSteps, 0 To NoSteps)
time step = Expiry / NoSteps
DiscountFactor = Exp(-IntRate * time step)
temp1 = Exp((IntRate + Volatility * Volatility) _
                                  * time step)
temp2 = 0.5 * (DiscountFactor + temp1)
u = temp2 + Sqr(temp2 * temp2 - 1)
d = 1 / u
p = (Exp(IntRate * time step) - d) / (u - d)

S(0, 0) = Asset
For n = 1 To NoSteps
    For j = n To 1 Step -1
        S(j, n) = u * S(j - 1, n - 1)
    Next j
        S(0, n) = d * S(0, n - 1)
Next n

For j = 0 To NoSteps
    V(j, NoSteps) = Payoff(S(j, NoSteps), Strike)
Next j

For n = NoSteps To 1 Step -1
    For j = 0 To NoSteps - 1
        V(j, n - 1) = max((p * V(j + 1, n) _
         + (1 - p) * V(j, n)) _
         * DiscountFactor, Payoff(S(j, n - 1), Strike))
    Next j
Next n
USPrice = V(0, 0)
End Function
```

15.21 THE CONTINUOUS-TIME LIMIT

Equation (15.2) and the Black–Scholes equation (5.6) are more closely related than they may at first seem. Recalling that the Black–Scholes equation is in continuous time, we examine (15.2) as $\delta t \to 0$.

First of all, we have chosen

$$u \sim 1 + \sigma\sqrt{\delta t}$$

and

$$v \sim 1 - \sigma\sqrt{\delta t}$$

Next we write

$$V = V(S, t), \quad V^+ = V(uS, t + \delta t) \quad \text{and} \quad V^- = V(vS, t + \delta t).$$

Expanding these expressions in Taylor series for small δt and substituting into (15.1) we find that

$$\Delta \sim \frac{\partial V}{\partial S} \quad \text{as} \quad \delta t \to 0.$$

Thus the binomial delta becomes, in the limit, the Black–Scholes delta.

Similarly, we can substitute the expressions for V, V^+ and V^- into (15.2) to find

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

This is the Black–Scholes equation. Again, the drift rate μ has disappeared from the equation.

15.22 NO ARBITRAGE IN THE BINOMIAL, BLACK–SCHOLES AND 'OTHER' WORLDS

With the binomial discrete-time model, as with the Black–Scholes continuous-time model, we have been able to eliminate uncertainty in the value of a portfolio by a judicious choice of a hedge. In both cases we find that it does not matter how the underlying asset moves, the resulting value of the portfolio is the same. This is especially clear in the above binomial model.

This hedging is only possible in these two simple, popular models. For consider a trivial generalization, the trinomial random walk.

In Figure 15.35 we see a trinomial random walk. In Figure 15.36 we see portfolio values if we try to hedge.

What happens if we try to hedge an option under this scenario? As before, we can 'hedge' with $-\Delta$ of the underlying but this time we would like to choose Δ so that the value of the portfolio (of one option and $-\Delta$ of the asset) is the same at time $t + \delta t$ no matter to which value the asset moves. In other words, we want the portfolio to have the same value for all *three* possible outcomes. Unfortunately, we cannot choose a value for Δ that ensures this to be the case. This amounts to solving two equations (first portfolio value = second portfolio value = third portfolio value) with just one unknown (the delta). Hedging is not possible in the trinomial world. Indeed, perfect hedging, and thus the application of the 'no-arbitrage principle'

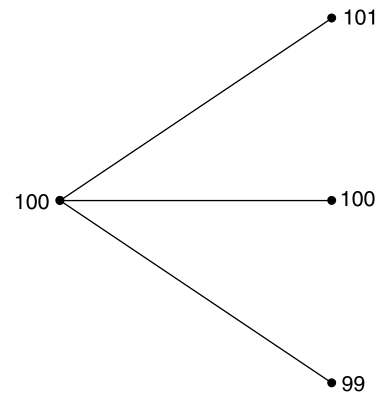


Figure 15.35 The trinomial tree.

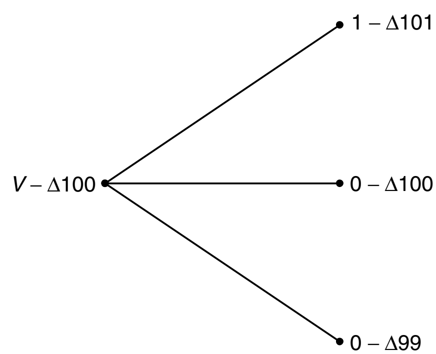


Figure 15.36 Is hedging possible?

is only possible in the two special cases, the Black–Scholes continuous time/continuous asset world, and the binomial world. And in the far more complex 'real' world, delta hedging is *not* possible.[3]

15.23 SUMMARY

In this chapter I described the basics of the binomial model, deriving pricing equations and algorithms for both European- and American-style exercise. The method can be extended in many ways, to incorporate dividends, to allow Bermudan exercise, to value path-dependent contracts and to price contracts depending on other stochastic variables such as interest rates. I have not gone into the method in any detail for the simple reason that the binomial method is just a simple version of an explicit finite-difference scheme. As such it will be discussed

[3] Is it good for the popular models to have such an unrealistic property? These models are at least a good *starting* point. I find it fascinating that the two most popular modeling 'worlds,' the Black–Scholes and the binomial, are the only two in which hedging is possible. I think this says more about mathematical modelers than it does about the nature of the financial markets.

in depth in Part Six. Finite-difference methods have an obvious advantage over the binomial method; they are far more flexible.

FURTHER READING

- The original binomial concept is due to Cox, Ross & Rubinstein (1979).
- Almost every book on options describes the binomial method in more depth than I do. One of the best is Hull (2005) who also describes its use in the fixed-income world.

APPENDIX: ANOTHER PARAMETERIZATION

The three constants u, v and p are chosen to give the binomial walk the same drift and standard deviation as that given by the stochastic differential equation (3.7). Having only these two equations for the three parameters gives us one degree of freedom in this choice. This degree of freedom is often used to give the random walk the further property that after an up and a down movement (or a down followed by an up) the asset returns to its starting value, S.[4] This gives us the requirement that

$$v(uS) = u(vS) = S$$

i.e.

$$uv = 1. \tag{15.4}$$

Our starting point, the lognormal random walk,

$$dS = \mu S\,dt + \sigma S\,dX$$

has the solution, found in Section 4.14.2,

$$S(t) = S(0)e^{\left(\mu-\frac{1}{2}\sigma^2\right)t+\sigma\phi\sqrt{t}}$$

where ϕ is a standardized Normal random variable.

For the binomial random walk to have the correct drift over a time period of δt we need

$$puS + (1-p)vS = SE\left[e^{\left(\mu-\frac{1}{2}\sigma^2\right)\delta t+\sigma\phi\sqrt{\delta t}}\right] = Se^{\mu\,\delta t},$$

i.e.

$$pu + (1-p)v = e^{\mu\,\delta t}.$$

Rearranging this equation we get

$$p = \frac{e^{\mu\delta t} - v}{u - v}. \tag{15.5}$$

[4] Other choices are possible. For example, sometimes the probability of an up move is set equal to the probability of a down move i.e. $p = 1/2$.

Then for the binomial random walk to have the correct variance we need (details omitted)

$$pu^2 + (1-p)v^2 = e^{(2\mu+\sigma^2)\delta t}. \tag{15.6}$$

Equations (15.4), (15.5) and (15.6) can be solved to give

$$u = \tfrac{1}{2}\left(e^{-\mu\,\delta t} + e^{(\mu+\sigma^2)\delta t}\right) + \tfrac{1}{2}\sqrt{\left(e^{-\mu\,\delta t} + e^{(\mu+\sigma^2)\delta t}\right)^2 - 4}.$$

Approximations that are good enough for most purposes are

$$u \approx 1 + \sigma\,\delta t^{1/2} + \tfrac{1}{2}\sigma^2\delta t,$$
$$v \approx 1 - \sigma\,\delta t^{1/2} + \tfrac{1}{2}\sigma^2\delta t$$

and

$$p \approx \frac{1}{2} + \frac{\left(\mu - \frac{1}{2}\sigma^2\right)\delta t^{1/2}}{2\sigma}.$$

Of course, if this is being used for pricing options, you must replace the μ with r everywhere.

CHAPTER 16
how accurate is the Normal approximation?

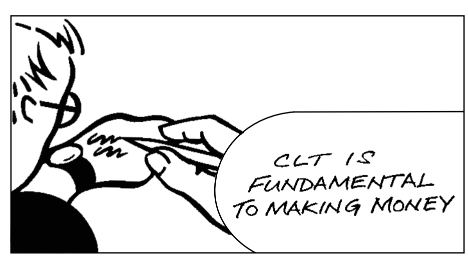
In this Chapter...

- why the Normal distribution is so popular
- how fat the tails really are
- what dropping the Normal assumption entails

16.1 INTRODUCTION

Without a shadow of a doubt the assumption of Normally distributed returns is one of the most important assumptions in quantitative finance. It allows us to make enormous advances because it comes with a lot of relatively easy-to-use analytical tools. Yet the Normal distribution has often been criticized for being unrealistic in its description of large events: stock crashes. The Normal distribution vastly underestimates their probability. In this short chapter we look at why the distribution is popular, see some simple statistics associated with tail events and look at alternatives.

16.2 WHY WE LIKE THE NORMAL DISTRIBUTION: THE CENTRAL LIMIT THEOREM



The Central Limit Theorem (CLT) states: 'Given a distribution with a mean m and variance s^2, the sampling distribution of the mean approaches a Normal distribution with a mean of m and a variance of s^2/N as N increases.'

By 'sampling distribution of the mean' is meant

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

where the X_i are all drawn from the initial distribution.

In layman's terms, if you add up lots of random numbers all drawn from the same 'building block' distribution then you get a Normal distribution. And this works for any building-block distributions (except for some 'small print' which we'll see in a moment). This explains why the Normal distribution is important in practice; it occurs whenever a distribution comes from adding up lots of random numbers. Perhaps stock price daily returns should be Normal since you 'add up' thousands of returns during each day.

And since the Normal distribution only has the two parameters, the mean and the variance, it follows that the skew and kurtosis etc. of the building-block distribution don't much matter to the final distribution.

The 'small print' are the conditions under which the Central Limit Theorem is valid. These conditions are:

- The building-block distributions must be identical (you aren't allowed to draw from different distributions each time).
- Each draw from the building-block distribution must be independent from other draws.
- The mean and standard deviation of the building-block distribution must both be finite.

There are generalizations of the CLT in which these conditions are weakened, but we won't go into those here.

16.3 NORMAL VERSUS LOGNORMAL

I often ask new students what distribution is assumed by the Black–Scholes model for the asset return. The answer (before I have taught them 'properly') is usually equally likely to be either Normal or lognormal. But then I get the same answers when I ask them what is the distribution assumed for the asset return.

You will know that the simple assumption for returns is that they are Normal and that, provided the parameters drift and volatility are constant, the resulting distribution for the asset is lognormal.

Here is a quick way of demonstrating and explaining lognormality that relies only on the Central Limit Theorem.

Start with a stock price with value S_0. Add a random return R_1 to this to get the stock price, S_1, at the next time step:

$$S_1 = S_0(1 + R_1).$$

After the second time step, and a random return of R_2, the stock price is

$$S_2 = S_0(1 + R_1)(1 + R_2).$$

After N time steps we have

$$S_N = S_0 \prod_{i=1}^{N} (1 + R_i). \tag{16.1}$$

What is the distribution for the stock price S_N?

We can use the Central Limit Theorem to answer that question quite easily. Take logarithms of (16.1) to get

$$\log(S_N) = \log(S_0) + \sum_{i=1}^{N} \log(1 + R_i).$$

Since R_i is random, it follows that $\log(1 + R_i)$ is random, so here we are adding up many random numbers. If the R_is are all drawn from the same distribution (and the other conditions for the CLT hold) and N is large, then this sum is approximately Normal. And that's what lognormal means. A random variable is lognormally distributed if the logarithm of it is Normally distributed. So S_N is lognormally distributed.

16.4 DOES MY TAIL LOOK FAT IN THIS?

There is evidence, and lots of it, that tails of returns distributions are fat. Take the probability density function for the daily returns on the S&P index since 1980. In Figure 16.1 is plotted the empirical distribution (scaled to have zero mean and standard deviation of one) and also the standardized Normal distribution. This is a typical plot of any financial data, whether it is an index, stock, exchange rate, etc. The empirical peak is higher than the Normal distribution and the tails are both fatter (although it is difficult to see that in the figure). Now, the high peak

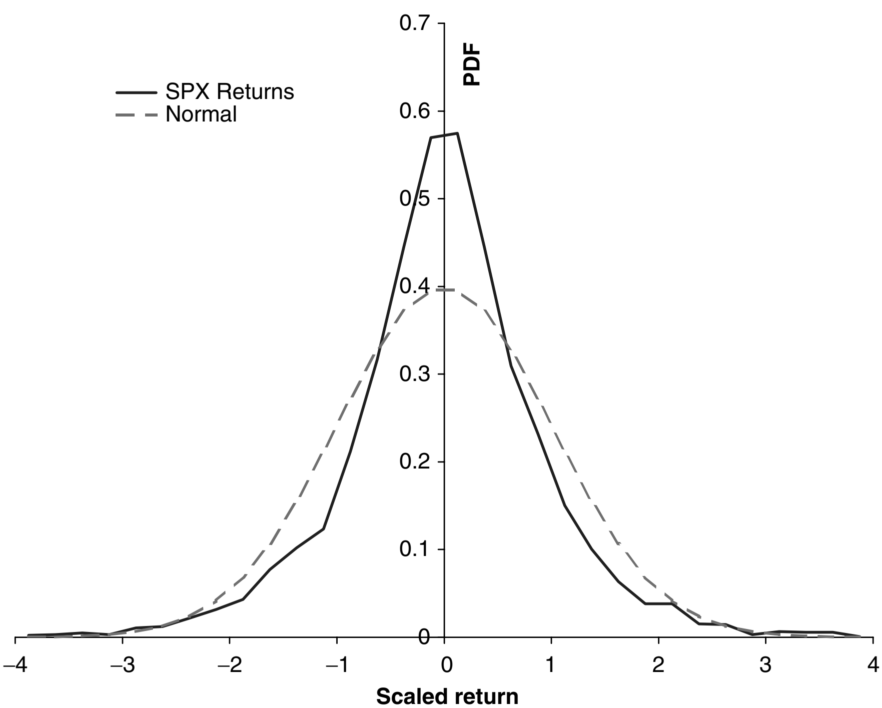


Figure 16.1 The standardized probability density functions for SPX returns and the Normal distribution.

doesn't matter so much but the tails are very, very important. Let's look at some very simple statistics.

We are going to work with the famous stock market crash of 19th October 1987. On that day the SP500 fell 20.5%. We will ask the question: 'What is the probability of a 20% one-day fall in the SP500?' We will look at the empirical data and the theoretical data.

16.4.1 Probability of a 20% SPX Fall: Empirical

Since we are working with 24 years of daily data, we could argue that empirically the probability of a 20% fall in the SPX is one in 24×252, or 0.000165. We could be far more sophisticated than that, and use ideas from Extreme Value Theory, but we will be content with that as a ball-park figure.

16.4.2 Probability of a 20% SPX Fall: Theoretical

To get a theoretical estimate, based on Normal distributions, we must first estimate the daily standard deviation for SPX returns. Over that period it was 0.0106, equivalent to an average volatility of 16.9%. What is the probability of a 20% or more fall when the standard deviation is 0.0106? This is a staggeringly small $1.8\ 10^{-79}$. That is just once every $2\ 10^{76}$ years.

Empirical answer: Once every 25 years. Theoretical answer: Once every $2\ 10^{76}$ years. That's how bad the Normal-distribution assumption is in the tails.

16.5 USE A DIFFERENT DISTRIBUTION, PERHAPS

That all sounds like a very compelling reason to dismiss the Normal distribution as being a poor model of returns. Perhaps we should be more scientific and work with more realistic distributions. That certainly is one option. The problem with working with 'more realistic' distributions is that they have properties that are somewhat difficult to handle. For example, the distributions that seem to fit returns the best are soooo fat tailed that their standard deviation is infinite (Table 16.1).

Such an observation fits nicely with the above conditions on the CLT. If the stock return from trade to trade has infinite standard deviation then we can't expect daily returns to be Normally distributed.

But you can imagine what hurdles that presents us with. Standard deviation is seen in classical theory as a measure of risk, it even has a catchy name, volatility, (when annualized) and its own symbol, σ. Throwing away such theory is not something to be done lightly. If standard deviation doesn't exist it follows that delta hedging is impossible, risk preferences need to be modeled and the option pricing equation becomes a more complicated partial *integro* differential equation, where the 'integro' part comes from a relationship between option values at all stock prices. Instead of the relatively nice local Black–Scholes equation which is in terms of differential calculus, we need a global model that includes integrals as well. The Further Reading section will give you some pointers as to who is active in this field of research.

Table 16.1 Normal distributions versus fat-tailed distributions.

Normal	Fat tail
Math easy	Math hard
Underestimates crashes	Good estimate of crashes
Practitioner approach	Scientist approach
Standard Deviation $\propto$ Volatility	Standard Deviation = ∞
Returns $\propto \delta t^{1/2}$	Returns $\propto \delta t^{1/2+}$
Can delta hedge	Can't, must accept risk
Risk preferences don't matter	Need to bring in preferences
Local models, derivatives only	Global, integrals

There are other ways to model fat tails that don't require infinite standard deviations and we shall look at them in Part Five.

16.6 SERIAL AUTOCORRELATION

Another reason why the Normal distribution might not be relevant is if there is any Serial Autocorrelation in stock price returns from trade to trade, or day to day. Serial Autocorrelation means the correlation between the return one day and the return the previous day, for example. It might be the case that an up move is more likely to be followed by another up move than by a down move. That would be positive serial autocorrelation.

Again there is evidence that there is such autocorrelation, perhaps not that strong on average, but over certain periods, especially intra day, the effect is enough to scupper the Normal distribution.

Very little has been written about serial autocorrelation in stock price returns, and almost nothing about pricing derivatives in such a framework. But we shall have a go at this subject in Chapter 65.

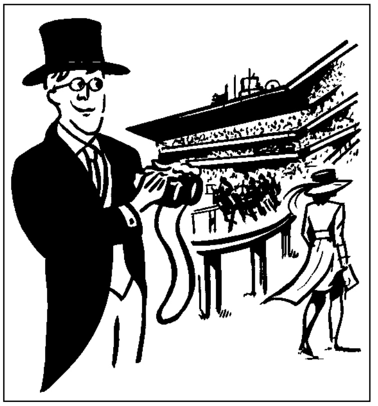
16.7 SUMMARY

My personal preference is for using the assumption of Normal distributions most of the time, and treat tail events separately. By that I mean always keep the thought that a stock may plummet dramatically right at the front of your mind. Take precautions against such moves by, for example, buying tail protection such as out-of-the-money puts, or by diversifying your portfolio; don't have all your money in a small number of stocks. We'll see how to examine market crashes when all stocks simultaneously experience tail events in Chapter 43.

FURTHER READING

- Jim Gatheral has written loads of great stuff on pricing with fat-tailed distributions.
- See Joshi (2003) for details of the Variance Gamma model. This model uses the idea of a random time to give fat tails, first introduced by Madan & Seneta (1990) and developed by Madan, Carr & Change (1998).
- See Kyprianou, Schoutens & Wilmott (2005) for Lévy processes and exotic option pricing.

CHAPTER 17
investment lessons from blackjack and gambling



In this Chapter...

- the rules of blackjack
- blackjack strategy and card counting
- the Kelly criterion and money management
- no arbitrage in horse racing

17.1 INTRODUCTION

When I lecture on portfolio management and the mathematics of investment decisions I often start off with a description of the card game blackjack. It is a very simple game, one that most people are familiar with, perhaps by the name of pontoon, 21 or *vingt et un*. The rules are easy to remember, each hand lasts a very short time, the game is easily learned by children and could well give them their first taste of gambling. For without this gambling element there is little point in playing blackjack.

Since the rules are simple and the probabilities can be analyzed, blackjack is also the perfect game to learn about risk, return and money management and, perhaps most importantly, to help you learn what type of gambler you are. Are you risk averse or a risk seeker? This is an important question for anyone who later will work in banking and may be gambling with OPM, other people's money.

Despite blackjack being perfect for learning the basics of financial risk and return, and despite bank training managers liking the idea of people being trained in risk management via this game, I am always asked by those training managers to change the title of my lecture. 'You can't call your lecture "Investment Lessons from Blackjack and Gambling," we'll get into trouble with [regulator goes here].' This is a bit silly. Anyone who doesn't think that investment and gambling share the same roots is silly. I can even go so far as to say that most professional gamblers that I know have a better understanding of risk, return and money management than most of the risk managers I know.

In this chapter we will see some of the ideas that these professional gamblers use.

17.2 THE RULES OF BLACKJACK

Players at blackjack sit around a kidney-shaped table, with the dealer standing opposite. A bird's eye view is shown in Figure 17.1.

Before any cards are dealt, the player must place his bet in front of his table position. The dealer deals two cards to each of the players, and two to himself (one of the dealer's cards is dealt face up and the other face down). This is the state of play shown in the figure. Court cards (kings, queens and jacks) count as 10, ace counts as either one or 11 and all other cards are counted at their face value. The value of the ace is chosen by the player.

The aim of the game for the player is to hold a card count greater than that of the dealer without exceeding 21 (going 'bust').

If the player's first two cards are an Ace and a 10-count card he has what is known as 'blackjack' or a natural. If he gets a natural with his first two cards the player wins, unless the dealer also has a natural, in which case it is a standoff or tie (a 'push') and no money changes hands. A winning natural pays the player 3 to 2.

Working clockwise around the table from his immediate left the dealer asks each player in turn whether they want to hit or stand. 'Hit' means to draw another card. 'Stand' means no more cards are taken. If the player hits and busts, his wager is lost. The player can keep taking cards until he is satisfied with his count or busts.

The player also has other decisions to make.

He is also allowed to double the bet on his first two cards and draw one additional card only. This is called 'doubling down.'

If the first two cards a player is dealt are a pair, he may split them into two separate hands, bet the same amount on each and then play them as two distinct hands. This is called 'splitting pairs.' Aces can receive only one additional card. After splitting, Ace + 10 counts as 21 and not as blackjack. If the dealer's up card is an ace, the player may take insurance, a bet not exceeding one half of his original bet. If the dealer's down card is a 10-count card, the player wins 2 to 1. Any other card means a win for the dealer.

It is sometimes permitted to 'surrender' your bet. When permitted, a player may give up his first two cards and lose only one half of his original bet.

The dealer has no decisions to make. He must always follow very simple rules when it comes to hitting or standing. He must draw on 16 and stand on 17. In some casinos, the dealer is required to draw on soft 17 (a hand in which an Ace counts as 11, not one). Regardless of the total the player has, the dealer must play this way.

Figure 17.1 Blackjack table layout.

In a tie no money is won or lost, but the bet stays on the table for the next round.

Rules differ subtly from casino to casino, as do the number of decks used.

The casino has an advantage over the player and so, generally speaking, the casino will win in the long run. The advantage to the dealer is that the player can go bust, losing his bet immediately, even if the dealer later busts. This asymmetry is the key to the house's edge. The key to the player's edge, which we will be exploiting shortly, is that he can vary both his bets and his strategy. The first published strategy for winning at blackjack was published by Ed Thorp in 1962 in his book *Beat the Dealer*. In this book Professor Thorp explained that the key ingredients to winning at blackjack were

- the strategy: Knowing when to hit or stand, doubledown etc. This will depend on what cards you are holding and the dealer's upcard;
- information: Knowing the approximate makeup of the remaining cards in the deck, some cards favor the player and others the dealer;
- money management: How to bet, when to bet small and when to bet large.

17.3 BEATING THE DEALER

The first key is in having the optimal strategy. That means knowing whether to hit or stand. You're dealt an eight and a four and the dealer's showing a six, what do you do? The optimal strategy involves knowing when to split pairs, double down (double your bet in return for only taking one extra card), or draw a new card. Thorp used a computer simulation to calculate the best strategies by playing thousands of blackjack hands. In his best-selling book *Beat the Dealer* Thorp presented tables like the one in Figure 17.2 showing the best strategies.

But the optimal strategy is still not enough, without the second key.

You've probably heard of the phrase 'card counter' and conjured up images of Doc Holliday in a ten-gallon hat. The truth is more mundane. Card counting is not about memorizing entire decks of cards but keeping track of the type and percentage of cards remaining in the deck during your time at the blackjack table. Unlike roulette, blackjack has 'memory.' What happens during one hand depends on the previous hands and the cards that have already been dealt out.

A deck that is rich in low cards, twos to sixes, is good for the house. Recall that the dealer must take a card when he holds sixteen or less; the high frequency of low-count cards increases his chance of getting close to 21 without busting. For example, take out all the fives from a single deck and the player has an advantage of 3.3 per cent! On the other hand, a deck rich in 10-count cards (10s and court cards) and Aces is good for the player, increasing the chances of either the dealer busting or the player getting a natural (21 with two cards) for which he gets paid at odds of three to two. In the simplest case, card counting means keeping a rough mental count of the percentage of aces and 10s, although more complex systems are possible for the really committed. When the deck favors the player he should increase his bet, when the deck is against him he should lower his bet. (And this bet variation must be done sufficiently subtly so as not to alert the dealers or pit bosses.)

One of the simplest card-counting techniques is to perform the following simple calculation in your head as the cards are being dealt. With a fresh deck(s) start from zero, and then for every Ace and 10 that is dealt subtract one; for every 2–6 add one. The larger the count, divided by an estimate of the number of cards left in the deck, the better are your chances of winning. You perform this mental arithmetic as the cards are being dealt around the table.

YOUR HAND \ DEALER'S UPCARD	2	3	4	5	6	7	8	9	10	A
17+	S	S	S	S	S	S	S	S	S	S
16	S	S	S	S	S	H	H	H	H	H
15	S	S	S	S	S	H	H	H	H	H
14	S	S	S	S	S	H	H	H	H	H
13	S	S	S	S	S	H	H	H	H	H
12	H	H	S	S	S	H	H	H	H	H
11	D	D	D	D	D	D	D	D	D	H
10	D	D	D	D	D	D	D	D	H	H
9	H	D	D	D	D	H	H	H	H	H
5-8	H	H	H	H	H	H	H	H	H	H
A, 8-10	S	S	S	S	S	S	S	S	S	S
A, 7	S	D	D	D	D	S	S	H	H	H
A, 6	H	D	D	D	D	H	H	H	H	H
A, 5	H	H	D	D	D	H	H	H	H	H
A, 4	H	H	D	D	D	H	H	H	H	H
A, 3	H	H	H	D	D	H	H	H	H	H
A, 2	H	H	H	D	D	H	H	H	H	H
A, A; 8, 8	SP	SP	SP	SP	SP	SP	SP	SP	SP	SP
10, 10	S	S	S	S	S	S	S	S	S	S
9, 9	SP	SP	SP	SP	SP	S	SP	SP	S	S
7, 7	SP	SP	SP	SP	SP	SP	H	H	H	H
6, 6	SP	SP	SP	SP	SP	H	H	H	H	H
5, 5	D	D	D	D	D	D	D	D	H	H
4, 4	H	H	H	SP	SP	H	H	H	H	H
3, 3	SP	SP	SP	SP	SP	SP	H	H	H	H
2, 2	SP	SP	SP	SP	SP	SP	H	H	H	H

Figure 17.2 The basic blackjack strategy.

In *Beat the Dealer*, Ed Thorp published his ideas and the results of his ‘experiments.’ He combined the card-counting idea, money management techniques (such as the Kelly criterion, below) and the optimal play strategy to devise a system that can be used by anyone to win at this casino game. ‘The book that made Las Vegas change the rules,’ as it says on the cover, and probably the most important gambling book ever, was deservedly in the *New York Times* and *Time* bestseller lists, selling more than 700,000 copies.

Although passionate about probability and gambling – he plays blackjack to relax – even Ed himself could not face the requirements of being a professional gambler. ‘The activities weren’t intellectually challenging along that life path. I elected not to do that.’

Once on a film set, Paul Newman asked him how much he could make at blackjack. Ed told him $300,000 a year. ‘Why aren’t you out there doing it?’ Ed’s response was that he could make a lot more doing something else, with the same effort, and with much nicer working conditions and a much higher class of people. Truer words were never spoken. Ed Thorp took his knowledge of probability, his scientific rigor and his money management skills to the biggest casino of them all, the stock market.

17.3.1 Summary of Winning at Blackjack

- If you play blackjack with no strategy you will lose your money quickly. If your strategy is to copy the dealer’s rules then there is a house edge of between five and six percent.

- The best strategy involves knowing when to hit or stand, when to split, double down, take insurance etc. This decision will be based on the two cards you hold and the dealer's face up card. If you play the best strategy you can cut the odds down to about evens.
- To win at blackjack takes patience and the ability to count cards.
- If you follow the optimal strategy and simultaneously bet high when the deck is favorable, and low otherwise, then you will win in the long run.

What does this have to do with investing?

Over the next two sections we will see how to use estimates of the odds (from card counting in blackjack, say, or statistical analysis of stock price returns) to manage our money optimally.

17.4 THE DISTRIBUTION OF PROFIT IN BLACKJACK

Let's introduce some notation for the distribution of winnings at blackjack. ϕ is a random variable denoting the outcome of a bet. There will be probabilities associated with each ϕ. Suppose that μ is the mean and σ the standard deviation of ϕ.

In blackjack ϕ will take discrete values:

$$\phi = -1, \quad \text{player loses,}$$

$$\phi = 0, \quad \text{a `push,'}$$

$$\phi = 1, \quad \text{player wins,}$$

$$\phi = 3/2, \quad \text{player gets a `natural.'}$$

The distribution is shown (schematically) in Figure 17.3.

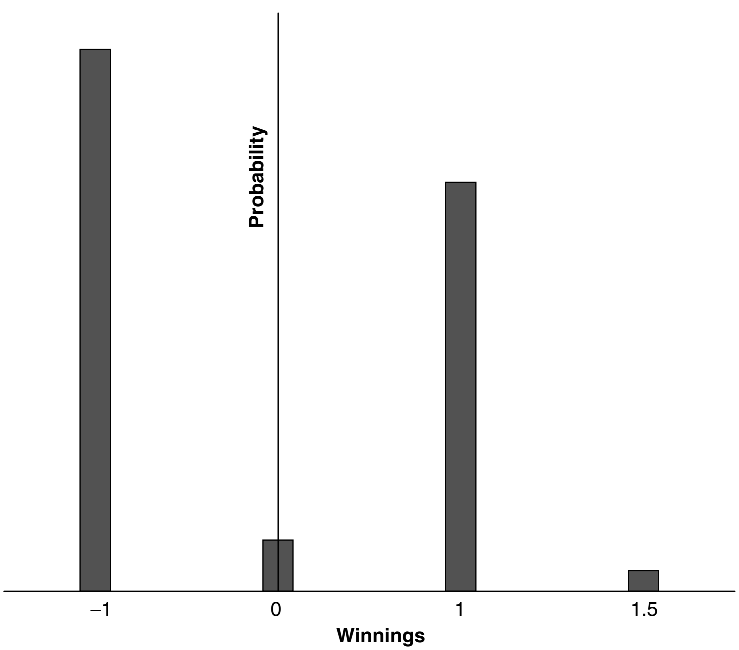
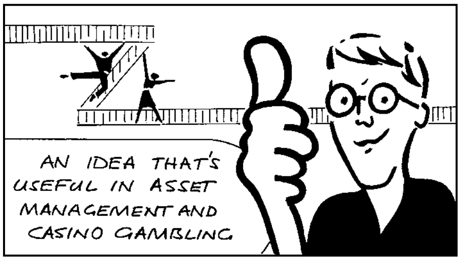


Figure 17.3 The blackjack probability density function (schematic).



17.5 THE KELLY CRITERION

To get us into the spirit of asset choice and money management, consider the following real-life example. You have \$1000 to invest and the only investment available to you is in a casino playing blackjack.

If you play blackjack with no strategy you will lose your money quickly. The odds, as ever, are in favor of the house. If your strategy is to copy the dealer's rules then there is a house edge of between 5 and 6%. This is because when you bust you lose, even if the dealer busts later. There is, however, an optimal strategy. The best strategy involves knowing when to hit or stand, when to split, double down, take insurance (pretty much never) etc. This decision will be based on the two cards you hold and the dealer's face up card. If you play the best strategy, you can cut the odds down to about evens, the exact figure depending on the rules of the particular casino.

To win consistently at blackjack takes two things: patience and the ability to count cards. The latter only means keeping track of, for example, the number of aces and ten-count cards left in the deck. Aces and tens left in the deck improve your odds of winning. If you follow the optimal strategy and simultaneously bet high when there are a lot of aces and tens left, and low otherwise, then you will in the long run do well. If there are any casino managers reading this, I'd like to reassure them that I have never mastered the technique of card counting, so its not worth them banning me. On the other hand, I always seem to win, but that may just be selective memory.

The following is a description of the **Kelly criterion**. It is a very simple way to optimize your bets or investments so as to maximize your long-term average growth rate. This is the subject of *money management*. This technique is not specific to blackjack, although I will continue to use this as a concrete example, but can be used with any gambling game or investment where you have a positive edge and have some idea of the real probabilities of outcomes. The idea has a long and fascinating history, all told in the book by Poundstone (2005). In that book you will also be able to read how the idea has divided the economics community from the gambling community.

We are going to use the ϕ notation for the outcome of a hand of blackjack, but since each hand is different we will add a subscript. So ϕ_i means the outcome of the ith hand.

Suppose I bet a fraction f of my \$1000 on the first hand of blackjack, how much will I have after the hand? The amount will be

$$1000\,(1 + f\phi_1),$$

where the subscript '1' denotes the first hand.

On to the second and subsequent hands. I will consistently bet a constant fraction f of my holdings each hand, so that after two hands I have an amount[1]

$$1000\,(1 + f\phi_1)(1 + f\phi_2).$$

After M hands I have

$$1000\,\Pi_{i=1}^{M}(1 + f\phi_i).$$

How should I choose the amount f?

[1] This is not quite what one does when counting cards, since one will change the amount f.

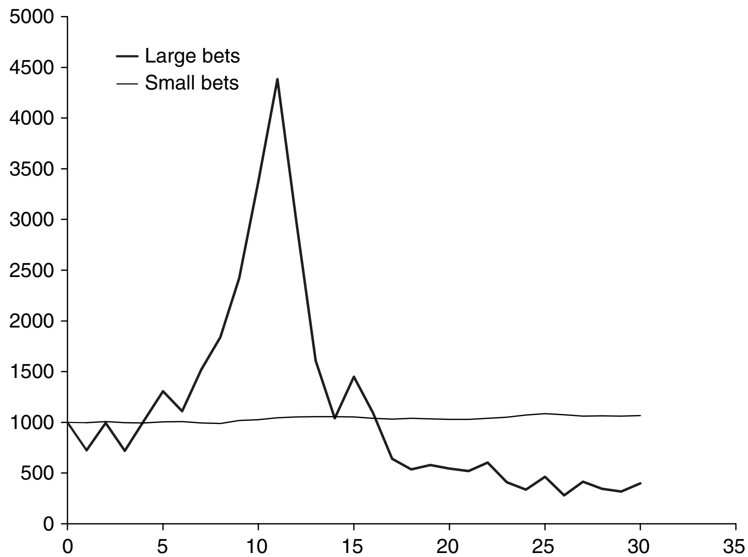


Figure 17.4 Big bets and small bets.

If f is large, one say, then I will certainly eventually lose all of my money, even if the expected return is positive. If f is very small then it will take me a very long time to make any sizeable profit. These two extremes are demonstrated in Figure 17.4.

Perhaps there is some middle ground where I will make a reasonable profit without risking it all. This is indeed the case; let's do the math.

I am going to choose the fraction f *to maximize my expected long-term growth rate*. There are other strategies that you can adopt, such as expected wealth maximization. Each goal has its pros and cons.

Expected growth maximization is an obvious goal for a hedge fund since average growth is a number that all potential investors will look at.

This growth rate is given by

$$\frac{1}{M}\log\left(1000\,\Pi_{i=1}^{M}(1+f\phi_i)\right)=\frac{1}{M}\sum_{i=1}^{M}\log\left(1+f\phi_i\right)+\frac{1}{M}\log(1000).$$

Assuming that the outcome of each hand is independent,[2] then the expected value of this is

$$E\left[\log\left(1+f\phi_i\right)\right],$$

ignoring the scaling factor log(1000).

Expanding the argument of the logarithm in Taylor series, we get

$$E\left[f\phi_i-\tfrac{1}{2}f^2\phi_i^2+\cdots\right].$$

Now assuming that the mean is small but that the standard deviation is not, we find that the expected long-term growth rate is approximately

$$f\mu-\tfrac{1}{2}f^2\sigma^2. \tag{17.1}$$

[2] An assumption not strictly true for blackjack, of course.

This is maximized by the choice

$$f^* = \frac{\mu}{\sigma^2},$$

giving an expected growth rate of

$$\frac{\mu^2}{2\sigma^2}$$

per hand.

If $\mu > 0$ then $f > 0$ and we stand to make a profit, in the long term. If $\mu < 0$, as it is for roulette or if you follow a naive blackjack strategy, then you should invest a negative amount, i.e. own the casino. (If you must play roulette, put all your money you would gamble in your lifetime on a color, and play once. Not only do you stand an almost 50% chance of doubling your money, you will gain an invaluable reputation as a serious player.)

The long-run growth rate maximization and the optimal amount to invest is called the Kelly criterion.

In Figure 17.5 is shown the function given in Equation (17.1) for the expected long-term growth rate. This example uses $\mu = 0.01$ and $\sigma = 1$. In this figure you will see the optimal betting fraction is the value for f which maximizes the function; this is the Kelly criterion. To the left of this is conservative betting, to the right is crazy. I say crazy because going beyond the optimal fraction increases your volatility of winnings, and decreases their expectation. At 'twice Kelly' your expected winnings are zero, and beyond that they become negative.

Given that in practice you rarely know the odds accurately it makes sense to bet conservatively, in case you accidentally stray into the crazy zone. For that reason many people use 'half

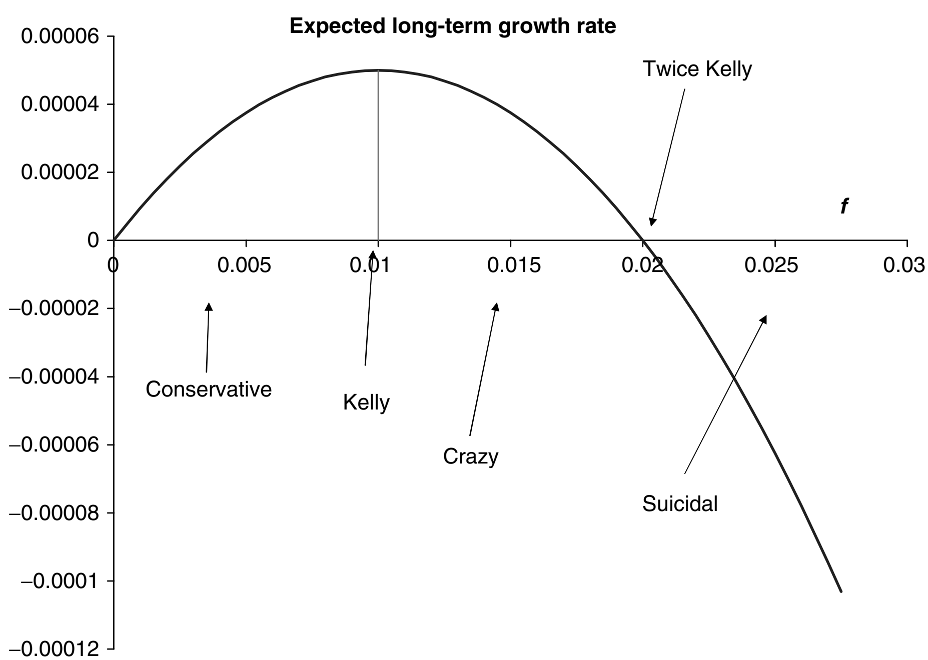


Figure 17.5 The expected long-term growth rate as a function of investment fraction.

Kelly,' that is a fraction that is half of the Kelly fraction. This halves your volatility, keeps you nicely away from the crazy zone, yet only decreases your expected growth rate by 25%.

If you can play M times in an evening you would expect a total growth of

$$\frac{\mu^2 M}{2\sigma^2}, \tag{17.2}$$

using full Kelly.

This illustrates one possible way of choosing a portfolio, which asset to invest in (blackjack) and how much to invest (f^*). Faced with other possible investments, then you could argue in favor of choosing the one with the highest (17.2), depending on the mean of the return, its standard deviation and how often the investment opportunity comes your way. These ideas are particularly important to the technical analyst or chartist who trades on the basis of signals such as golden crosses, saucer bottoms, and head and shoulder patterns. Not only do the risk and return of these signals matter, but so does their frequency of occurrence.

The Kelly criterion is about maximizing expected long-term growth. Many people believe this to be a quite aggressive strategy leading to possible large downturns. There are many other quantities to optimize, of course, so the Kelly criterion might not be the right choice for you. You could, for example, choose to minimize downturn, or maximize some risk-adjusted return, as discussed in the next chapter.

17.6 CAN YOU WIN AT ROULETTE?

Let's go back to Ed Thorp, but a few years before his work on blackjack.

In spring 1955 Ed Thorp was in his second year of graduate physics at UCLA. At tea time one Sunday he got to chatting with colleagues about how to make 'easy money.' The conversation turned to gambling, and roulette in particular. Was it possible to predict, at least with some exploitable degree of accuracy, the outcome of a spin of the wheel? Some of his colleagues, the ones in the know, were certain that the roulette wheels were manufactured so precisely that there were no imperfections that could be discerned, never mind exploited. But Ed's counter to that was simple, if the wheels are so perfect you should be able to predict, using simple Newtonian principles, the path of the ball and its final resting place.

Ed got to work on this problem in the late 1950s, playing around with a cheap miniature roulette wheel, filming and timing the revolutions. He met up with Claude Shannon, the father of information theory in 1959, originally to discuss his blackjack results, but the conversation soon turned to other games and roulette in particular. Shannon was fascinated. Shortly afterwards they met up at Shannon's house, the basement of which was packed with mechanical and engineering gadgets, the perfect playground for further roulette experiments.

Ed and Shannon together took the roulette analysis to greater heights, investing \$1,500 in a full-size professional wheel. They calibrated a simple mathematical model to the experiments, to try to predict the moment when the spinning ball would fall into the waiting pockets. From their model they were able to predict any single number with a standard deviation of 10 pockets. This converts to a 44 per cent edge on a bet on a single number. Betting on a specific octant gave them a 43 per cent advantage.

From November 1960 until June 1961 Ed and Shannon designed and built the world's first wearable computer. The twelve transistors, cigarette-pack sized computer was fed data by switches operated by their big toes. One switch initialized the computer and the other was for timing the rotation of the ball and rotor. The computer predictions were heard by the computer wearer as one of eight tones via an earpiece. (Ed and Shannon decided that the best bet was on octants rather than single numbers since the father of information theory knew that faced with n options individuals take a time $a + b\log(n)$ to make a decision.)

This computer was tested out in Las Vegas in the summer of 1961. But for problems with broken wires and earpieces falling out, the trip was a success. Similar systems were later built for the Wheel of Fortune which had an even greater edge, an outstanding 200 per cent. On 30th May 1985 Nevada outlawed the use of any device for predicting outcomes or analyzing probabilities or strategies.

17.7 HORSE RACE BETTING AND NO ARBITRAGE

Several times we have seen how the absence of arbitrage opportunities leads to the idea of risk-neutral pricing. The value of an option can be interpreted as the present value of the expected payoff, with the expectation being with respect to the risk-neutral asset price path. In this context risk neutral just means that the asset price increases with a growth rate that is the same as the risk-free interest rate. In other words, what we really believe that the asset price is going to do in the future (in terms of its growth rate) is irrelevant. We don't even need to know the growth rate of an asset to price its options, only its volatility.

Something related happens in the world of sports betting.

17.7.1 Setting the Odds in a Sporting Game

In a horse race, football or baseball game the odds are set not to reflect the real probabilities of a horse or a team winning but to reflect the betting that has occurred. Depending on how the betting goes, the odds will be set so that the house/bookie cannot lose. For example, in a soccer match between England and Germany the Germans are more likely to win, but the patriotic English will bet more heavily on England (presumably). The odds given by the bookies will reflect this betting and make it look like England is more likely to win.

Of course, in Germany the situation is reversed. The best bet would be on Germany, but placed in England, and one on England placed in Germany. In practice, however, bookies in one country would lay off their bets on bookies in other countries so all bookies have roughly the same odds. Otherwise there would be straightforward arbitrage opportunities.

Therefore it's unlikely for there to be a sure-fire bet (unless the bookie has made a mistake, the race is fixed, or you can find two or more bookies that aren't directly or indirectly laying off their bets on each other).

But you can win, on average. By exploiting the difference between the real probability of a horse winning and the odds you can get. (There are differences between real odds and what you get paid in all casino games, but it's only in blackjack that this can be exploited.)

17.7.2 The Mathematics

Suppose that there are N horses in a race, with an amount W_i bet on the ith horse. The odds set by the bookie are $q_i : 1$. This means that if you bet 1 on horse i you will lose the 1 if the horse loses, but will take home $q_i + 1$ if the horse wins, your original 1 plus a further q_i. How does the bookie set the odds to ensure he never loses?

The total takings before the race is

$$\sum_{i=1}^{N} W_i.$$

If horse j wins the bookie has to pay out

$$(q_j + 1)W_j.$$

All that the bookie has to do is to ensure that

$$\sum_{i=1}^{N} W_i \geq (q_j + 1)W_j,$$

or equivalently

$$q_j \leq \frac{\sum_{i=1}^{N} W_i}{W_j} - 1 \quad \text{for all } j.$$

Nothing too complicated.

But see how the odds have been chosen to reflect the betting. Nowhere was there any mention of the likelihood of horse j winning.

17.8 ARBITRAGE

Suppose the bookie made an error when setting the odds. How could you determine whether there was an arbitrage opportunity? (Don't forget that only positive bets are allowed, there's no going short here.)

Let's introduce some more notation. The w_i are the bets that you place. (We can forget about the wagers made by everyone else, the Ws.) Let's assume that your total wager is 1, so that

$$\sum_{i=1}^{N} w_i = 1. \tag{17.3}$$

The amount you win is

$$(q_j + 1)w_j \tag{17.4}$$

if horse j is the winner.

Can you find a w_i for all i such that they add up to 1, are all positive and that expression (17.4) is positive for all j? If you can there is an arbitrage opportunity.

The requirement that (17.4) is positive can be written as

$$w_j \geq \frac{1}{q_j + 1}. \tag{17.5}$$

Can we find positive ws such that (17.3) and (17.5) hold? This is very easy to visualize, at least when there are two or three horses. Let's look at the two-horse race.

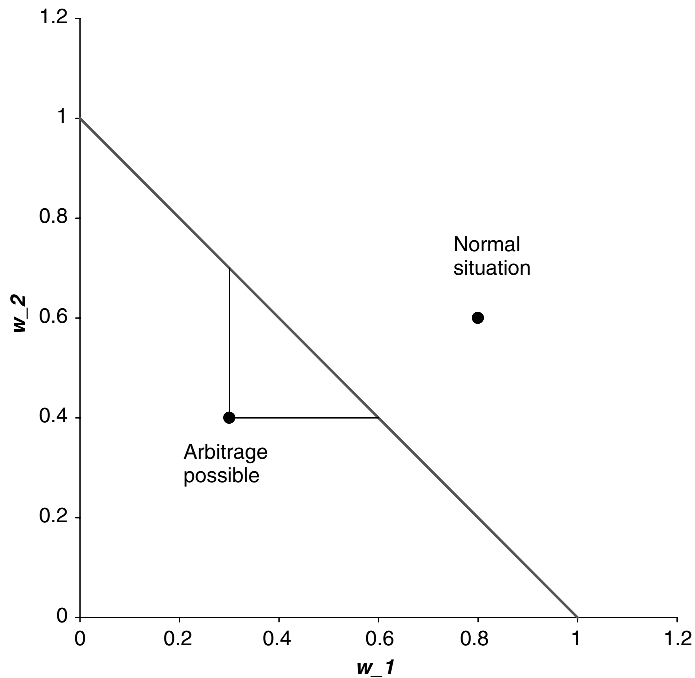


Figure 17.6 Arbitrage in a two-horse race.

In Figure 17.6 the axes represent the amount of the wager on each of the two horses. The line shows the constraint (17.3). The wagers must lie on this line. The two dots mark the point

$$\left(\frac{1}{q_1+1}, \frac{1}{q_2+1}\right) \tag{17.6}$$

in each of two situations. One dot is the typical situation where there is no arbitrage opportunity and the other dot does have an associated arbitrage opportunity. Let's see the details.

To find an arbitrage opportunity we must find a pair (w_1, w_2) lying on the line such that each coordinate is greater than a certain quantity, depending on the qs. Plot the point (17.6) and draw a line vertically up, and another line horizontally to the right, as shown in the figure, emanating from the dot.

Does the quadrant defined by these two lines include any of the line? If not, as would be the case with the higher dot, then there is no arbitrage possible. If some of the line is included then arbitrage is possible.

17.8.1 How Best to Profit from the Opportunity?

There's a simple test to see whether we are in an arbitrage situation. In general, if

$$\sum_{i=1}^{N} \frac{1}{q_i+1} \geq 1$$

then there is no arbitrage. If the sum is less than 1, there is an arbitrage.

You can benefit from the arbitrage by placing wagers w_i such that they lie on the part of the line encompassed by the quadrant. Which part of the line, though, is up to you. By that I mean that you must make some statement about what you are trying to achieve or optimize in the arbitrage. One possibility is to look at the worst-case scenario and maximize the payback in that case. Alternatively, specify real probabilities for each of the horses winning.

17.9 HOW TO BET

We saw how odds are established by bookies. We even saw how to spot arbitrage opportunities. In practice, of course, you could spend a lifetime looking for arbitrage opportunities that rarely occur in real life. Now we are going to see if we can exploit the difference between the odds as set by the bookie and the odds that you estimate. Remember, the odds set by the bookie are really determined by the wagers placed, which are more to do with irrational sentiment ('I'm going to bet on this horse 'cos it's got the same name as the pet rat I had when I was a child') than with a cold-hearted estimation of the probabilities.

We need some more notation. Let's use p_i as the probability of the ith horse winning the race. This is supposed to be the real probability, not the bookie's probability. Obviously, the odds must sum to 1:

$$\sum_{i=1}^{N} p_i = 1.$$

If we wager w_i on the ith horse then we expect to make

$$m = \left(\sum_{i=1}^{N} p_i w_i (q_i + 1) \right) - 1 \tag{17.7}$$

This is under the assumption that the total wager, the sum of all the ws, is 1. An obvious goal is to make this quantity positive; we want to get a positive return on average. But there may be many ways to make this positive. How do we decide which way is best?

Another quantity we might want to look at is the standard deviation of winnings. This is given by

$$\sqrt{\sum_{i=1}^{N} p_i \left(w_i (q_i + 1) - 1 - m \right)^2} \tag{17.8}$$

This measures the dispersion of winnings about the average, and is often interpreted as a measure of risk. If this were zero our profit or loss would be a sure thing.

See Table 17.1 for an example.

How should you bet? The following calculations are easily done on a spreadsheet.

Scenario 1: Maximize expected return

Since you place no premium on reducing risk you should bet everything on the horse that maximizes

$$p_i (q_i + 1).$$

Table 17.1 Odds and probabilities in a horse race.

Horse	Bookie's odds	Your estimate of probability	Wager
Nijinsky	5	0.2	
Red Rum	6	0.2	
Oxo	1	0.1	
Red Marauder	1	0.1	
Gay Lad	2	0.1	
Roquefort	2	0.1	
Red Alligator	2	0.1	
Shergar	2	0.1	

Table 17.2 Maximizing expectation.

Horse	Bookie's odds	Your estimate of probability	Wager
Nijinsky	5	0.2	0
Red Rum	6	0.2	1
Oxo	1	0.1	0
Red Marauder	1	0.1	0
Gay Lad	2	0.1	0
Roquefort	2	0.1	0
Red Alligator	2	0.1	0
Shergar	2	0.1	0

Table 17.3 Minimizing standard deviation.

Horse	Bookie's odds	Your estimate of probability	Wager
Nijinsky	5	0.2	0.063062
Red Rum	6	0.2	0.054068
Oxo	1	0.1	0.189203
Red Marauder	1	0.1	0.189246
Gay Lad	2	0.1	0.126108
Roquefort	2	0.1	0.126108
Red Alligator	2	0.1	0.126108
Shergar	2	0.1	0.126108

In this case, that is Red Rum. The expected return is 40% with a standard deviation of 280%. A very risky bet (see Table 17.2).

Scenario 2: Minimize standard deviation

An interesting strategy.

I say 'interesting' because this strategy results in zero standard deviation, and a return of −62%. In other words, a guaranteed loss! (See Table 17.3).

Table 17.4 Maximize return divided by standard deviation.

Horse	Bookie's odds	Your estimate of probability	Wager
Nijinsky	5	0.2	0.459016
Red Rum	6	0.2	0.540984
Oxo	1	0.1	0
Red Marauder	1	0.1	0
Gay Lad	2	0.1	0
Roquefort	2	0.1	0
Red Alligator	2	0.1	0
Shergar	2	0.1	0

Scenario 3: Maximize return divided by standard deviation

A strategy that seeks to benefit from a positive expectation but with a smaller risk. For mathematical reasons (the Central Limit Theorem) this is a natural strategy. The solution is given in Table 17.4.

The expected return is now 31% with a standard deviation of 164%.

17.10 SUMMARY

The mathematics of gambling is almost identical to the mathematics of 'investing.' The main difference between gambling and investing is that the parameters are usually easier to measure with gambling games. If you can't cope with the mathematics (and the emotional roller coaster ride) of gambling then you shouldn't be working in a bank ;-)

FURTHER READING

- See Kelly's original 1956 paper.
- The classic reference texts on blackjack are by Thorp (1962) and Wong (1981).
- The gripping story of John Kelly, Claude Shannon, Ed Thorp and a cast of many intriguing characters can be found in Poundstone (2005).

CHAPTER 18
portfolio management

In this Chapter...

- Modern Portfolio Theory and the Capital Asset Pricing Model
- optimizing your portfolio
- alternative methodologies such as cointegration
- how to analyze portfolio performance

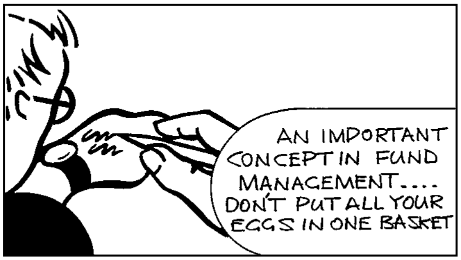
18.1 INTRODUCTION

The theory of derivative pricing is a theory of deterministic returns: We hedge our derivative with the underlying to eliminate risk, and our resulting risk-free portfolio then earns the risk-free rate of interest. Banks make money from this hedging process; they sell something for a bit more than it's worth and hedge away the risk to make a guaranteed profit.

But not everyone is hedging. Fund managers buy and sell assets (including derivatives) with the aim of beating the bank's rate of return. In so doing they take risk. In this chapter I explain some of the theories behind the risk and reward of investment. Along the way I show the benefits of diversification, how the return and risk on a portfolio of assets is related to the return and risk on the individual assets, and how to optimize a portfolio to get the best value for money.

For the most part, the assumptions are as follows.

- We hold a portfolio for 'a single period,' examining the behavior after this time.
- During this period returns on assets are Normally distributed.
- The return on assets can be measured by an expected return (the drift) for each asset, a standard deviation of return (the volatility) for each asset and correlations between the asset returns.



18.2 DIVERSIFICATION

In this section I introduce some more notation, and show the effects of diversification on the return of the portfolio.

We hold a portfolio of N assets. The value today of the ith asset is S_i and its random return is R_i over our time horizon T. The Rs are Normally distributed with mean $\mu_i T$ and standard deviation $\sigma_i\sqrt{T}$. The correlation between the returns on the ith and jth assets is ρ_{ij} (with $\rho_{ii} = 1$). The parameters μ, σ and ρ correspond to the drift, volatility and correlation that we are used to. Note the scaling with the time horizon.

If we hold w_i of the ith asset, then our portfolio has value

$$\Pi = \sum_{i=1}^{N} w_i S_i.$$

At the end of our time horizon the value is

$$\Pi + \delta\Pi = \sum_{i=1}^{N} w_i S_i (1 + R_i).$$

We can write the relative change in portfolio value as

$$\frac{\delta\Pi}{\Pi} = \sum_{i=1}^{N} W_i R_i, \tag{18.1}$$

where

$$W_i = \frac{w_i S_i}{\sum_{i=1}^{N} w_i S_i}.$$

The weights W_i sum to one.

From (18.1) it is simple to calculate the expected return on the portfolio

$$\mu_\Pi = \frac{1}{T} E\left[\frac{\delta\Pi}{\Pi}\right] = \sum_{i=1}^{N} W_i \mu_i \tag{18.2}$$

and the standard deviation of the return

$$\sigma_\Pi = \frac{1}{\sqrt{T}} \sqrt{\mathrm{var}\left[\frac{\delta\Pi}{\Pi}\right]} = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} W_i W_j \rho_{ij} \sigma_i \sigma_j}. \tag{18.3}$$

In these, we have related the parameters for the individual assets to the expected return and the standard deviation of the entire portfolio.

18.2.1 Uncorrelated Assets

Suppose that we have assets in our portfolio that are uncorrelated, $\rho_{ij} = 0,\ i \neq j$. To make things simple assume that they are equally weighted so that $W_i = 1/N$. The expected return on the portfolio is represented by

$$\mu_\Pi = \frac{1}{N}\sum_{i=1}^{N}\mu_i,$$

the average of the expected returns on all the assets, and the volatility becomes

$$\sigma_\Pi = \sqrt{\frac{1}{N^2}\sum_{i=1}^{N}\sigma_i^2}.$$

This volatility is $O(N^{-1/2})$ since there are N terms in the sum. As we increase the number of assets in the portfolio, the standard deviation of the returns tends to zero. It is rather extreme to assume that all assets are uncorrelated but we will see something similar when I describe the Capital Asset Pricing Model below; diversification reduces volatility without hurting expected return.

I am now going to refer to volatility or standard deviation as **risk**, a bad thing to be avoided (within reason), and the expected return as **reward**, a good thing that we want as much of as possible.

18.3 MODERN PORTFOLIO THEORY

We can use the above framework to discuss the 'best' portfolio. The definition of 'best' was addressed very successfully by Nobel Laureate Harry Markowitz. His model provides a way of defining portfolios that are **efficient**. An efficient portfolio is one that has the highest reward for a given level of risk, or the lowest risk for a given reward. To see how this works imagine that there are four assets in the world, A, B, C and D with reward and risk as shown in Figure 18.1 (ignore E for the moment). If you could buy any one of these (but as yet you are not allowed more than one), which would you buy? Would you choose D? No, because it has the same risk as B but less reward. It has the same reward as C but for a higher risk. We can rule out D. What about B or C? They are both appealing when set against D, but against each other it is not so clear. B has a higher risk, but gets a higher reward. However, comparing them both with A we see that there is no contest. A is the preferred choice. If we introduce asset E with the same risk as B and a higher reward than A, then we cannot objectively say which out of A and E is the better; this is a subjective choice and depends on an investor's **risk preferences**.

Now suppose that I have the two assets A and E of Figure 18.2, and I am allowed to combine them in my portfolio, what effect does this have on my risk/reward?

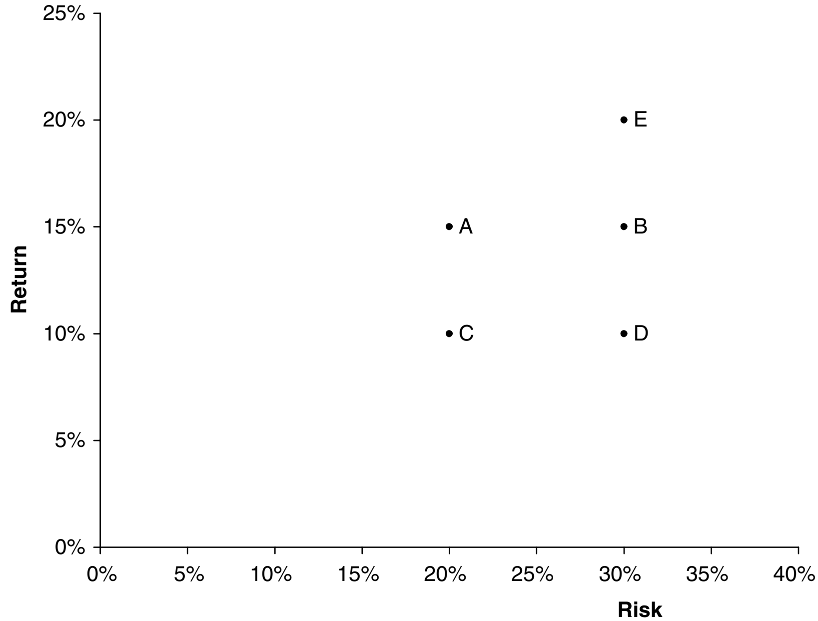


Figure 18.1 Risk and reward for five assets.

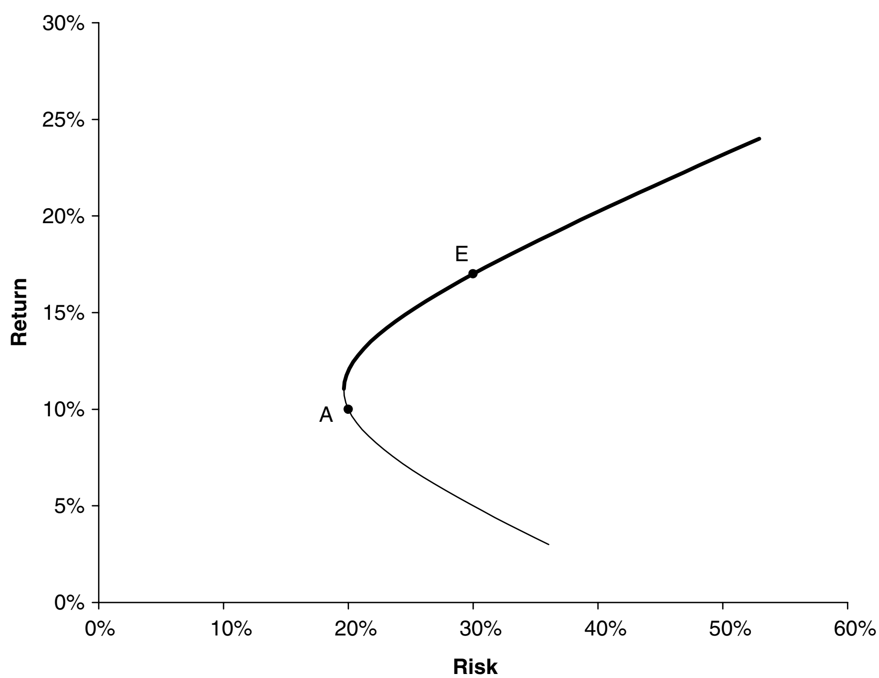


Figure 18.2 Two assets and any combination.

From (18.2) and (18.3) we have

$$\mu_\Pi = W\mu_A + (1 - W)\mu_E$$

and

$$\sigma_\Pi^2 = W^2\sigma_A^2 + 2W(1 - W)\rho\sigma_A\sigma_E + (1 - W)^2\sigma_E^2.$$

Here W is the weight of asset A and, remembering that the weights must add up to one, the weight of asset E is $1 - W$.

As we vary W, so the risk and the reward change. The line in risk/reward space that is parameterized by W is a hyperbola, as shown in Figure 18.2. The part of this curve in bold is efficient, and is preferable to the rest of the curve. Again, an individual's risk preferences will say where he wants to be on the bold curve. When one of the volatilities is zero the line becomes straight. Anywhere on the curve between the two dots requires a long position in each asset. Outside this region, one of the assets is sold short to finance the purchase of the other. Everything that follows assumes that we can sell short as much of an asset as we want. The results change slightly when there are restrictions.

If we have many assets in our portfolio we no longer have a simple hyperbola for our possible risk/reward profiles; instead we get something like that shown in Figure 18.3. This figure now uses *all* of A, B, C, D and E, not just the A and E. Even though B, C and D are not individually appealing they may well be useful in a portfolio, depending how they correlate, or not, with other investments. In this figure we can see the **efficient frontier** marked in bold. Given any choice of portfolio we would choose to hold one that lies on this efficient frontier.

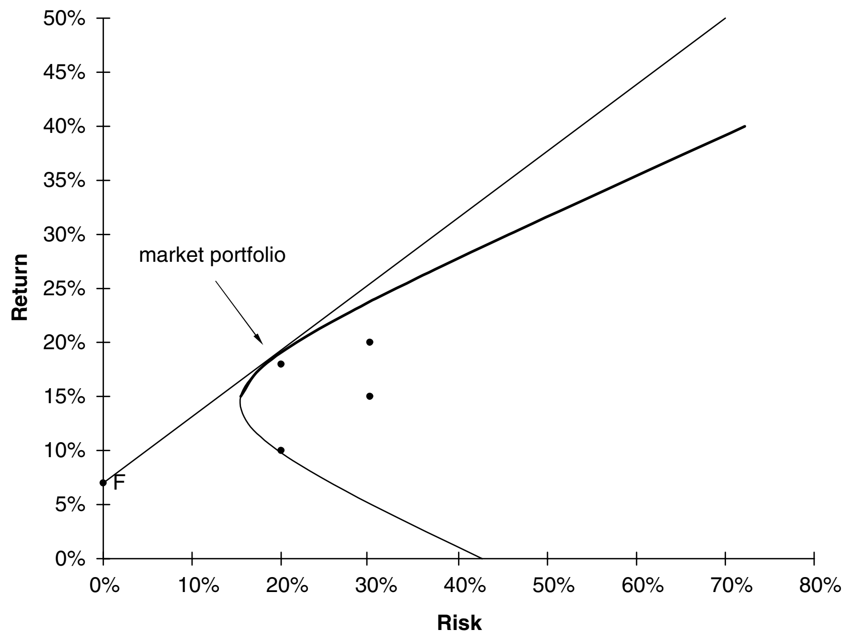


Figure 18.3 Portfolio possibilities and the efficient frontier.

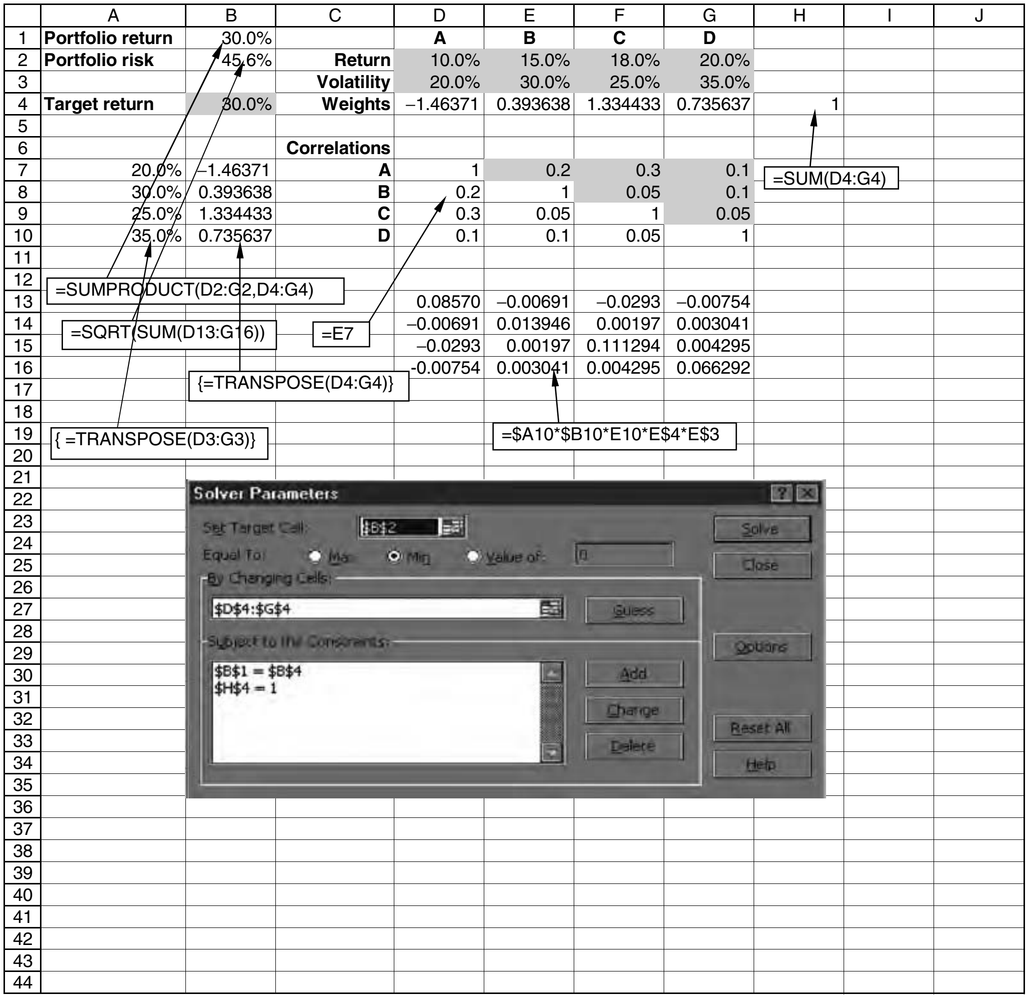


Figure 18.4 Spreadsheet for calculating one point on the efficient frontier.

The calculation of the risk for a given return is demonstrated in the spreadsheet in Figure 18.4. This spreadsheet can be used to find the efficient frontier if it is used many times for different target returns.

18.3.1 Including a Risk-free Investment

A risk-free investment earning a guaranteed rate of return r would be the point F in Figure 18.3. If we are allowed to hold this asset in our portfolio, then since the volatility of this asset is zero, we get the new efficient frontier which is the straight line in the figure. The portfolio for which the straight line touches the original efficient frontier is called the **market portfolio**. The straight line itself is called the **capital market line**.[1]

[1] In the risk-neutral world they think that all investments lie on the horizontal line going through the point $(0, r)$.

18.4 WHERE DO I WANT TO BE ON THE EFFICIENT FRONTIER?

Having found the efficient frontier we want to know whereabouts on it we should be. This is a personal choice, the efficient frontier is objective, given the data, but the 'best' position on it is subjective.

The following is a way of interpreting the risk/reward diagram that may be useful in choosing the best portfolio.

The return on portfolio Π is Normally distributed because it is comprised of assets which are themselves Normally distributed. It has mean μ_Π and standard deviation σ_Π (I have ignored the dependence on the horizon T).

The slope of the line joining the portfolio Π to the risk-free asset is

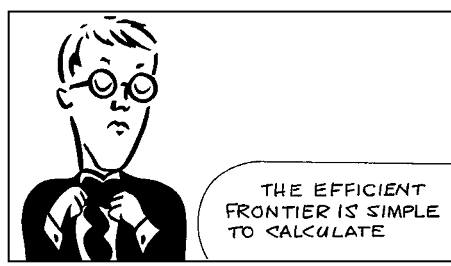
$$s = \frac{\mu_\Pi - r}{\sigma_\Pi}.$$

This is an important quantity; it is a measure of the likelihood of Π having a return that exceeds r. If $C(\cdot)$ is the cumulative distribution function for the standardized Normal distribution then $C(s)$ is the probability that the return on Π is at least r. More generally

$$C\left(\frac{\mu_\Pi - r^*}{\sigma_\Pi}\right)$$

is the probability that the return exceeds r^*. This suggests that if we want to minimize the chance of a return of less than r^* we should choose the portfolio from the efficient frontier set Π_{eff} with the largest value of the slope

$$\frac{\mu_{\Pi_{\text{eff}}} - r^*}{\sigma_{\Pi_{\text{eff}}}}.$$



Conversely, if we keep the slope of this line fixed at s then we can say that with a confidence of $C(s)$ we will lose no more than

$$\mu_{\Pi_{\text{eff}}} - s\sigma_{\Pi_{\text{eff}}}.$$

Our portfolio choice could be determined by maximizing this quantity. These two strategies are shown schematically in Figure 18.5.

Neither of these methods give satisfactory results when there is a risk-free investment among the assets and there are unrestricted short sales, since they result in infinite borrowing.

Another way of choosing the optimal portfolio is with the aid of a **utility function**. This approach is popular with economists. In Figure 18.6 I show **indifference curves** and the efficient frontier. The curves are called by this name because they are meant to represent lines along which the investor is indifferent to the risk/reward trade-off. An investor wants high return, and low risk. Faced with portfolios A and B in the figure, he sees A with low return and low risk, but B has a better reward at the cost of greater risk. The investor is indifferent between these two. However, C is better than both of them, being on a preferred curve.

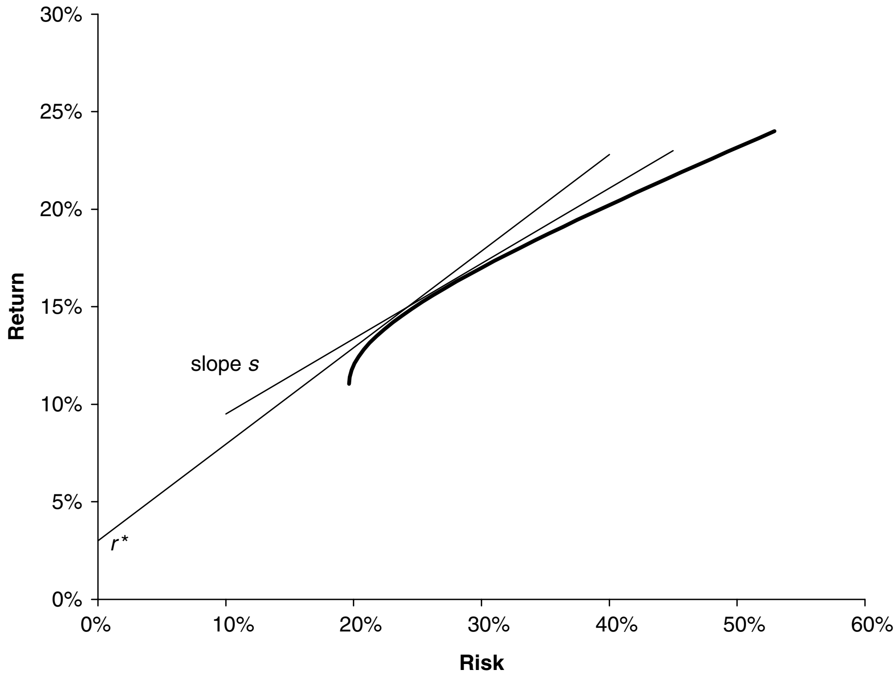


Figure 18.5 Two simple ways for choosing the best efficient portfolio.

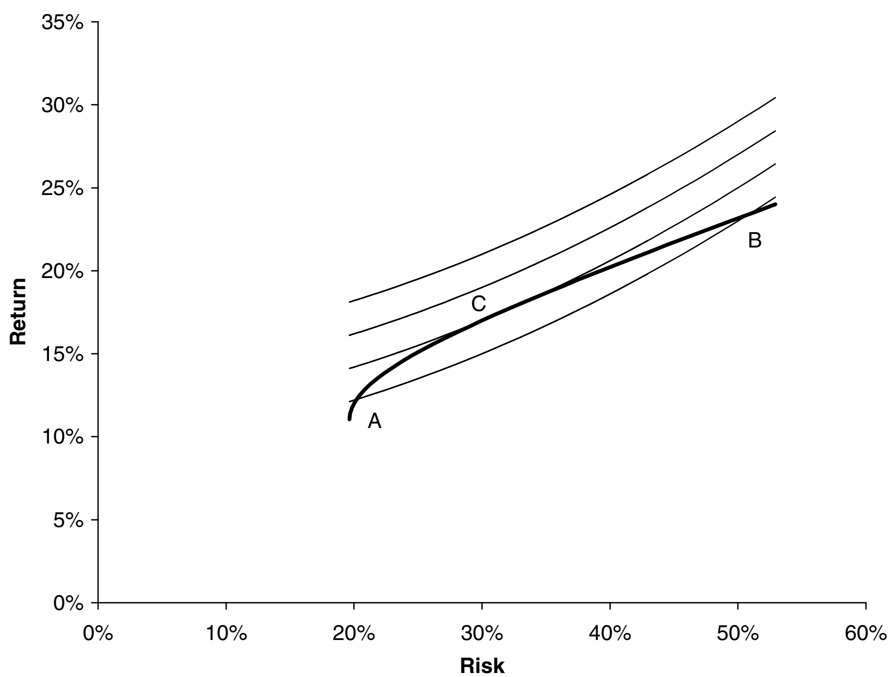
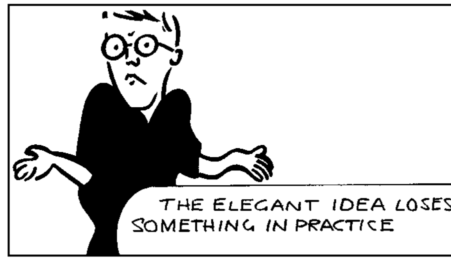


Figure 18.6 The efficient frontier and indifference curves.

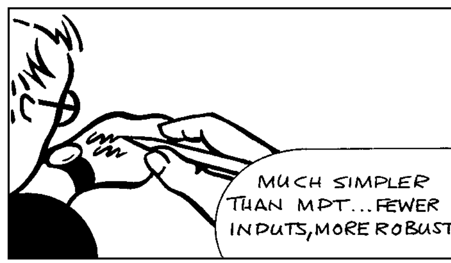
18.5 MARKOWITZ IN PRACTICE



The inputs to the Markowitz model are expected returns, volatilities and correlations. With N assets this means $N + N + N(N-1)/2$ parameters. Most of these cannot be known accurately (do they even exist?); only the volatilities are at all reliable. Having input these parameters, we must optimize over all weights of assets in the portfolio: Choose a portfolio risk and find the weights that make the return on the portfolio a maximum subject to this volatility. This is a very time-consuming process computationally unless one only has a small number of assets.

The problem with the practical implementation of this model was one of the reasons for development of the simpler model of the next section.

18.6 CAPITAL ASSET PRICING MODEL



Before discussing the **Capital Asset Pricing Model** or **CAPM** we must introduce the idea of a security's beta. The beta, β_i, of an asset relative to a portfolio M is the ratio of the covariance between the return on the security and the return on the portfolio to the variance of the portfolio. Thus

$$\beta_i = \frac{\mathrm{Cov}[R_i R_M]}{\mathrm{Var}[R_M]}.$$

18.6.1 The Single-index Model

I will now build up a **single-index model** and describe extensions later. I will relate the return on all assets to the return on a representative index, M. This index is usually taken to be a wide-ranging stock market index in the single-index model. We write the return on the ith asset as

$$R_i = \alpha_i + \beta_i R_M + \epsilon_i.$$

Using this representation we can see that the return on an asset can be decomposed into three parts: A constant drift, a random part common with the index M and a random part uncorrelated with the index, ϵ_i. The random part ϵ_i is unique to the ith asset, and has mean zero. Notice how all the assets are related to the index M but are otherwise completely uncorrelated. In Figure 18.7 is shown a plot of returns on Walt Disney stock against returns on the S&P500; α and β can be determined from a linear regression analysis. The data used in this plot ran from January 1985 until almost the end of 1997.

The expected return on the index will be denoted by μ_M and its standard deviation by σ_M. The expected return on the ith asset is then

$$\mu_i = \alpha_i + \beta_i \mu_M$$

and the standard deviation

$$\sigma_i = \sqrt{\beta_i^2 \sigma_M^2 + e_i^2}$$

where e_i is the standard deviation of ϵ_i.

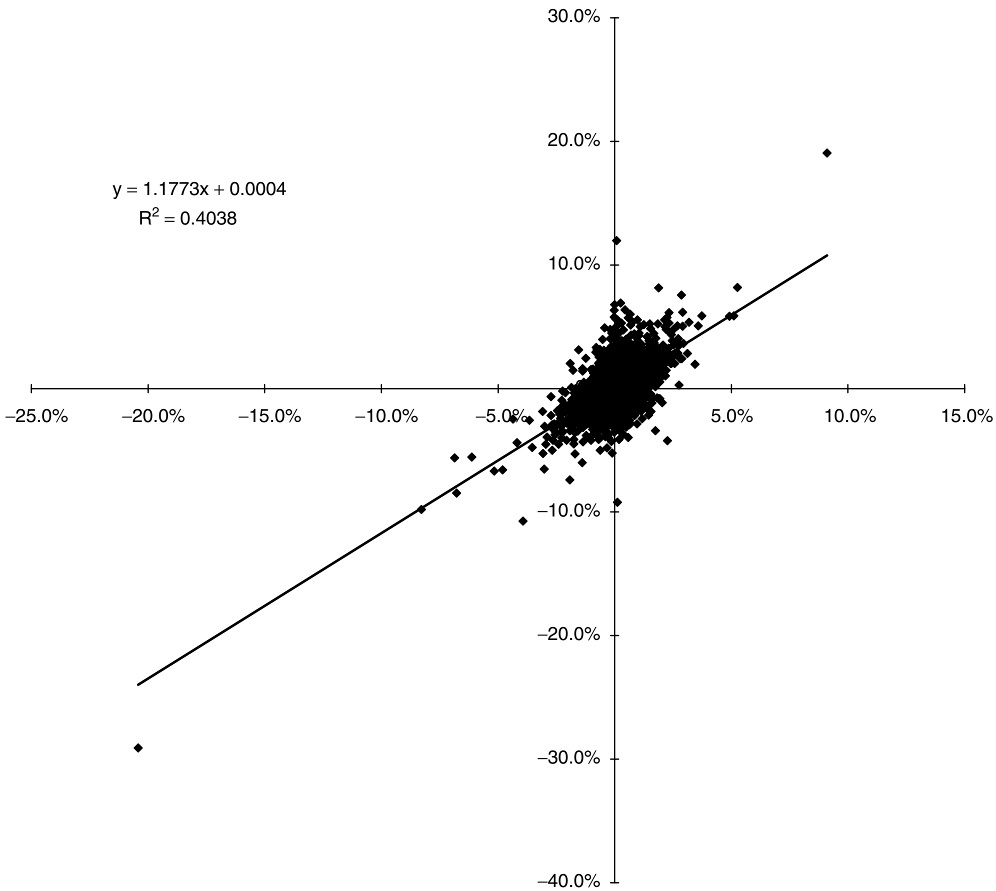


Figure 18.7 Returns on Walt Disney stock against returns on the S&P500.

If we have a portfolio of such assets then the return is given by

$$\frac{\delta\Pi}{\Pi} = \sum_{i=1}^{N} W_i R_i = \left(\sum_{i=1}^{N} W_i \alpha_i\right) + R_M \left(\sum_{i=1}^{N} W_i \beta_i\right) + \sum_{i=1}^{N} W_i \epsilon_i.$$

From this it follows that

$$\mu_\Pi = \left(\sum_{i=1}^{N} W_i \alpha_i\right) + E[R_M] \left(\sum_{i=1}^{N} W_i \beta_i\right).$$

Let us write

$$\alpha_\Pi = \sum_{i=1}^{N} W_i \alpha_i \quad \text{and} \quad \beta_\Pi = \sum_{i=1}^{N} W_i \beta_i,$$

so that

$$\mu_\Pi = \alpha_\Pi + \beta_\Pi E[R_M] = \alpha_\Pi + \beta_\Pi \mu_M.$$

Similarly the risk in Π is measured by

$$\sigma_\Pi = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} W_i W_j \beta_i \beta_j \sigma_M^2 + \sum_{i=1}^{N} W_i^2 e_i^2}.$$

If the weights are all about the same, N^{-1}, then the final terms inside the square root are also $O(N^{-1})$. Thus this expression is, to leading order as $N \to \infty$,

$$\sigma_\Pi = \left|\sum_{i=1}^{N} W_i \beta_i\right| \sigma_M = |\beta_\Pi| \sigma_M.$$

Observe that the contribution from the uncorrelated ϵs to the portfolio vanishes as we increase the number of assets in the portfolio: The risk associated with the ϵs is called **diversifiable risk**. The remaining risk, which is correlated with the index, is called **systematic risk**.

18.6.2 Choosing the Optimal Portfolio

The principal is the same as the Markowitz model for optimal portfolio choice. The only difference is that there are a lot fewer parameters to be input, and the computation is a lot faster.

The procedure is as follows. Choose a value for the portfolio return μ_Π. Subject to this constraint, minimize σ_Π. Repeat this minimization for different portfolio returns to obtain the efficient frontier. The position on this curve is then a subjective choice.

18.7 THE MULTI-INDEX MODEL

The model presented above is a single-index model. The idea can be extended to include further representative indices. For example, as well as an index representing the stock market one might include an index representing bond markets, an index representing currency markets or even an economic index if it is believed to be relevant in linking assets. In the multi-index model we write each asset's return as

$$R_i = \alpha_i + \sum_{j=1}^{n} \beta_{ji} R_j + \epsilon_i,$$

where there are n indices with return R_j. The indices can be correlated to each other. Similar results to the single-index model follow.

It is usually not worth having more than three or four indices. The fewer the parameters, the more robust will be the model. At the other extreme is the Markowitz model with one index per asset.

18.8 COINTEGRATION

Whether you use MPT or CAPM you will always worry about the accuracy of your parameters. Both of these methods are only as accurate as the input data, CAPM being more reliable than MPT generally speaking, because it has fewer parameters.

There is another method that is gaining popularity, and which I will describe here briefly. It is unfortunately a complex technique requiring sophisticated statistical analysis (to do it properly) but which at its core makes a lot of sense. Instead of asking whether two series are correlated we ask whether they are **cointegrated**.

Two stocks may be perfectly correlated over short timescales yet diverge in the long run, with one growing and the other decaying. Conversely, two stocks may follow each other, never being more than a certain distance apart, but with any correlation, positive, negative or varying. If we are delta hedging then maybe the short timescale correlation matters, but not if we are holding stocks for a long time in an unhedged portfolio. To see whether two stocks stay close together we need a definition of **stationarity**. A time series is stationary if it has finite and constant mean, standard deviation and autocorrelation function. Stocks, which tend to grow, are not stationary. In a sense, stationary series do not wander too far from their mean.

We can see the difference between stationary and non-stationary with our first coin-tossing experiment. The time series given by 1 every time we throw a head and -1 every time we throw a tail is stationary. It has a mean of zero, a standard deviation of 1 and an autocorrelation function that is zero for any non-zero lag. But what if we add up the results, as we might do if we are betting on each toss? This time series is non-stationary. This is because the standard deviation of the sum grows like the square root of the number of throws. The mean may be zero but the sum is wandering further and further away from that mean.

Testing for the stationarity of a time series X_t involves a linear regression to find the coefficients a, b and c in

$$X_t = aX_{t-1} + b + ct.$$

If it is found that $|a| > 1$ then the series is unstable. If $-1 \leq a < 1$ then the series is stationary. If $a = 1$ then the series is non-stationary. As with all things statistical, we can only say that our value for a is accurate with a certain degree of confidence. To decide whether we have got a stationary or non-stationary series requires us to look at the Dickey–Fuller statistic to estimate the degree of confidence in our result. From this point on the subject of cointegration gets complicated.

How is this useful in finance? Even though individual stock prices might be non-stationary it is possible for a linear combination (i.e. a portfolio) to be stationary. Can we find λ_i, with $\sum_{i=1}^{N} \lambda_i = 1$, such that

$$\sum_{i=1}^{N} \lambda_i S_i$$

is stationary? If we can, then we say that the stocks are cointegrated.

For example, suppose we find that the S&P500 is cointegrated with a portfolio of 15 stocks. We can then use these 15 stocks to **track the index**. The error in this tracking portfolio will have constant mean and standard deviation, and so should not wander too far from its average. This is clearly easier than using all 500 stocks for the tracking (when, of course, the tracking error would be zero).

We don't have to track the index, we could track anything we want, such as $e^{0.2t}$ to choose a portfolio that gets a 20% return. We could analyze the cointegration properties of two related stocks, Nike and Reebok, for example, to look for relationships. This would be pairs trading. Clearly there are similarities with MPT and CAPM in concepts such as means and standard deviations. The important difference is that cointegration assumes far fewer properties for the individual time series. Most importantly, volatility and correlation do not appear explicitly.

18.9 PERFORMANCE MEASUREMENT

If one has followed one of the asset allocation strategies outlined above, or just traded on gut instinct, can one tell how well one has done? Were the outstanding results because of an uncanny natural instinct, or were the awful results simply bad luck?

The ideal performance would be one for which returns outperformed the risk-free rate, but *in a consistent* fashion. Not only is it important to get a high return from portfolio management, but one must achieve this with as little randomness as possible.

The two commonest measures of 'return per unit risk' are the **Sharpe ratio** of 'reward to variability' and the **Treynor ratio** of 'reward to volatility'. These are defined as follows:

$$\text{Sharpe ratio} = \frac{\mu_\Pi - r}{\sigma_\Pi}$$

and

$$\text{Treynor ratio} = \frac{\mu_\Pi - r}{\beta_\Pi}.$$

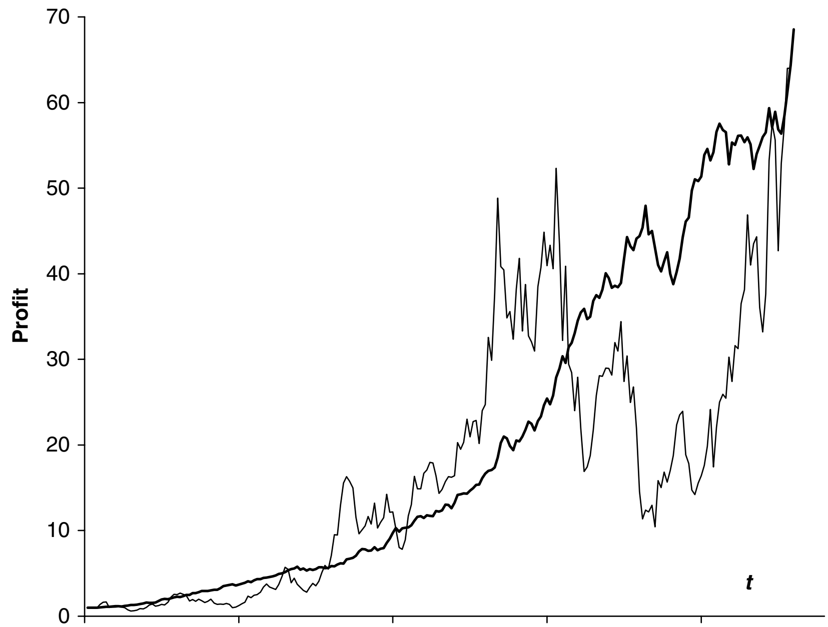


Figure 18.8 A good and a bad manager; same returns, different variability.

In these μ_Π and σ_Π are the *realized* return and standard deviation for the portfolio over the period. The β_Π is a measure of the portfolio's volatility. The Sharpe ratio is usually used when the portfolio is the whole of one's investment and the Treynor ratio when one is examining the performance of one component of the whole firm's portfolio, say. When the portfolio under examination is highly diversified the two measures are the same (up to a factor of the market standard deviation).

In Figure 18.8 we see the portfolio value against time for a good manager and a bad manager.

18.10 SUMMARY

Portfolio management and asset allocation are about taking risks in return for a reward. The questions are how to decide how much risk to take, and how to get the best return. But derivatives theory is based on not taking any risk at all, and so I have spent little time on portfolio management in the book. On the other hand, as I have stressed, there is so much uncertainty in the subject of finance that elimination of risk is well-nigh impossible and the ideas behind portfolio management should be appreciated by anyone involved in derivatives theory or practice. I have tried to give the flavor of the subject with only the easiest-to-explain mathematics; the following sources will prove useful to anyone wanting to pursue the subject further.

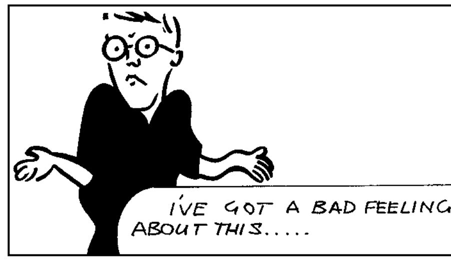
FURTHER READING

- See Markowitz's original book for all the details of MPT, Markowitz (1959).
- One of the best texts on investments, including chapters on portfolio management, is Sharpe (1985).
- For a description of cointegration and other techniques in econometrics see Hamilton (1994) and Hendry (1995).
- See Farrell (1997) for further discussion of portfolio performance.
- I have not discussed the subject of continuous-time asset allocation (yet), but the elegant subject is explained nicely in the collection of Robert Merton's papers, Merton (1992).
- Transaction costs can have a big effect on portfolios that are supposed to be continuously rebalanced. See Morton & Pliska (1995) for a model with costs, and Atkinson & Wilmott (1995), Atkinson, Pliska & Wilmott (1997) and Atkinson & Al–Ali (1997) for asymptotic results.
- For a description of chaos-based methods in finance, and how they won the First International Non-linear Financial Forecasting Competition, see Alexander & Giblin (1997).
- For a review of current thinking in risk management see Alexander (1998).

CHAPTER 19
Value at Risk

In this Chapter...

- the meaning of VaR
- how VaR is calculated in practice
- some of the difficulties associated with VaR for portfolios containing derivatives



19.1 INTRODUCTION

It is the mark of a prudent investor, be they a major bank with billions of dollars' worth of assets or a pensioner with just a few hundred, that they have some idea of the possible losses that may result from the typical movements of the financial markets. Having said that, there have been well-publicized examples where the institution had no idea what might result from some of their more exotic transactions, often involving derivatives.

As part of the search for more transparency in investments, there has grown up the concept of Value at Risk as a measure of the possible downside from an investment or portfolio.

19.2 DEFINITION OF VALUE AT RISK

One of the definitions of **Value at Risk** (VaR), and the definition now commonly intended, is the following.

> Value at Risk is an estimate, with a given degree of confidence, of how much one can lose from one's portfolio over a given time horizon.

The portfolio can be that of a single trader, with VaR measuring the risk that he is taking with the firm's money, or it can be the portfolio of the entire firm. The former measure will be of interest in calculating the trader's efficiency and the latter will be of interest to the owners of the firm who will want to know the likely effect of stock market movements on the bottom line.

The degree of confidence is typically set at 95%, 97.5%, 99% etc. The time horizon of interest may be one day, say, for trading activities, or months for portfolio management. It is supposed to be the timescale associated with the orderly liquidation of the portfolio, meaning the sale of assets at a sufficiently low rate for the sale to have little effect on the market. Thus the VaR is an estimate of a loss that can be realized, not just a 'paper' loss.

As an example of VaR, we may calculate (by the methods to be described here) that over the next week there is a 95% probability that we will lose no more than \$10 m. We can write this as

$$\text{Prob}\left\{\delta V \le -\$10\,\text{m}\right\} = 0.05,$$

where δV is the change in the portfolio's value. (I use $\delta\cdot$ for 'the change in' to emphasize that we are considering changes over a finite time.) In symbols,

$$\text{Prob}\left\{\delta V \le -\text{VaR}\right\} = 1 - c,$$

where the degree of confidence is c, 95% in the above example.

VaR is calculated assuming normal market circumstances, meaning that extreme market conditions such as crashes are not considered, or are examined separately. Thus, effectively, VaR measures what can be expected to happen during the day-to-day operation of an institution.

The calculation of VaR requires at least having the following data: The current prices of all assets in the portfolio and their volatilities and the correlations between them. If the assets are traded we can take the prices from the market (**marking to market**). For OTC contracts we must use some 'approved' model for the prices, such as a Black–Scholes-type model; this is then **marking to model**. Usually, one assumes that the movement of the components of the portfolio are random and drawn from Normal distributions. I make that assumption here, but make a few general comments later on.

19.3 VaR FOR A SINGLE ASSET

Let us begin by estimating the VaR for a portfolio consisting of a single asset.

We hold a quantity Δ of a stock with price S and volatility σ. We want to know with 99% certainty what is the maximum we can lose over the next week. I am deliberately using notation similar to that from the derivatives world.

In Figure 19.1 is the distribution of possible returns over the time horizon of one week. How do we calculate the VaR? First of all we are assuming that the distribution is Normal. Since the time horizon is so small, we can reasonably assume that the mean is zero. The standard deviation of the stock price over this time horizon is

$$\sigma S \left(\frac{1}{52}\right)^{1/2},$$

since the timestep is $1/52$ of a year. Finally, we must calculate the position of the extreme left-hand tail of this distribution corresponding to $1\% = (100 - 99)\%$ of the events. We only need do this for the standardized Normal distribution because we can get to any other Normal distribution by scaling. Referring to Table 19.1, we see that the 99% confidence interval corresponds to

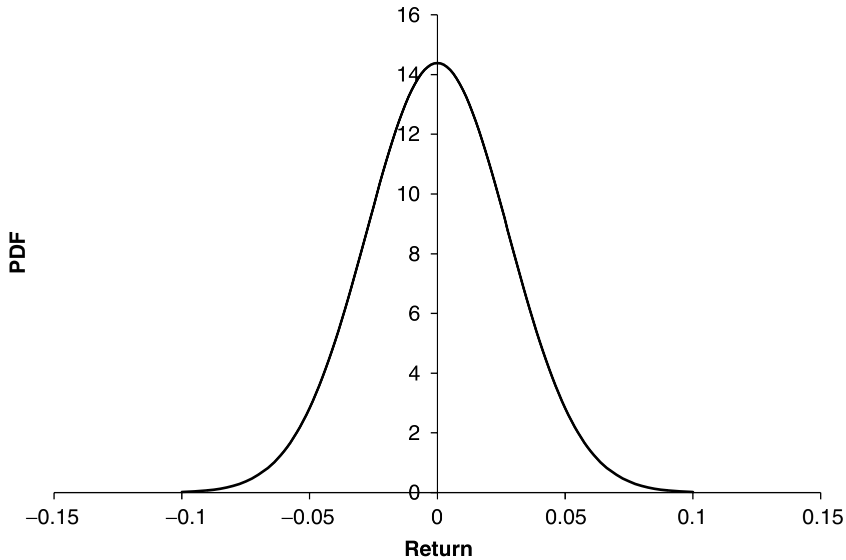


Figure 19.1 The distribution of future stock returns.

Table 19.1 Degree of confidence and the relationship with deviation from the mean.

Degree of confidence (%)	Number of standard deviations from the mean
99	2.326342
98	2.053748
97	1.88079
96	1.750686
95	1.644853
90	1.281551

2.33 standard deviations from the mean. Since we hold a number Δ of the stock, the VaR is given by $2.33\sigma\,\Delta S(1/52)^{1/2}$.

More generally, if the time horizon is δt and the required degree of confidence is c, we have

$$\mathrm{VaR} = -\sigma\,\Delta S(\delta t)^{1/2}\alpha(1-c), \tag{19.1}$$

where $\alpha(\cdot)$ is the inverse cumulative distribution function for the standardized Normal distribution, shown in Figure 19.2.

In (19.1) we have assumed that the return on the asset is Normally distributed *with a mean of zero*. The assumption of zero mean is valid for short time horizons: The standard deviation of the return scales with the square root of time but the mean scales with time itself. For longer time horizons, the return is shifted to the right (one hopes) by an amount proportional to the time horizon. Thus for longer timescales, expression (19.1) should be modified to account for

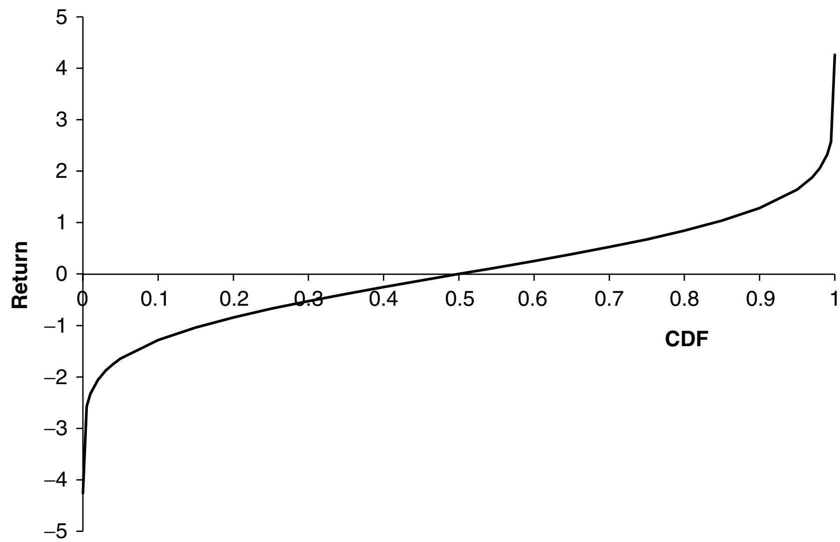
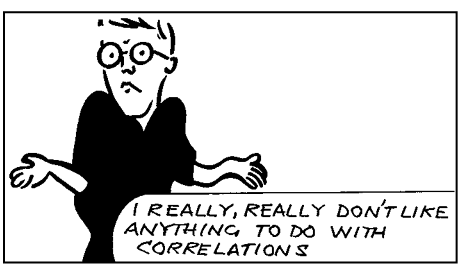


Figure 19.2 The inverse cumulative distribution function for the standardized Normal distribution.

the drift of the asset value. If the rate of this drift is μ then (19.1) becomes

$$\mathrm{VaR} = \Delta S\left(\mu\ \delta t - \sigma\ \delta t^{1/2}\alpha(1-c)\right).$$

Note that I use the *real* drift rate and not the *risk-neutral*. I shall not worry about this drift adjustment for the rest of this chapter.



19.4 VaR FOR A PORTFOLIO

If we know the volatilities of all the assets in our portfolio and the correlations between them then we can calculate the VaR for the whole portfolio.

If the volatility of the ith asset is σ_i and the correlation between the ith and jth assets is ρ_{ij} (with $\rho_{ii} = 1$), then the VaR for the portfolio consisting of M assets with a holding of Δ_i of the ith asset is

$$-\alpha(1-c)\delta t^{1/2}\sqrt{\sum_{j=1}^{M}\sum_{i=1}^{M}\Delta_i\Delta_j\sigma_i\sigma_j\rho_{ij}S_iS_j}.$$

You will recognize this, apart from the multiplicative factor at the front, as being the same as the standard deviation of a stock portfolio's return from Chapter 18.

Several obvious criticisms can be made of this definition of VaR:[1] Returns are not Normal, volatilities and correlations are notoriously difficult to measure, and it does

[1] VaR is like your star sign. You have to tell your investors/risk people what your VaR is, just like you have to tell the person you are chatting up your star sign. You don't necessarily believe there is meaning in either.

not allow for derivatives in the portfolio. We discuss the first criticism later; I now describe in some detail ways of incorporating derivatives into the calculation.

19.5 VaR FOR DERIVATIVES

The key point about estimating VaR for a portfolio containing derivatives is that, even if the change in the underlying *is* Normal, the essential non-linearity in derivatives means that the change in the derivative can be far from Normal. Nevertheless, if we are concerned with very small movements in the underlying, for example over a very short time horizon, we may be able to approximate for the sensitivity of the portfolio to changes in the underlying by the option's delta. For larger movements we may need to take a higher-order approximation. We see these approaches and pitfalls next.

19.5.1 The Delta Approximation

Consider for a moment a portfolio of derivatives with a single underlying, S. The sensitivity of an option, or a portfolio of options, to the underlying is the delta, Δ. If the standard deviation of the distribution of the underlying is $\sigma S\ \delta t^{1/2}$ then the standard deviation of the distribution of the option position is

$$\sigma S\ \delta t^{1/2}\Delta.$$

Δ must here be the delta of the whole position, the sensitivity of all of the relevant options to the particular underlying, i.e. the sum of the deltas of all the option positions on the same underlying.

It is but a small, and obvious, step to the following estimate for the VaR of a portfolio containing options:

$$-\alpha(1-c)\delta t^{1/2}\sqrt{\sum_{j=1}^{M}\sum_{i=1}^{M}\Delta_i\Delta_j\sigma_i\sigma_j\rho_{ij}S_iS_j}.$$

Here Δ_i is the rate of change of the *portfolio* with respect to the ith asset.

19.5.2 Which Volatility Do I Use?

For a single underlying, the delta approximation to VaR depends on the standard deviation

$$\sigma S\ \delta t^{1/2}\Delta. \tag{19.2}$$

As we've seen many times, there are different types of volatility.[2] So which volatility goes into this formula?

Suppose you have an estimate for actual volatility, σ, and it differs from implied volatility, $\tilde{\sigma}$. Which goes into expression (19.2)? The answer is both. The delta represents how the option value will vary as the stock price varies, and this is governed by the market's pricing of the options. Therefore the delta must be the usual formula using $\tilde{\sigma}$. However, the σ in front represents the movement in the stock, its real movement, and should therefore be the actual volatility. Subtle.

[2] ... fortunately or unfortunately depending on whether you make money from the differences.

19.5.3 The Delta-Gamma Approximation

The delta approximation is satisfactory for small movements in the underlying. A better approximation may be achieved by going to higher order and incorporating the gamma or convexity effect.

I demonstrate this by example. Suppose that our portfolio consists of an option on a stock. The relationship between the change in the underlying, δS, and the change in the value of the option, δV, is

$$\delta V = \frac{\partial V}{\partial S}\delta S + \tfrac{1}{2}\frac{\partial^2 V}{\partial S^2}(\delta S)^2 + \frac{\partial V}{\partial t}\delta t + \cdots.$$

Since we are assuming that

$$\delta S = \mu S\, \delta t + \sigma S\, \delta t^{1/2}\phi,$$

where ϕ is drawn from a standardized Normal distribution, we can write

$$\delta V = \frac{\partial V}{\partial S}\sigma S\, \delta t^{1/2}\phi + \delta t\left(\frac{\partial V}{\partial S}\mu S + \tfrac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2\phi^2 + \frac{\partial V}{\partial t}\right) + \cdots.$$

This can be rewritten as

$$\delta V = \Delta\sigma S\, \delta t^{1/2}\phi + \delta t\left(\Delta\mu S + \tfrac{1}{2}\Gamma\sigma^2 S^2\phi^2 + \Theta\right) + \cdots. \tag{19.3}$$

To leading order, the randomness in the option value is simply proportional to that in the underlying. To the next order there is a deterministic shift in δV due to the deterministic drift of S and the theta of the option. More importantly, however, the effect of the gamma is to introduce a term that is non-linear in the random component of δS.

In Figure 19.3 are shown three pictures. First, there is the assumed distribution for the change in the underlying. This is a Normal distribution with standard deviation $\sigma S\, \delta t^{1/2}$, drawn in bold in the figure. Second, is shown the distribution for the change in the option assuming the delta approximation only. This is a Normal distribution with standard deviation $\Delta\sigma S\, \delta t^{1/2}$. Finally, there is the distribution for the change in the underlying assuming the delta/gamma approximation.

From this figure we can see that the distribution for the delta/gamma approximation is far from Normal. In fact, because expression (19.3) is quadratic in ϕ, δV must satisfy the following constraint

$$\delta V \geq -\frac{\Delta^2}{2\Gamma} \quad \text{if} \quad \Gamma > 0$$

or

$$\delta V \leq -\frac{\Delta^2}{2\Gamma} \quad \text{if} \quad \Gamma < 0.$$

The extreme value is attained when

$$\phi = -\frac{\Delta}{\sigma S\Gamma\, \delta t^{1/2}}.$$

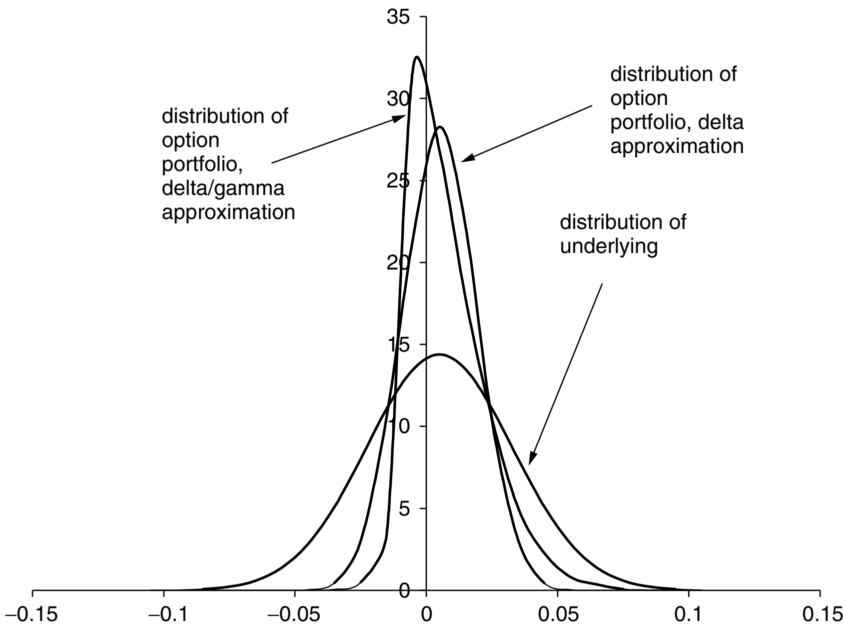


Figure 19.3 A Normal distribution for the change in the underlying (bold), the distribution for the change in the option assuming the delta approximation (another Normal distribution) and the distribution for the change in the option assuming the delta/gamma approximation (definitely not a Normal distribution).

The question to ask is then: 'Is this critical value for ϕ in the part of the tail in which we are interested?' If it is not then the delta approximation may be satisfactory, otherwise it will not be. If we cannot use an approximation we may have to run simulations using valuation formulae.

One obvious conclusion to be drawn is that positive gamma is good for a portfolio and negative gamma is bad. With a positive gamma the downside is limited, but with a negative gamma it is the upside that is limited.

19.5.4 Use of Valuation Models

The obvious way around the problems associated with non-linear instruments is to use a simulation for the random behavior of the underlyings and then use valuation formulae or algorithms to deduce the distribution of the changes in the whole portfolio. This is the ultimate solution to the problem but has the disadvantage that it can be very slow. After all, we may want to run tens of thousands of simulations but if we must solve a multifactor partial differential equation each time then we find that it will take far too long to calculate the VaR.

19.5.5 Fixed-income Portfolios

When the asset or portfolio has interest rate dependence then it is usual to treat the yield to maturity on each instrument as the Normally distributed variable. Yields on different instruments are then suitably correlated. The relationship of price to change in yield is via duration (and convexity at higher order). So our fixed-income asset can be thought of as a derivative of the yield. The VaR is then estimated using duration in place of delta (and convexity in place of gamma) in the obvious way.

19.6 SIMULATIONS

The two simulation methods described in this book are **Monte Carlo**, based on the generation of Normally distributed random numbers, and **bootstrapping** using actual asset price movements taken from historical data.

Within these two simulation methods, there are two ways to generate future scenarios, depending on the timescale of interest and the timescale for one's model or data. If one is interested in a horizon of one year and one has a model or data for returns with this same horizon, then this is easily used to generate a distribution of future scenarios. On the other hand, if the model or data is for a shorter timescale, a stochastic differential equation or daily data, say, and the horizon is one year, then the model must be used to build up a one-year distribution by generating whole year-long paths of the asset. This is more time consuming but is important for path-dependent contracts when the whole path taken must obviously be modeled.

Remember, the simulation must use *real* returns and not *risk-neutral*.

19.6.1 Monte Carlo

Monte Carlo simulation is the generation of a distribution of returns and/or asset price paths by the use of random numbers. This subject is discussed in great depth in Chapter 80. The technique can be applied to VaR using numbers, ϕ, drawn from a Normal distribution, to build up a distribution of future scenarios. For each of these scenarios use some pricing methodology to calculate the value of a portfolio (of the underlying asset and its options) and thus directly estimate its VaR.

19.6.2 Bootstrapping

Another method for generating a series of random movements in assets is to use historical data. Again, there are two possible ways of generating future scenarios: A one-step procedure if you have a model for the distribution of returns over the required time horizon, or a multi-step procedure if you only have data/model for short periods and want to model a longer time horizon.

The data that we use will consist of daily returns, say, for all underlying assets going back several years. The data for each day are recorded as a vector, with one entry per asset.

Suppose we have real time-series data for N assets and that our data are daily data stretching back four years, resulting in approximately 1000 daily *returns* for each asset. We are going to use these returns for simulation purposes. This is done as follows.

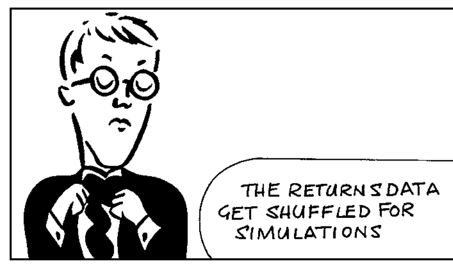
Assign an 'index' to each daily change. That is, we assign 1000 numbers, one for each vector of returns. To visualize this, imagine writing the returns for all of the N assets on the blank pages of a notebook. On page 1 we write the changes in asset values that occurred from 8th July 1998 to 9th July 1998. On page 2 we do the same, but for the changes from 9th July to 10th July 1998. On page 3 ... from 10th to 11th July etc. We will fill 1000 pages if we have

1000 datasets. Now, draw a number from 1 to 1000, uniformly distributed; it is 534. Go to page 534 in the notebook. Change all assets from today's value by the vector of returns given on the page. Now draw another number between 1 and 1000 at random and repeat the process. Increment this new value again using one of the vectors. Continue this process until the required time horizon has been reached. This is one realization of the path of the assets. Repeat this simulation to generate many, many possible realizations to get an accurate distribution of all future prices.

By this method we generate a distribution of possible future scenarios based on historical data.

Note how we keep together all asset changes that happen on a certain date. By doing this we ensure that we capture any correlation that there may be between assets.

This method of bootstrapping is very simple to implement. The advantages of this method are that it naturally incorporates any correlation between assets, and any non-Normality in asset price changes. It does not capture any autocorrelation in the data, but then neither does a Monte Carlo simulation in its basic form. The main disadvantage is that it requires a lot of historical data that may correspond to completely different economic circumstances than those that currently apply.



In Figure 19.4 is shown the daily historical returns for several stocks and the 'index' used in the random choice.

19.7 USE OF VaR AS A PERFORMANCE MEASURE

One of the uses of VaR is in the measurement of performance of banks, desks or individual traders. In the past, 'trading talent' has been measured purely in terms of profit; a trader's bonus is related to that profit. This encourages traders to take risks; think of tossing a coin with you receiving a percentage of the profit but without the downside (which is taken by the bank), how much would you bet? A better measure of trading talent might take into account the risk in such a bet, and reward a good return-to-risk ratio. The ratio

$$\frac{\text{Return in excess of risk-free}}{\text{volatility}} = \frac{\mu - r}{\sigma},$$

the **Sharpe ratio**, is such a measure. Alternatively, use VaR as the measure of risk and profit/loss as the measure of return:

$$\frac{\text{daily P\&L}}{\text{daily VaR}}.$$

19.8 INTRODUCTORY EXTREME VALUE THEORY

More modern techniques for estimating tail risk use **Extreme Value Theory**. The idea is to represent more accurately the outer limits of returns distributions since this is where the most important risk is. Throw Normal distributions away; their tails are far too thin to capture the frequent market crashes (and rallies). I won't go into the details here; a whole book could be written on this subject (and has been, Embrechts, Klüppelberg and Mikosch 1997, a very good book).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	**Index**	**Prob.**	**TELEBRAS**	**ELETROBRAS**	**PETROBRAS**	**CVDR**	**USIMINAS**	**YPF**	**TAR**	**TEO**	**TGS**	**PEREZ**	**TELMEX**	**TELEVISA**
2	1	0.001	–0.0152210	0.0180185	0.0118345	–0.0240975	–0.0111733	0.0355909	0.0185764	0.0071943	0.0121214	–0.0182190	–0.0235305	–0.0153849
3	2	0.001	0.0091604	0.0072214	–0.0046130	0.0001039	–0.0226244	0.0207620	0.0121953	0.0106953	0.0000000	–0.0026393	0.0327898	0.0746615
4	3	0.001	–0.0498546	–0.0397883	–0.0324353	–0.0246926	–0.0115608	0.0068260	–0.0121953	–0.0106953	0.0119762	–0.0067506	–0.0139213	–0.0144930
5	4	0.001	–0.0357762	–0.0494712	–0.0358569	–0.1001323	–0.0602100	–0.0068260	–0.0123458	–0.0254097	0.0000000	–0.0144520	–0.0284379	-0.0371790
6	5	0.001	0.0033058	–0.0204013	0.0112413	0.0264184	0.0114388	–0.0068729	0.0061920	0.0036697	–0.0240976	0.0096238	0.0000000	0.0000000
7	6	0.001	0.0424306	0.0260111	–0.0255043	–0.0022909	0.0004155	0.0067340	0.0452054	0.0438519	0.0115608	0.0140863	0.0421608	0.0932575
8	7	0.001	0.0279317	0.0006216	0.0006216	0.0006216	–0.0234759	0.0040080	–0.0208341	0.0000000	–0.0114287	0.0000000	0.0123205	–0.0065147
9	8	0.001	–0.0139801	–0.0383626	–0.0346393	0.0161040	0.0122250	0.0000000	0.0259755	0.0176996	0.0114287	–0.0020299	0.0439192	0.0384663
10	9	0.001	–0.0052219	–0.0036563	0.0403377	0.0142520	0.0229896	0.0063898	0.0101524	0.0085349	0.0114287	–0.0019638	0.0000000	0.0171924
11	10	0.001	–0.0130693	0.0000000	–0.0116961	–0.0047282	0.0112996	0.0000000	0.0100503	0.0112677	0.0112996	0.0081304	–0.0267702	–0.0229895
12	11	0.001	–0.0106656	–0.0148639	0.0069306	–0.0048541	–0.0114030	0.0000000	–0.0100503	–0.0169496	–0.0112996	–0.0166294	–0.0116960	0.0057971
13	12	0.001	0.0079035	0.0184167	–0.0023392	–0.0240975	0.0112996	0.0000000	–0.0050633	–0.0085837	0.0335227	–0.0170073	0.0039139	0.0228581
14	13	0.001	0.0000000	–0.0072225	–0.0284021	0.0001035	–0.0226248	0.0000000	–0.0257084	–0.0380719	–0.0110498	–0.0063762	0.0000000	0.0112361
15	14	0.001	0.0000000	–0.0042230	0.0242792	–0.0004134	–0.0126835	–0.0061920	0.0100001	0.0057471	0.0109291	0.0144183	0.0411581	0.0745331
16	15	0.001	–0.0475377	–0.0138620	–0.0251059	–0.0444106	–0.0254146	–0.0062305	–0.0150379	–0.0057471	0.0000000	0.0300138	–0.0080972	0.0000000
17	16	0.001	0.0297428	0.0102652	0.0114445	0.0714950	–0.0008256	0.0415490	0.0277796	0.0309303	0.0408220	–0.0035537	0.0419109	0.0475023
18	17	0.001	–0.0037045	–0.0192060	–0.0248187	0.0078541	–0.0129914	–0.0116280	–0.0045977	0.0052771	0.0000000	0.0007187	–0.0075758	–0.0105264
19	18	0.001	–0.0208341	–0.0205823	–0.0264421	–0.0021947	–0.0267507	0.0231224	0.0091744	–0.0052771	–0.0304592	0.0152095	–0.0269766	–0.0268113
20	19	0.001	0.0568874	0.0828276	0.1052577	0.0477150	0.0612661	0.0271019	0.0753494	0.0746435	0.0388398	0.0599762	0.0498324	0.0438026
21	20	0.001	0.0655911	0.0209154	0.0209922	0.0420456	0.0224756	–0.0173917	0.0165293	0.0223058	0.0000000	0.0232360	0.0036036	0.0194181
22	21	0.001	–0.0091109	0.0309987	0.0615672	–0.0428678	0.0221646	0.0058309	0.0122201	0.0145988	0.0392207	0.0139353	–0.0592766	–0.0342332
23	22	0.001	0.0231470	–0.0036810	–0.0003084	0.0158210	0.0099482	0.0342891	0.0160646	0.0264121	0.0204089	0.0208262	0.0679507	0.0386522
24	23	0.001	0.0157998	–0.0014713	–0.0122804	–0.0094697	–0.0014375	0.0277026	0.0504962	0.0440396	0.0492710	0.0285456	–0.0073260	0.0280392
25	24	0.001	–0.0045147	–0.0054003	–0.0008614	–0.0020500	0.0379553	–0.0054795	–0.0191210	–0.0136988	0.0000000	0.0101176	–0.0073801	–0.0233111
26	25	0.001	0.0270287	0.0275012	0.0583711	0.0001022	0.0361421	0.0280130	–0.0038536	0.0069849	0.0000000	0.0034104	–0.0109690	–0.0432968
27	26	0.001	-0.0089286	0.0068362	0.0038686	0.0248968	–0.0087868	0.0000000	–0.0116506	–0.0093241	0.0000000	0.0110554	–0.0036832	–0.0044346
28	27	0.001	0.0133632	0.0232183	0.0512933	0.0202027	0.0436750	–0.0167135	–0.0117880	–0.0285055	–0.0099503	–0.0080555	0.0000000	0.0088496
29	28	0.001	–0.0110341	0.0417395	0.0176995	0.0506932	0.0085108	0.0000000	–0.0240012	–0.0268636	0.0000000	0.0005324	0.0000000	0.0043956
30	29	0.001	–0.0044544	–0.0289144	0.0069770	–0.0002423	–0.0002043	–0.0056023	0.0000000	0.0024907	0.0000000	–0.0037901	0.0000000	–0.0277796
31	30	0.001	0.0044544	0.0001022	–0.0344781	–0.0075315	0.0086859	0.0056023	0.0202027	0.0049628	0.0204089	0.0147792	0.0038241	0.0277796
32	31	0.001	–0.0044544	–0.0065956	–0.0225751	0.0151071	–0.0173439	0.0111112	–0.0040080	0.0024722	0.0396091	0.0090663	–0.0115164	–0.0184337
33	32	0.001	0.0088685	–0.0067409	0.0284793	0.0204063	0.0261791	0.0096270	0.0043197	–0.0052632	0.0000000	0.0184492	–0.0076923	0.1048796
34	33	0.001	–0.0178163	–0.0084174	0.0036982	0.0183391	–0.0186661	–0.0242436	–0.0350913	–0.0240332	–0.0202027	–0.0280518	–0.0234386	–0.1242977
35	34	0.001	–0.0205219	–0.0174839	–0.0199091	–0.0849882	–0.0120831	–0.0061539	–0.0134834	–0.0108697	0.0000000	–0.0075047	0.0039448	0.0097562
36	35	0.001	–0.0092167	0.0172806	0.0323986	0.0769500	–0.0145534	–0.0313505	–0.0368705	–0.0221003	0.0202027	–0.0096608	–0.0320027	–0.0600180
37	36	0.001	–0.0256977	–0.0470936	–0.0464976	–0.0325785	–0.0460766	–0.0128207	–0.0189579	–0.0340942	-0.0100503	–0.0547012	–0.0081633	0.0102565
38	37	0.001	0.0303167	0.0398480	0.0188954	0.0204067	0.0278547	0.0625204	0.0421608	0.0396653	0.0000000	0.0471675	0.0322609	0.0497615
39	38	0.001	0.0000000	0.0178185	0.0003050	0.0122735	0.0093548	0.0000000	0.0091325	0.0082988	0.0298530	0.0320480	0.0039604	–0.0097562
40	39	0.001	–0.0113079	–0.0069686	0.0628009	–0.0039604	0.0266682	0.0180185	–0.0091325	–0.0109690	0.0194181	0.0310527	–0.0039604	–0.0148151
41	40	0.001	–0.0479236	–0.0464835	–0.0264332	0.0000000	–0.0266682	–0.0180185	–0.0279088	–0.0198027	–0.0096619	–0.0065198	–0.0119762	–0.0150379
42	41	0.001	–0.0374621	–0.0690578	–0.0427517	–0.0537785	–0.0563828	0.0000000	–0.0094787	–0.0229895	–0.0196085	–0.0109291	–0.0161947	0.0050378

Figure 19.4 Spreadsheet showing bootstrap data.

19.8.1 Some EVT Results

Distribution of maxima/minima

If X_i are independent, identically distributed random variables and

$$x = \max(X_1, X_2, \dots, X_n)$$

then the distribution of x converges to

$$\exp\left(-\left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi}\right).$$

When $\xi = 0$ this is a Gumbel distribution, when $\xi < 0$ it is a Weibull and when $\xi > 0$ a Frechet. Frechet is the one of interest in finance because it is associated with fat tails.

Peaks over threshold

Consider the probability that loss exceeds u by an amount y (given that the threshold u *has* been exceeded):

$$F_u(y) = P(X - u \le y | X > u).$$

This can be approximated by a Generalized Pareto Distribution:

$$1 - \left(1 + \frac{\xi X}{\beta}\right)^{-1/\xi}$$

For heavy tails we have $\xi > 0$ in which case not all moments exist:

$$E[X^k] = \infty \quad \text{for} \quad k \ge 1/\xi.$$

The parameters in the models are fitted by Maximum Likelihood Estimation, using historical data for example, and from that we can extrapolate to the future.

Example

(From Alexander McNeil, 1998).

Fit a Frechet distribution to the 28 annual maxima from 1960 to October *16*th 1987, the business day before the big one. Now calculate probability of various returns. For example, 50-year return level being the level which on average should only be exceeded in one year every 50 years. Result: 24%. And then the next day, October 19th: 20.4%. Note that in the dataset the largest fall had been 'just' 6.7%.

19.9 COHERENCE

Some measures of risk make more sense than others. Artzner, Delbaen, Eber & Heath (1997) have stated some properties that sensible risk measures ought to have; they call such sensible measures 'coherent.' If we use $\rho(X)$ to denote this risk measure for a set of outcomes X then the necessary properties are as follows.

1. Sub-additivity: $\rho(X+Y) \leq \rho(X) + \rho(Y)$. This just says that if you add two portfolios together the total risk can't get any worse than adding the two risks separately. Indeed, there may be cancellation effects or economies of scale that will make the risk better.
2. Monotonicity: If $X \leq Y$ for each scenario then $\rho(X) \leq \rho(Y)$. Pretty obviously, if one portfolio has better values than another under all scenarios then its risk will be better.
3. Positive homogeneity: For all $\lambda > 0$, $\rho(\lambda X) = \lambda\rho(X)$. Double your portfolio then you double your risk.
4. Translation invariance: For all constant c, $\rho(X + c) = \rho(X) - c$. Think of just adding cash to a portfolio.

Do the common measures of risk satisfy these reasonable criteria? Generally speaking, surprisingly and rather unfortunately not! For example, the classical VaR violates sub-additivity.

19.10 SUMMARY

The estimation of downside potential in any portfolio is clearly very important. Not having an idea of this could lead to the disappearance of a bank, and has. In practice, it is more important to the managers in banks, and not the traders. What do they care if their bank collapses as long as they can walk into a new job?

I have shown the simplest ways of estimating Value at Risk, but the subject gets much more complicated. 'Complicated' is not the right word, 'messy' and 'time-consuming' are better. And currently there are many software vendors, banks and academics touting their own versions in the hope of becoming the market standard. In Chapter 42 we'll see a few of these in more detail.

FURTHER READING

- See Lawrence (1996) for the application of the Value at Risk methodology to derivatives.
- See Chew (1996) for a wide ranging discussion of risk management issues and details of important real-life VaR 'crises.'
- See Jorion (1997) for further information about the mathematics of Value at Risk.
- The allocation of bank capital is addressed in Matten (1996).
- Alexander & Leigh (1997) and Alexander (1997a) discuss the estimation of covariance matrices for VaR.
- Artzner, Delbaen, Eber & Heath (1997) discuss the properties that a sensible VaR model must possess.
- Lillo, Mantegna, Bouchaud & Potters (2002) introduce the concept of 'variety' as a measure of the dispersion of stocks.

CHAPTER 20
forecasting the markets?

In this Chapter. . .

- some of the commonly used technical methods for predicting market direction
- some modern approaches to modeling markets and their microstructure

20.1 INTRODUCTION

People have been making predictions about the future since the dawn of time. And predicting the future of the financial markets has been especially popular. Despite the claims of many 'legendary' investors it is not clear whether there is any validity in any of the methods they use, or whether the claims are examples of survivor bias. The big losers tend to keep quiet.

Humans just seem to like determinism, and have difficulty, especially at an early age, handling the random and meaningless. It goes back to caveman times. Ug and his friend Og were standing in the entrance to their cave when along came a saber-toothed tiger. The tiger went for Og and dragged him off for his supper. It just so happened that Ug was scratching his left ear when this happened. Now every time that Ug sees a saber-toothed tiger he scratches his left ear, just to be on the safe side. Now maybe there was no connection between what Ug was doing when Og was dragged off, but perhaps there was, and there's no harm in playing it safe in the future.

In this chapter we look at some of the traditional methods for determining trends, technical analysis, and also some of the more recent methods, often emanating from physics. I won't be describing some of the more dubious ideas, such as astrology, but then we Scorpios tend to be sceptical.[1]

20.2 TECHNICAL ANALYSIS

Technical analysis is a way of predicting future price movements based only on observing the past history of prices. This price history may also include other quantities such as volume

[1] I did know a trader once who always had astrological charts on his computer. I assumed they were just a screensaver, used to hide the trader's actual strategy while he was away from his desk. Shortly after he was fired we all discovered that actually . . .

of trade. These methods contrast with **fundamental analysis** in which prediction is made based on an examination of the factors underlying the stock or other instrument. This may include general economic or political analysis, or analysis of factors specific to the stock, such as the effect of global warming on snowfall in the Alps, if one is concerned with a travel company. In practice, most traders will use a combination of both technical and fundamental analysis.

Technical analysis is also called **charting** because the graphical representation of prices etc. plays an important part. Technical analysis is thought to be particularly good for timing market moves; fundamental analysis may get the direction right, but not necessarily when the move will happen.

20.2.1 Plotting

The simplest chart types just join together the prices from one day to the next, with time along the horizontal axis. These are the sort of plots we have seen throughout this book. Sometimes a logarithmic scale is used for the vertical price axis to represent return rather than absolute level. Later on we'll see some more complicated types of plotting. Sometimes you will see trading volume on the same graph; this is also used for prediction but I won't go into any details here, see Figure 20.1.

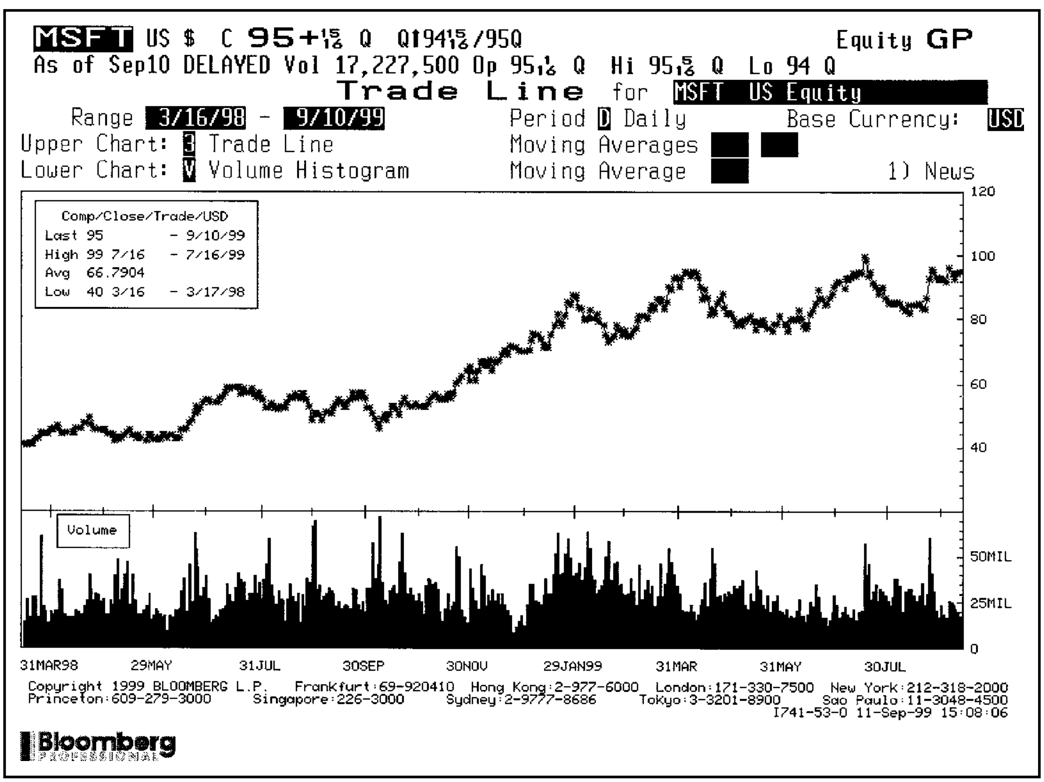


Figure 20.1 Price and volume. Source: Bloomberg L.P.

20.2.2 Support and Resistance

Resistance is a price level which an asset seems to have difficulty rising above. This may be a previously realized highest value, or it may be a psychologically important (round) number. **Support** is a level below which an asset price seems to be reluctant to fall. There may be sufficient demand at this low price to stop it falling any further. Examples of support and resistance are shown in Figure 20.2.

When a support or resistance level finally breaks it is said to do so quite dramatically.

20.2.3 Trendlines

Similar to support and resistance are **trendlines**. These are formed by joining together successive peaks and/or troughs in the price history to form a rising or falling support or resistance level. An example is shown in Figure 20.3.

20.2.4 Moving Averages

Moving averages are calculated in many ways. Different time windows can be used, or even exponentially-weighted averages can be calculated. Moving averages are supposed to distill out the basic trend in a price by smoothing the random noise.

Sometimes two moving averages are calculated, say a ten-day and a 250-day average. The crossing of these two would signify a change in the underlying trend and a time to buy or sell.

Although I'm not the greatest fan of technical analysis, there is some evidence that there may be predictive power in moving averages.

Figure 20.4 shows a Bloomberg screen with Microsoft share price, five- and 15-day moving averages.

20.2.5 Relative Strength

The **relative strength index** is the percentage of up moves in the last N days. A number higher than 70% is said to be overbought and are therefore likely to fall and below 30% is said to be oversold and should rise.

20.2.6 Oscillators

An **oscillator** is another indicator of over/underbought conditions. One way of calculating it is as follows.

Define k by

$$100 \times \frac{\text{Current close} - \text{lowest over } n \text{ periods}}{\text{Highest over } n \text{ periods} - \text{lowest over } n \text{ periods}}.$$

Now take a moving average of the last three days, say. This average is plotted against time and any move outside the range 30–70% could be an indication of a move in the asset. See Figure 20.5.

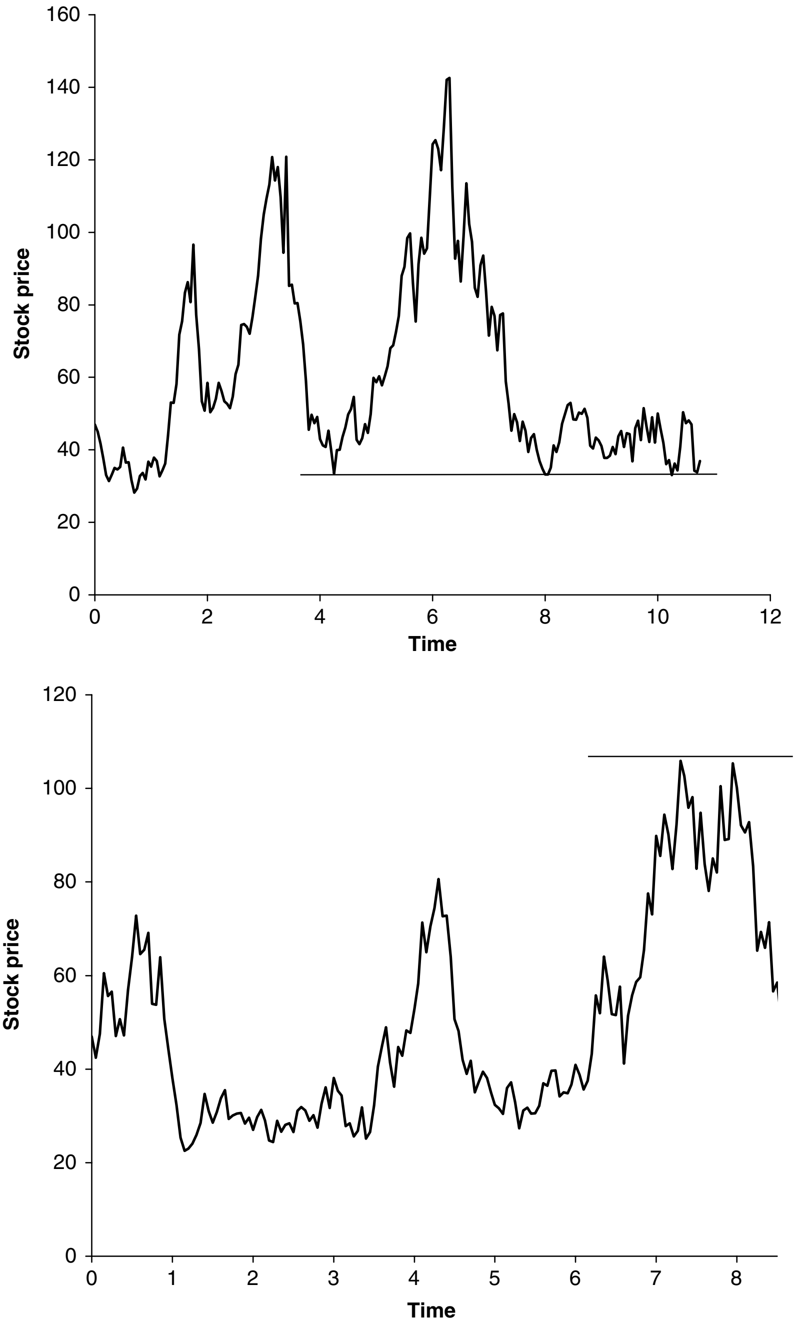


Figure 20.2 Support and resistance.

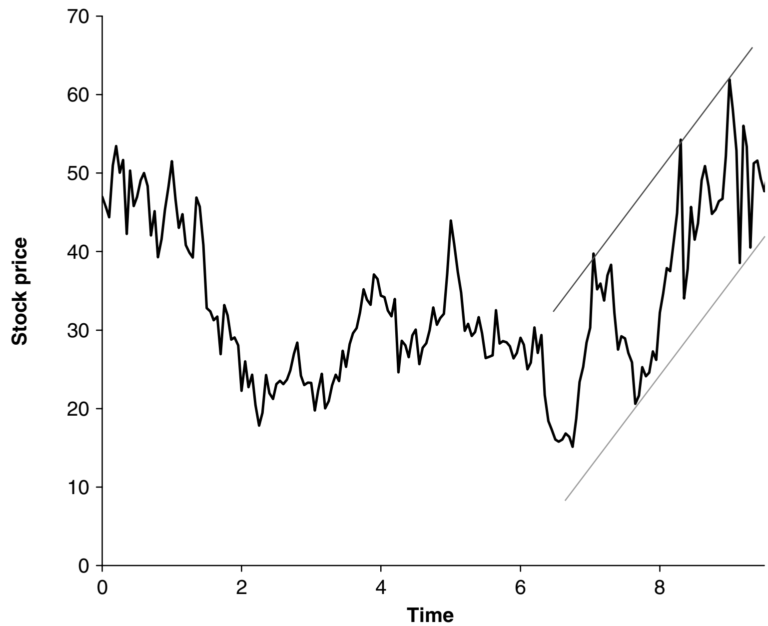


Figure 20.3 A trending stock.

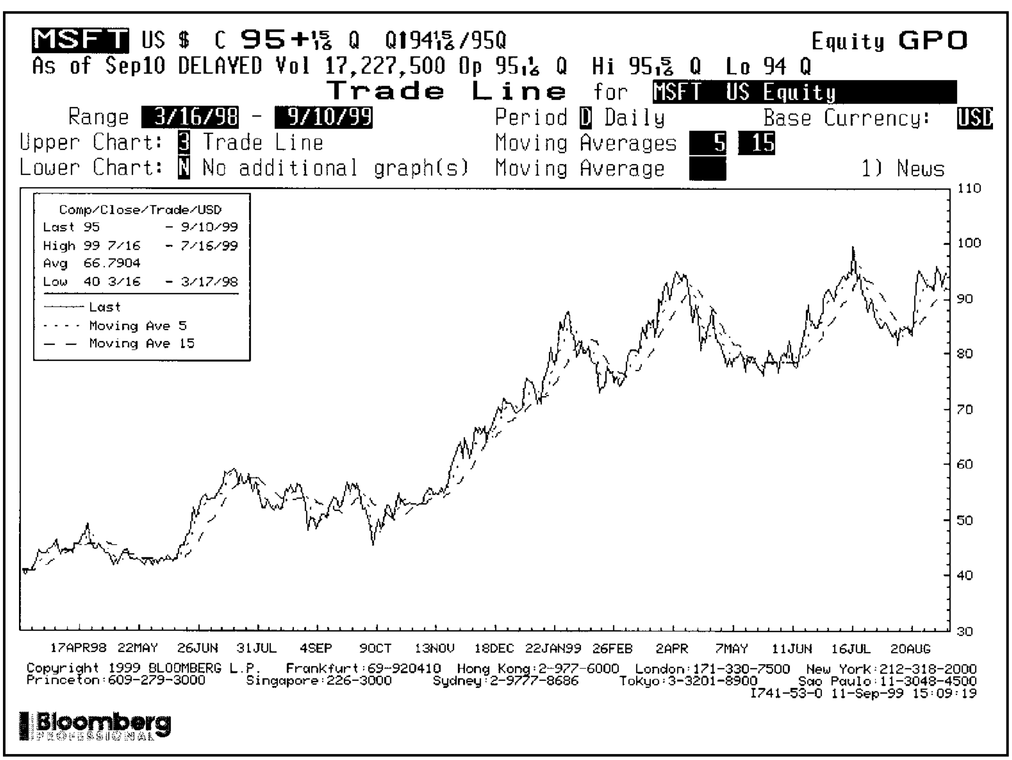


Figure 20.4 Two moving averages. Source: Bloomberg L.P.

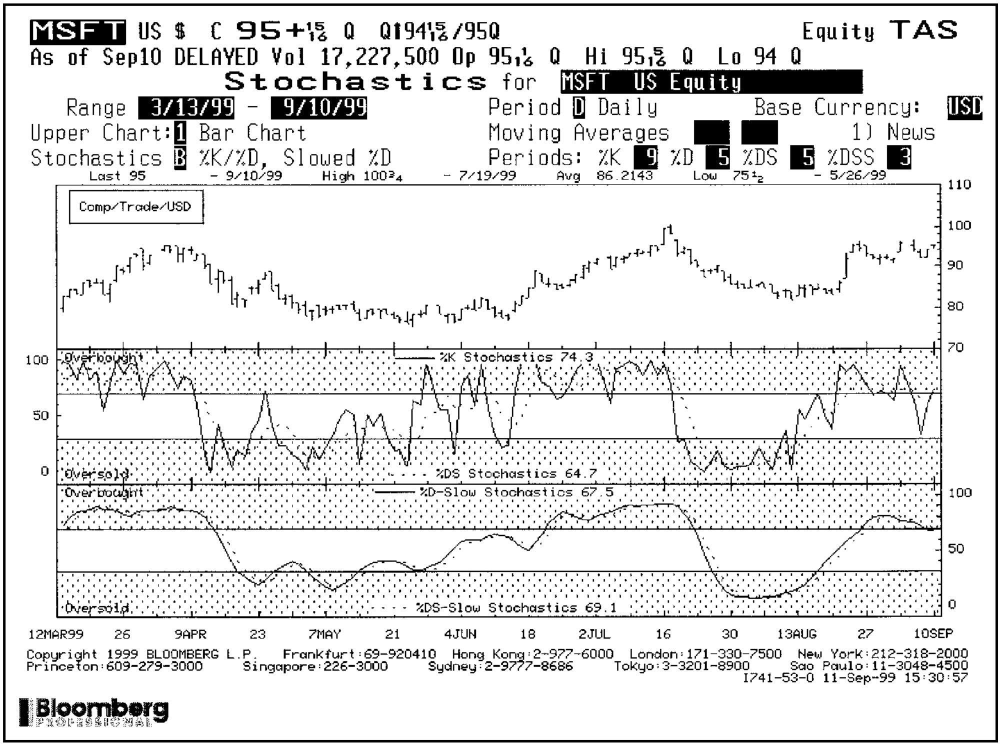


Figure 20.5 Oscillator. Source: Bloomberg L.P.

20.2.7 Bollinger Bands

Bollinger Bands are plots of a specified number of standard deviations above and below a specified moving average, see Figure 20.6.

20.2.8 Miscellaneous Patterns

As well as the ‘quantitative’ side of charting there is also the ‘artistic’ side. Practitioners say that certain patterns anticipate certain future moves. It’s rather like your grandmother reading tea leaves.

Head and shoulders is a common pattern and is best described with reference to Figure 20.7. There are a left and a right shoulder with the head rising above. Following on from the right shoulder should be a dramatic decline in the asset price.

This pattern is supposed to be one of the most reliable predictors. It is also seen in an upside-down formation.

Saucer tops and bottoms are also known as **rounding tops** and **bottoms**. They are the result of a gradual change in supply and demand. The shape is generally fairly symmetrical as the prices rises and falls. These patterns are quite rare. They contain no information about the strength of the new trend. See Figure 20.8.

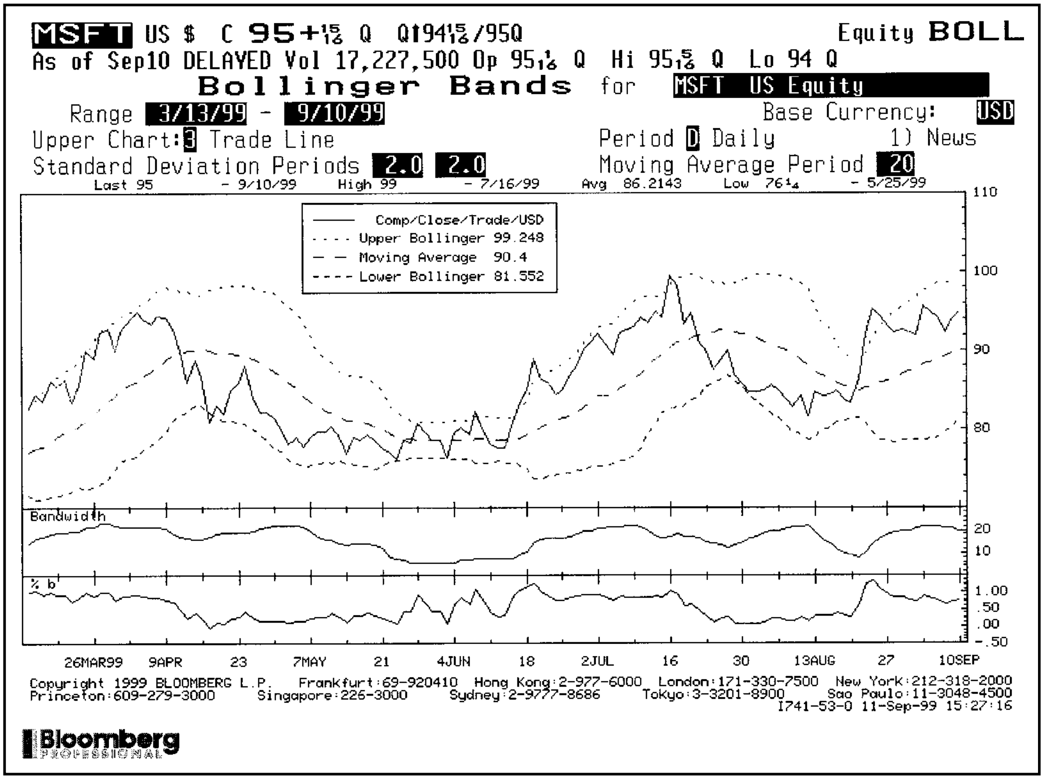


Figure 20.6 Bollinger Bands. Source: Bloomberg L.P.

Double and triple tops and bottoms are quite rare patterns, the triple being even rarer than the double. The double top looks like an 'M' and a double bottom like a 'W.' The triple top is similar but with three peaks, as shown in Figure 20.9. The key point about the peaks and troughs is that they should all be at approximately the same level.

20.2.9 Japanese Candlesticks

Japanese candlesticks contain more information than the simple plots described so far. They record the opening and closing prices as well as the day's high and low. A rectangle is drawn extending from the close to the open, and is colored white if close is above open and black if close is below open (see Figure 20.10). The high-low range is marked by a continuous line.

Certain combinations of candlesticks, appearing consecutively have special meanings and names like 'Hanging Man' and 'Upside Gap Two Crows.' See Figure 20.11 for candlesticks in action. On this Letter chart are shown 'HR' = Bearish Harami, 'D' = Doji (representing indecision), 'BH' Bullish Harami, 'EL' = Bearish Engulfing Line, and 'H' = Hanging Man (representing reversal after a trend).

Figure 20.12 shows some of the possible candlestick shapes and their interpretation.

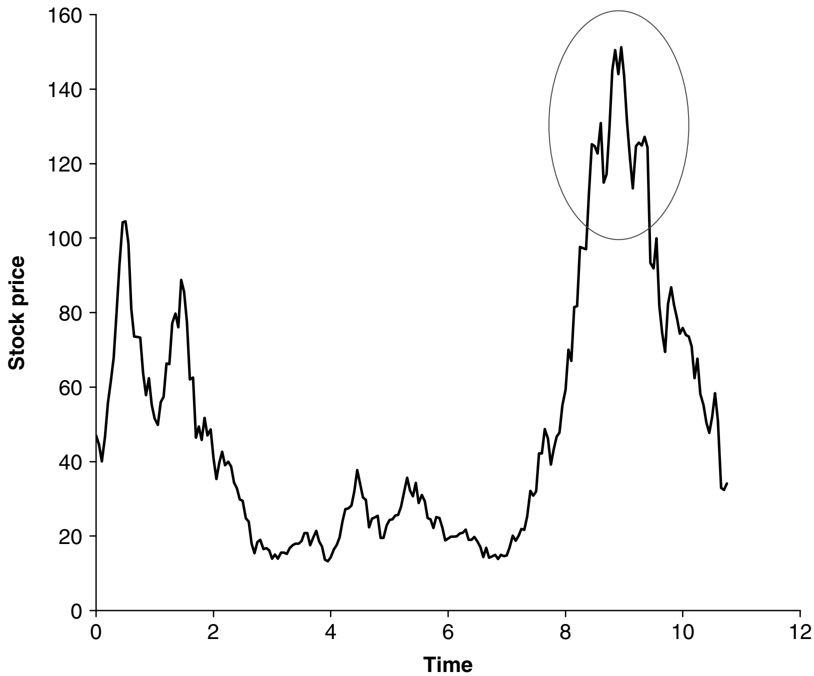


Figure 20.7 Head and shoulders.

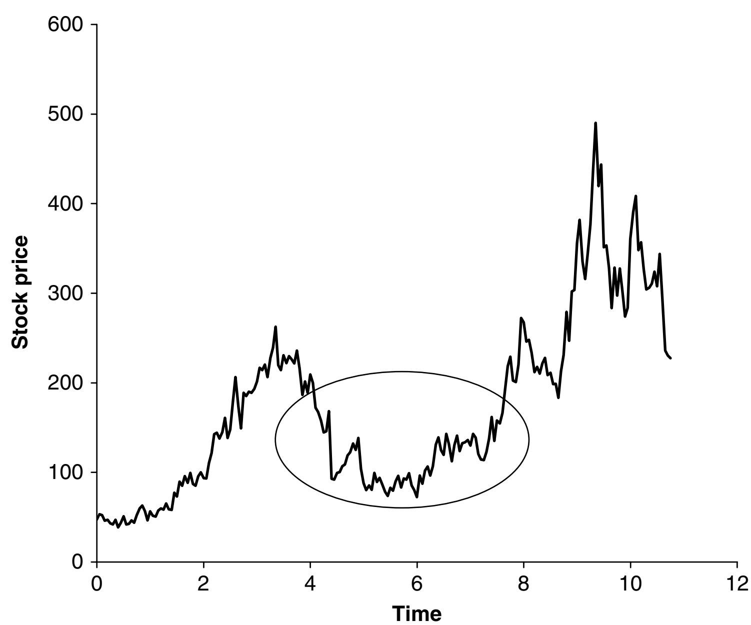


Figure 20.8 Saucer bottom.

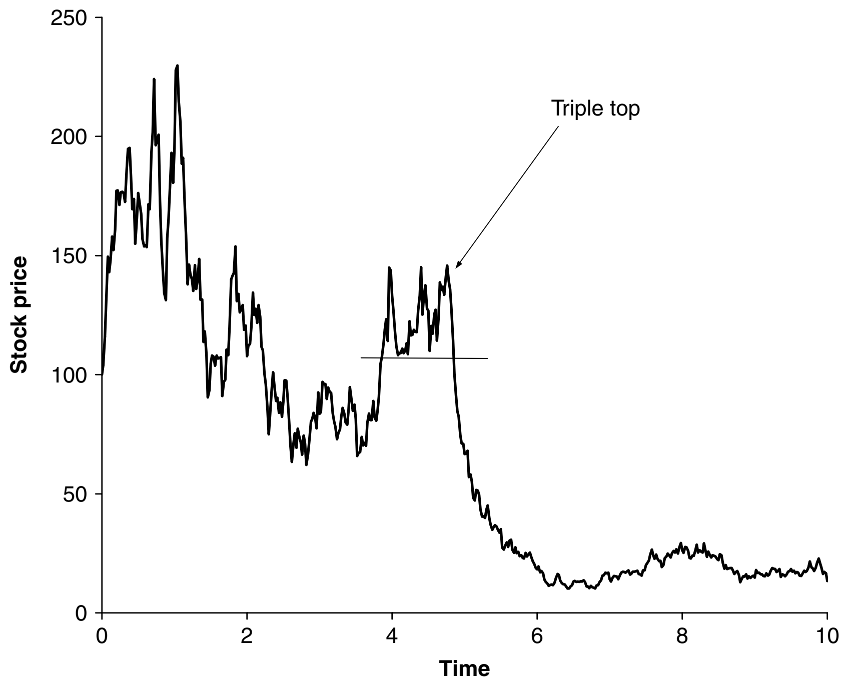


Figure 20.9 A triple top.

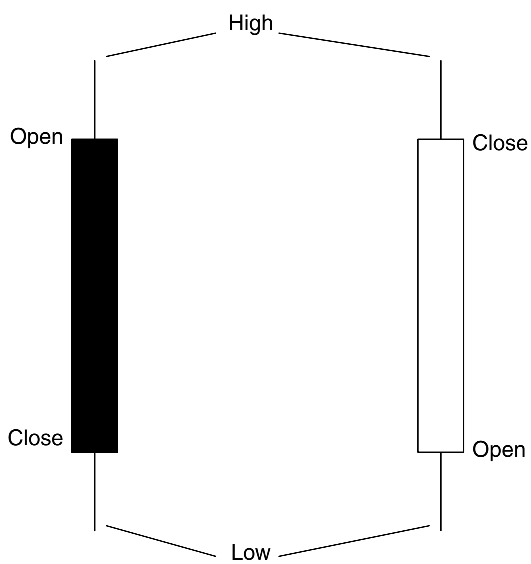


Figure 20.10 Japanese candlesticks.

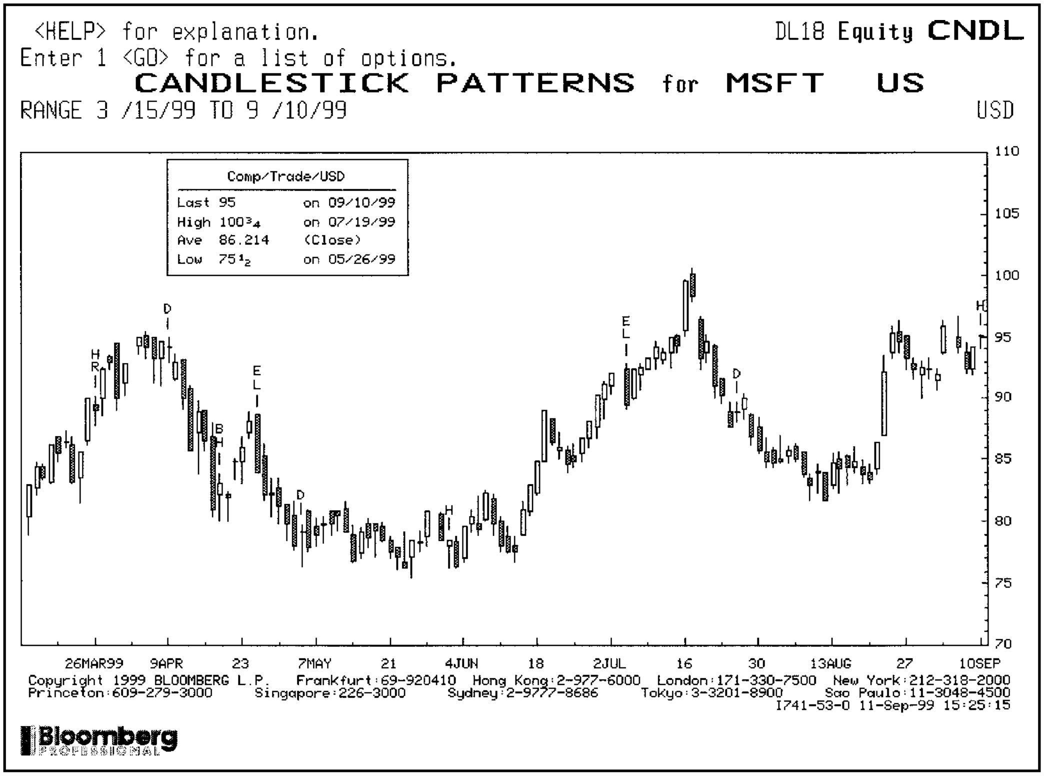
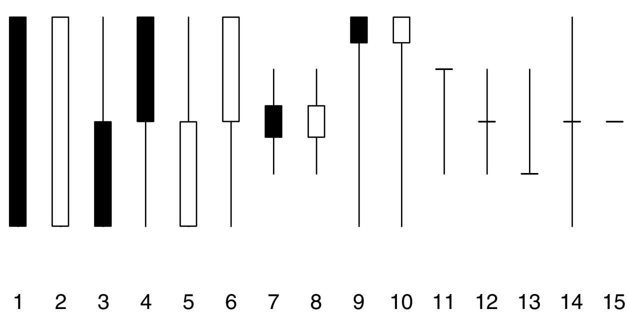


Figure 20.11 A candlestick chart. Source: Bloomberg L.P.



1 Extremely bearish
2 Extremely bullish
3 Bearish
4 Bearish
5 Bullish
6 Bullish
7 Neutral
8 Neutral
9 In downtrend, bullish; in uptrend, bearish
10 In downtrend, bullish; in uptrend, bearish
11 A turning period
12 A turning period
13 End of downtrend
14 A turning period
15 Possible turning period

Figure 20.12 The meanings of the various candlesticks.

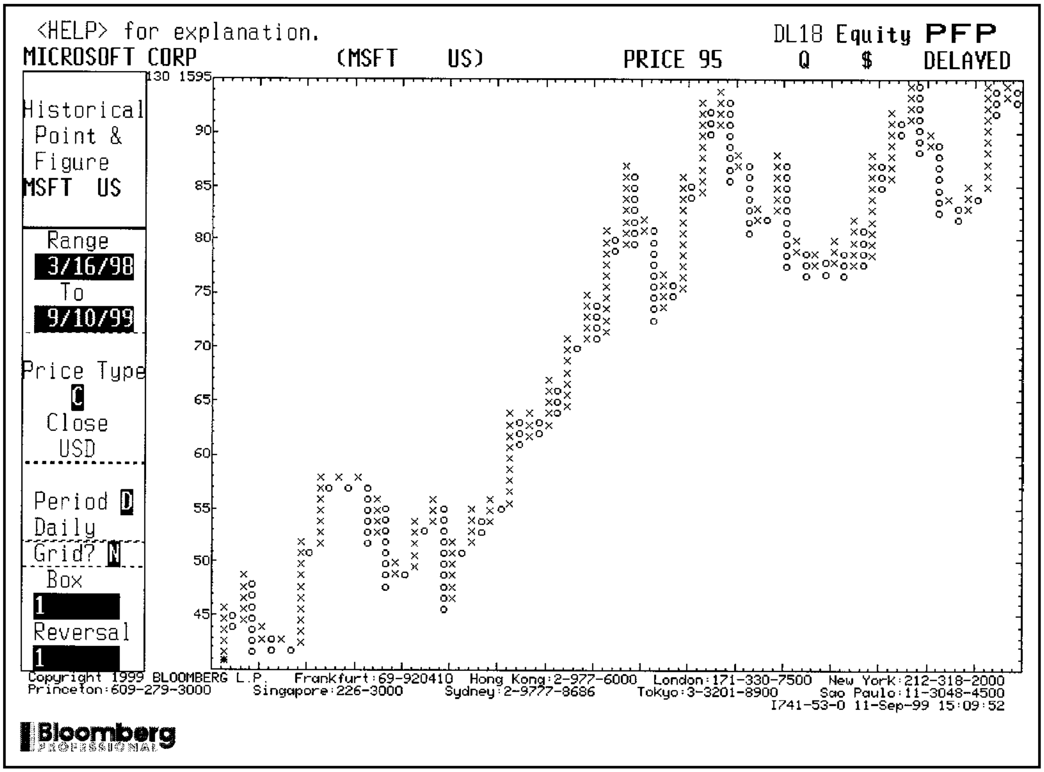


Figure 20.13 A point and figure chart of Microsoft. Source: Bloomberg L.P.

20.2.10 Point and Figure Charts

Point and figure charts are different from the charts described above in that they do not have any explicit timescale on the horizontal axis. Figure 20.13 is an example of a point and figure chart. Each box on the chart represents a prespecified asset price move. The boxes are a way of discretizing asset price moves, instead of discretizing in time. For each consecutive asset price rise of the box size draw an ‘X’ in the box, in a rising column, one above the other. When this uptrend finishes, and the asset falls, start putting ‘O’ in a descending column, to the right of the previous rising Xs.

- A long column of Xs denotes demand exceeding supply.
- A long column of Os denotes supply exceeding demand.
- Short up and down columns denote a balance of supply and demand.

20.3 WAVE THEORY

As well as plotting and spotting trends in price movements there have been some theories for price prediction based on market cycles or waves. Below, I briefly mention a couple.

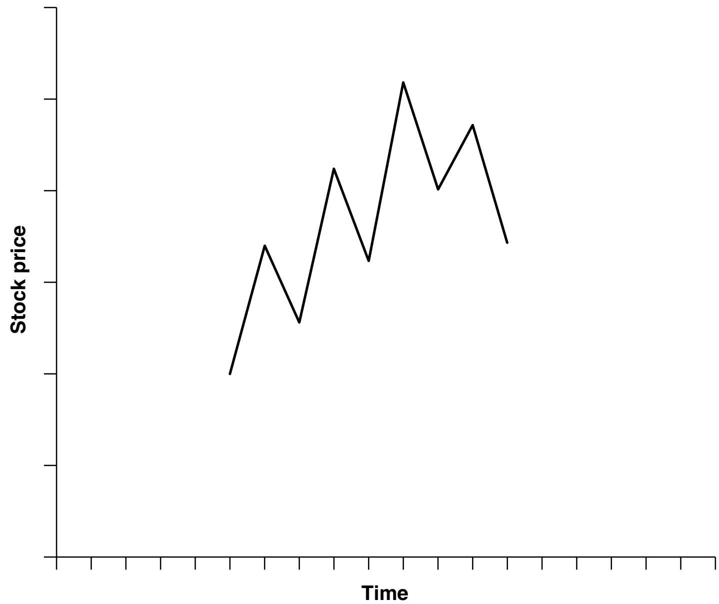


Figure 20.14 Elliot waves.

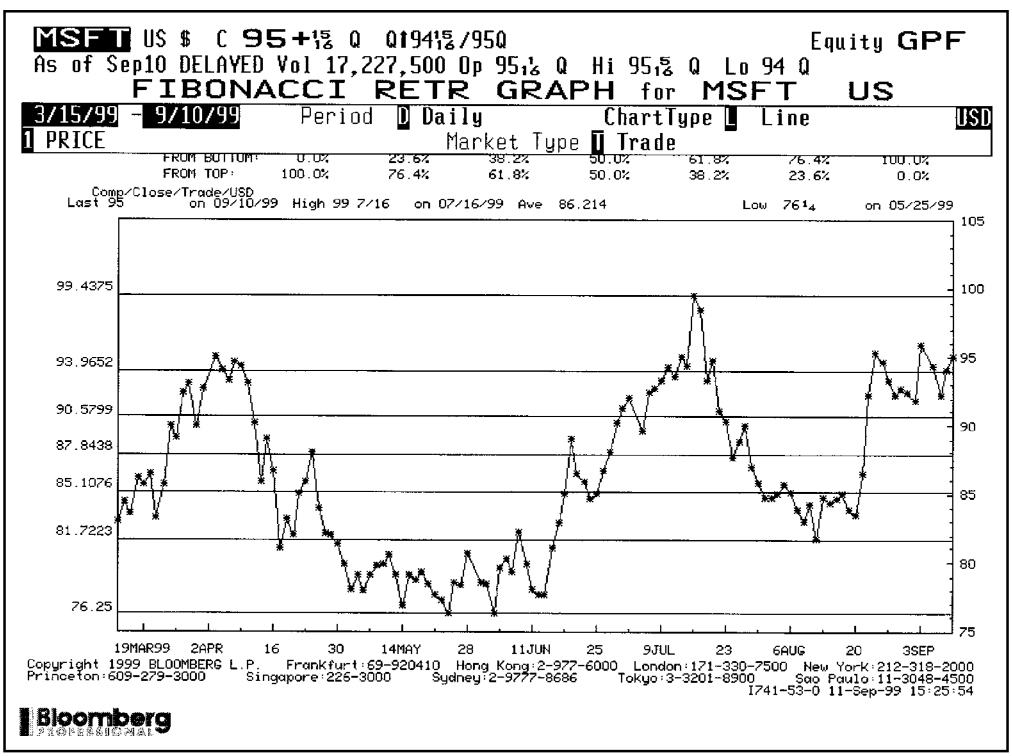


Figure 20.15 Fibonacci lines. Source: Bloomberg L.P.

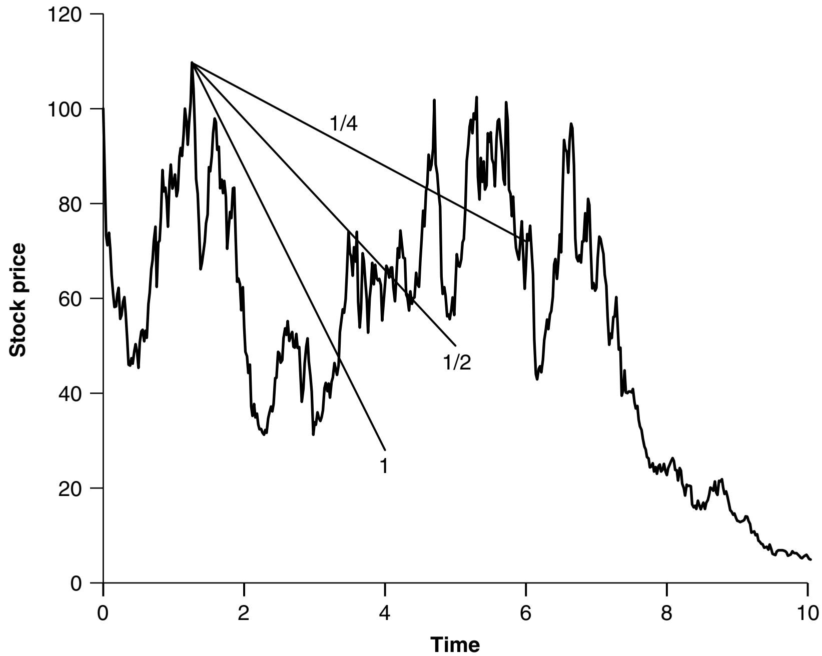


Figure 20.16 Gann charts.

20.3.1 Elliott Waves and Fibonacci Numbers

Ralph N. Elliott observed repetitive patterns, waves or cycles in prices. Roughly speaking, there are supposed to be five points in a bullish wave and then three in a bearish one (see Figure 20.14).[2] Within this **Elliott wave theory** there is also supposed to be some predictive ability in terms of the sizes of the peaks in each wave. For some reason, the ratio of peaks in a trend are supposed to be fairly constant; the ratio of second peak to first should be approximately 1.618 and of the third to the second 2.618. Unfortunately, the number 1.618 is approximately the **Golden ratio** of the ancient Greeks; $\frac{1}{2}(\sqrt{5}+1)$. It is also the ratio of successive numbers in the **Fibonacci series** given by $a_n = a_{n-1} + a_{n-2}$ for large n. I say, unfortunately, because people extrapolate wildly from this. And if it's a coincidence then ... Figure 20.15 shows the key levels coming from the Fibonacci series.

20.3.2 Gann Charts

Figure 20.16 shows a Gann chart. The lines all have slopes which are fractions of the slope of the lowest line. Need I say more?

20.4 OTHER ANALYTICS

There's an almost endless number of ways that chartists analyze data. I'll mention just a couple more before moving on.

[2] In Brownian Motion there are, of course, an infinite number of peaks and troughs in any period.

Volume is simply the number of contracts traded in a given period. A rising price and high volume means a strong, upwardly trending market. But a rising price with low volume could be a sign that the market is about to turn.

Open interest is the number of still outstanding futures contracts, those which have not been closed out. Because there are equal numbers of buyers and sellers, open interest does not necessarily give any directional info, but an increase in open interest can mean that an existing trend is strong.

20.5 MARKET MICROSTRUCTURE MODELING

The financial markets are made up of many types of players. There are the 'producers' who manufacture or produce or sell various goods and who may be involved in the financial markets for hedging. There are the 'speculators' who try and spot trends in the market, to exploit them and make money. These speculators may be using technical analysis methods, such as those described above, or fundamental analysis, whereby they examine the records and future plans of firms to determine whether stocks are under- or overpriced. Almost all traders use technical analysis at some time. Then there are the market makers who buy and sell financial instruments, holding them for a very short time, often seconds, and profit on bid-offer spreads.

There have been many attempts to model the interaction of these agents, sometimes in a game theoretic way, to try and model the asset price movements that in this book we have taken for granted. For example, can the dynamics induced by the actions of a combination of these three types of agent result in Brownian motion and lognormal random walks?

Below are just a very few examples of work in this area.

20.5.1 Effect of Demand on Price

Buying and selling assets moves their prices. Market makers respond to demand by increasing price, and reduce prices when the market is selling. If one can model the relationship between demand and price then it should be possible to analyze the effect that various types of technical trading rule have on the evolution of prices, and eventually to model the dynamics of prices.

A common starting point is to assume that there are two types of trader and one market maker. One trader follows a technical trading rule such as watching a moving average and the other is a **noise trader** who randomly buys or sells.

Interesting results follow from such models. For example

- trend followers can induce patterns in asset price time series;
- these artificially induced patterns can only be exploited for gain by someone following a suitably different trend;
- the more people following the same trend as you, the more money you will lose.

There are good reasons for there being genuine trends in the market: There is a slow diffusion of information from the knowledgeable to the less knowledgeable; The piece-by-piece secret acquisition of a company will gradually move a stock price upwards.

On the other hand, if there is no genuine reason for a trend, if it is simply a case of trend followers begetting a trend, then it may be beneficial to be a contrarian.

20.5.2 Combining Market Microstructure and Option Theory

Arbitrage does exist; many people make money from its existence. Yet the action of arbitragers will, via a demand/price relationship, remove the arbitrage. But there will be a timescale associated with this removal. What is the optimal way to exploit the arbitrage opportunity while knowing that your actions will to some extent be self-defeating?

20.5.3 Imitation

Another approach to market microstructure modeling is based on the true observation that people copy each other. In these models there are a number of traders who act partly in response to private information about a stock, partly randomly as noise traders, and partly to imitate their nearest neighbors. These models can result in market bubbles or market crashes.

20.6 CRISIS PREDICTION

There has been some work on analyzing data over various timescales to determine the likelihood of a market crash. Some ideas from earthquake modeling have been used to derive a 'Richter'-like measure of market moves. Of course, an effective predictor of market crashes could either:

- increase the chance or size of a crash as everyone panics, or
- reduce the chance or size of the crash since everyone gets advance warning and can calmly and logically act accordingly.

20.7 SUMMARY

I started out in finance many years ago plotting all of the technical indicators. I was not very successful at it. I could only get directions right for those assets with obvious seasonality effects, such as some commodities.

There is only one technical indicator that I believe in. There is definitely a strong correlation between hemlines and the state of the economy. The shorter the skirts, the better the economy.

FURTHER READING

- The book on technical analysis written by the news agency Reuters (1999) is excellent, as is Meyers (1994).
- Farmer & Joshi (2000) discuss and model trend following, and the creation of trends. They also demonstrate properties of the relationship between demand and price that prevent arbitrage.
- Bhamra (2000) has worked on imitation in financial markets.
- Olsen & Associates (`www.olsen.ch`) are currently working in the area of crisis modeling and prediction.

- Johnson, Hui & Lo (1999) models self-organized segregation of traders, and concludes that cautious traders perform poorly.
- The above is only a brief description of a very few examples from an expanding field. See O'Hara (1995) for a wide-ranging discussion of market microstructure models.
- Bernstein (1998) has a whole chapter on the Golden Ratio.
- Elton & Gruber (1995) describe the efficient market hypothesis and criticize technical analysis.
- Prast (2000a,b) discusses 'herding' in the financial markets.
- A mathematical model leading to support and resistance can be found in Osband (2004).

CHAPTER 21

a trading game

In this Chapter...

- A simple game simulating trading in options

21.1 INTRODUCTION

A lot of people reading this book will never have traded options or even stocks. In this chapter I describe a very simple trading game so that a group of people can try out their skill without losing their shirts. The game is based on that by one of my ex-students, David Epstein.

21.2 AIMS

The aims of this game are to familiarize students with the basic market-traded derivative contracts and to promote an understanding of the concepts involved in trading, such as bid, offer, arbitrage and liquidity.

21.3 OBJECT OF THE GAME

To make more money than your opponents. After the final round of trading, each player sums up their profits and losses. The player who has made the most profit is the winner.

21.4 RULES OF THE GAME

1. One person (possibly a lecturer) is the game organizer and in charge of choosing the types of contracts available for trading, the number and length of the trading rounds, judging any disputes and jollying the game along during slack periods.
2. The trading game takes place over a number of rounds. At the end of each round, a six-sided die is thrown. After the last round, the 'share price' is deemed to be the sum of all of the die rolls.
3. Traded contracts may include some or all of forwards, calls and puts at the discretion of the organizer (Figure 21.1). The organizer must also decide what exercise prices are available for call or put options.

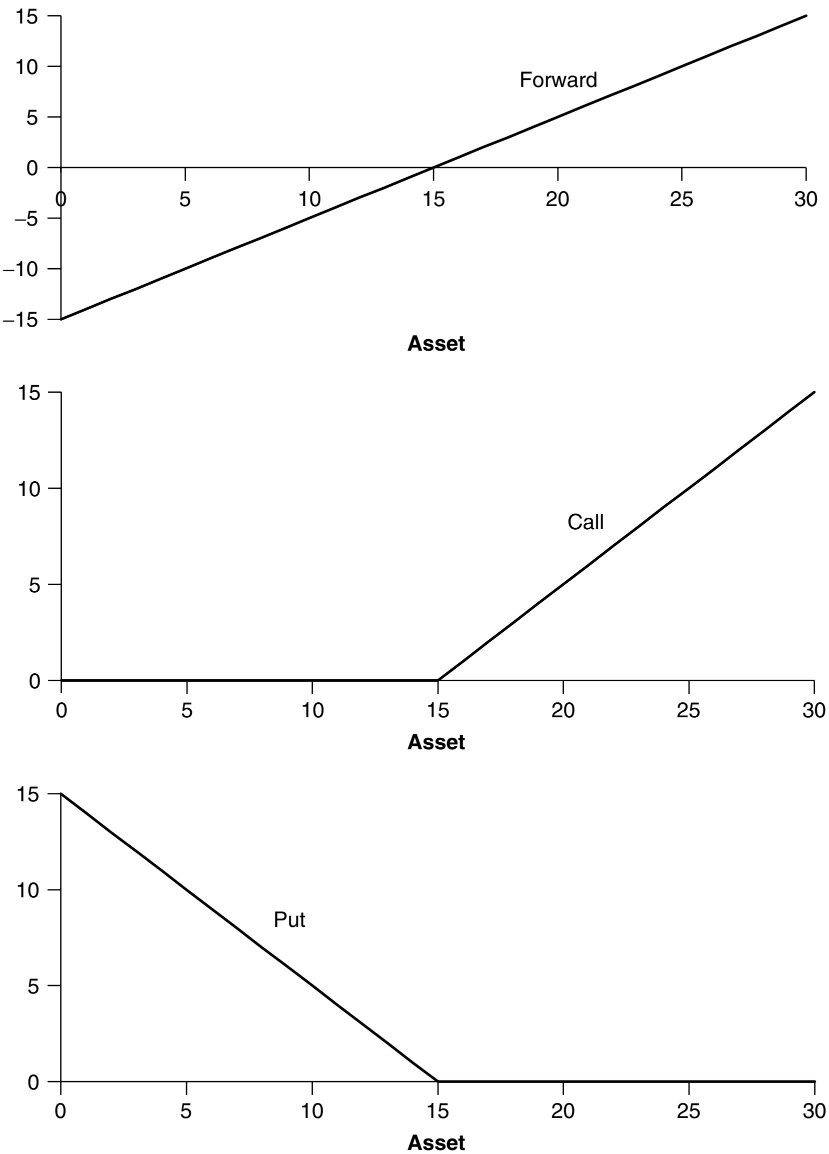


Figure 21.1 Available contracts.

4. All contracts expire at the end of the final round. The settlement value of each traded contract can then be determined by substituting the share price into the appropriate formula. A player's profits and losses on each trade can then be calculated and the resultant profit/loss is their final score.
5. During a round, players can offer to buy or sell any of the traded contracts. If another player chooses to take them up on their offer, then the deal is agreed and both parties must record the transaction on their trading sheet.

6. A deal on a contract must include the following information:

   - Forward: Forward price and quantity
   - Call or Put: Type of option (call or put), exercise price, cost and quantity

   The organizer chooses the types of contracts available and the strike prices.

   The forward price or option cost and the quantity in a deal are chosen by the players.

For beginners, play three games in succession, with the following structures:

1. Play with just the forward contract.
2. Play with the forward contract and the call option with exercise price 15.
3. Play with the forward and the call and put options with exercise price 15.

All three games take place over five rounds, each five minutes in length.

21.5 NOTES

1. Depending on the level of prior knowledge of the players, the organizer may need to explain the characteristics of the various traded contracts. It will be instructive to emphasize that the forward contract has no cost initially.
2. There will probably be times when the organizer has to act as a 'market-maker' and promote trading, for instance, asking the group if anyone wants to buy shares or at what price someone is willing to do so.
3. For more advanced students, consider introducing some of the following ideas:

   - Increase the number of rounds.
   - Decrease the length of each round.
   - Include extra calls and puts with different exercise prices or which either come into existence or expire at different times. You must fix the details of these extra contracts in advance of the game.
   - Include other contracts e.g. Asian options or Barriers.
   - Include a second die for a second underlying share price.

4. Including futures with 'daily' marking to market can be tried, but slows down the game. Nevertheless, it does illustrate the importance of margin, especially if the students have a limit on how much 'in debt' they are allowed to become.

21.6 HOW TO FILL IN YOUR TRADING SHEET

21.6.1 During a Trading Round

In the 'Contract' column of Figure 21.2, fill in the specifications of the instrument that you have bought/sold. Specify the forward price or exercise price if applicable (e.g. if there is more than one contract of this type in the game).

In the 'Buy/Sell' column, fill in whether you have bought or sold the contract and the quantity.

In the 'Cost per contract' column, fill in the cost of a single contract.

Trading sheet

Contract	Buy/Sell	Cost per contract	Total cost	Settlement value	Profit/loss

Figure 21.2 The Trading Game — designed by David Epstein, 1999.

21.6.2 At the End of the Game

In the 'Settlement Value' column, fill in the value of a single contract with the final share price.

In the 'Profit/Loss per contract' column, fill in the profit/loss for a single contract.

In the 'Total Profit/Loss' fill in the total profit/loss for the trade (= profit/loss × quantity).

Example

During a round, your transactions are:

Buy 10 call options, with exercise price 20, at a cost of \$2 each. Sell 1 put option, with exercise price 15, at a cost of \$1. Buy 5 forwards, with forward price 19.

Your trading sheet should be filled in as below:

Contract	Buy/Sell	Cost per contract	Total cost	Settlement value	Profit/loss
Call 20	Buy 10	2			
Put 15	Sell 1	1			
Forward 19	Buy 5	—			

At the end of the game, the final share price is 21. Consequently, the trading sheet is completed as follows:

Contract	Buy/Sell	Cost per contract	Total cost	Settlement value	Profit/loss
Call 20	Buy 10	2	$21 - 20 = 1$	$1 - 2 = -1$	$-1 \times 10 = -10$
Put 15	Sell 1	1	0	$1 - 0 = +1$	$+1 \times 1 = +1$
Forward 19	Buy 5	—	$21 - 19 = 2$	$+2$	$+2 \times 5 = +10$

The total profit and loss for the trader is therefore $-10 + 1 + 10 = +1$.

Remember that:

- If you buy a contract, your profit/loss = settlement value − cost per contract
- If you sell a contract, your profit/loss = cost per contract − settlement value

PART TWO
exotic contracts and path dependency

The second part of the book builds upon both the mathematics and the finance of Part One. The mathematical tools are extended to examine and model path dependency. The financial contracts that we look at now include path-dependent contracts such as barriers, Asians and lookbacks.

Chapter 22: An Introduction to Exotic and Path-dependent Derivatives This is just an overview of exotic options in which I try to classify various kinds of exotics. I describe several important features to look out for.

Chapter 23: Barrier Options The commonest type of path-dependent option is the barrier option. It is only weakly path-dependent (a concept I will explain) and slots very easily into the Black–Scholes framework.

Chapter 24: Strongly Path-dependent Derivatives Some options are harder to price for technical reasons; they are strongly path-dependent. However, with a little bit of ingenuity, not much, we can price these contracts quite easily.

Chapter 25: Asian Options Asian options depend on the realized average of an asset price path; they are considered in some depth.

Chapter 26: Lookback Options Lookback options depend on the realized maximum or minimum of the asset price path. They are explained in depth.

Chapter 27: Derivatives and Stochastic Control Some recent exotic contracts have an important element of optimal decision-making about them. In other words, the option holder has to make some decisions during the life of the contract. This may involve when to do something, such as exercise, or what quantity of the underlying to trade, or which cashflows to take. To analyze many of these contracts requires some knowledge of the mathematics of stochastic control.

Chapter 28: Miscellaneous Exotics A chapter in which I describe some more tricks of the trade for derivatives valuation.

Chapter 29: Equity and FX Term Sheets Here I present a selection of term sheets, some analysis, some mathematics and in some cases, Visual Basic code. (Although you may not fully appreciate the code until you have completed the part of the book on numerical methods).

CHAPTER 22
an introduction to exotic and path-dependent derivatives

In this Chapter...

- how to classify options according to important features
- how to think about derivatives in a way that makes it easy to compare and contrast different contracts
- the names and contract details for many basic types of exotic options

22.1 INTRODUCTION

The contracts we have seen so far are the most basic, and most important, derivative contracts but they only hint at the features that can be found in the more interesting products. In this chapter I prepare the way for the complex products I will be discussing in the next few chapters.

Exotic options are interesting for several reasons. They are harder to price, sometimes being very model-dependent. The risks inherent in the contracts are usually more obscure and can lead to unexpected losses. Careful hedging becomes important, whether delta hedging or some form of static hedging to minimize cashflows. Actually, how to hedge exotics is all that really matters. A trader may have a good idea of a reasonable price for an instrument, either from experience or by looking at the prices of similar instruments. But he may not be so sure about the risks in the contract or how to hedge them away successfully.

It is an impossible task to classify all options. The best that we can reasonably achieve is a rough characterization of the most popular of the features to be found in derivative products. I list some of these features in this chapter and give several examples. In the following few chapters I go into more detail in the description of the options and their pricing and hedging. The features that I describe now are time dependence, cashflows, strong path dependence and weak path dependence, dimensionality, the 'order' of an option, and finally options with embedded decisions.

We are going to continue with the Black–Scholes theme for the moment and show how to price and hedge exotics in their framework. Their assumptions will be relaxed in later chapters.

One approach that I will not really be taking is to try to decompose an exotic into a portfolio of vanillas. If such a decomposition is exact then the contract was not exotic in the first place. (Riddle: When can you decompose an exotic into a portfolio of vanillas? Answer: When it isn't exotic.) You may get an idea for an approximate price and how it may be possible to hedge an exotic statically by such a means but ultimately you will probably have to face a full and proper mathematical analysis of the exotic features in a contract.

22.2 OPTION CLASSIFICATION

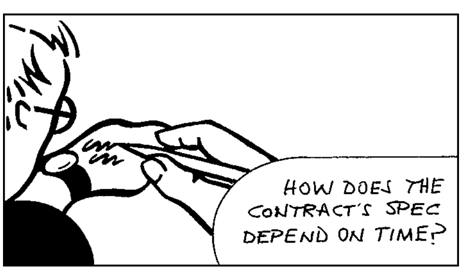
In order to figure out how to price and hedge exotic options I have found it incredibly helpful to classify them according to six criteria or features. I can't overemphasize how useful this has been to me. These six features are unashamedly mathematical in nature, having nothing whatsoever to do with the name of the option or what is the underlying. Being mathematical in nature they very quickly give you the following information.

- What kind of pricing method should best be used.
- Whether you can re-use some old code.
- How long it will take you to code it up.
- How fast it will eventually run.

The six features to look out for are

1. Time dependence
2. Cashflows
3. Path dependence
4. Dimensionality
5. Order
6. Embedded decisions

Some of these classes can be broken down further, as shown in Figure 22.1. Let's now look at these features one by one, in increasing order of interest.



22.3 TIME DEPENDENCE

We have seen time dependence in parameters, and have shown how to apply the Black–Scholes formulae when interest rates, dividends and volatility vary in time (in a known, deterministic, way). Those are features of our model. But here we are more concerned with time dependence in the option contract. For example, early exercise might only be permitted on certain dates or during certain periods. This intermittent early exercise is a characteristic of **Bermudan options**. Such contracts are referred to as time-inhomogeneous.

Time dependence is first on our list of features, since it is not all that earth shattering. Probably the only reason for the importance of time dependence at all is that it requires us to be a little bit careful with any numerical method we employ. Inevitably when solving

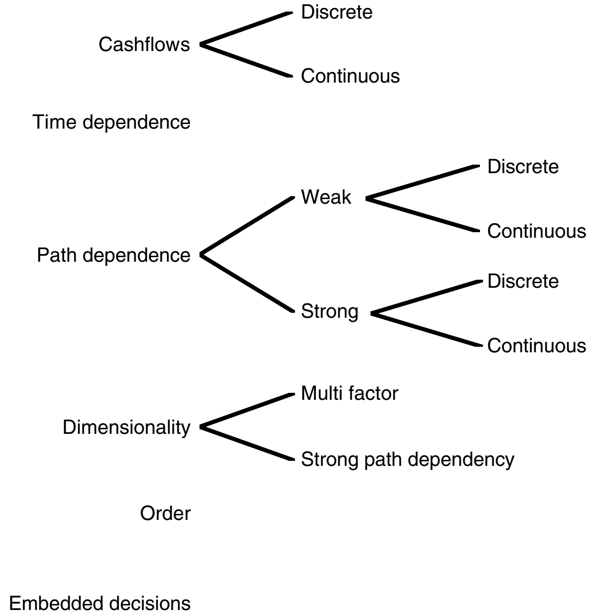


Figure 22.1 Option classification chart.

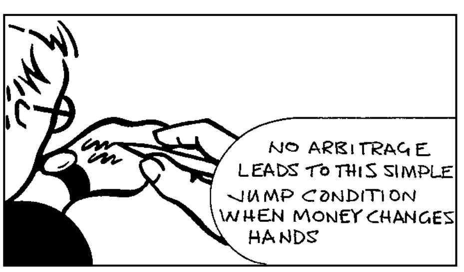
for an option price via numerical methods we end up needing to do some discretization of time. If the contract has time dependence then we may have restrictions imposed on our discretization.

- Time dependence in an option contract means that our numerical discretization may have to be lined up to coincide with times at, or periods during, which something happens.
- This means that our code will have to keep track of time, dates etc. This is not difficult, just annoying.

22.4 CASHFLOWS

Imagine a contract that pays the holder an amount q at time t_0. The contract could be a bond and the payment a coupon. If we use $V(t)$ to denote the contract value (ignoring any dependence on any underlying asset) and t_0^- and t_0^+ to denote just before and just after the cashflow date then simple arbitrage considerations lead to

$$V(t_0^-) = V(t_0^+) + q.$$



This is a **jump condition**. The value of the contract jumps by the amount of the cashflow. If this were not the case then there would be an arbitrage opportunity. The behavior of the contract value across the payment date is shown in Figure 22.2.

If the contract is contingent on an underlying variable so that we have $V(S, t)$ then we can accommodate cashflows that

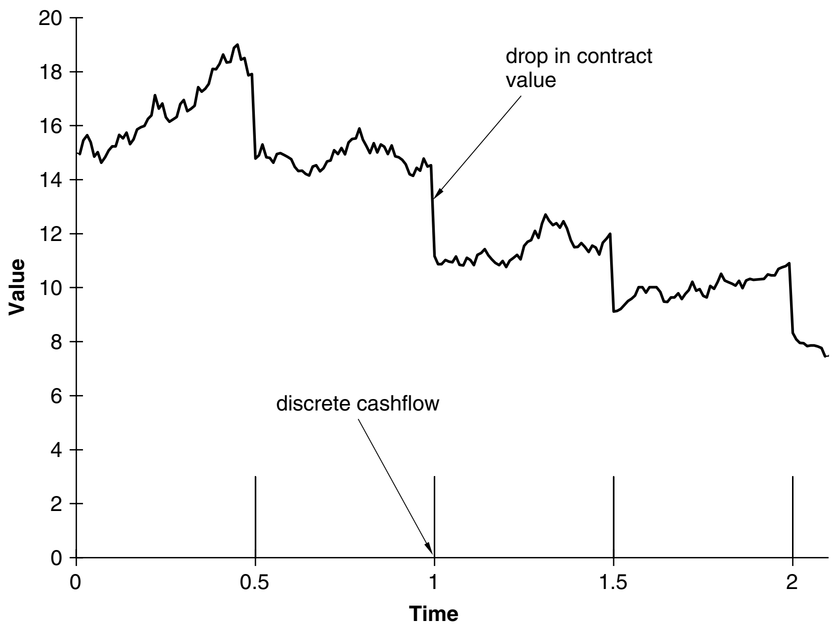


Figure 22.2 A discrete cashflow and its effect on a contract value.

depend on the level of the asset S, i.e. we could have $q(S)$. Furthermore, this also allows us to lump all our options on the same underlying into one large portfolio. Then, across the expiry of each option, there will be a jump in the value of our whole portfolio of the amount of the payoff for that option.

There is one small technical requirement here, the cashflow must be a deterministic function of time and the underlying asset. For example, the contract holder could receive a payment of S^2, for some asset with price S. However, the above argument would not be valid if, for example, the cashflow depended on the toss of a coin; one dollar is received if heads is thrown and nothing otherwise. The jump condition does not necessarily apply, because the cashflow is not deterministic.

If the cashflow is not deterministic the modeling is not so straightforward. There is no 'no arbitrage' argument to appeal to, and the result could easily depend on an individual's risk preferences. Nevertheless, we could say, for example, that the jump condition would be that the change in value of the contract would be the *expected* value of the cashflow:

$$V(t_0^-) = V(t_0^+) + E[q].$$

Such a condition would not, however, allow for the risk inherent in the uncertain cashflow.

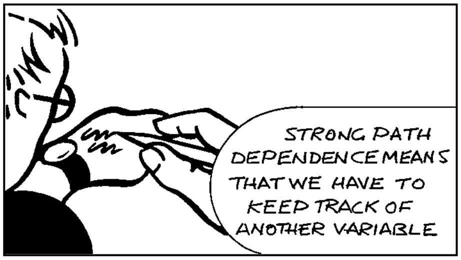
That was an example of a **discrete cashflow**. You may also see **continuous cashflows** in option contracts. The term sheet may specify something, for example, along the lines of 'the holder receives \$1 every day that the stock price is below \$80.' That would effectively be a continuous cashflow. When the cashflow is paid continuously then we no longer have jump conditions, instead we modify the basic Black–Scholes equation to add a source term. We'll see examples later.

- When a contract has a discretely paid cashflow you should expect to have to apply jump conditions. This also means that the contract has time dependence, see above.
- Continuously paid cashflows mean a modification, although rather simple, to the governing equation.

22.5 PATH DEPENDENCE

Many options have payoffs that depend on the path taken by the underlying asset, and not just the asset's value at expiration. These options are called **path dependent**. Path dependency comes in two varieties, strong and weak.

22.5.1 Strong Path Dependence



Of particular interest, mathematical and practical, are the strongly path-dependent contracts. These have payoffs that depend on some property of the asset price path in addition to the value of the underlying at the present moment in time; in the equity option language, we cannot write the value as $V(S, t)$. The contract value is a function of at least one more independent variable. This is best illustrated by an example.

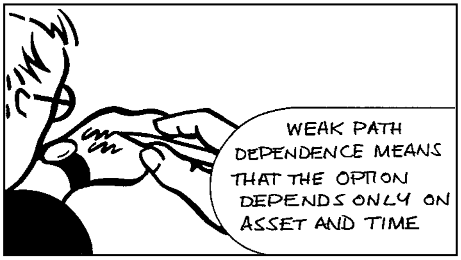
The Asian option has a payoff that depends on the average value of the underlying asset from inception to expiry. We must keep track of more information about the asset price path than simply its present position. The extra information that we need is contained in the 'running average.' This is the average of the asset price from inception until the present, when we are valuing the option. No other information is needed. This running average is then used as a new independent variable, the option value is a function of this as well as the usual underlying and time, and a derivative of the option value with respect to the running average appears in the governing equation.

There are many such contracts in existence, and later I show how to put many of them into the same general framework.

Strong path dependency comes in two forms, **discretely sampled** and **continuously sampled**, depending on whether a discrete subset of asset prices is used or a continuous distribution of them.

- Strong path dependency means that we have to work in higher dimensions. A consequence of this is that our code may take longer to run.

22.5.2 Weak Path Dependence



A simple example of a contract with weak path dependence is the **barrier**. Barrier (or knock-in, or knock-out) options are triggered by the action of the underlying hitting a prescribed value at some time before expiry. For example, as long as the asset remains below 150, the contract will have a call payoff at expiry. However, should the asset reach this level before expiry

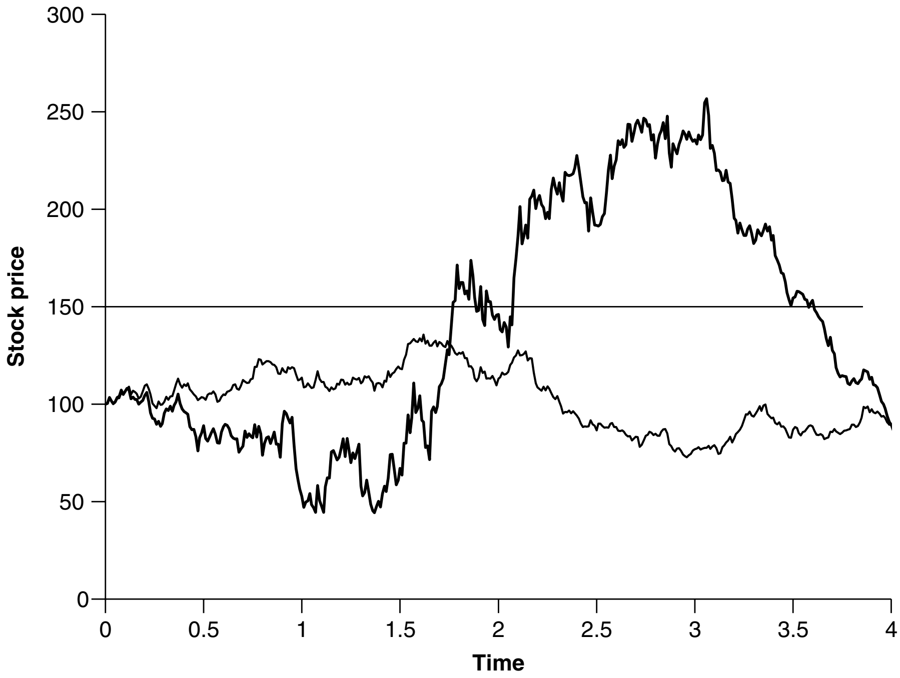
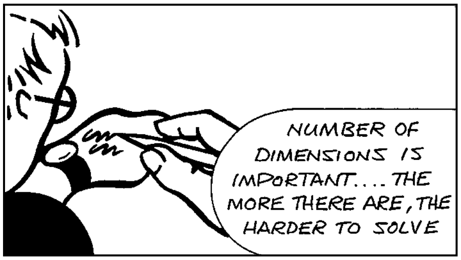


Figure 22.3 Two paths, having the same value at expiry, but with completely different payoffs.

then the option becomes worthless; the option has 'knocked out.' This contract is clearly path-dependent, for consider the two paths in Figure 22.3; one has a payoff at expiry because the barrier was not triggered, the other is worthless, yet both have the same value of the underlying at expiry.

We shall see in Chapter 23 that such a contract is only weakly path-dependent: we still solve a partial differential equation in the two variables, the underlying and time. And that is the difference, mathematically speaking, between strong and weak path dependency. A weakly path-dependent contract does not require us to introduce an extra variable to handle the path dependency. Again, we can imagine discrete and continuous versions.

- Weak path dependency means that we *don't* have to work in higher dimensions, so our code should be pretty fast.



22.6 DIMENSIONALITY

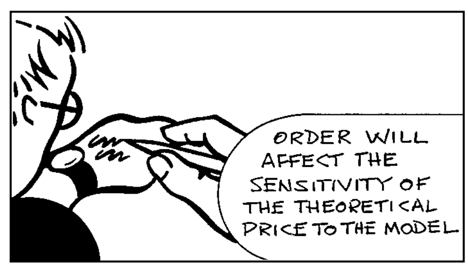
Dimensionality refers to the number of underlying independent variables. The vanilla option has two independent variables, S and t, and is thus two dimensional. The weakly path-dependent contracts have the same number of dimensions as their non-path-dependent cousins, i.e. a barrier call option has the same two dimensions as a vanilla call. For these contracts the roles of the asset dimension and the time dimension are quite different from each other, as discussed in Chapter 6 on the diffusion equation. This is because the governing equation, the Black–Scholes equation, contains a second asset-price derivative but only a first time derivative.

We can have two types of three-dimensional problem. The first type is the strongly path-dependent contract. We will see examples of these in later chapters. Typically, the new independent variable is a measure of the path-dependent quantity on which the option is contingent. The new variable may be the average of the asset price to date, say. In this case, derivatives of the option value with respect to this new variable are only of the first order. Thus the new variable acts more like another time-like variable.

The second type of three-dimensional problem occurs when we have a second source of randomness, such as a second underlying asset. We might, for example, have an option on the maximum of two equities. Both of these underlyings are stochastic, each with a volatility, and there will be a correlation between them. In the governing equation we will see a second derivative of the option value with respect to each asset. We say that there is diffusion in both S_1 and S_2.

- Higher dimensions means longer computing time.
- The number of dimensions we have also tells us what kind of numerical method to use. High dimensions mean that we probably want to use Monte Carlo; low means finite difference.

22.7 THE ORDER OF AN OPTION



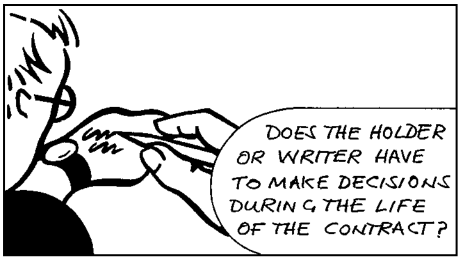
The next classification that we make is the **order** of an option. Not only is this a classification but the idea also introduces fundamental modeling issues.

The basic, vanilla options are of first order. Their payoffs depend only on the underlying asset, the quantity that we are *directly* modeling. Other, path-dependent, contracts can still be of first order if the payoff depends only on properties of the asset price path. ‘Higher order’ refers to options whose payoff, and hence value, is contingent on the value of *another* option. The obvious second-order options are compound options, for example, a call option giving the holder the right to buy a put option. The compound option expires at some date T_1 and the option on which it is contingent, expires at a later time T_2. Technically speaking, such an option is weakly path dependent. The *theoretical* pricing of such a contract is straightforward as we shall see.

From a practical point of view, the compound option raises some important modeling issues: The payoff for the compound option depends on the *market* value of the underlying option, and not on the theoretical price. If you hold a compound option, and want to exercise the first option then you must take possession of the underlying option. If that option is worth less than you think it should (because your model says so) then there is not much you can do about it. High-order option values are very sensitive to the basic pricing model and should be handled with care. This issue of not only modeling the underlying, but also modeling what the market does (regardless of whether it is ‘correct’ or not) will be seen in other parts of this book.

- When an option is second or higher order we have to solve for the first-order option, first. We thus have a layer cake; we must work on the lower levels and the results of those feed into the higher levels.

- This means that computationally we have to solve more than one problem to price our option.



22.8 EMBEDDED DECISIONS

We have seen early exercise in the American option problem. Early exercise is a common feature of other contracts, perhaps going by other names. For example, the conversion of convertible bonds, Chapter 33, is mathematically identical to the early exercise of an American option. The key point about early exercise is that the holder of this valuable right should ideally act *optimally*, i.e. they must decide *when* to exercise or convert. In the partial differential equation framework that has been set up, this optimality is achieved by solving a free boundary problem, with a constraint on the option value, together with a smoothness condition. It is this smoothness condition, that the derivative of the option value with respect to the underlying is continuous, that ensures optimality i.e. maximization of the option value with respect to the exercise or conversion strategy. It is perfectly possible for there to be more than one early-exercise region.[1]

Holding an American option you are faced with the decision whether and when to exercise your rights. The American option is the most common contract that contains within it a decision feature. Other contracts require more subtle and interesting decisions to be made. We'll be seeing several examples of these later and I'll mention just the one now.

The passport option, discussed in depth in Chapter 28, is an option on a trading account. You buy and sell some asset; if you are in profit on the expiry of the option you keep the money, if you have made a loss it is written off. The decisions to be made here are when to buy, sell or hold, and how much to buy, sell or hold.

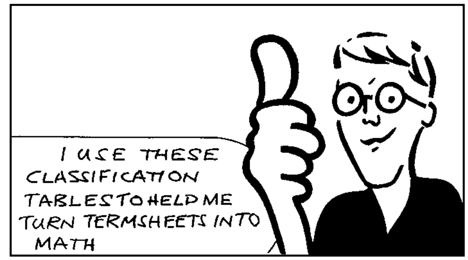
When a contract has embedded decisions you need an algorithm for deciding how that decision will be made. That algorithm amounts to assuming that the holder of the contract acts to *make the option value as high as possible for the delta-hedging writer*. The pricing algorithm then amounts to searching across all possible holder decision strategies for the one that maximizes the option value. That sounds hard, but approached correctly is actually remarkably straightforward, especially if you use the finite-difference method. The justification for seeking the strategy that maximizes the value is that the writer cannot afford to sell the option for anything less, otherwise he would be exposed to 'decision risk.'

When the option writer or issuer is the one with the decision to make then the value is based on seeking the strategy that minimizes the value.

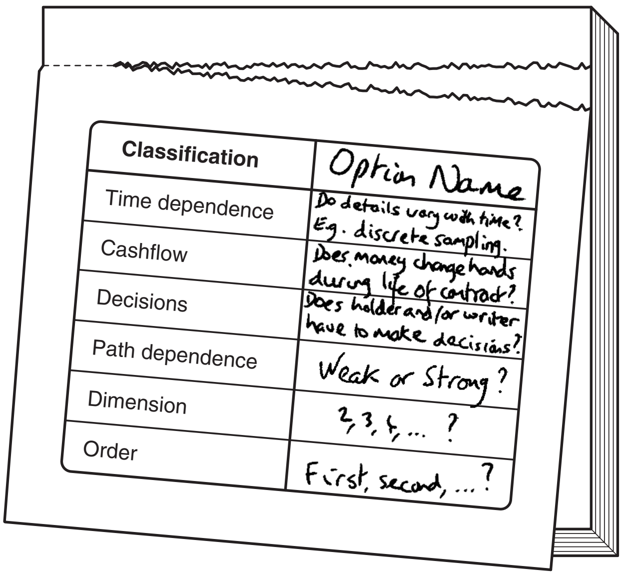
- Decision features mean that we'd really like to price via finite differences.
- The code will contain a line in which we seek the best price, so watch out for $\geq$ or $\leq$ signs.

[1] One rarely mentioned aspect of American options and, generally speaking, contracts with early exercise-type characteristics, is that they are path-dependent. Whether the owner of the option still holds the option at expiry depends on whether or not he has exercised the option, and thus on the path taken by the underlying. For American-type options this path dependence is weak, in the sense that the partial differential equation to be solved has no more independent variables than a similar, but European, contract.

22.9 **CLASSIFICATION TABLES**



Watch out for tables like the following for the classification of special contracts.



22.10 **EXAMPLES OF EXOTIC OPTIONS**

There now follow some basic examples, just to get you into the swing of things. Over the remainder of Part Two we will see many, many more contracts, of increasing complexity.

22.10.1 Compounds and Choosers

Compound and **chooser options** are simply options on options. The compound option gives the holder the right to buy (call) or sell (put) another option. Thus we can imagine owning a call on a put, for example. This gives us the right to buy a put option for a specified amount on a specified date. If we exercise the option then we will own a put option which gives us the right to sell the underlying. This compound option is second order because the compound option gives us rights over another derivative. Although the Black–Scholes model can theoretically cope with second-order contracts it is not so clear that the model is completely satisfactory in practice; when we exercise the contract we get an option at the market price, not at our theoretical price.

In the Black–Scholes framework the compound option is priced as follows. There are two steps: First price the underlying option and then price the compound option. Suppose that the underlying option has a payoff of $F(S)$ at time T, and that the compound option can be exercised at time $T_{Co} < T$ to get $G(V(S, T_{Co}))$ where $V(S, t)$ is the value of the underlying option. Step one is to price the underlying option i.e. to find $V(S, t)$. This satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \quad \text{with} \quad V(S, T) = F(S).$$

Solve this problem so that you have found $V(S, T_{Co})$. This is the (theoretical) value of the underlying option at time T_{Co}, which is the time at which you can exercise your compound option. Now comes the second step, to value the compound option. The value of this is $Co(S, t)$ which satisfies

$$\frac{\partial\ Co}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2\ Co}{\partial S^2} + rS\frac{\partial\ Co}{\partial S} - r\ Co = 0 \quad \text{with} \quad Co(S, T_{Co}) = G(V(S, T_{Co})).$$

As an example, if we have a call on a call with exercise prices E for the underlying and E_{Co} for the compound option, then we have

$$F(S) = \max(S - E, 0) \quad \text{and} \quad G(V) = \max(V - E_{Co}, 0).$$

In Figure 22.4 is shown the value of a vanilla call option at the time of expiry of a put option on this call. This is obviously some time before the expiry of the underlying call. In the same figure is the payoff for the put on this option. This is the final condition for the Black–Scholes partial differential equation.

It is possible to find analytical formulae for the price of basic compound options in the Black–Scholes framework when volatility is constant. These formulae involve the cumulative distribution function for a bivariate Normal variable. However, because of the second-order nature of compound options and thus their sensitivity to the precise nature of the asset price random walk, these formulae are dangerous to use in practice. Practitioners use either a stochastic volatility model or an implied volatility surface, two subjects I cover in later chapters.

Chooser options are similar to compounds in that they give the holder the right to buy a further option. With the chooser option the holder can choose whether to receive a call or a

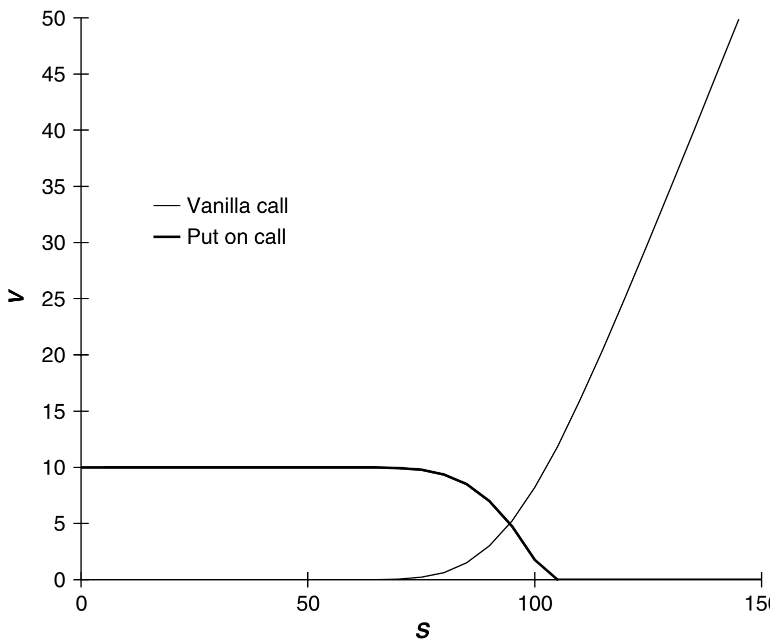


Figure 22.4 The value of a vanilla call option some time before expiry and the payoff for a put on this option.

put, for example. Generally, we can write the value of the chooser option as $Ch(S, t)$ and the value of the underlying options as $V_1(S, t)$ and $V_2(S, t)$ (or more). Now

$$\frac{\partial\, Ch}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2\, Ch}{\partial S^2} + rS\frac{\partial\, Ch}{\partial S} - r\, Ch = 0,$$

$$\frac{\partial V_1}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + rS\frac{\partial V_1}{\partial S} - rV_1 = 0$$

and

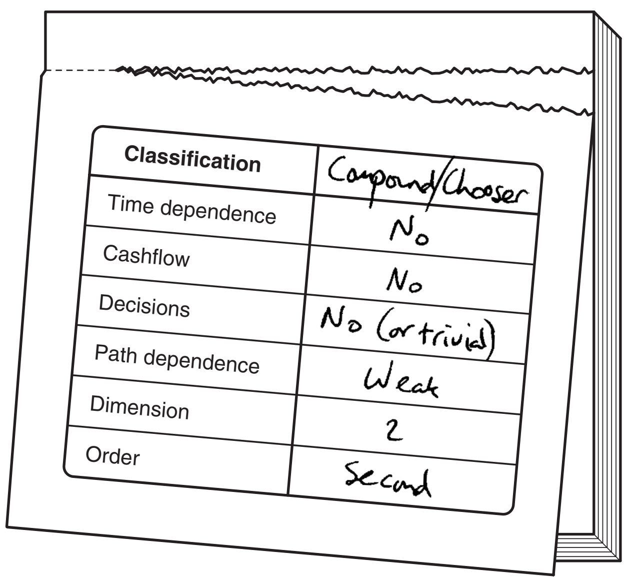
$$\frac{\partial V_2}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_2}{\partial S^2} + rS\frac{\partial V_2}{\partial S} - rV_2 = 0.$$

Final conditions are the usual payoffs for the underlying options at their expiry dates and

$$Ch(S, T_{Ch}) = \max(V_1(S, T_{Ch}) - E_1, V_2(S, T_{Ch}) - E_2, 0),$$

with the obvious notation.

The practical problems with pricing choosers are the same as for compounds.



Option classification table for Compounds and Choosers.

In Figure 22.5 is shown the values of a vanilla call and a vanilla put some time before expiry. In the same figure is the payoff for a call on the best of these two options (less an exercise price). This is the final condition for the Black–Scholes partial differential equation.

Figures 22.6 and 22.7 show the Bloomberg screens for valuing chooser options.

Extendible options are very, very similar to compounds and choosers. At some specified time the holder can choose to accept the payoff for the original option or to extend the option's life and even change the strike. Sometimes it is the writer who has these powers of extension. The reader has sufficient knowledge to be able to model these contracts in the Black–Scholes framework.

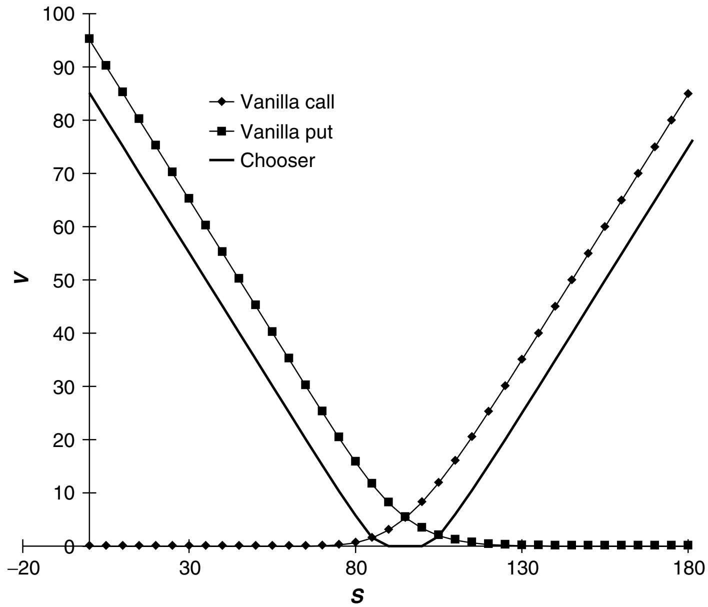


Figure 22.5 The value of a vanilla call option and a vanilla put option some time before expiry and the payoff for the best of these two.

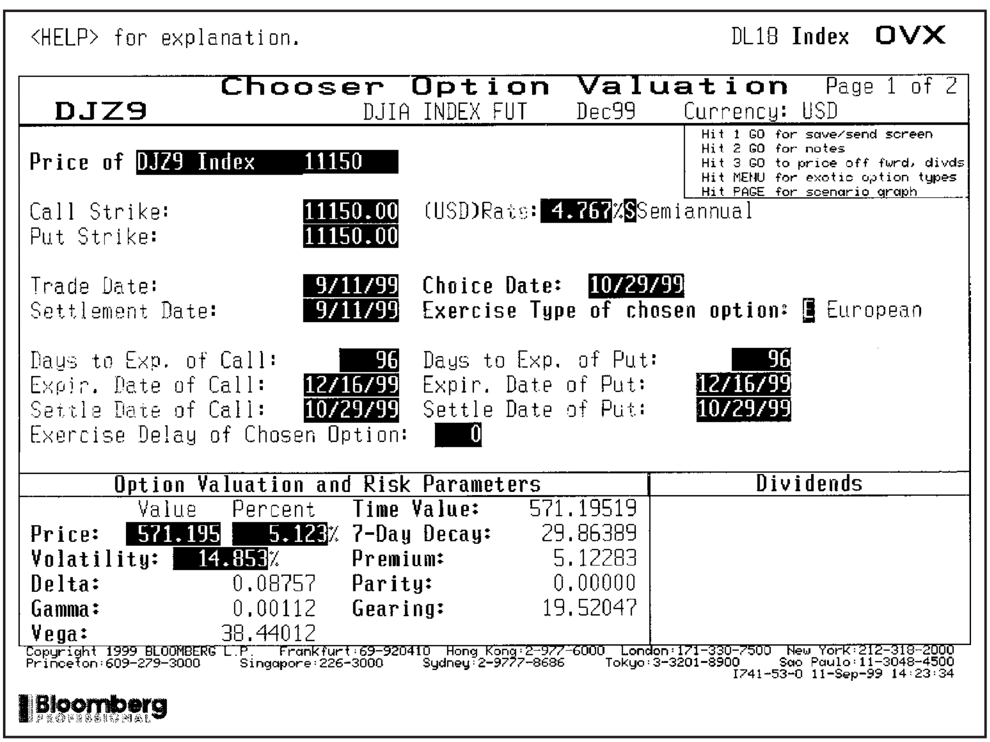


Figure 22.6 Bloomberg chooser option valuation screen. Source: Bloomberg L.P.

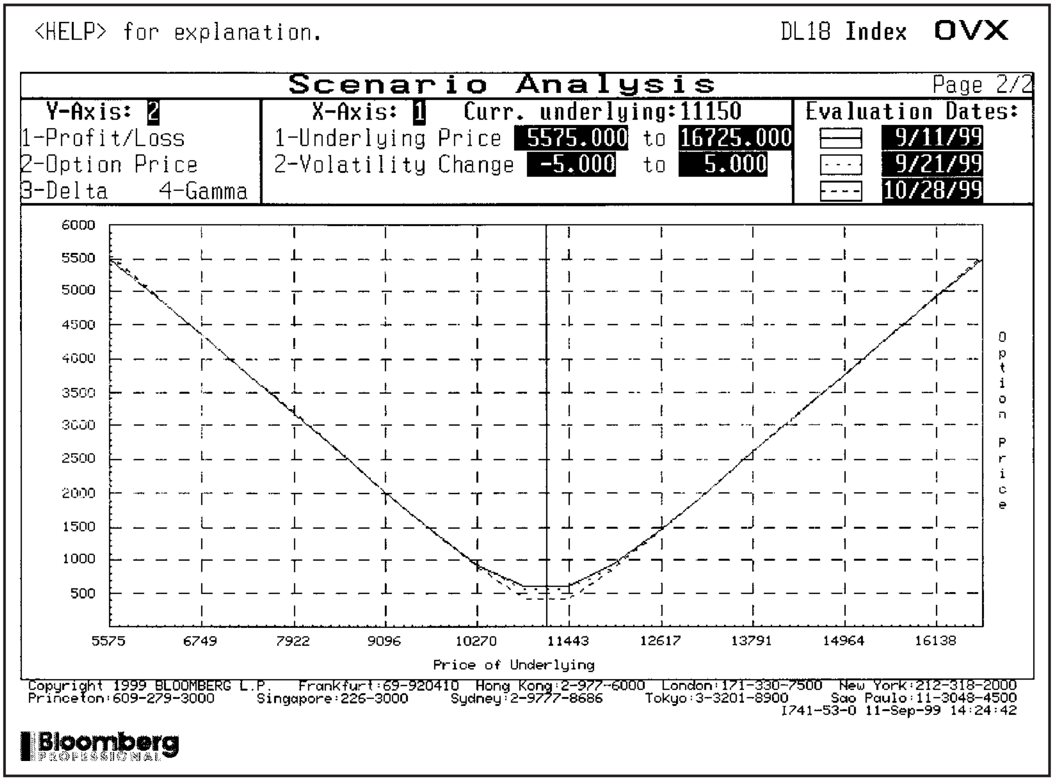
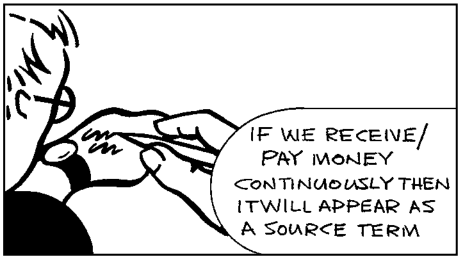


Figure 22.7 Bloomberg scenario analysis for a chooser. Source: Bloomberg L.P.

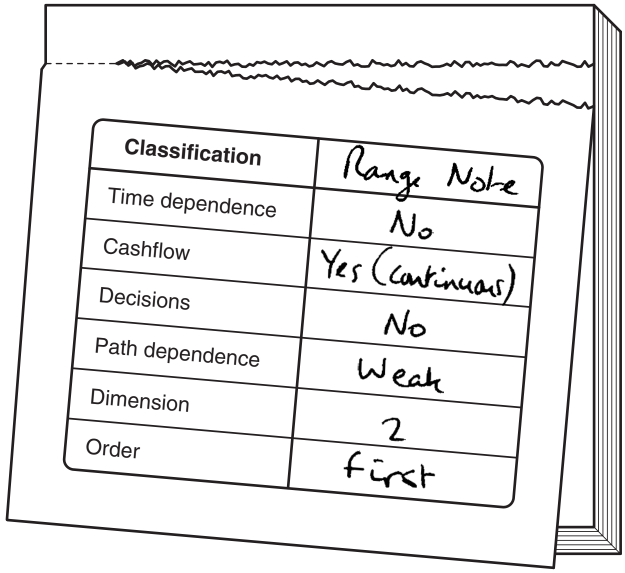
22.10.2 Range Notes

Range notes are very popular contracts, existing on the 'lognormal' assets such as equities and currencies, and as fixed-income products. In its basic, equity derivative, form the range note pays at a rate of L all the time that the underlying lies within a given range, $S_l \leq S \leq S_u$. That is, for every dt that the asset is in the range you receive $L\,dt$. Introducing $\mathcal{I}(S)$ as the function taking the value 1 when $S_l \leq S \leq S_u$ and zero otherwise, the range note satisfies

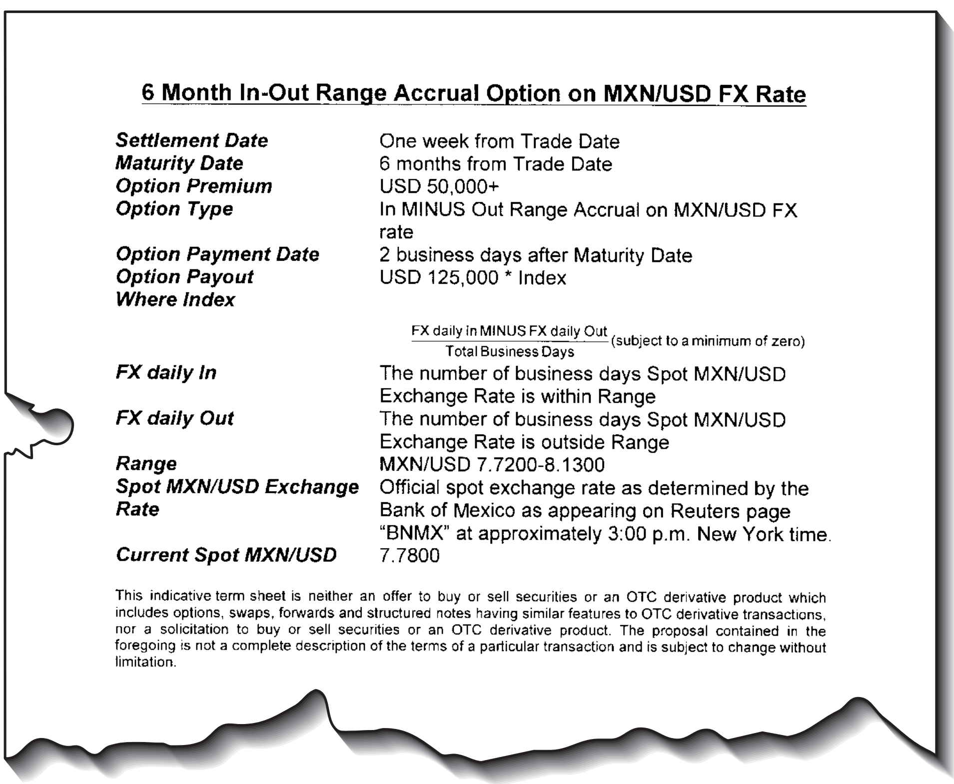
$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + L\mathcal{I}(S) = 0.$$



In Figure 22.8 is shown the term sheet for a range note on the Mexican peso–US dollar exchange rate. This contract pays out the positive part of the difference between number of days the exchange rate is inside the range less the number of days outside the range. This payment is received at expiry. (This contract is subtly different, and more complicated than the basic range note described above. Why? When you have finished Part Two you should be able to price this contract.)



Option classification table for a Range Note.



6 Month In-Out Range Accrual Option on MXN/USD FX Rate

Settlement Date	One week from Trade Date
Maturity Date	6 months from Trade Date
Option Premium	USD 50,000+
Option Type	In MINUS Out Range Accrual on MXN/USD FX rate
Option Payment Date	2 business days after Maturity Date
Option Payout	USD 125,000 * Index
Where Index	
	$\frac{\text{FX daily In MINUS FX daily Out}}{\text{Total Business Days}}$ (subject to a minimum of zero)
FX daily In	The number of business days Spot MXN/USD Exchange Rate is within Range
FX daily Out	The number of business days Spot MXN/USD Exchange Rate is outside Range
Range	MXN/USD 7.7200-8.1300
Spot MXN/USD Exchange Rate	Official spot exchange rate as determined by the Bank of Mexico as appearing on Reuters page "BNMX" at approximately 3:00 p.m. New York time.
Current Spot MXN/USD	7.7800

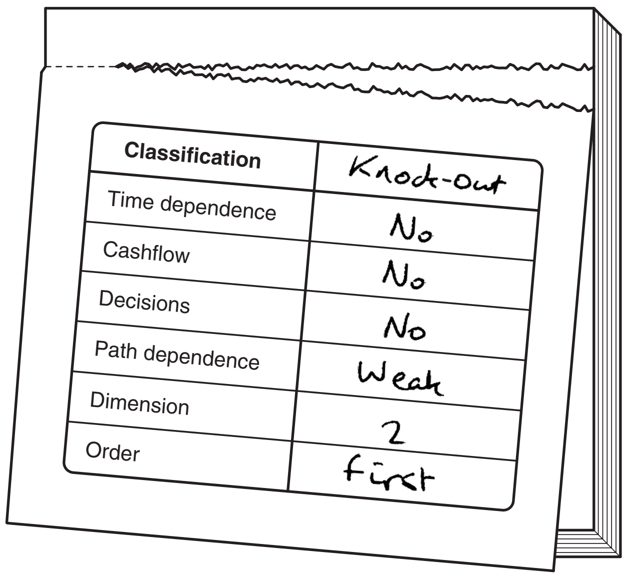
This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 22.8 Term sheet for an in-out range accrual note on MXN/USD.

22.10.3 Barrier Options

Barrier options have a payoff that is contingent on the underlying asset reaching some specified level before expiry. The critical level is called the barrier; there may be more than one. Barrier options are weakly path dependent. Barrier options are discussed in depth in Chapter 23.

Barrier options come in two main varieties, the 'in' barrier option (or **knock-in**) and the 'out' barrier option (or **knock-out**). The former only have a payoff if the barrier level is reached before expiry and the latter only have a payoff if the barrier is *not* reached before expiry. These contracts are weakly path dependent, meaning that the price depends only on the current level of the asset and the time to expiry. They satisfy the Black–Scholes equation, with special boundary conditions as we shall see.

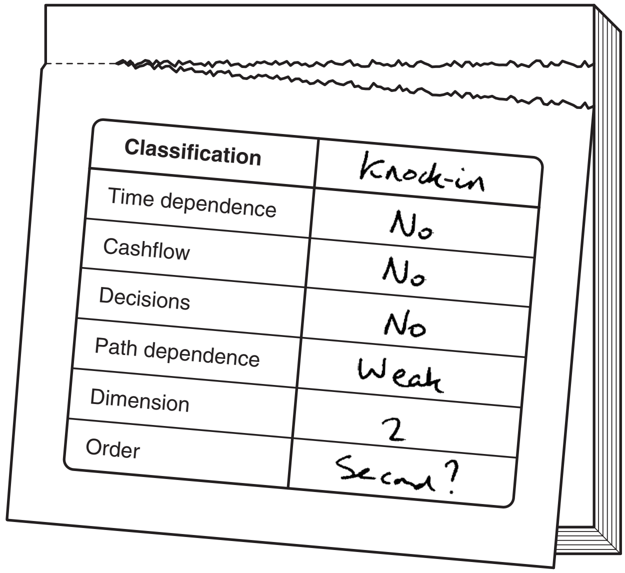


Classification option table for Knock-out.

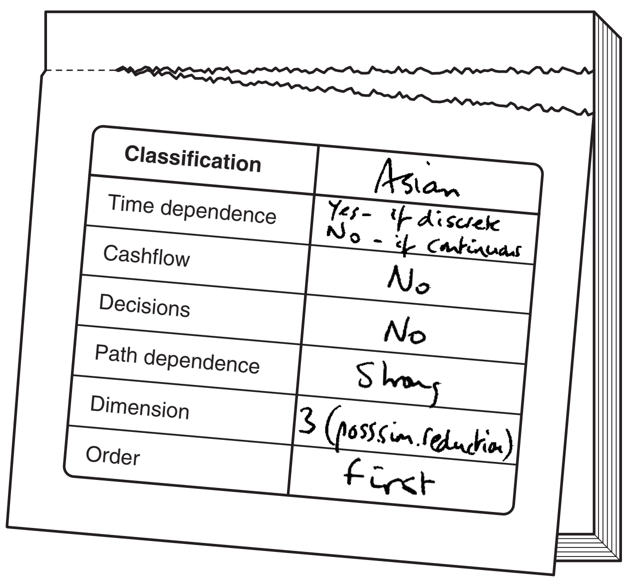
22.10.4 Asian Options

Asian options have a payoff that depends on the average value of the underlying asset over some period before expiry. They are the first strongly path-dependent contract we examine. They are strongly path dependent because their value prior to expiry depends on the path taken and not just on where they have reached. Their value depends on the *average to date* of the asset. This average to date will be very important to us; we introduce something like it as a new state variable. We shall see how to derive a partial differential equation for the value of this Asian contract, but now the differential equation will have *three* independent variables.

The average used in the calculation of the option's payoff can be defined in many different ways. It can be an arithmetic average or a geometric average, for example. The data could



Classification option table for Knock-in.



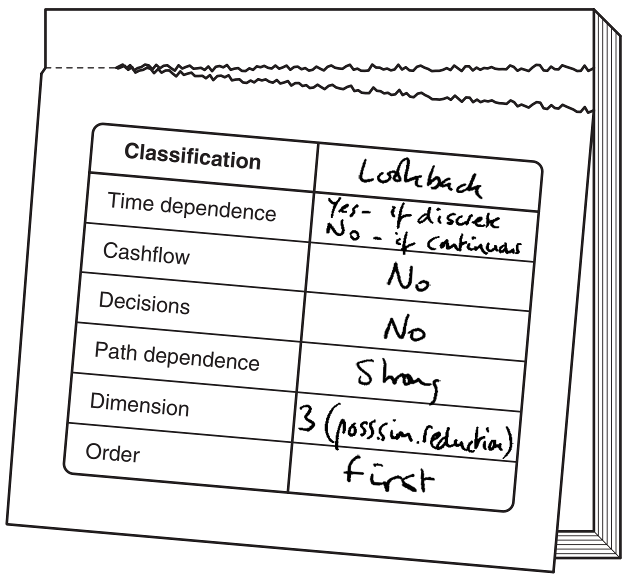
Classification option table for Asian Options.

be continuously sampled, so that every realized asset price over the given period is used. More commonly, for practical and legal reasons, the data are usually sampled discretely; the calculated average may only use every Friday's closing price, for example. We shall see in Chapter 25 how to price contracts with a wide range of definitions for the average and with either continuous or discrete sampling.

22.10.5 Lookback Options

Lookback options have a payoff that depends on the realized maximum or minimum of the underlying asset over some period prior to expiry. An extreme example, which captures the flavor of these contracts, is the option that pays off the difference between that maximum realized value of the asset and the minimum value over the next year. Thus it enables the holder to buy at the lowest price and sell at the highest, every trader's dream. Of course, this payoff comes at a price. And for such a contract that price would be very high.

Again the maximum or minimum can be calculated continuously or discretely, using every realized asset price or just a subset. In practice the maximum or minimum is measured discretely.



Classification option table for Lookback.

22.11 **SUMMARY OF MATH/CODING CONSEQUENCES**

Classification	Examples	Consequences
Time dependence	Bermudan exercise, discrete sampling, ...	Must keep track of time in code
Cashflow	Swap, instalments, ...	Jump in option value/Source term in pde
Path dependence	Barrier, Asian, lookback, ...	If strong path dependency, need extra dimension
Dimension	Strongly path-dependent, multi asset, ...	Monte Carlo may be better than finite difference
Order	Compounds, in barriers, ...	Solve lower level option(s) first and input into higher
Decisions	American, passport, chooser, ...	Finite difference better than Monte Carlo, 'optimize'

22.12 SUMMARY

This chapter suggests ways to think about derivative contracts that make their analysis simpler. To be able to make comparisons between different contracts is a big step forward in understanding them. After digesting this chapter, and the next few, you will be able to tell very quickly whether a particular contract is easy or difficult to price and hedge. And you will know whether the Black–Scholes framework is suitable, or whether it may be dangerous to apply it directly.

In this chapter we also began to look at some rather more complicated contracts than we have seen so far. We examine some of these contracts in depth in the next few chapters, considering them from both a theoretical and a practical viewpoint.

FURTHER READING

- Geske (1979) discusses the valuation of compound options.
- See Taleb (1997) for more details of classifications of the type I have described. This book is an excellent, and entertaining read.
- The book by Zhang (1997) is a discussion of many types of exotic options, with many formulae.
- See Kyprianou, Schoutens & Wilmott (2005) for Lévy processes and exotic option pricing.

CHAPTER 23
barrier options

In this Chapter...

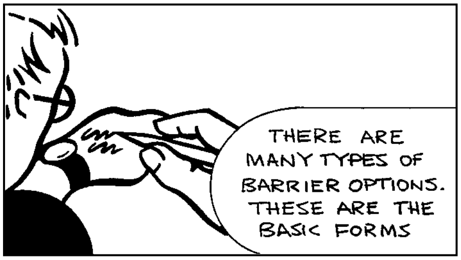
- the different types of barrier options
- how to price many barrier contracts in the partial differential equation framework
- some of the practical problems with the pricing and hedging of barriers

23.1 INTRODUCTION

I mentioned barrier options briefly in the previous chapter. But in this chapter we study them in detail, from both a theoretical and a practical perspective. **Barrier options** are path-dependent options. They have a payoff that is dependent on the realized asset path via its level; certain aspects of the contract are triggered if the asset price becomes too high or too low. For example, an up-and-out call option pays off the usual $\max(S - E, 0)$ at expiry unless at any time previously the underlying asset has traded at a value S_u or higher. In this example, if the asset reaches this level (from below, obviously) then it is said to 'knock out,' becoming worthless. Apart from 'out' options like this, there are also 'in' options which only receive a payoff if a level is reached, otherwise they expire worthless.

Barrier options are popular for a number of reasons. Perhaps the purchaser uses them to hedge very specific cashflows with similar properties. Usually, the purchaser has very precise views about the direction of the market. If he wants the payoff from a call option but does not want to pay for all the upside potential, believing that the upward movement of the underlying will be limited prior to expiry, then he may choose to buy an up-and-out call. It will be cheaper than a similar vanilla call, since the upside is severely limited. If he is right and the barrier is not triggered he gets the payoff he wanted. The closer that the barrier is to the current asset price then the greater the likelihood of the option being knocked out, and thus the cheaper the contract.

Conversely, an 'in' option will be bought by someone who believes that the barrier level will be realized. Again, the option is cheaper than the equivalent vanilla option.



23.2 DIFFERENT TYPES OF BARRIER OPTIONS

There are two main types of barrier option:

- The **out option**, this only pays off if a level is *not* reached. If the barrier is reached then the option is said to have **knocked out**.
- The **in option**, this pays off as long as a level is reached before expiry. If the barrier is reached then the option is said to have **knocked in**.

Then we further characterize the barrier option by the position of the barrier relative to the initial value of the underlying:

- If the barrier is above the initial asset value, we have an **up** option.
- If the barrier is below the initial asset value, we have a **down** option.

Finally, we describe the payoff received at expiry:

- The payoffs are all the usual suspects, call, put, binary, etc.

USD/MXN Double Knock-Out Note

Principal Amount	USD 10,000,000
Issuer	XXXX
Maturity	6 months from Trade Date
Issue Price	100%
Coupon	If the USD/MXN spot exchange rate trades above the Upper Barrier or below the Lower Barrier at any time during the term of the Note: Zero Otherwise: $400\%\times\max\left(0,\frac{8.2500-FX}{FX}\right)$ where FX is the USD/MXN spot exchange rate at Maturity
Redemption Amount	100%
Upper Barrier Level	8.2500
Lower Barrier Level	7.4500

This indicative termsheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 23.1 Term sheet for a USD/MXN double knock-out note.

The above classifies the commonest barrier options. In all of these contracts the position of the barrier could be time dependent. The level may begin at one level and then rise, say. Usually the level is a piecewise-constant function of time.

Another style of barrier option is the **double barrier**. Here there is both an upper and a lower barrier, the first above and the second below the current asset price. In a double 'out' option the contract becomes worthless if *either* of the barriers is reached. In a double 'in' option one of the barriers must be reached before expiry, otherwise the option expires worthless. Other possibilities can be imagined; one barrier is an 'in' and the other an 'out,' at expiry the contract could have either an 'in' or an 'out' payoff.

Sometimes a **rebate** is paid if the barrier level is reached. This is often the case for 'out' barriers in which case the rebate can be thought of as cushioning the blow of losing the rest of the payoff. The rebate may be paid as soon as the barrier is triggered or not until expiry.

In Figure 23.1 is shown the term sheet for a double knock-out option on the Mexican peso, US dollar exchange rate. The upper barrier is set at 8.25 and the lower barrier at 7.45. If the exchange rate trades inside this range until expiry then there is a payment. This is a very vanilla example of a barrier contract.

23.3 PRICING METHODOLOGIES

23.3.1 Monte Carlo Simulation

Pricing via Monte Carlo simulation is simple in principle:

- the value of an option is the present value of the expected payoff under a risk-neutral random walk.

The pricing algorithm:

1. Simulate the risk-neutral random walk starting at today's value of the asset over the required time horizon. This gives one realization of the underlying price path.
2. For this realization calculate the option payoff.
3. Perform many more such realizations over the time horizon.
4. Calculate the average payoff over all realizations.
5. Take the present value of this average, this is the option value.

Advantages of Monte Carlo pricing

- It is easy to code
- It is hard to make mistakes in the coding

Disadvantages of Monte Carlo pricing

- More work is needed to get the greeks
- It can be slow since tens of thousands of simulations are needed to get an accurate answer

23.3.2 Partial Differential Equations

Barrier options are path-dependent. Their payoff, and therefore value, depends on the path taken by the asset up to expiry.

- Yet that dependence is weak. We only have to know whether or not the barrier has been triggered, we do not need any other information about the path.

23.4 PRICING BARRIERS IN THE PARTIAL DIFFERENTIAL EQUATION FRAMEWORK

Barrier options are path-dependent, yet that dependence is weak. This is in contrast to some of the contracts we will be seeing shortly, such as the Asian option, that are strongly path-dependent. I use $V(S, t)$ to denote the value of the barrier contract *before the barrier has been triggered*. This value still satisfies the Black–Scholes equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

The details of the barrier feature come in through the specification of the boundary conditions.

23.4.1 'Out' Barriers

If the underlying asset reaches the barrier in an 'out' barrier option then the contract becomes worthless. This leads to the boundary condition

$$V(S_u, t) = 0 \quad \text{for } t < T,$$

for an up-barrier option with the barrier level at $S = S_u$. We must solve the Black–Scholes equation for $0 \le S \le S_u$ with this condition on $S = S_u$ and a final condition corresponding to the payoff received if the barrier is not triggered. For a call option we would have

$$V(S, T) = \max(S - E, 0).$$

If we have a down-and-out option with a barrier at S_d then we solve for $S_d \le S < \infty$ with

$$V(S_d, t) = 0,$$

and the relevant final condition at expiry.

The boundary conditions are easily changed to accommodate rebates. If a rebate of R is paid when the barrier is hit then

$$V(S_d, t) = R.$$

23.4.2 'In' Barriers

An 'in' option only has a payoff if the barrier is triggered. If the barrier is not triggered then the option expires worthless

$$V(S, T) = 0.$$

The value in the option is in the potential to hit the barrier. If the option is an up-and-in contract then on the upper barrier the contract must have the same value as a vanilla contract:

$$V(S_u, t) = \text{value of vanilla contract, a function of } t.$$

Using the notation $V_v(S, t)$ for value of the equivalent vanilla contract (a vanilla call, if we have an up-and-in call option) then we must have

$$V(S_u, t) = V_v(S_u, t) \quad \text{for } t < T.$$

A similar boundary condition holds for a down-and-in option.

The contract we receive when the barrier is triggered is a derivative itself, and therefore the 'in' option is a second-order contract. However, since an 'in' barrier option plus the same 'out' barrier option is a vanilla option, we can actually reduce the pricing problem to first order.

In solving for the value of an 'in' option completely numerically we must solve for the value of the vanilla option first, before solving for the value of the barrier option. The solution therefore takes roughly twice as long as the solution of the 'out' option.[1]

23.4.3 Some Formulae when Volatility is Constant

When volatility is constant we can solve for the theoretical price of many types of barrier contract. Some examples are given here and lots more can be found at the end of the chapter. (However, such formulae are rarely used in practice for reasons to be discussed below.)

I continue to use $V_v(S, t)$ for the value of the equivalent vanilla contract.

Down-and-out call option

As the first example, consider the down-and-out call option with barrier level S_d below the strike price E. The function $V_v(S, t)$ is the Black–Scholes value of a vanilla option with the same maturity and payoff as our barrier option. The value of the down-and-out option is then given by

$$V(S, t) = V_v(S, t) - \left(\frac{S}{S_d}\right)^{1-(2r/\sigma^2)} V_v\left(\frac{S_d^2}{S}, t\right).$$

Let us confirm that this is indeed the solution. First, does it satisfy the Black–Scholes equation? Clearly, the first term on the right-hand side does. The second term does also. Actually, if we have any solution, V_{BS}, of the Black–Scholes equation it is easy to show that

$$S^{1-(2r/\sigma^2)} V_{BS}\left(\frac{X}{S}, t\right)$$

is also a solution for any X.

[1] And, of course, the vanilla option must be solved for $0 \le S < \infty$.

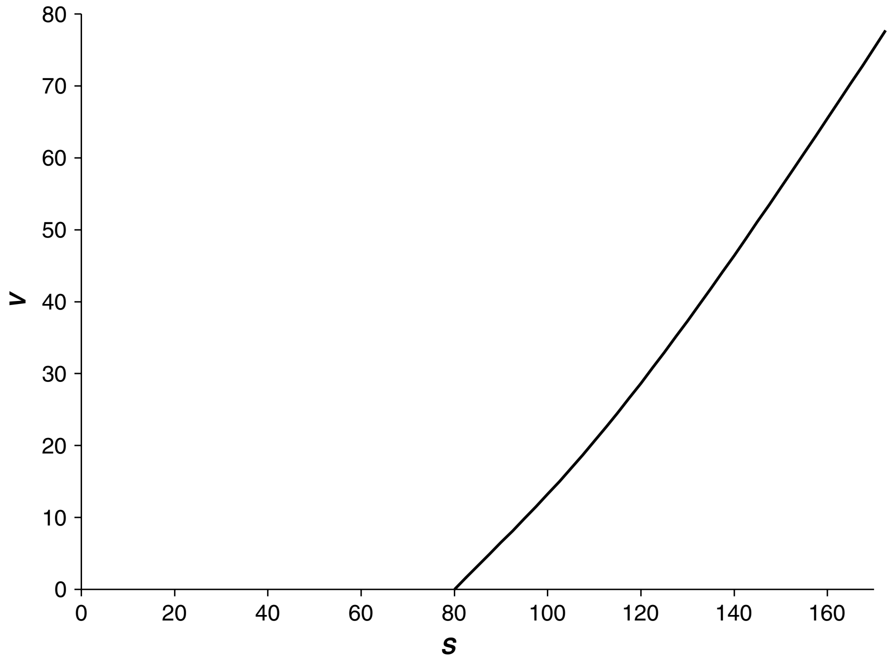


Figure 23.2 Value of a down-and-out call option.

What about the condition that the option value must be zero on $S = S_d$? Substitute $S = S_d$ in the above to confirm that this is the case. And the final condition? Since $S_d^2/S < E$ for $S > S_d$ the value of $V_v(S_d^2/S, T)$ is zero. Thus the final condition is satisfied.

The value of this option is shown as a function of S in Figure 23.2.

Down-and-in call option

In the absence of any rebates the relationship between an 'in' barrier option and an 'out' barrier option (with same payoff and same barrier level) is very simple:

$$\text{in} + \text{out} = \text{vanilla}.$$

If the 'in' barrier is triggered then so is the 'out' barrier, so whether or not the barrier is triggered we still get the vanilla payoff at expiry.

Thus, the value of a down-and-in call option is

$$V(S,t) = \left(\frac{S}{S_d}\right)^{1-(2r/\sigma^2)} V_v\left(\frac{S_d^2}{S}, t\right).$$

The value of this option is shown as a function of S in Figure 23.3. Also shown is the value of the vanilla call. Note that the two values coincide at the barrier.

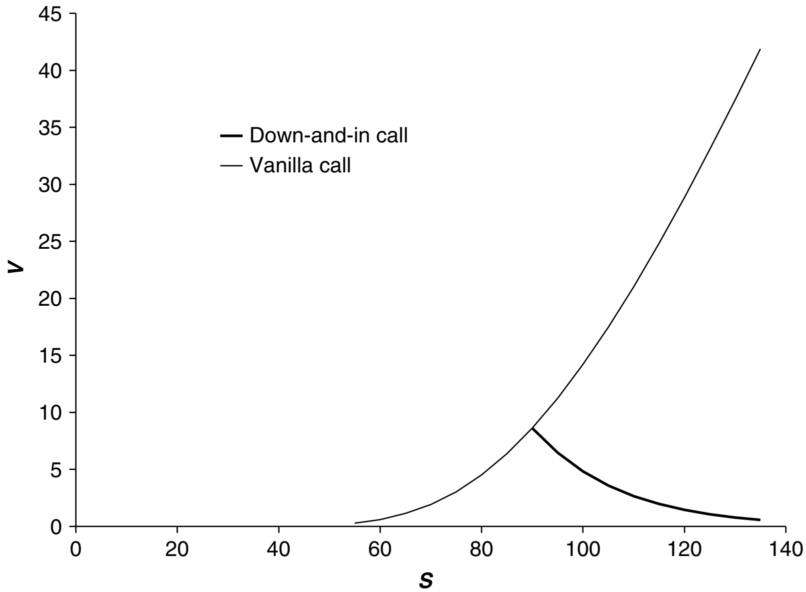


Figure 23.3 Value of a down-and-in call option.

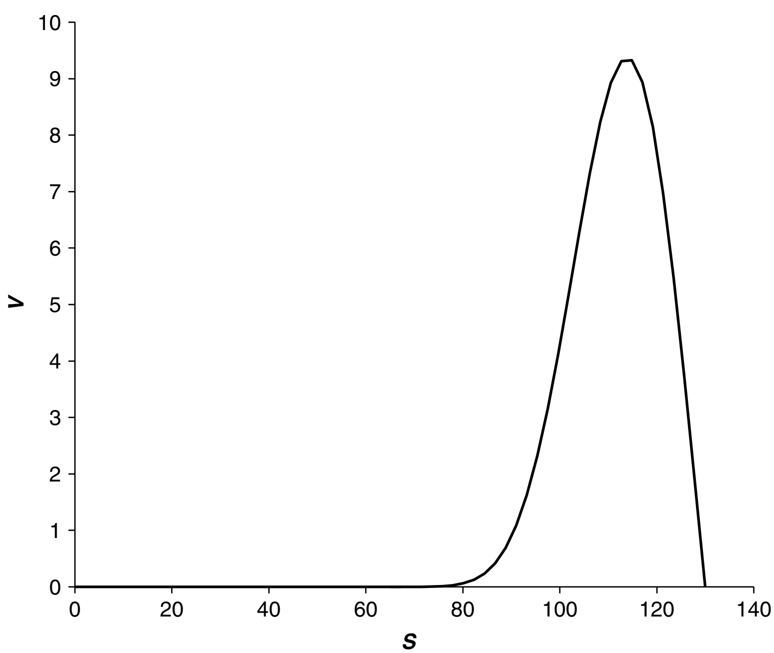


Figure 23.4 Value of an up-and-out call option.

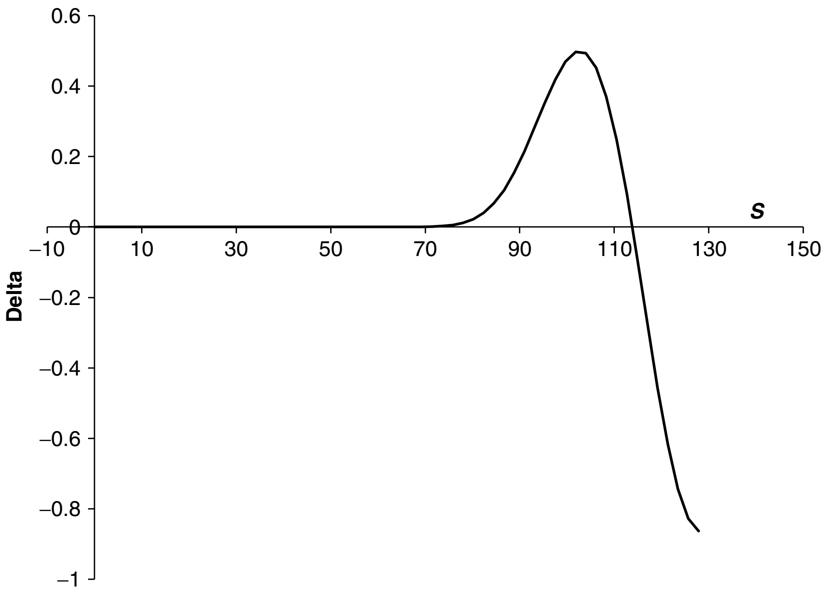
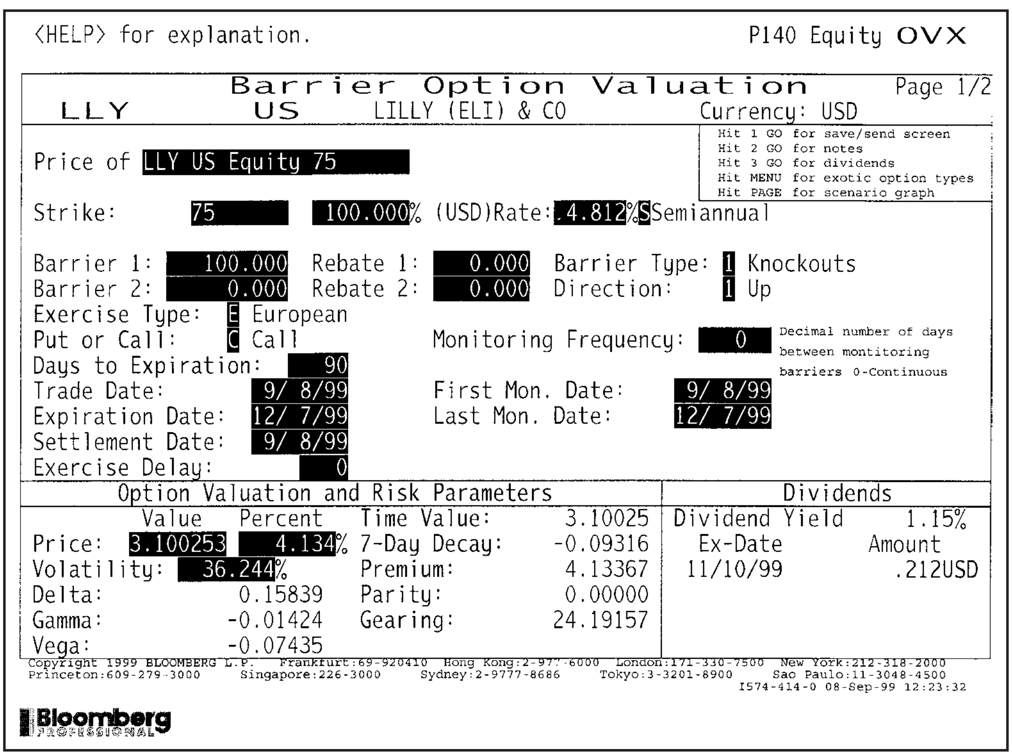


Figure 23.5 Delta of an up-and-out call option.



<HELP> for explanation. P140 Equity OVX

Barrier Option Valuation Page 1/2

LLY US LILLY (ELI) & CO Currency: USD

Hit 1 GO for save/send screen
Hit 2 GO for notes
Hit 3 GO for dividends
Hit MENU for exotic option types
Hit PAGE for scenario graph

Price of LLY US Equity 75

Strike: 75 100.000% (USD)Rate: 4.812% S Semiannual

Barrier 1: 100.000 Rebate 1: 0.000 Barrier Type: 1 Knockouts
Barrier 2: 0.000 Rebate 2: 0.000 Direction: 1 Up
Exercise Type: E European
Put or Call: C Call Monitoring Frequency: 0 Decimal number of days between montitoring barriers 0-Continuous
Days to Expiration: 90
Trade Date: 9/ 8/99 First Mon. Date: 9/ 8/99
Expiration Date: 12/ 7/99 Last Mon. Date: 12/ 7/99
Settlement Date: 9/ 8/99
Exercise Delay: 0

Option Valuation and Risk Parameters

	Value	Percent		
			Time Value:	3.10025
Price:	3.100253	4.134%	7-Day Decay:	-0.09316
Volatility:	36.244%		Premium:	4.13367
Delta:	0.15839		Parity:	0.00000
Gamma:	-0.01424		Gearing:	24.19157
Vega:	-0.07435			

Dividends

Dividend Yield	1.15%
Ex-Date	Amount
11/10/99	.212USD

Copyright 1999 BLOOMBERG L.P. Frankfurt:69-920410 Hong Kong:2-977-6000 London:171-330-7500 New York:212-318-2000
Princeton:609-279-3000 Singapore:226-3000 Sydney:2-9777-8686 Tokyo:3-3201-8900 Sao Paulo:11-3048-4500
I574-414-0 08-Sep-99 12:23:32

Bloomberg PROFESSIONAL

Figure 23.6 An up-and-out call again. Source: Bloomberg L.P.

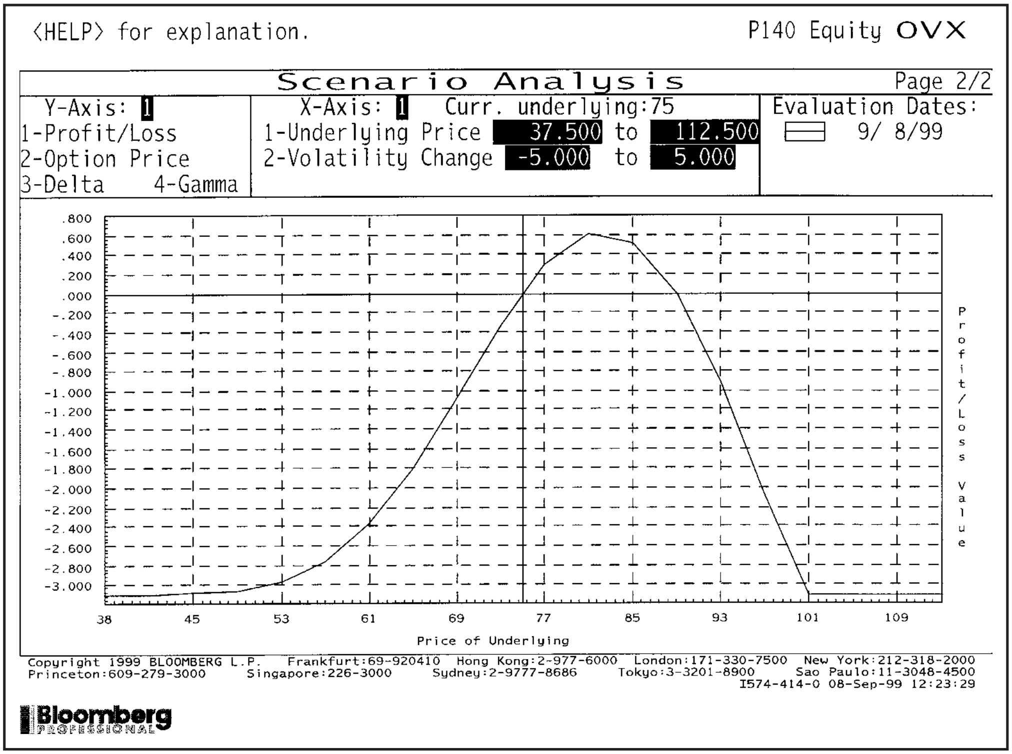


Figure 23.7 Profit/loss for an up-and-out call. Source: Bloomberg L.P.

Up-and-out call option

The barrier S_u for an up-and-out call option must be above the strike price E (otherwise the option would be valueless). This makes the solution for the price more complicated, and I just quote it here. The value of an up-and-out call option is

$$S\left(N(d_1) - N(d_3) - b(N(d_6) - N(d_8))\right) - Ee^{-r(T-t)}\left(N(d_2) - N(d_4) - a(N(d_5) - N(d_7))\right),$$

where $N(\cdot)$ is the cumulative distribution function for a standardized Normal variable and a, b and the ds are given at the end of the chapter.

The value of this option is shown as a function of S in Figure 23.4. In Figure 23.5 is shown the delta.

Formulae can be found for many barrier options (assuming volatility is constant), see the Appendix at the end of this chapter. When there are two barriers the solution can often be found by Fourier series, see Chapter 6.

Figure 23.6 shows the Bloomberg barrier option calculator and Figure 23.7 shows the option *profit/loss* against asset price.

23.4.4 Some More Examples

Figures 23.8–23.13 are all taken from Bloomberg, who use the formulae explained above, and below, for the pricing.

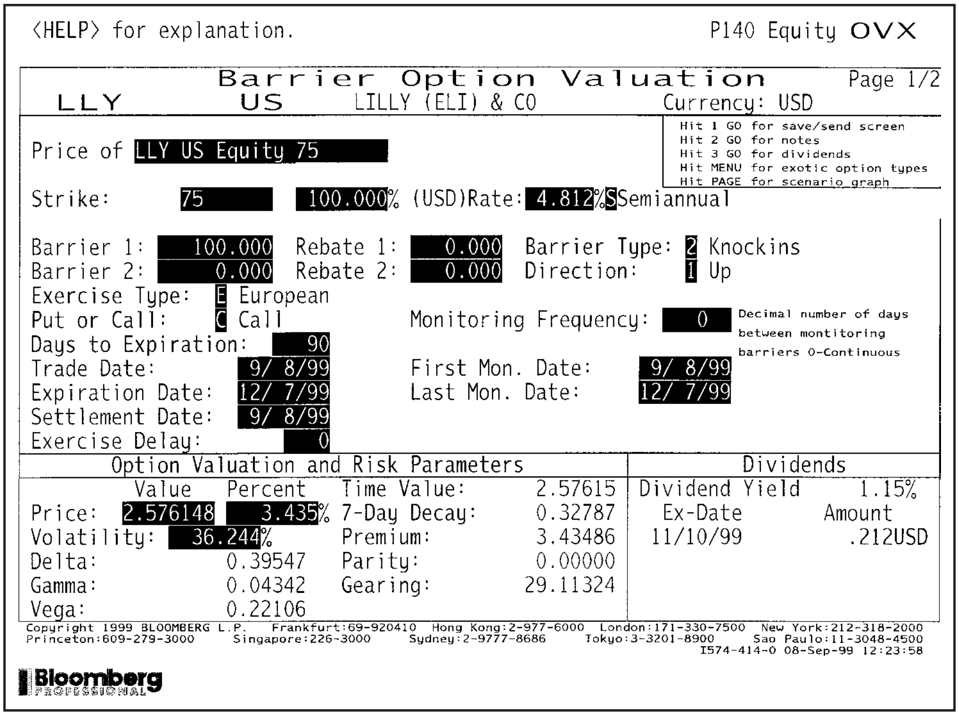


Figure 23.8 Calculator for an up-and-in call. Source: Bloomberg L.P.

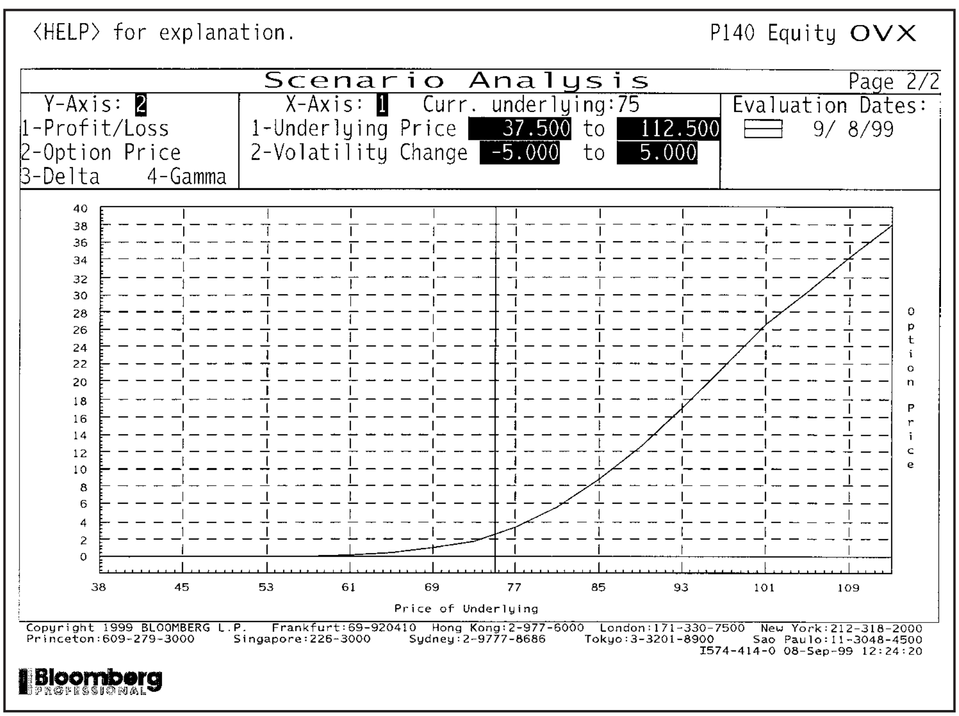


Figure 23.9 Value of an up-and-in call. Source: Bloomberg L.P.

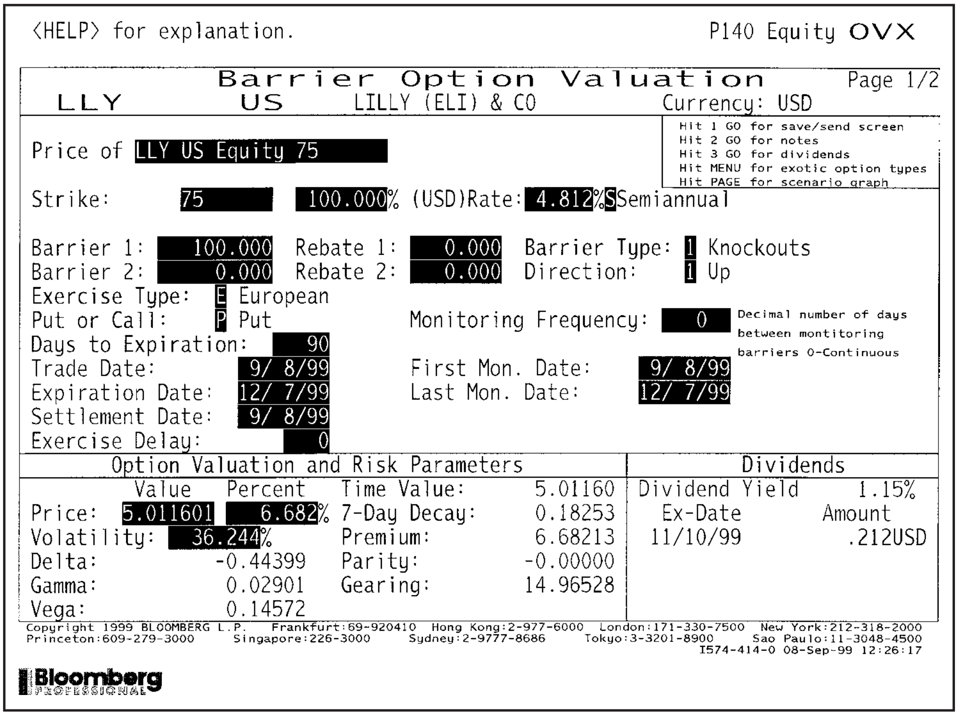


Figure 23.10 Calculator for an up-and-out put. Source: Bloomberg L.P.

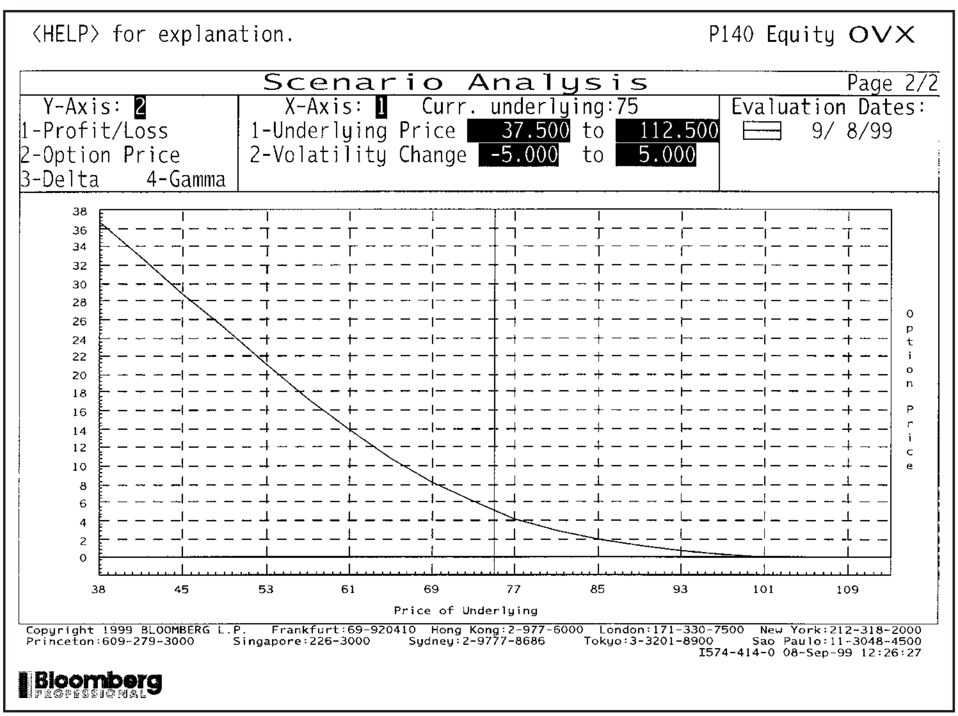


Figure 23.11 Value of an up-and-out put. Source: Bloomberg L.P.

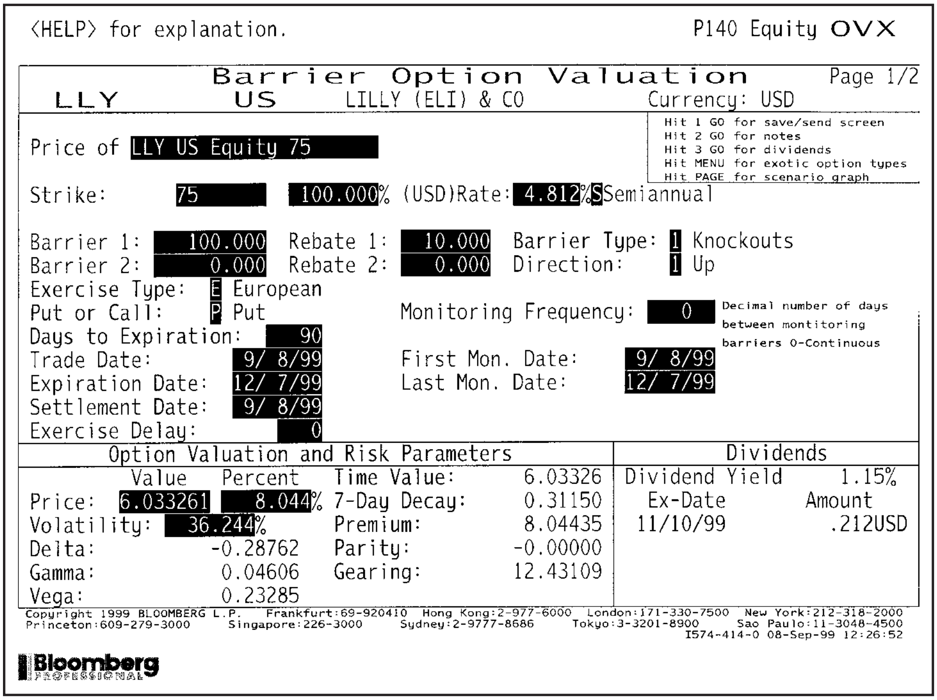


Figure 23.12 Calculator for an up-and-out put with a rebate on the upper barrier. Source: Bloomberg L.P.

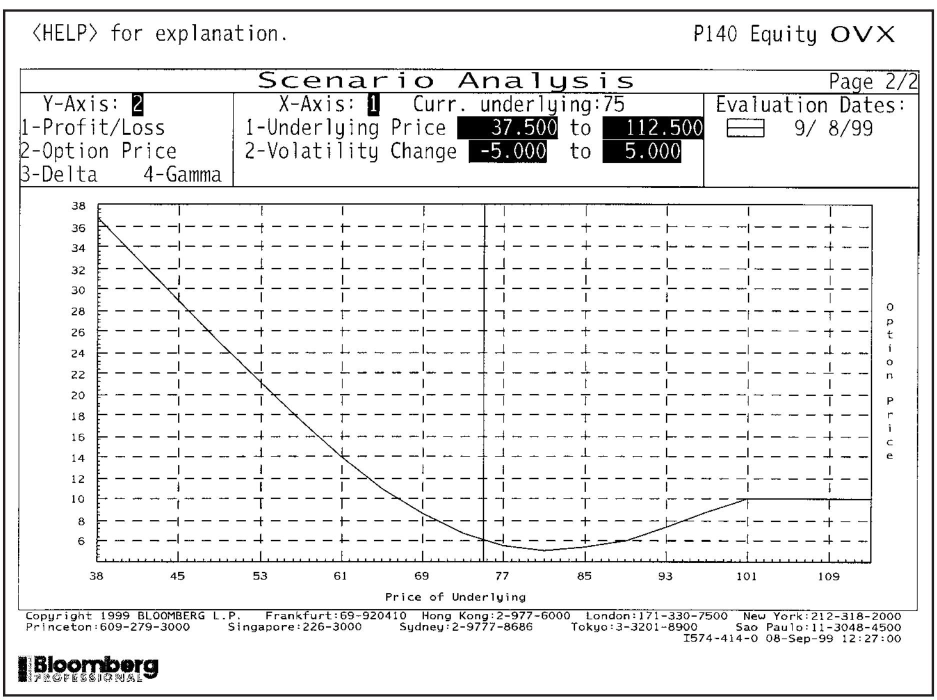


Figure 23.13 Value of an up-and-out put with a rebate on the upper barrier. Source: Bloomberg L.P.

23.5 OTHER FEATURES IN BARRIER-STYLE OPTIONS

Not so long ago barrier options were exotic, the market for them was small and few people were comfortable pricing them. Nowadays they are heavily traded and it is only the contracts with more unusual features that can rightly be called exotic. Some of these features are described below.

23.5.1 Early Exercise

It is possible to have American-style early exercise. The contract must specify what the payoff is if the contract is exercised before expiry. As always, early exercise is a simple constraint on the value of the option.

In Figure 23.14 is the term sheet for a Knock-out Instalment Premium Option on the US dollar, Japanese yen exchange rate. This knocks out if the exchange rate ever goes above 140. If the option expires without ever hitting this level there is a vanilla call payoff. I mention this contract in the section on early exercise because it has a similar feature. To keep the contract alive the holder must pay in instalments every month. We saw this instalment feature in Chapter 9 were it was likened to American exercise. The question is when to stop paying the instalments? This can be done optimally.

USD/JPY KO Instalment-Premium Option

Notional Amount	USD 50,000,000
Option Type	133.25 (ATMS) USD Put/JPY Call with KO and Instalment Premium
Maturity	6 months from Trade Date
Knockout Mechanism	If, at any time from Trade Date to Maturity, the USD/JPY spot rate trades in the interbank market at or above JPY 140.00 per USD, the option will automatically be cancelled, with no further rights or obligations arising for the parties thereto.
Upfront Premium	JPY 1.50 per USD
Instalments	JPY 1.50 per USD, payable monthly from Trade Date (5 instalments)
Instalment Mechanism	As long as the instalments continue to be paid, the option will be kept alive, but the Counterparty has the right to cease paying the instalments and to thereby let the option be cancelled at any time.
Spot Reference	JPY 133.25 per USD

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 23.14 Term sheet for a USD/JPY knock-out instalment premium option.

23.5.2 The Intermittent Barrier

The position of the barrier(s) can be time-dependent. A more extreme version of a time-dependent barrier is to have a barrier that disappears altogether for specified time periods. These options are called **protected** or **partial** barrier options. An example is shown in Figure 23.15.

There are two types of such contract. In one the barrier is triggered as long as the asset price is beyond the barrier on days on which the barrier is active. The solution of this problem is shown schematically in Figure 23.16.

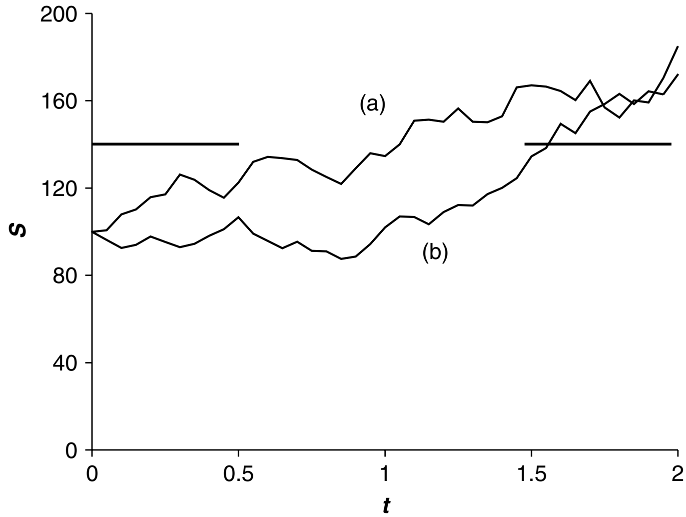


Figure 23.15 The intermittent barrier. Two varieties: (a) Barrier triggered if asset outside barrier on active days; (b) Barrier only triggered by asset price crossing barrier.

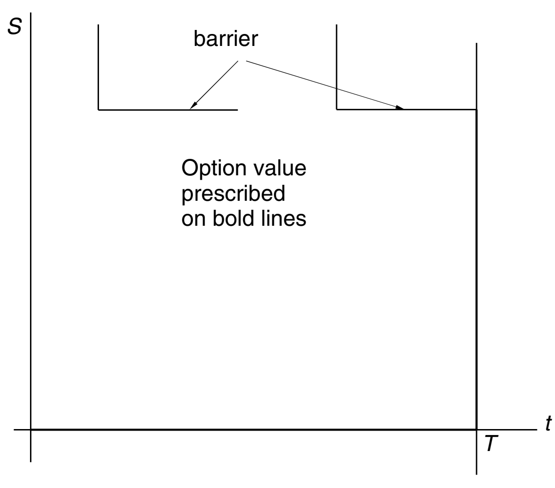


Figure 23.16 The intermittent barrier: Barrier triggered if asset outside barrier on active days. Solution procedure.

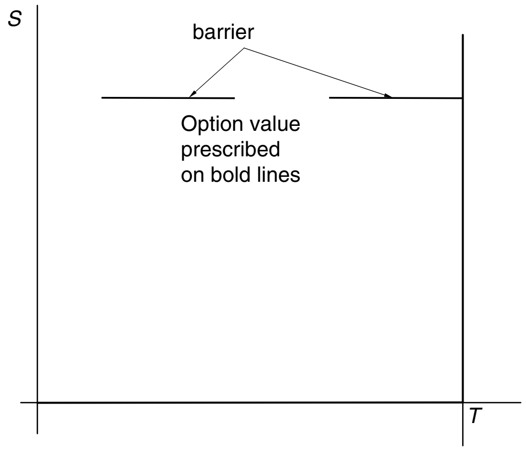


Figure 23.17 The intermittent barrier: Barrier only triggered by asset price crossing barrier. Solution procedure.

The second type of intermittent barrier is only triggered if the asset path crosses the barrier on an active day. The barrier will not be triggered if the asset weaves its way through the barriers. The solution of this problem is shown schematically in Figure 23.17.

23.5.3 Repeated Hitting of the Barrier

The double barrier that we have seen above can be made more complicated. Instead of only requiring one hit of either barrier we could insist that *both* barriers are hit before the barrier is triggered.

This contract is easy to value. Observe that the first time that one of the barriers is hit the contract becomes a vanilla barrier option. Thus on the two barriers we solve the Black–Scholes equation with boundary conditions that our double barrier value is equal to an up-barrier option on the lower barrier and a down-barrier option on the upper barrier. This makes this barrier option second order, since we have to solve for another contract first.

In Chapter 28 we will see the related Parisian option, the payoff of which depends on the time that the asset has been beyond the barrier.

23.5.4 Resetting of Barrier

Another type of barrier contract that can be priced by the same two- (or more) step procedure as 'in' barriers is the reset barrier. When the barrier is hit the contract turns into another barrier option with a different barrier level. The contract may be time-dependent in the sense that if the barrier is hit before a certain time we get a new barrier option, and if it is hit after a certain time we get the vanilla payoff.

Related to these contracts are the **roll-up** and **roll-down options**. These begin life as vanilla options, but if the asset reaches some predefined level they become a barrier option. For example, with a roll-up put if the roll-up strike level is reached the contract becomes an up-and-out put with the roll-up strike being the strike of the barrier put. The barrier level will then be at a prespecified level.

23.5.5 Outside Barrier Options

Outside or **rainbow barrier options** have payoffs or a trigger feature that depends on a second underlying. Thus the barrier might be triggered by one asset, with the payoff depending on the other. These products are clearly multi-factor contracts.

23.5.6 Soft Barriers

The **soft barrier option** allows the contract to be gradually knocked in or out. The contract specifies two levels, an upper and a lower. In the knock-out option a proportion of the contract is knocked out depending on the distance that the asset has reached between the two barriers. For example, suppose that the option is an up and out with a soft barrier range of 100 to 120. If the maximum asset value reached before expiry is 105 then 5/20 or 25% of the payoff is lost.

23.5.7 Parisian Options

Parisian options have barriers that are triggered only if the underlying has been beyond the barrier level for more than a specified time. This additional feature reduces the possibility of manipulation of the trigger event and makes the dynamic hedging easier. However, this new feature also increases the dimensionality of the problem and so we must wait for a few chapters before we can analyze this contract fully. In Chapter 28 we will see the Parisian option in detail.

23.5.8 The Emergency Exit

The **Emergency Exit** is a feature that can be found in any exotic contract. It is an escape clause allowing you to exit a position under prescribed conditions. For example, an emergency exit in a contract might specify that you can 'escape' at any time before expiration and receive a rebate. This is very much like an American exercise feature and is mathematically modeled as such. If the rebate is R then we would have

$$V \geq R.$$

It is possible for R to be time-dependent, so the escape clause is only active for set times, and/or to be S dependent, so that you can only escape when the stock price is at certain levels.

23.6 FIRST-EXIT TIME

The path dependency in a barrier option arises because the option payoff depends on whether or not the barrier has been triggered. The *value* can be interpreted as the present value of the risk-neutral expected payoff but the likelihood of the barrier being triggered before expiry only has meaning if we calculate the probability using the *real* random walk for the asset. For an up-and-in barrier option, the probability of the barrier being triggered before expiry, $Q(S,t)$, is given by the solution of

$$\frac{\partial Q}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 Q}{\partial S^2} + \mu S \frac{\partial Q}{\partial S} = 0$$

with

$$Q(S, T) = 0 \quad \text{and} \quad Q(S_u, t) = 1,$$

as discussed in Chapter 10. Because we are using the real process for S this problem contains the real drift rate μ. The expected time $u(S)$ before the level S_u is hit from below is the solution of

$$\tfrac{1}{2}\sigma^2 S^2 \frac{d^2u}{dS^2} + \mu S \frac{du}{dS} = -1$$

i.e.

$$\frac{1}{\frac{1}{2}\sigma^2 - \mu} \log\left(\frac{S}{S_u}\right),$$

but only for $2\mu > \sigma^2$. If $2\mu < \sigma^2$ then the expected first-exit time is infinite.

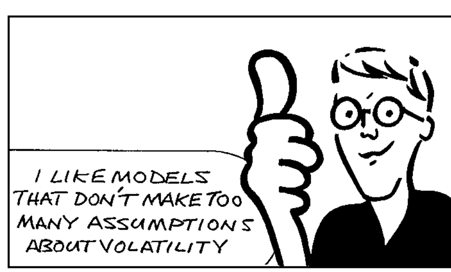
Calculations like these can be used by the speculator who has a view on the direction of the underlying, believing that the barrier will or will not be triggered. His view can be quantified as a probability, for example. Or he can determine whether the first-exit time is greater or less than the remaining time to maturity.

The hedger will also find such calculations useful. As we discuss below, delta hedging barrier options is notoriously difficult and usually they are statically hedged as well to some extent. The choice of the static hedge may be influenced by the *real* time at which the barrier is expected to be triggered.

23.7 MARKET PRACTICE: WHAT VOLATILITY SHOULD I USE?

Practitioners do not price contracts using a single, constant volatility. Let us see some of the pitfalls with this, and then see what practitioners do.

In Figure 23.18 we see a plot of the value of an up-and-out call option using three different volatilities, 15%, 20% and 25%. I have chosen three very different values to make a point. If we are uncertain about the value of the volatility (as we surely are) then which value do we use to price the contract? Observe that at approximately $S = 100$ the option value seems to be insensitive to the volatility; the vega is zero. It almost looks as if it doesn't really matter what volatility we use. This is highly misleading.



Ask the question, Do I believe that volatility will be one of 15%, 20% or 25%, and will be fixed at that level? Or do I believe that volatility could move around between 15% and 25%? Clearly the latter is closer to the truth. But the measurement of vega, and the plots in Figure 23.18, assume that volatility is fixed until expiry. If we are concerned with playing it safe we should assume that the behavior of volatility will be that which gives us the lowest value if we are buying the contract. The worst outcome for volatility is for it to be low

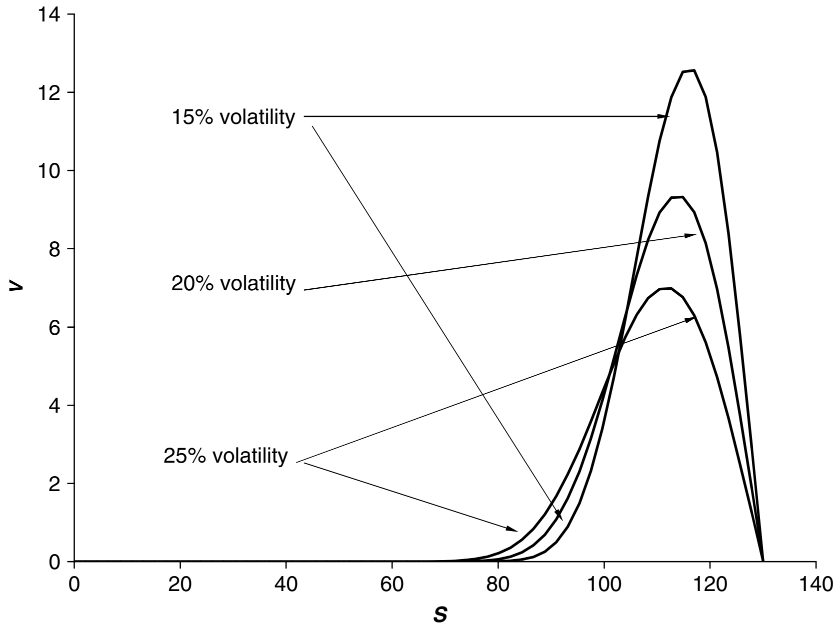


Figure 23.18 Theoretical up-and-out call price with three different volatilities.

around the strike price, and high around the barrier. Financially, this means that if we are near the strike we get a small payoff, but if we are near the barrier we are likely to hit it. Mathematically, the 'worst' choice of volatility path depends on the sign of the gamma at each point. If gamma is positive then low volatility is bad, if gamma is negative then high volatility is bad. A better way to price options when the volatility is uncertain is described in Chapter 52. When the gamma is not single-signed, the measurement of vega can be meaningless. Barrier options with non-single-signed gamma include the up-and-out call, down-and-out put and many double-barrier options.

Figures 23.19 through 23.22 show the details of a double knockout put contract, its price versus the underlying, its gamma versus the underlying and its price versus volatility. This is a contract with a gamma that changes sign as can be seen from Figure 23.21. You must be very careful when pricing such a contract as to what volatility to use. Suppose you wanted to know the implied volatility for this contract when the price was 3.2, what value would you get? Refer to Figure 23.22.

To accommodate problems like this, practitioners have invented a number of 'patches.' One is to use two different volatilities in the option price. For example, one can calculate implied volatilities from vanilla options with the same strike, expiry and payoff as the barrier option and also from American-style one-touch options with the strike at the barrier level. The implied volatility from the vanilla option contains the market's estimate of the value of the payoff, but including all the upside potential that the call has but which is irrelevant for the up-and-out option. The one-touch volatility, however, contains the market's view of the likelihood of the barrier level being reached. These two volatilities can be used to price an up-and-out call by observing that an 'out' option is the same as a vanilla minus an 'in' option. Use the vanilla volatility to price the vanilla call and the one-touch volatility to price the 'in' call.

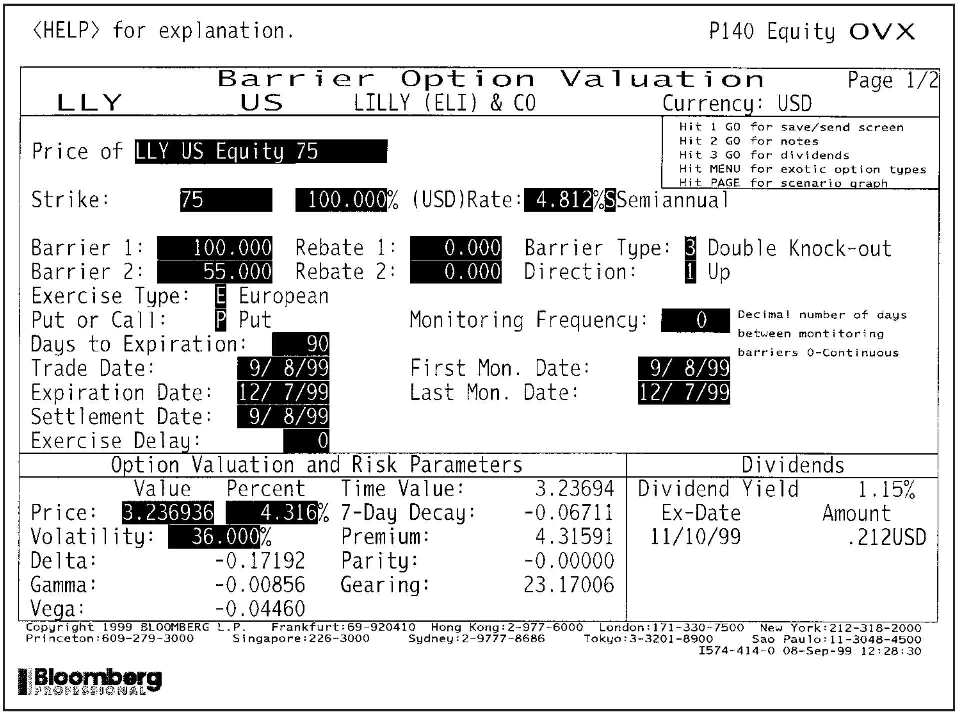


Figure 23.19 Details of a double knockout put. Source: Bloomberg L.P.

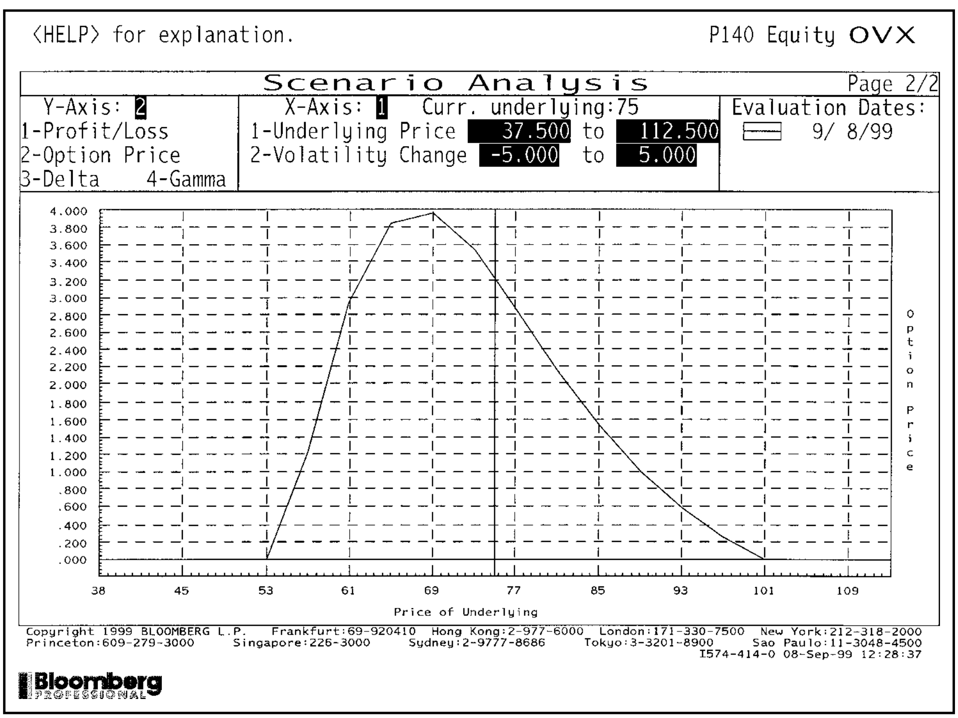


Figure 23.20 Price of the double knockout put. Source: Bloomberg L.P.

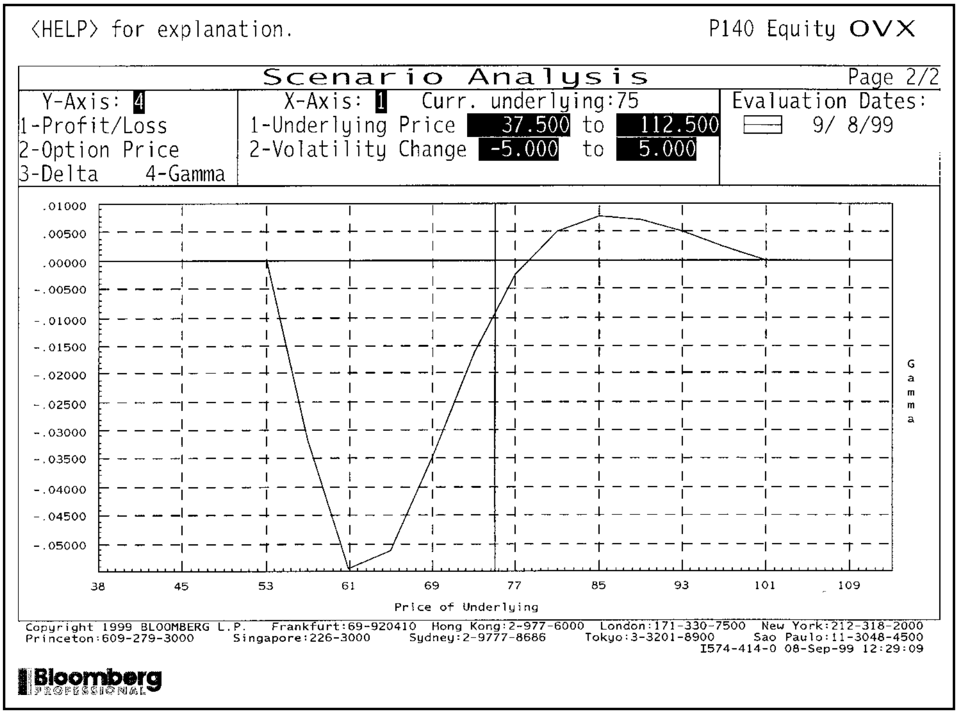


Figure 23.21 Gamma of the double knockout put. Source: Bloomberg L.P.

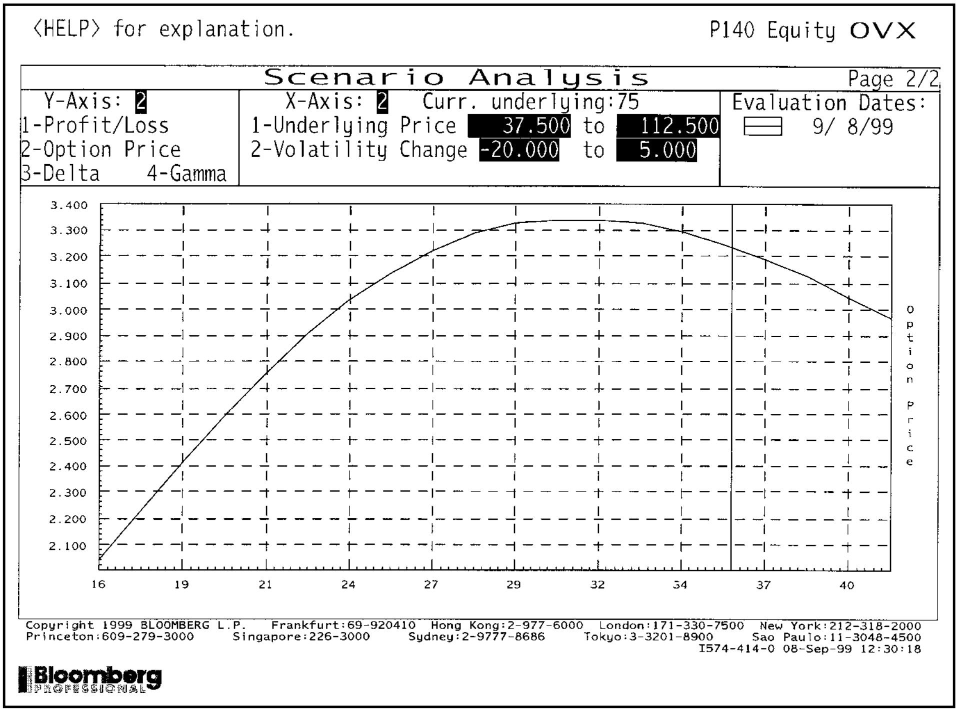


Figure 23.22 Option price versus volatility for the double knockout put. Source: Bloomberg L.P.

Another practitioner approach to the pricing is to use a volatility surface, implied from market prices of all traded vanilla contracts. This is then employed in a binomial tree or finite-difference scheme to price the barrier option *consistently* across instruments. This is the subject of Chapter 50. Stochastic volatility models are also commonly used for pricing barriers, see Chapter 51. There is no standard model for pricing barriers, hence the use in practice of several different models. Barrier options are sufficiently simple and common that you feel everyone ought to be able to agree on a price, and margins can be quite tight. However, they are not yet sufficiently liquid that the market will price them for you.

23.8 HEDGING BARRIER OPTIONS

Barrier options have discontinuous delta at the barrier. For a knock-out, the option value is continuous, decreasing approximately linearly towards the barrier then being zero beyond the barrier. This discontinuity in the delta means that the gamma is instantaneously infinite at the barrier. Delta hedging through the barrier is virtually impossible, and certainly very costly. This raises the issue of whether there are improvements on delta hedging for barrier options.

There have been a number of suggestions made for ways to hedge barrier options *statically*. These methods try to mimic as closely as possible the value of a barrier option with vanilla calls and puts, or with binary options. In Chapter 60 I describe a couple of ways of statically hedging barrier options with traded vanilla options. A very common practice for hedging a short up-and-out call is to buy a long call with the same strike and expiry. If the option does knock out then you are fortunate in being left with a long call position.

I now describe another simple but useful technique, based on the **reflection principle** and **put-call symmetry**. This technique only really works if the barrier and strike lie in the correct order, as we shall see. The method gives an approximate hedge only.

The simplest example of put-call symmetry is actually put-call parity. At all asset levels we have

$$V_C - V_P = S - Ee^{-r(T-t)},$$

where E is the strike of the two options, and C and P refer to call and put. Suppose we have a down-and-in call, how can we use this result? To make things simple for the moment, let's have the barrier and the strike at the same level. Now hedge our down-and-in call with a short position in a vanilla put with the same strike. If the barrier is reached we have a position worth

$$V_C - V_P.$$

The first term is from the down-and-in call and the second from the vanilla put. This is exactly the same as

$$S - Ee^{-r(T-t)} = E(1 - e^{-r(T-t)}),$$

because of put-call parity and since the barrier and the strike are the same. If the barrier is not touched then both options expire worthless. If the interest rate were zero then we would have a perfect hedge. If rates are non-zero what we are left with is a one-touch option with small and time-dependent value on the barrier. Although this leftover cashflow is non-zero, it is small, bounded and more manageable than the original cashflows.

Now suppose that the strike and the barrier are distinct. Let us continue with the down-and-in call, now with barrier below the strike. The static hedge is not much more complicated than the previous example. All we need to know is the relationship between the value of a call option with strike E when $S = S_d$ and a put option with strike S_d^2/E. It is easy to show from the formulae for calls and puts that if interest rates are zero, the value of this call at $S = S_d$ is equal to a number E/S_d of the puts, valued at S_d. We would therefore hedge our down-and-in call with E/S_d puts struck at S_d^2/E. Note that the geometric average of the strike of the call and the strike of the put is the same as the barrier level; this is where the idea of 'reflection' comes in. The strike of the hedging put is at the reflection in the barrier of the call's strike. When rates are non-zero there is some error in this hedge, but again it is small and manageable, decreasing as we get closer to expiry. If the barrier is not touched then both options expire worthless (the strike of the put is below the barrier remember).

If the barrier level is above the strike, matters are more complicated since if the barrier is touched we get an in-the-money call. The reflection principle does not work because the put would also be in the money at expiry if the barrier is not touched.

In Chapter 60 we see how to hedge contracts statically by matching payoffs around a boundary. This technique is particularly suited to barrier options.

23.9 SLIPPAGE COSTS

The delta of a barrier option is discontinuous at the barrier, whether it is an in- or an out-option. This presents a particular problem to do with **slippage** or **gapping**. Should the underlying move significantly as the barrier is triggered it is likely that it will not be possible to hedge continuously through the barrier. For example, if the contract is knocked out then one finds oneself with a $-\Delta$ holding of the underlying that should have been offloaded sooner. This can have a significant effect on the hedging costs.

It is not too difficult to allow for the *expected* slippage costs, and all that is required is a slight modification to the apparent barrier level.

At the barrier we hold $-\Delta$ of the underlying. The value of this position is $-\Delta X$, since $S = X$ is the barrier level. Suppose that the asset moves by a small fraction k before we can close out our asset position, or equivalently, that there is a transaction charge involved in closing.[2] We thus lose

$$-k\Delta X$$

on the trigger event.

Now refer to Figure 23.23 where we'll look at the specific example of a down-and-out option. Because we lose $-k\Delta X$ we should use the boundary condition

$$V(X, t) = -k\Delta X.$$

After a little bit of Taylor series, and since $\Delta = \partial V/\partial S$, we find that this is approximately the same as

$$V((1 + k)X, t) = 0.$$

In other words, we should apply the boundary condition at a slightly higher value of S and so slightly reduce the option's value.

[2] Much more about this in Chapter 48.

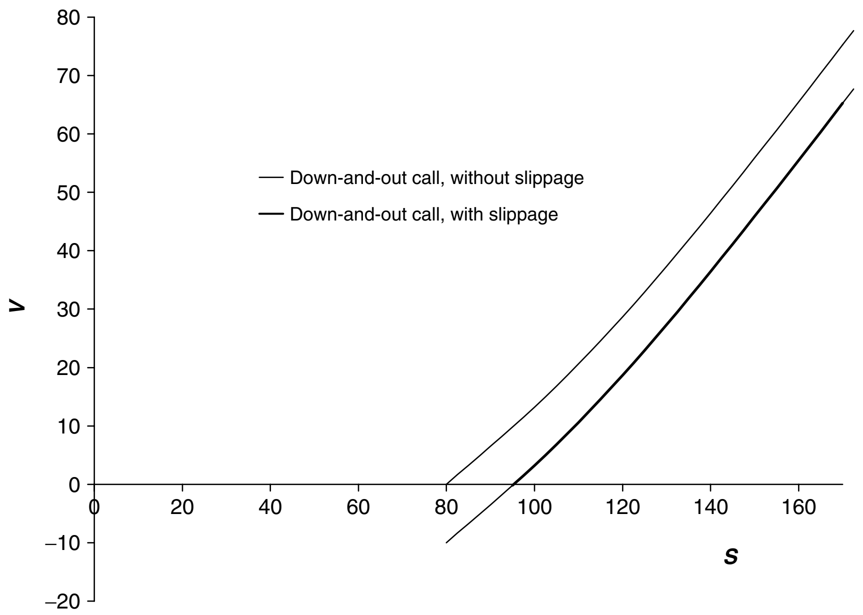


Figure 23.23 Incorporating slippage.

23.10 SUMMARY

In this chapter we have seen a description of many types of barrier option. We have seen how to put these contracts into the partial differential equation framework. Many of these contracts have simple pricing formulae. Unfortunately, the extreme nature of these contracts make them very difficult to hedge in practice and in particular, they can be very sensitive to the volatility of the underlying. Worse still, if the gamma of the contract changes sign we cannot play safe by adding a spread to the volatility. Practitioners seem to be most comfortable statically hedging as much of the barrier contract as possible using traded vanilla options and pricing the residual using a full implied volatility surface. The combination of these two principles is crucial. If one were to use a volatility surface without statically hedging then one could make matters worse; the volatility surface implied from vanillas may turn out to give the barrier option an inaccurate value. Less dangerous, but still not ideal, is the static hedging of the barrier option with vanillas and then using a single volatility to price the barrier. If both of these concepts are used together there is an element of consistency across the pricing.

FURTHER READING

- Many of the original barrier formulae are due to Reiner & Rubinstein (1991).
- The formulae above are explained in Taleb (1997) and Haug (1998). Taleb discusses barrier options in great detail, including the reality of hedging that I have only touched upon.
- The article by Carr (1995) contains an extensive literature review as well as a detailed discussion of protected barrier options and rainbow barrier options.

- See Derman, Ergener & Kani (1997) for a full description of the static replication of barrier options with vanilla options.
- See Carr (1994) for more details of put-call symmetry.
- See Haug (2002) for the pricing of barrier options that depend on two underlying assets.
- More closed-form solutions can be found in Banerjee (2003).

APPENDIX: MORE FORMULAE

In the following I use $N(\cdot)$ to denote the cumulative distribution function for a standardized Normal variable. The dividend yield on stocks or the foreign interest rate for FX are denoted by q. Also

$$a = \left(\frac{S_b}{S}\right)^{-1+(2(r-q)/\sigma^2)},$$

$$b = \left(\frac{S_b}{S}\right)^{1+(2(r-q)/\sigma^2)},$$

where S_b is the barrier position (whether S_u or S_d should be obvious from the example),

$$d_1 = \frac{\log(S/E) + (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = \frac{\log(S/E) + (r - q - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_3 = \frac{\log(S/S_b) + (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_4 = \frac{\log(S/S_b) + (r - q - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_5 = \frac{\log(S/S_b) - (r - q - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_6 = \frac{\log(S/S_b) - (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_7 = \frac{\log(SE/S_b^2) - (r - q - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_8 = \frac{\log(SE/S_b^2) - (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

Up-and-out call

$$Se^{-q(T-t)}\left(N(d_1) - N(d_3) - b(N(d_6) - N(d_8))\right)$$
$$- Ee^{-r(T-t)}\left(N(d_2) - N(d_4) - a(N(d_5) - N(d_7))\right).$$

Up-and-in call

$$Se^{-q(T-t)}\left(N(d_3)+b(N(d_6)-N(d_8))\right)-Ee^{-r(T-t)}\left(N(d_4)+a(N(d_5)-N(d_7))\right).$$

Down-and-out call

1. $E > S_b$:

$$Se^{-q(T-t)}\left(N(d_1)-b(1-N(d_8))\right)-Ee^{-r(T-t)}\left(N(d_2)-a(1-N(d_7))\right).$$

2. $E < S_b$:

$$Se^{-q(T-t)}\left(N(d_3)-b(1-N(d_6))\right)-Ee^{-r(T-t)}\left(N(d_4)-a(1-N(d_5))\right).$$

Down-and-in call

1. $E > S_b$:

$$Se^{-q(T-t)}b(1-N(d_8))-Ee^{-r(T-t)}a(1-N(d_7)).$$

2. $E < S_b$:

$$\begin{aligned}&Se^{-q(T-t)}\left(N(d_1)-N(d_3)+b(1-N(d_6))\right)\\&\quad-Ee^{-r(T-t)}\left(N(d_2)-N(d_4)+a(1-N(d_5))\right).\end{aligned}$$

Down-and-out put

$$\begin{aligned}&-Se^{-q(T-t)}\left(N(d_3)-N(d_1)-b(N(d_8)-N(d_6))\right)\\&\quad+Ee^{-r(T-t)}\left(N(d_4)-N(d_2)-a(N(d_7)-N(d_5))\right).\end{aligned}$$

Down-and-in put

$$\begin{aligned}&-Se^{-q(T-t)}\left(1-N(d_3)+b(N(d_8)-N(d_6))\right)\\&\quad+Ee^{-r(T-t)}\left(1-N(d_4)+a(N(d_7)-N(d_5))\right).\end{aligned}$$

Up-and-out put

1. $E > S_b$:

$$-Se^{-q(T-t)}\left(1-N(d_3)-bN(d_6)\right)+Ee^{-r(T-t)}\left(1-N(d_4)-aN(d_5)\right).$$

2. $E < S_b$:

$$-Se^{-q(T-t)}\left(1-N(d_1)-bN(d_8)\right)+Ee^{-r(T-t)}\left(1-N(d_2)-aN(d_7)\right).$$

Up-and-in put

1. $E > S_b$:

$$-Se^{-q(T-t)}\left(N(d_3)-N(d_1)+bN(d_6)\right)+Ee^{-r(T-t)}\left(N(d_4)-N(d_2)+aN(d_5)\right).$$

2. $E < S_b$:

$$-Se^{-q(T-t)}bN(d_8) + Ee^{-r(T-t)}aN(d_7).$$

The following charts (Figures 23.24–23.35) show each of the above types of barrier option, as well as the underlying vanilla option.

Note that with Out options the value of the barrier option ‘hugs’ the vanilla, except that it must be zero at the barrier. With In options, the barrier value hugs zero except that it becomes the vanilla value at the barrier.

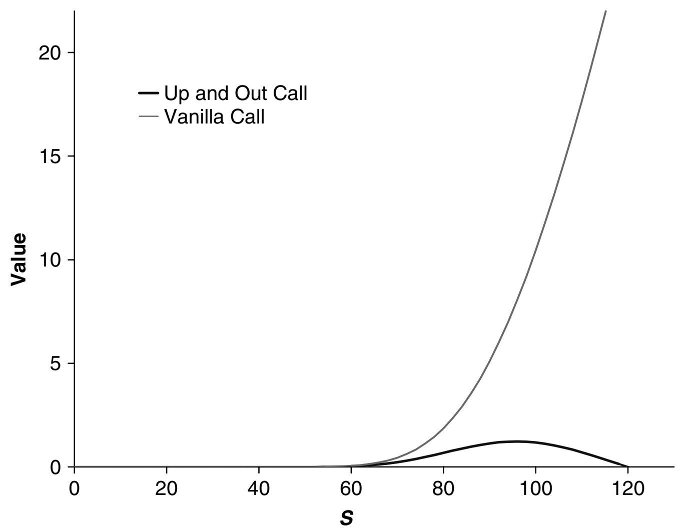


Figure 23.24 Up-and-out call. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 120$.

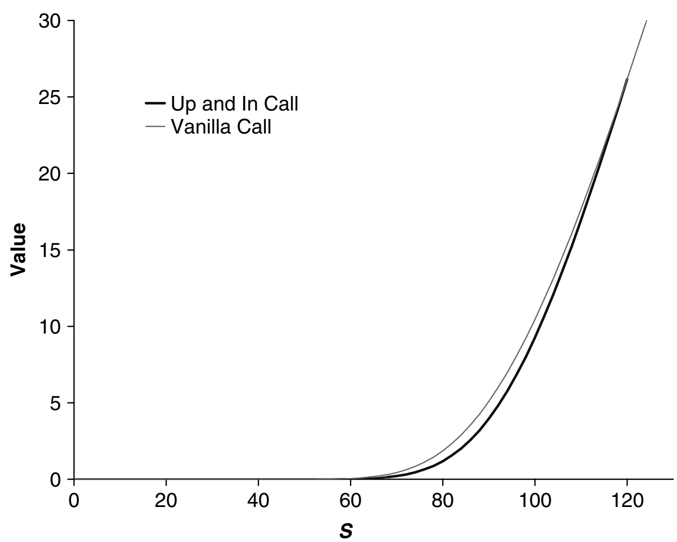


Figure 23.25 Up-and-in call. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 120$.

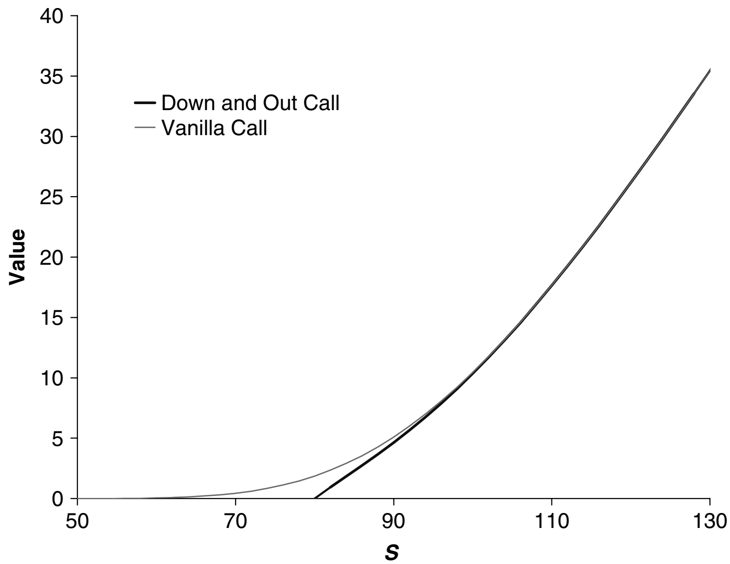


Figure 23.26 Down-and-out call. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 80$.

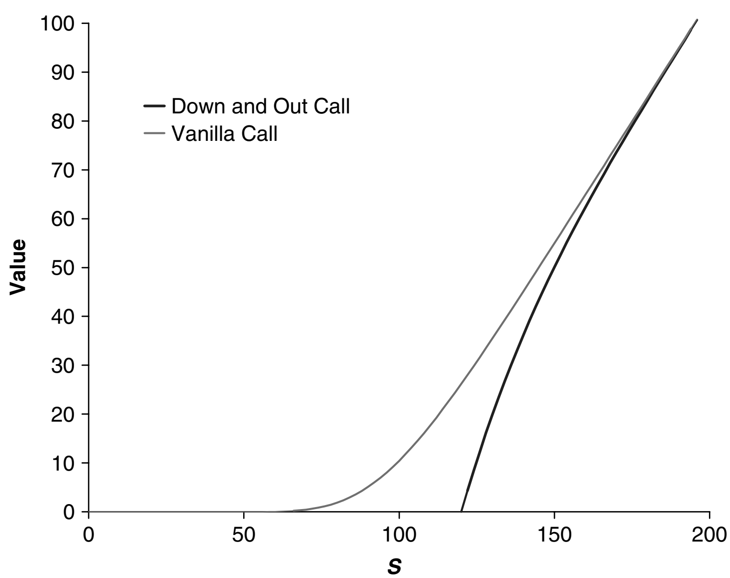


Figure 23.27 Down-and-out call. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 120$.

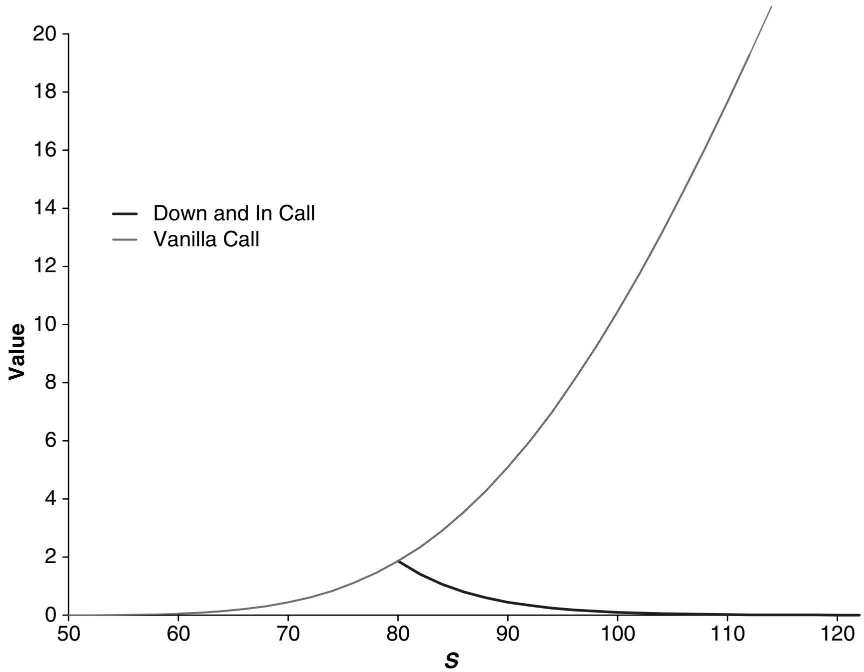


Figure 23.28 Down-and-in call. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 80$.

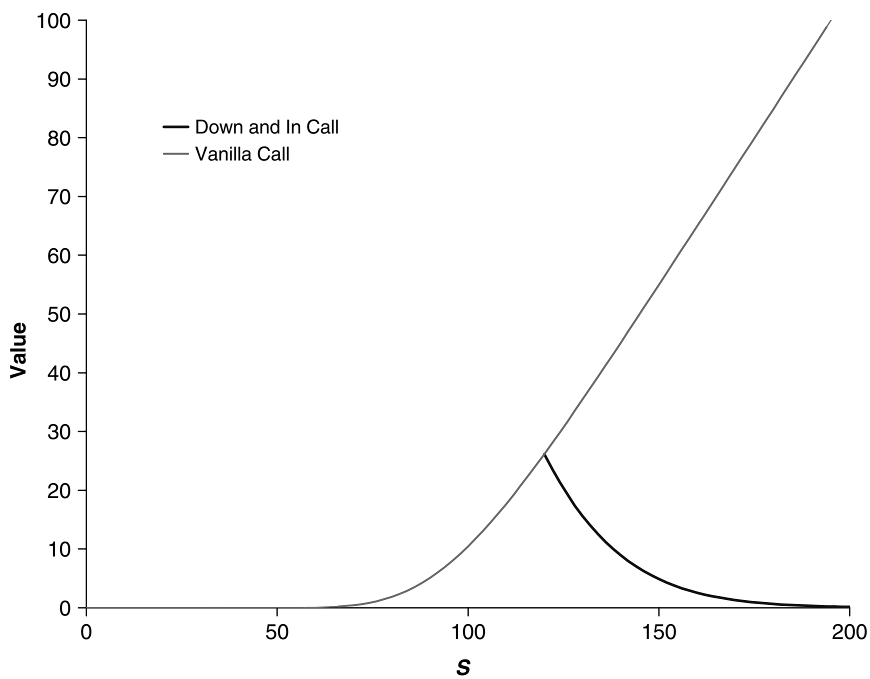


Figure 23.29 Down-and-in call. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 120$.

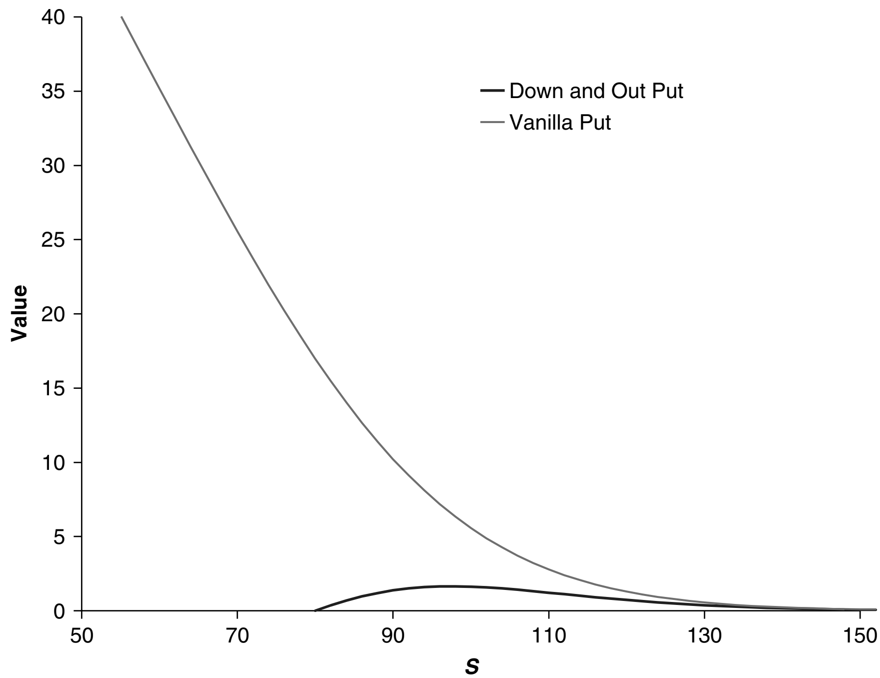


Figure 23.30 Down-and-out put. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 80$.

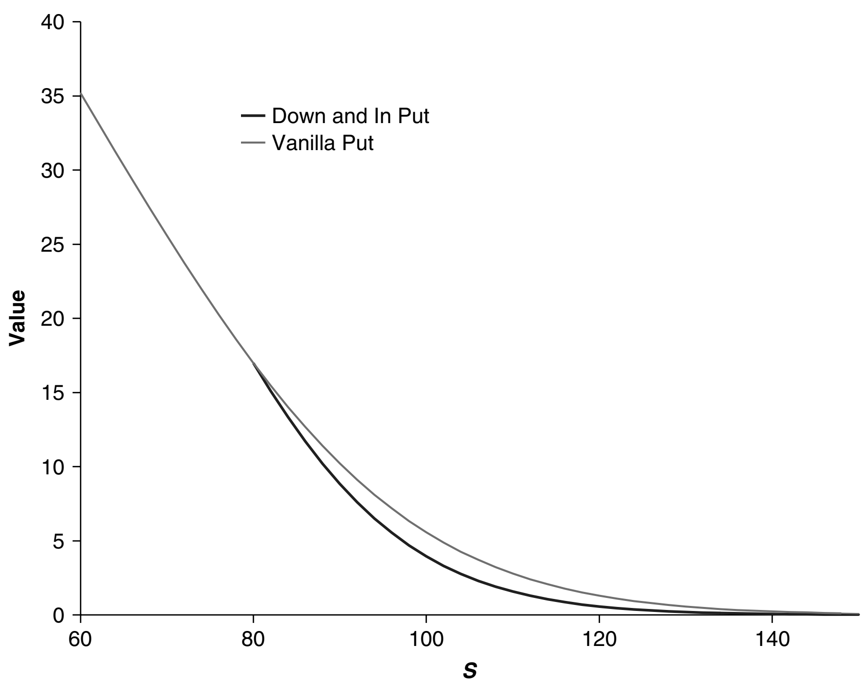


Figure 23.31 Down-and-in put. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 80$.

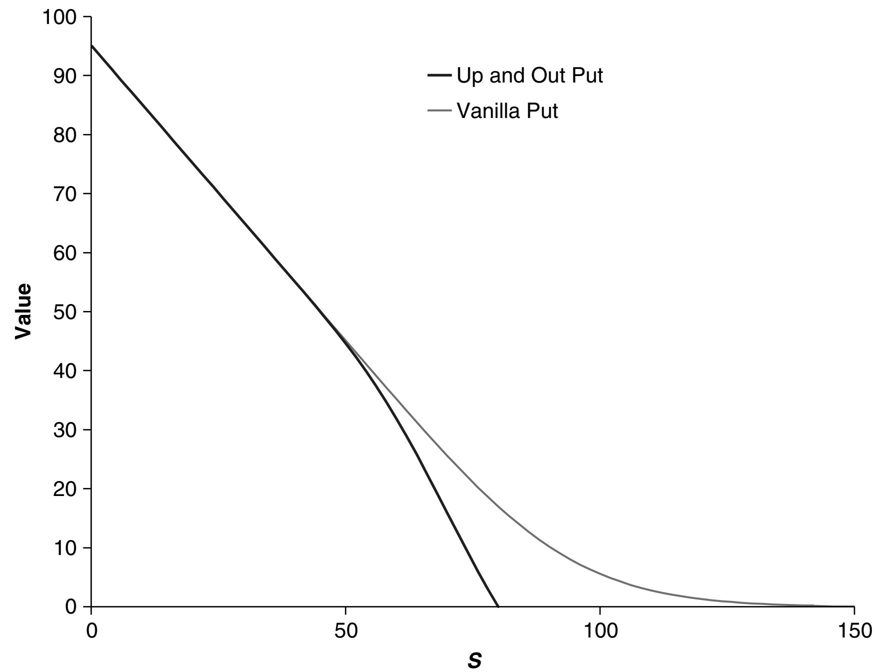


Figure 23.32 Up-and-out put. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 80$.

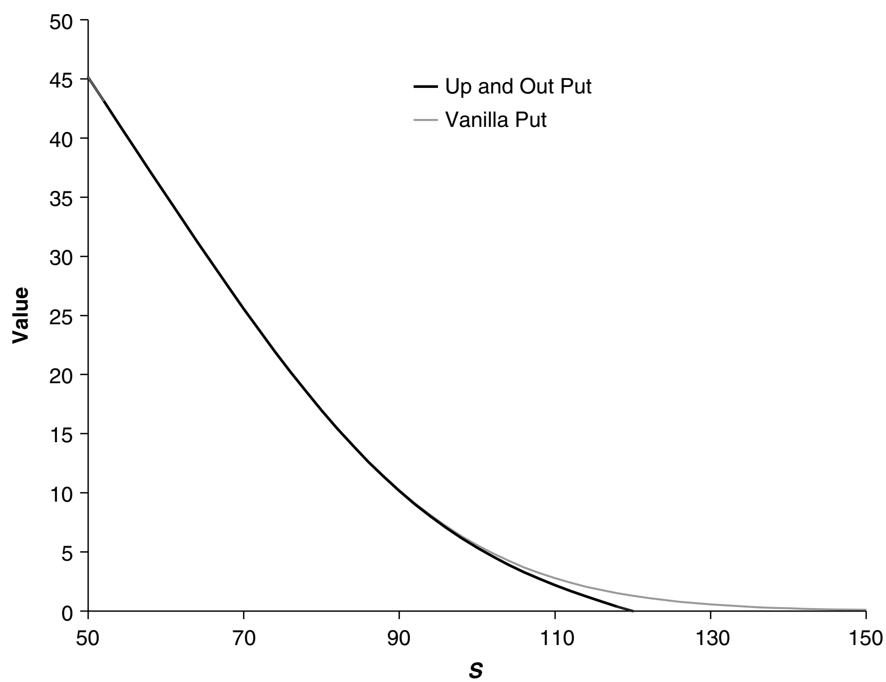


Figure 23.33 Up-and-out put. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 120$.

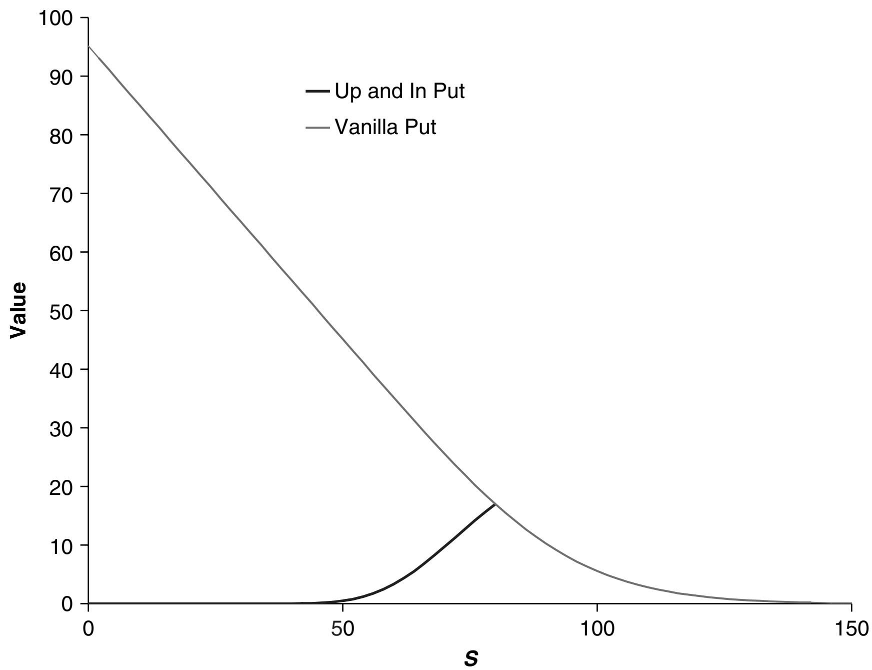


Figure 23.34 Up-and-in put. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 80$.

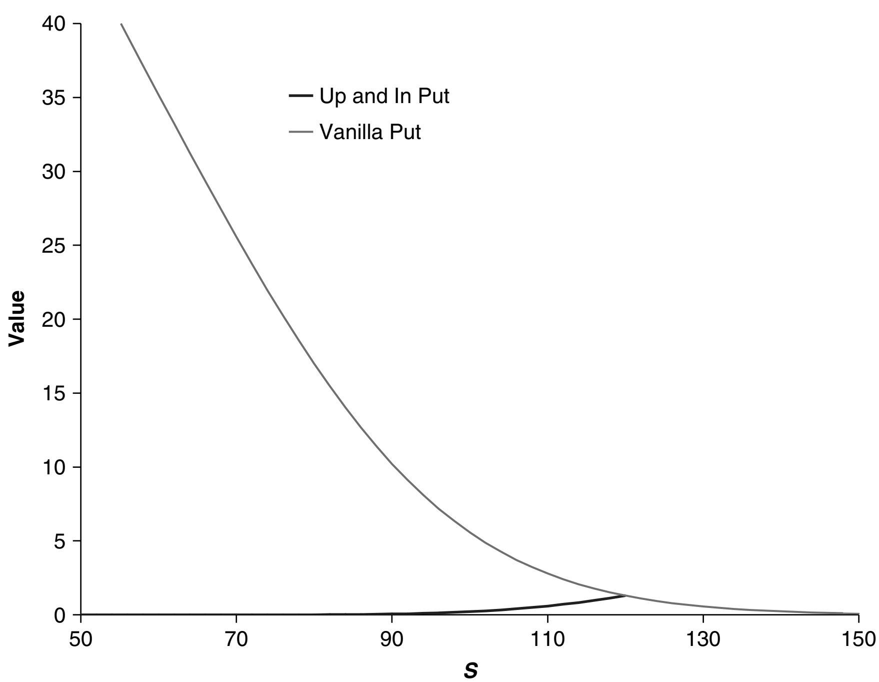
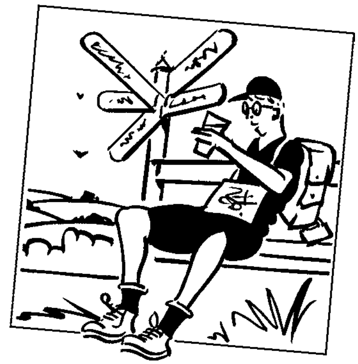


Figure 23.35 Up-and-in put. $\sigma = 0.2$, $r = 0.05$, $q = 0$, $E = 100$, $T = 1$ and $S_b = 120$.

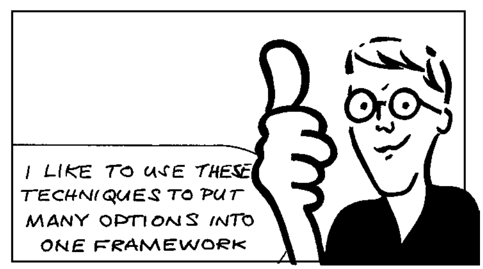
CHAPTER 24

strongly path-dependent derivatives



In this Chapter...

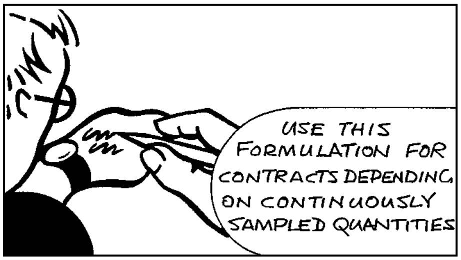
- strong path dependence
- pricing many strongly path-dependent contracts in the Black–Scholes partial differential equation framework
- how to handle both continuously sampled and discretely sampled paths
- jump conditions for differential equations



24.1 INTRODUCTION

To be able to turn the valuation of a derivative contract into the solution of a partial differential equation is a big step forward. The partial differential equation approach is one of the best ways to price a contract because of its flexibility and because of the large body of knowledge that has grown up around the fast and accurate numerical solution of these problems. This body of knowledge was, in the main, based around the solution of differential equations arising in physical applied mathematics but is starting to be used in the financial world.

In this chapter I show how to generalize the Black–Scholes analysis, delta hedging and no arbitrage, to the pricing of many more derivative contracts, specifically contracts that are strongly path-dependent. I will describe the theory in the abstract, giving brief examples occasionally, but saving the detailed application to specific contracts until later chapters.



24.2 PATH-DEPENDENT QUANTITIES REPRESENTED BY AN INTEGRAL

Imagine a contract that pays off at expiry, T, an amount that is a function of the path taken by the asset between time zero and expiry. Let us suppose that this path-dependent quantity can be represented by an integral of some function of the asset over the period zero to T:

$$I(T) = \int_0^T f(S, \tau)\, d\tau.$$

This is not such a strong assumption; in particular most of the path-dependent quantities in exotic derivative contracts, such as averages, can be written in this form with a suitable choice of $f(S, t)$.

We are thus assuming that the payoff is given by

$$P(S, I)$$

at time $t = T$.

Prior to expiry we have information about the possible final value of S (at time T) in the present value of S (at time t). For example, the higher S is today, the higher it will probably end up at expiry. Similarly, we have information about the possible final value of I in the value of the integral to date:

$$I(t) = \int_0^t f(S, \tau)\, d\tau. \tag{24.1}$$

As we get closer to expiry, so we become more confident about the final value of I.

One can imagine that the value of the option is therefore not only a function of S and t, but also a function of I; I will be our new independent variable, called a **state variable**. We see in the next section how this observation leads to a pricing equation. In anticipation of an argument that will use Itô's lemma, we need to know the stochastic differential equation satisfied by I. This could not be simpler. Incrementing t by dt in (24.1) we find that

$$dI = f(S, t)\, dt. \tag{24.2}$$

This obvious result is illustrated in Figure 24.1.

Observe that I is a smooth function (except at discontinuities of f) and from (24.2) we can see that its stochastic differential equation contains no stochastic terms.

24.2.1 Examples

An Asian option has a payoff that depends on the average of the asset price over some period. If that period is from time zero to expiry and the average is arithmetic then we write

$$I = \int_0^t S\, d\tau.$$

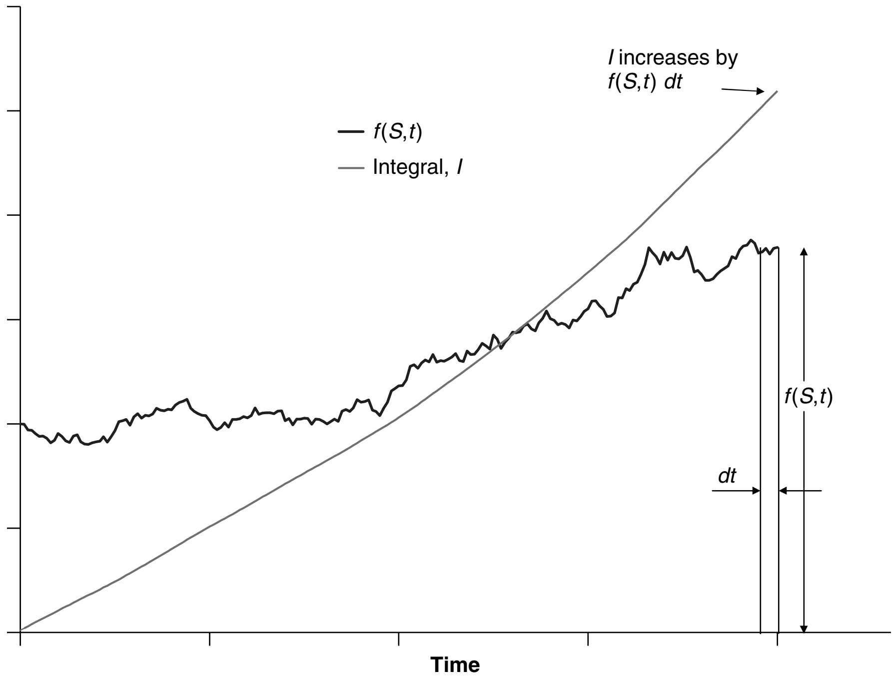


Figure 24.1 Incrementing I.

The payoff may then be, for example,

$$\max\left(\frac{I}{T}-S,0\right).$$

This would be an average strike put, of which more later.

If the average is geometric then we write

$$I=\int_0^t \log(S)\,d\tau.$$

As another example, imagine a contract that pays off according to the time that the asset has spent above a certain level. We would then introduce

$$I=\int_0^t \mathcal{H}(S-S_u)\,d\tau,$$

where $\mathcal{H}(\cdot)$ is the Heaviside function.

Now add a little twist to this: suppose that the contract pays off according to both the time that the asset has spent above a certain level and the square of its value. We would have

$$I=\int_0^t S^2\mathcal{H}(S_u-S)\,d\tau,$$

and then the option value would be a function of S, I and t.

See how many fancy options can be reinterpreted via such integrals. We are now ready to price some options.

24.3 CONTINUOUS SAMPLING: THE PRICING EQUATION

I will derive the pricing partial differential equation for a contract that pays some function of our new variable I. The value of the contract is now a function of the three variables, $V(S, I, t)$. Set up a portfolio containing one of the path-dependent options and short a number Δ of the underlying asset:

$$\Pi = V(S, I, t) - \Delta S.$$

The change in the value of this portfolio is given by

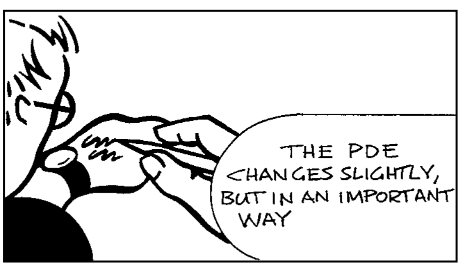
$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \frac{\partial V}{\partial I} dI + \left(\frac{\partial V}{\partial S} - \Delta \right) dS.$$

Choosing

$$\Delta = \frac{\partial V}{\partial S}$$

to hedge the risk, and using (24.2), we find that

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + f(S, t) \frac{\partial V}{\partial I} \right) dt.$$



This change is risk free, and thus earns the risk-free rate of interest r, leading to the pricing equation

$$\boxed{\begin{array}{l} \dfrac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 V}{\partial S^2} + f(S, t) \dfrac{\partial V}{\partial I} \\ \quad + rS \dfrac{\partial V}{\partial S} - rV = 0 \end{array}} \qquad (24.3)$$

This is to be solved subject to

$$V(S, I, T) = P(S, I).$$

This completes the formulation of the valuation problem. The obvious changes can be made to accommodate dividends on the underlying.

24.3.1 Example

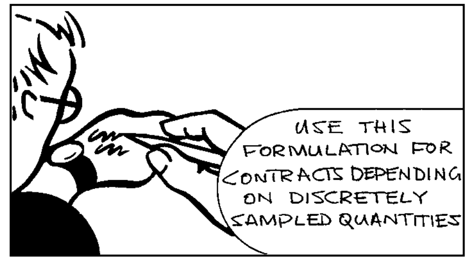
Continuing with the arithmetic Asian example, we have

$$I = \int_0^t S \, d\tau,$$

so that the equation to be solved is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + S \frac{\partial V}{\partial I} + rS \frac{\partial V}{\partial S} - rV = 0.$$

24.4 PATH-DEPENDENT QUANTITIES REPRESENTED BY AN UPDATING RULE



For practical and legal reasons path-dependent quantities are never measured continuously. There is a minimum time step between sampling of the path-dependent quantity. This time step may be small, one day, say, or much longer. From a practical viewpoint it is difficult to incorporate every single traded price into an average, for example. Data can be unreliable and the exact time of a trade may not be known accurately. From a legal point of view, to avoid disagreements over the value of the path-dependent quantity, it is usual to use only key prices, such as closing prices, that are, in a sense, guaranteed to be a genuine traded price. If the time between samples is small we can confidently use a continuous-sampling model; the error will be small. If the time between samples is long, or the time to expiry itself is short, we must build this into our model. This is the goal of this section.

I introduce the idea of an **updating rule**, an algorithm for defining the path-dependent quantity in terms of the current 'state of the world.' The path-dependent quantity is measured on the **sampling dates** t_i, and takes the value I_i for $t_i \leq t < t_{i+1}$. At the sampling date t_i the quantity I_{i-1} is updated according to a rule such as

$$I_i = F(S(t_i), I_{i-1}, i).$$

Note how, in this simplest example (which can be generalized), the new value of I is determined by only the old value of I and the value of the underlying on the sampling date, and the sampling date. This updating rule takes the part of the integral expression we had before when we were considering continuously sampled quantities. The updating rule (or integral) together with the payoff function effectively define the exotic option with which we are dealing.

When we come to concrete examples, as next, I shall change notation and use names for variables that are meaningful, I hope this does not cause any confusion.

24.4.1 Examples

We saw how to use the continuous running integral in the valuation of Asian options. But what if that integral is replaced by a discrete sum? In practice, the payoff for an Asian option depends on the quantity

$$I_M = \sum_{k=1}^{M} S(t_k),$$

where M is the total number of sampling dates. This is the discretely sampled sum. A more natural quantity to consider is

$$A_M = \frac{I_M}{M} = \frac{1}{M}\sum_{k=1}^{M} S(t_k), \tag{24.4}$$

because then the payoff for the discretely-sampled arithmetic average strike put is

$$\max\left(A_M - S, 0\right).$$

Can we write (24.4) in terms of an updating rule? Yes, easily: if we write

$$A_i = \frac{1}{i}\sum_{k=1}^{i} S(t_k)$$

then we have

$$A_1 = S(t_1), \quad A_2 = \frac{S(t_1) + S(t_2)}{2} = \tfrac{1}{2}A_1 + \tfrac{1}{2}S(t_2),$$
$$A_3 = \frac{S(t_1) + S(t_2) + S(t_3)}{3} = \tfrac{2}{3}A_2 + \tfrac{1}{3}S(t_3), \cdots$$

or generally

$$A_i = \frac{1}{i}S(t_i) + \frac{i-1}{i}A_{i-1}.$$

We will see how to use this for pricing in the next section. But first, another example.

The lookback option has a payoff that depends on the maximum or minimum of the realized asset price. If the payoff depends on the maximum sampled at times t_i then we have

$$I_1 = S(t_1), \quad I_2 = \max(S(t_2), I_1), \quad I_3 = \max(S(t_3), I_2) \cdots .$$

The updating rule is therefore simply

$$I_i = \max(S(t_i), I_{i-1}).$$

In Chapter 26, where we consider lookbacks in detail, we use the notation M_i for minimum or maximum.

How do we use these updating rules in the pricing of derivatives?

24.5 DISCRETE SAMPLING: THE PRICING EQUATION

Following the continuous-sampling case we can anticipate that the option value will be a function of three variables, $V(S, I, t)$. We derive the pricing equation in a heuristic fashion. The derivation can be made more rigorous, but there is little point since the conclusion is correct and so obvious.

The first step in the derivation is the observation that the stochastic differential equation for I is degenerate:

$$dI = 0.$$

This is because the variable I can only change at the discrete set of dates t_i. This is true if $t \neq t_i$ for any i. So provided we are not *on* a sampling date the quantity I is constant, the stochastic differential equation for I reflects this, and the pricing equation is simply the basic Black–Scholes equation:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

Remember, though, that V is still a function of *three* variables; I is effectively treated as a parameter.

How does the equation know about the path dependency? What happens *at* a sampling date?

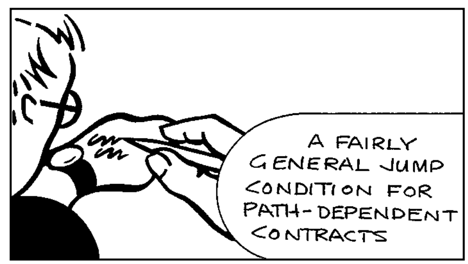
The answer to the latter question gives us the answer to the former. Across a sampling date nothing much happens. Across a sampling date the option value stays the same. As we get closer and closer to the sampling date we become more and more sure about the value that I will take according to the updating rule. Since the outcome on the sampling date is known and since *no money changes hands* there cannot be any jump in the value of the option. This is a simple application of the no arbitrage principle.

Across a sampling date the option value is continuous. If we introduce the notation t_i^- to mean the time infinitesimally before the sampling date t_i and t_i^+ to mean infinitesimally just after the sampling date, then continuity of the option value is represented mathematically by

$$V(S, I_{i-1}, t_i^-) = V(S, I_i, t_i^+).$$

In terms of the updating rule, we have

$$\boxed{V(S, I, t_i^-) = V(S, F(S, I, i), t_i^+)}$$



This is called a **jump condition**.

We call this a jump condition even though there is no jump in this case. (When money does change hands on a special date there will be a sudden change in the value of the option at that time, as discussed in Chapter 22.) If we follow the path of S in time we see that it is continuous. However, the path for I is discontinuous. There is a deterministic jump in I across the sampling date. If we were to plot V as a function of S and I just before and just after the sampling date we would see that *for fixed S and I* the option price would be discontinuous. But this plot would have to be interpreted correctly; $V(S, I, t)$ may be discontinuous as a function of S and I but V is continuous along each *realized* path of S and I.

In Figure 24.2 I show the relationship between before and after option values.

24.5.1 Examples

To price an arithmetic Asian option with the average sampled at times t_i solve the Black–Scholes equation for $V(S, A, t)$ with

$$V(S, A, t_i^-) = V\left(S, \frac{i-1}{i}A + \frac{1}{i}S, t_i^+\right),$$

and a suitable final condition representing the payoff.

To price a lookback depending on the maximum sampled at times t_i solve the Black–Scholes equation for $V(S, M, t)$ with

$$V(S, M, t_i^-) = V\left(S, \max(S, M), t_i^+\right). \qquad (24.5)$$

How this particular jump condition works is shown in Figure 24.3. The top right-hand plot is the S, M plane just after the sample of the maximum has been taken. Since the sample has

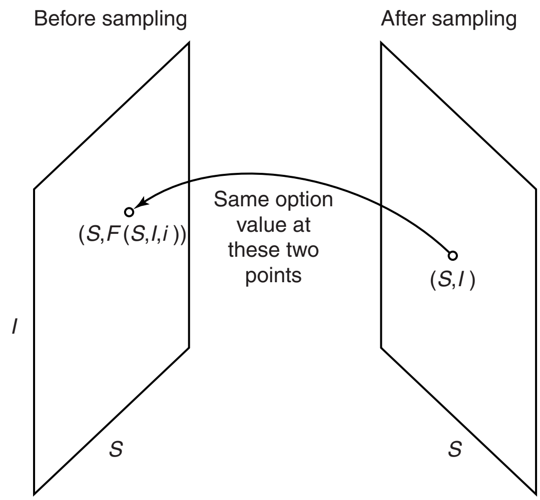


Figure 24.2 Representation of the jump condition.

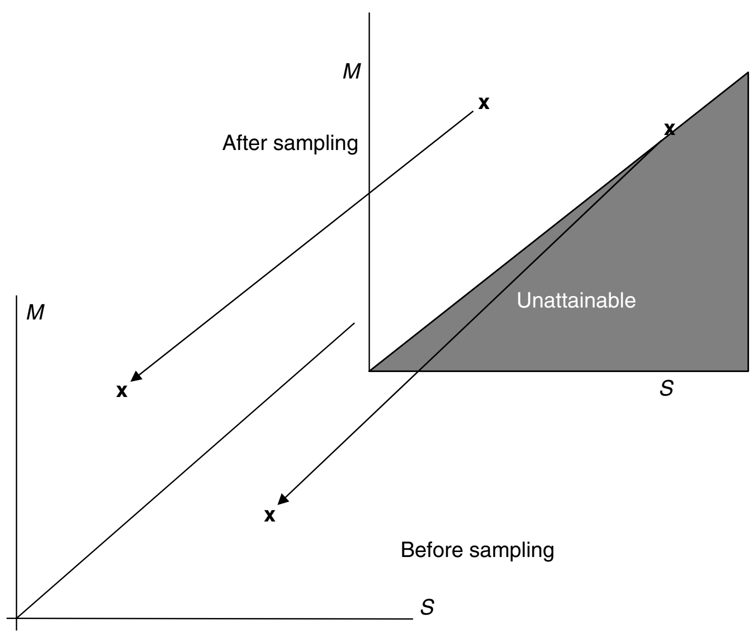


Figure 24.3 The jump condition for a lookback option.

just been taken the region $S > M$ cannot be reached; it is the region labeled 'Unattainable.' When we come to solve the Black–Scholes equation numerically in Part Six we will see how we work backwards in time, so that we will find the option value for time t_i^+ *before* the value for time t_i^-. So we will have found the option value $V(S, M, t_i^+)$ for all $S < M$. To find the option value just before the sampling we must apply the jump condition (24.5). Pictorially, this means that the option value at time t_i^- for $S < M$ is the same as the t_i^+ value, just follow the left-hand arrow in the figure. *However*, for $S > M$ (which is attainable before the sample is taken) the option value comes from the $S = M$ line at time t_i^+ for the same S value; after all, S is continuous. Now, just follow the right-hand arrow.

24.5.2 The Algorithm for Discrete Sampling

Since the path-dependent quantity, I, is updated discretely and is therefore constant between sampling dates, the partial differential equation for the option value between sampling dates is the Black–Scholes equation with I treated as a parameter. The algorithm for valuing an option on a discretely-sampled quantity is as follows:

- Working backwards from expiry, solve

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

  between sampling dates. (How to do this is shown in Part Six.) Stop when you get to the time step on which the sampling takes place.
- Then apply the appropriate jump condition across the current sampling date to deduce the option value immediately before the present sampling date using the calculated value of the option just after. Use this as your final condition for further time stepping of the Black–Scholes equation.
- Repeat this process as necessary to arrive at the current value of the option.

24.6 HIGHER DIMENSIONS

The methods outlined above are not restricted to a single path-dependent quantity. Any finite number of path-dependent variables can be accommodated, theoretically. Imagine a contract that pays off the difference between a continuous geometric and a continuous arithmetic average. To price this one would need to introduce I_g and I_a, defined by

$$I_g = \int_0^t \log(S)\,d\tau \quad \text{and} \quad I_a = \int_0^t S\,d\tau.$$

The solution would then be a function of four variables, $V(S, I_g, I_a, t)$. However, this is at the limit of practicality for a numerical solution of a partial differential equation. Unless there is a similarity solution, reducing the dimensionality of the problem, it may be better to consider Monte Carlo simulation.

The same thoughts apply to discrete sampling or a combination of discrete and continuous.

24.7 PRICING VIA EXPECTATIONS

In Chapter 10 I showed how we can value options in the Black–Scholes world by taking the present value of the expected payoff under a risk-neutral random walk. This approach applies perfectly well to all of the path-dependent options we have described or are going to describe. Simply simulate the random walk

$$dS = rS\,dt + \sigma S\,dX,$$

as will be discussed in Chapter 80, for many paths, calculate the payoff for each path—and this means calculating the value of the path-dependent quantity which is usually very simple to do—take the average payoff over all the paths and then take the present value of that average. That is the option fair value. Note that there is no μ in this, it is the risk-neutral random walk that must be simulated.

This is a very general and powerful technique, useful for path-dependent contracts for which either a partial differential equation approach is impossible or too high-dimensional. The only disadvantage is that it is harder to value American options in this framework.

24.8 EARLY EXERCISE

If you have found a partial differential equation formulation of the option problem then it is simple to incorporate the early exercise feature of American and Bermudan options. Simply apply the constraint

$$V(S, I, t) \geq P(S, I),$$

together with continuity of the delta of the option, where $P(S, I)$ is the payoff function (and it can also be time-dependent). This condition is to be applied at any time that early exercise is allowed. If you have found a partial differential equation formulation of the problem and it is in sufficiently low dimension then incorporating early exercise in the numerical scheme is a matter of adding a couple of lines of code, see Chapter 78.

24.9 SUMMARY

The basic theory has been explained above for the pricing of many path-dependent contracts in the partial differential equation framework. We have examined both continuously sampled and discretely sampled path-dependent quantities. In the next few chapters we discuss these matters in more detail, first with Asian and lookback options and then for a wider range of exotic contracts. The practical implementation of these models is described in Chapter 78.

FURTHER READING

- See Bergman (1985) for the early work on a unified partial differential equation framework for path-dependent contracts.
- The excellent book by Ingersoll (1987) also discusses the partial differential equation approach to pricing exotics.
- The general framework for exotics is described in a *Risk* magazine article by Dewynne & Wilmott (1994a) and also in Dewynne & Wilmott (1996).

CHAPTER 25
asian options

In this Chapter...

- many types of Asian option with a payoff that depends on an average
- the different types of averaging of asset prices that are used in defining the payoff
- how to price these contracts in the partial differential equation framework

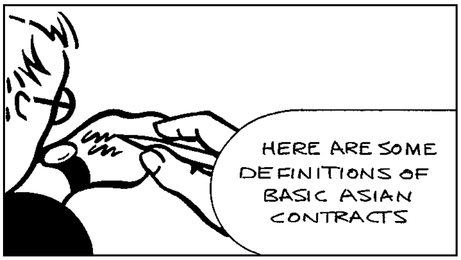
25.1 INTRODUCTION

Asian options give the holder a payoff that depends on the average price of the underlying over some prescribed period. This averaging of the underlying can significantly reduce the price of an Asian option compared with a similar vanilla contract. Anything that reduces the up-front premium in an option contract tends to make them more popular.

- The volatility of the average of a stock price over time is less than the volatility of the stock. This often results in the price of Asian options being less than equivalent vanilla options.
- The average is less exposed to sudden crashes or rallies in a stock price. All is not lost if you own a call and the stock plummets just before expiration.
- The average over time is harder to manipulate than a single stock price. Asians are of obvious use with easily manipulated, thinly traded underlyings.
- The average rate options can be used to lock in the price of a commodity or an exchange rate for those who have a continual and fairly predictable exposure to one of these over extended periods.
- End users of commodities or energy are often exposed to average prices over time, so Asian options are of obvious appeal. (Asian options are often considered to be the vanillas of the crude oil market.)

Asian options might be bought by someone with a stream of cashflows in a foreign currency, due to sales abroad for example, and who wants to hedge against fluctuations in the exchange rate. The Asian tail that we see later is designed to reduce exposure to sudden movements in the underlying just before expiry; some pension schemes have such a feature. Asian options were first successfully priced in 1987 by David Spaughton and Mark Standish of Bankers Trust. They were in Tokyo when they developed the first commercially used pricing formula for options linked to the average price of crude oil. Apart from that, the adjective 'Asian' has no significance.

In this chapter we find differential equations for the value of a wide variety of such Asian options. There are many ways to define an 'average' of the price, and we begin with a discussion of the obvious possibilities. We will see how to write the price as the solution of a partial differential equation in *three* variables: The underlying asset, time, and a new state variable representing the evolution of the average.



25.2 PAYOFF TYPES

Assuming for the moment that we have defined our average A, what sort of payoffs are common? As well as calls, puts etc. there is also the classification of **strike** and **rate**. These classifications work as follows. Take the payoff for a vanilla option, a vanilla call, say,

$$\max(S - E, 0).$$

Replace the strike price E with an average and you have an **average strike call**. This has payoff

$$\max(S - A, 0).$$

An **average strike put** thus has payoff

$$\max(A - S, 0).$$

Now take the vanilla payoff and instead replace the asset with its average; what you get is a rate option. For example, an **average rate call** has payoff

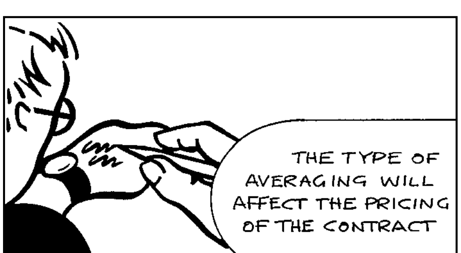
$$\max(A - E, 0)$$

and an **average rate put** has payoff

$$\max(E - A, 0).$$

The average rate options can be used to lock in the price of a commodity or an exchange rate for those who have a continual and fairly predictable exposure to one of these over extended periods.

The difference between calls and puts is simple from a pricing point of view, but the strike/rate distinction can make a big difference. Strike options are easier to value numerically.



25.3 TYPES OF AVERAGING

The precise definition of the average used in an Asian contract depends on two elements: How the data points are combined to form an average and which data points are used. The former means whether we have an arithmetic or geometric average or something more complicated. The latter means how many data points do we use in the average, all quoted prices, or just a subset, and over what time period.

25.3.1 Arithmetic or Geometric

The two simplest and obvious types of average are the **arithmetic average** and the **geometric average**. The arithmetic average of the price is the sum of all the constituent prices, equally weighted, divided by the total number of prices used. The geometric average is the *exponential* of the sum of all the *logarithms* of the constituent prices, equally weighted, divided by the total number of prices used. Another popular choice is the exponentially weighted average, meaning instead of having an equal weighting to each price in the average, the recent prices are weighted more than past prices in an exponentially decreasing fashion.

25.3.2 Discrete or Continuous

How much data do we use in the calculation of the average? Do we take every traded price or just a subset? If we take closely spaced prices over a finite time then the sums that we calculate in the average become integrals of the asset (or some function of it) over the averaging period. This would give us a **continuously sampled average**. More commonly, we only take data points that are reliable, using closing prices, a smaller set of data. This is called **discrete sampling**. This issue of continuous or discrete sampling was discussed in the previous chapter.

25.4 SOLUTION METHODS

As is often the case we usually want to choose between Monte Carlo and partial differential equation methods for pricing in practice. There are pros and cons to both of these.

25.4.1 Monte Carlo Simulation

In a nutshell,

- the value of an option is the present value of the expected payoff under a risk-neutral random walk.

The pricing algorithm:

1. Simulate the risk-neutral random walk starting at today's value of the asset over the required time horizon. This gives one realization of the underlying price path.
2. For this realization calculate the option payoff.
3. Perform many more such realizations over the time horizon.
4. Calculate the average payoff over all realizations.
5. Take the present value of this average; this is the option value.

Advantages of Monte Carlo pricing

- It is easy to code.
- It is hard to make mistakes in the coding.

Disadvantages of Monte Carlo pricing

- More work is needed to get the greeks.
- It can be slow since tens of thousands of simulations are needed to get an accurate answer.

Over the next few sections we will be looking at the partial differential equation approach. But first, its advantages.

Advantages of the partial differential equation approach

- It is faster than Monte Carlo.
- It is flexible, allowing for more advanced volatility models, for example.
- It copes very well with early exercise.

Disadvantage of the partial differential equation approach

- It is harder than Monte Carlo to code.

25.5 EXTENDING THE BLACK–SCHOLES EQUATION

25.5.1 Continuously sampled Averages

Figure 25.1 shows a realization of the random walk followed by an asset, in this case YPF, an Argentinian oil company, together with a continuously sampled running arithmetic average.

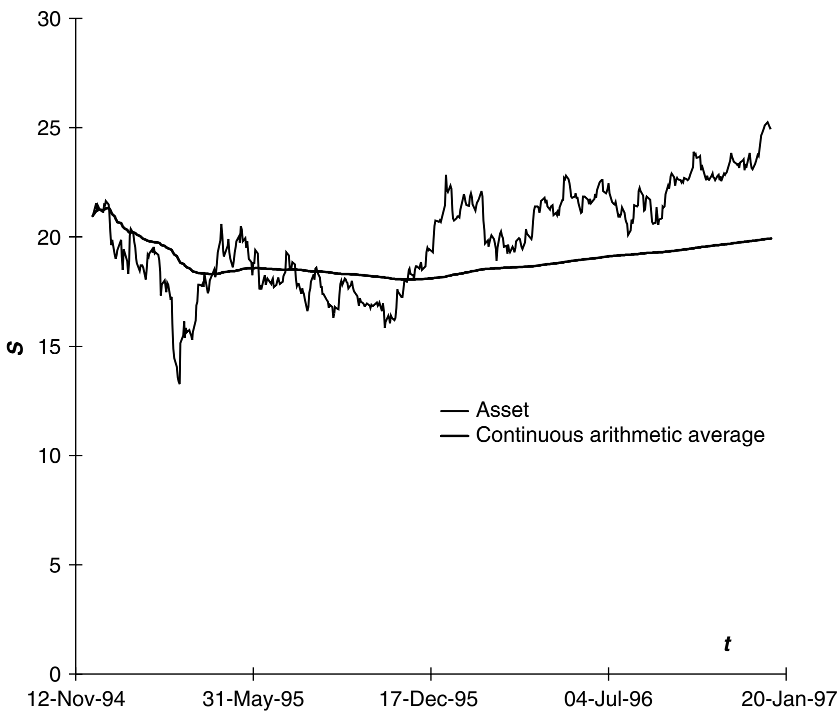


Figure 25.1 An asset price random walk and its continuously measured arithmetic running average.

This average is defined as

$$\frac{1}{t}\int_0^t S(\tau)\,d\tau.$$

If we introduce the new state variable

$$I=\int_0^t S(\tau)\,d\tau$$

then, following the analysis of Chapter 24, the partial differential equation for the value of an option contingent on this average is

$$\frac{\partial V}{\partial t}+S\frac{\partial V}{\partial I}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV=0.$$

Figure 25.2 shows a realization of an asset price random walk with the continuously sampled geometric running average.

The continuously sampled geometric average is defined to be

$$\exp\left(\frac{1}{t}\int_0^t \log S(\tau)\,d\tau\right).$$

To value an option contingent on this average we define

$$I=\int_0^t \log S(\tau)\,d\tau$$

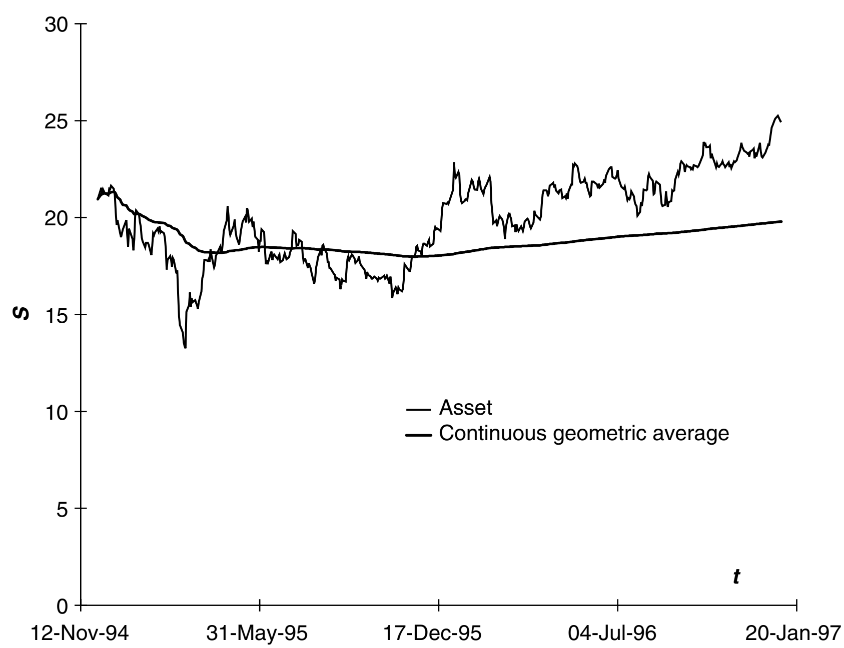


Figure 25.2 An asset price random walk and its continuous geometric running average.

and, following again the analysis of Chapter 24, the partial differential equation for the value of the option is

$$\frac{\partial V}{\partial t} + \log S \frac{\partial V}{\partial I} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

25.5.2 Discretely sampled Averages

Discretely sampled averages, whether arithmetic or geometric, fit easily into the framework established in Chapter 24. In Figures 25.3 and 25.4 are shown examples of a realized asset price and discretely sampled arithmetic and geometric averages respectively.

Above, we modeled the continuously sampled average as an integral. By a discretely sampled average we mean the sum, rather than the integral, of a finite number of values of the asset during the life of the option. Such a definition of average is easily included within the framework of our model.

If the sampling dates are t_i, $i = 1, \dots$ then the discretely sampled arithmetic averages are defined by

$$A_i = \frac{1}{i}\sum_{k=1}^{i} S(t_k).$$

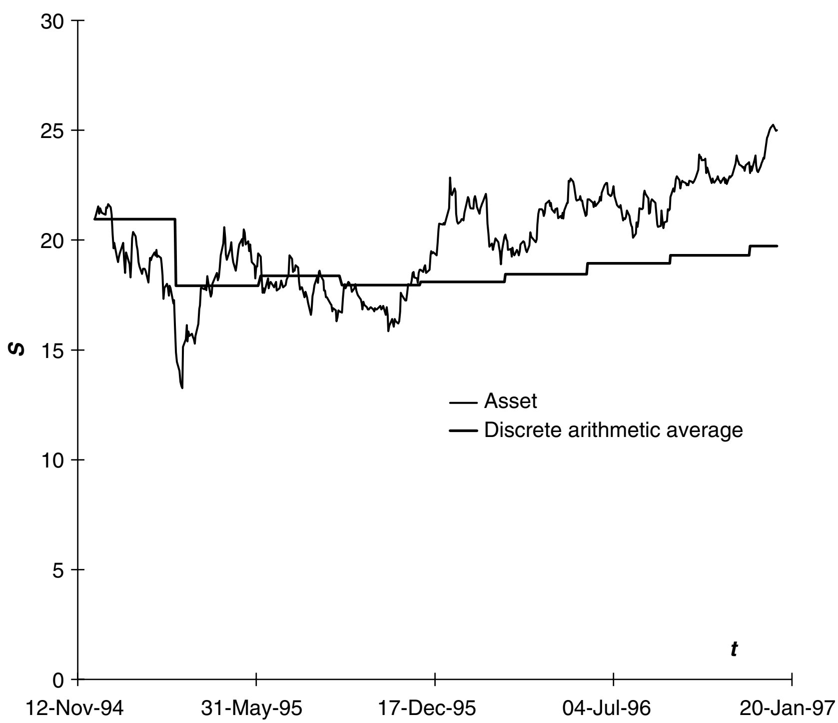


Figure 25.3 An asset price random walk and its discretely sampled arithmetic running average.

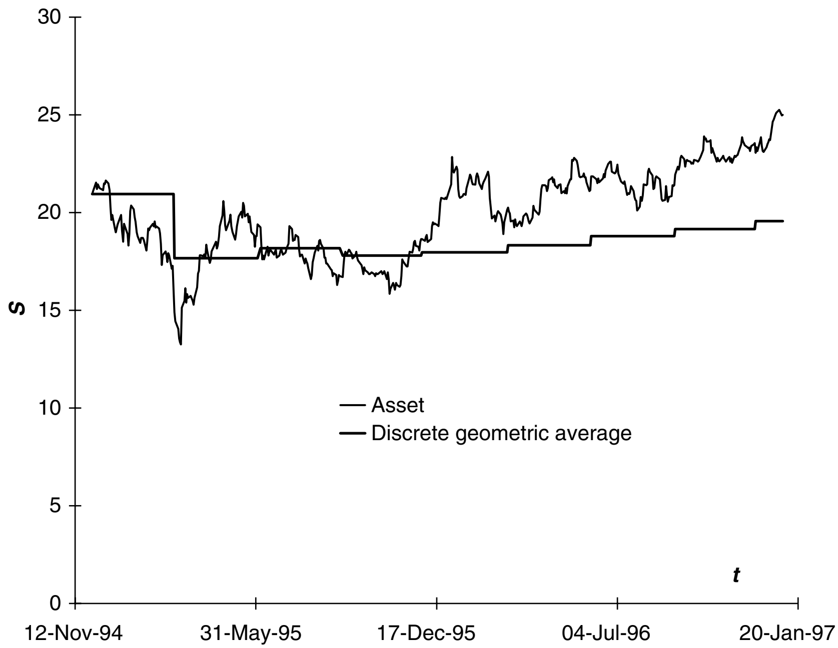


Figure 25.4 An asset price random walk and its discretely sampled geometric running average.

In particular

$$A_1 = S(t_1), \quad A_2 = \frac{S(t_1) + S(t_2)}{2} = \tfrac{1}{2}A_1 + \tfrac{1}{2}S(t_2),$$
$$A_3 = \frac{S(t_1) + S(t_2) + S(t_3)}{3} = \tfrac{2}{3}A_2 + \tfrac{1}{3}S(t_3), \cdots.$$

It is easy to see that these are equivalent to

$$A_i = \frac{i-1}{i}A_{i-1} + \frac{1}{i}S(t_i).$$

At the top of Figure 25.5 is shown a realized asset price path. Below that is the discretely sampled average. This is necessarily piecewise constant. At the bottom of the figure is the value of some option (it doesn't matter which). The option value must be continuous to eliminate arbitrage opportunities.

Using the results of Chapter 24, the jump condition for an Asian option with discrete arithmetic averaging is then simply

$$V(S, A, t_i^-) = V\left(S, \frac{i-1}{i}A + \frac{1}{i}S, t_i^+\right).$$

This is a result of the continuity of the option price across a sampling date i.e. no arbitrage.

Similarly the discretely sampled geometric average has the jump condition

$$V(S, A, t_i^-) = V\left(S, \exp\left(\frac{i-1}{i}\log(A) + \frac{1}{i}\log(S)\right), t_i^+\right)$$

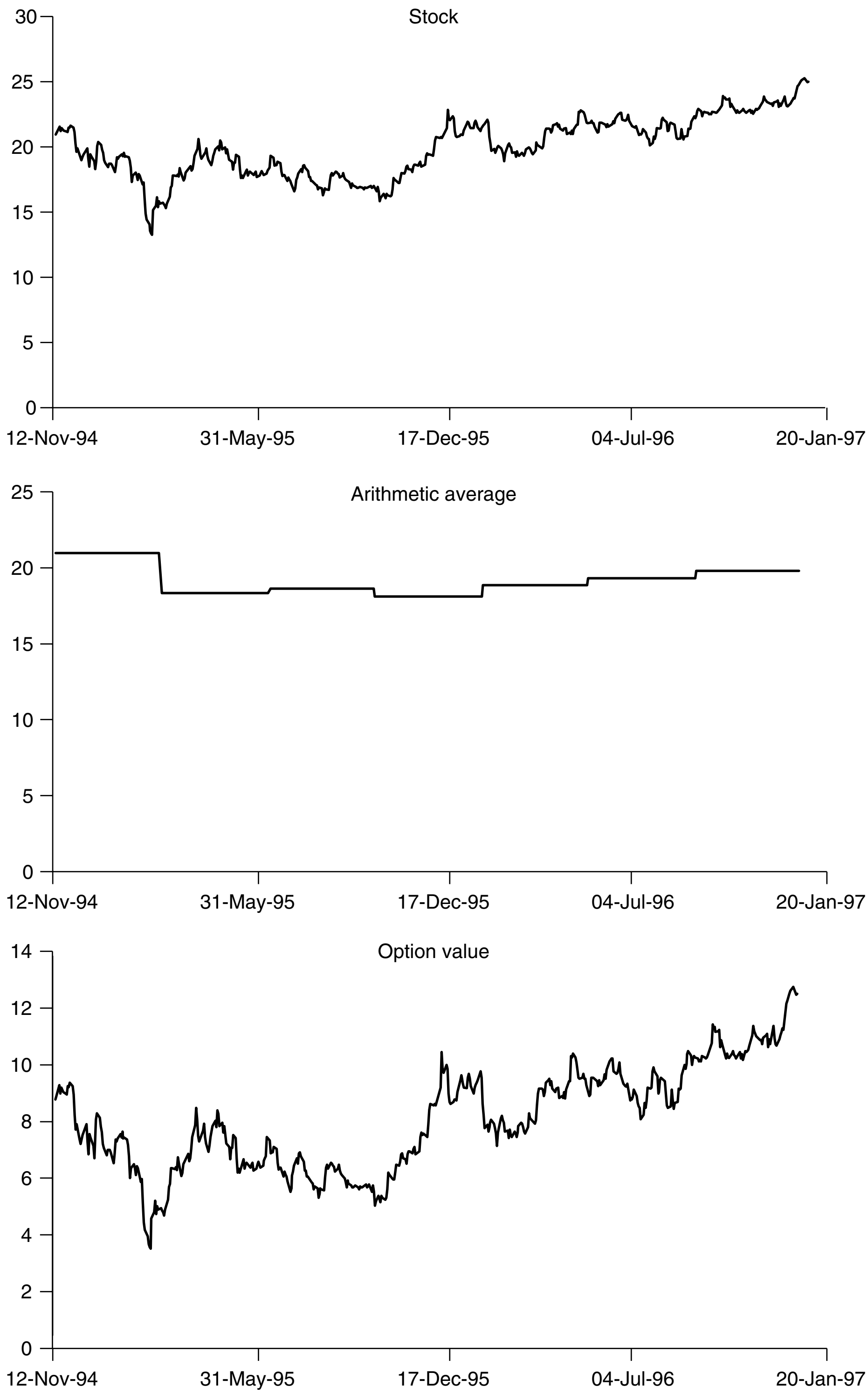


Figure 25.5 Top: An asset price random walk. Middle: Its discretely sampled arithmetic running average. Bottom: The option value.

since

$$A_i = \exp\left(\frac{i-1}{i}\log(A_{i-1}) + \frac{1}{i}\log(S(t_i))\right).$$

Figures 25.6 and 25.7 show details of an arithmetic average rate call with discrete sampling.

25.5.3 Exponentially Weighted and Other Averages

Simple modifications that are easily handled in the partial differential equation framework are the exponential average and the average up to a fixed time.

In the exponential continuously sampled arithmetic average just introduce the new variable

$$I = \lambda \int_{-\infty}^{t} e^{-\lambda(t-\tau)} S(\tau)\, d\tau$$

which satisfies

$$dI = \lambda(S - I)\, dt.$$

From this, the governing partial differential equation is obvious. The geometric equivalent is dealt with similarly.

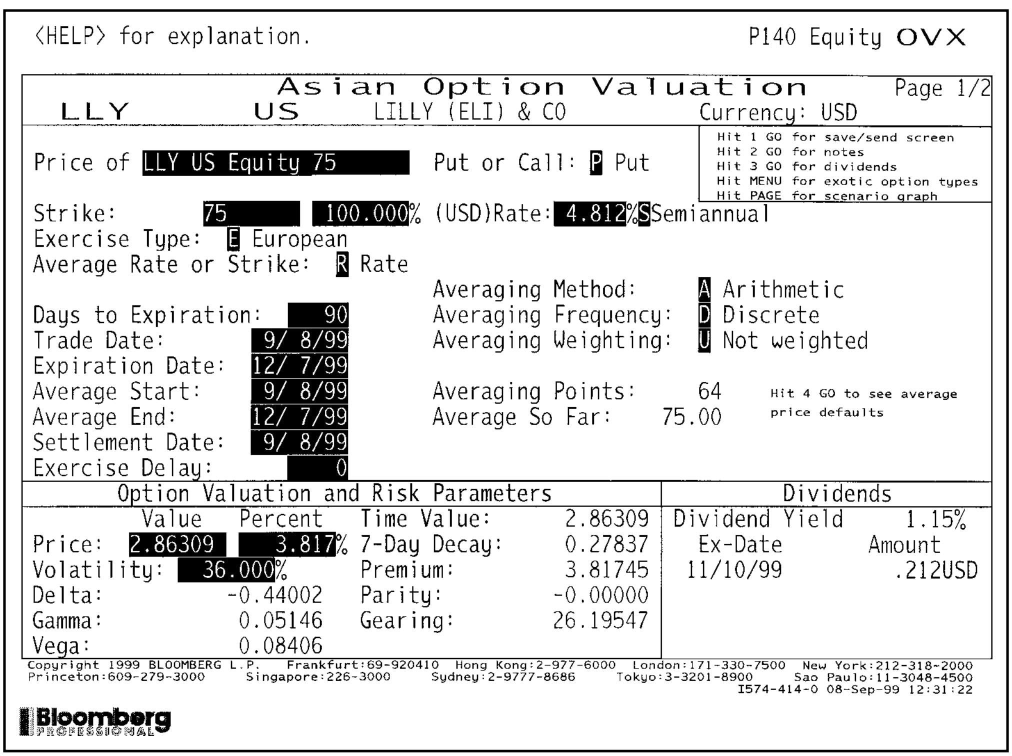


Figure 25.6 Asian option calculator and contract details. Source: Bloomberg L.P.

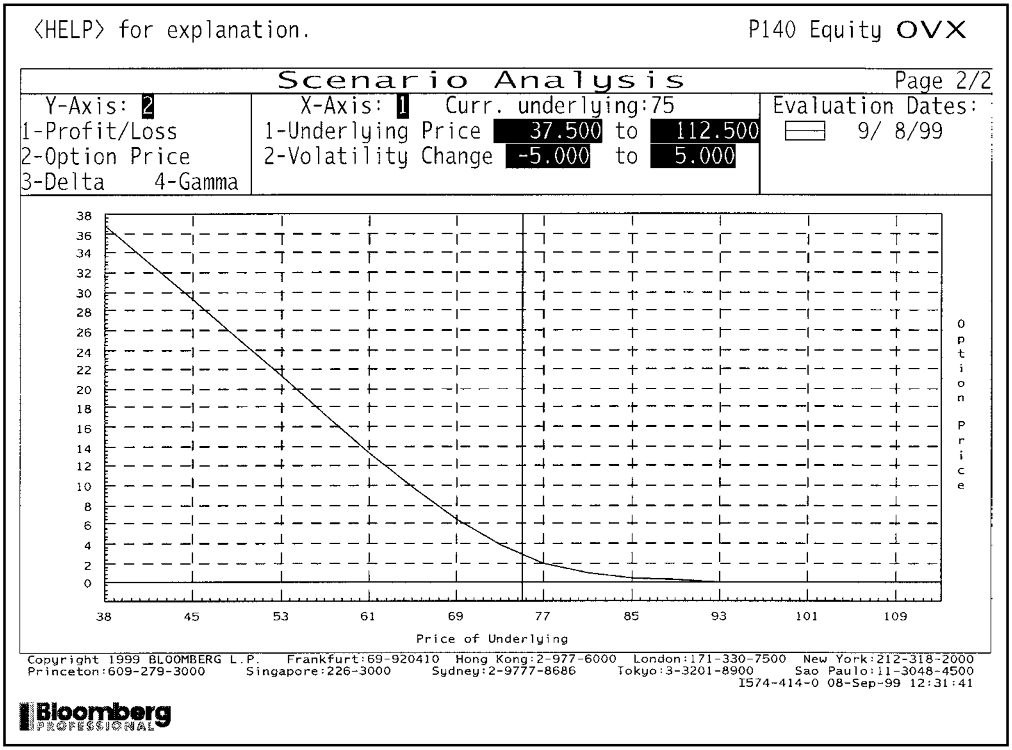


Figure 25.7 Asian option price. Source: Bloomberg L.P.

The reader can determine for himself the jump condition when the average is a discretely sampled exponential average.

When the average is only taken up to a fixed point, so that, for example, the payoff depends on

$$I = \int_0^{T_0} S(\tau)\, d\tau \quad \text{with} \quad T_0 < T,$$

then the new term in the partial differential equation (the derivative with respect to I) disappears for times greater that T_0. That is,

$$\frac{\partial V}{\partial t} + S\mathcal{H}(T_0 - t)\frac{\partial V}{\partial I} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

where $\mathcal{H}$ is the Heaviside function.

One type of contract that is *not* easily put into a partial differential equation framework with a finite number of underlyings is the moving window option. In this option, the holder can exercise early for an amount that depends on the average over the previous three months, say. The key point about this contract that makes it difficult is that the starting point of the averaging period is not fixed in time. As a result the stochastic differential equation for the path-dependent quantity cannot be written in terms of *local* values of the independent variables: All details of the path need to be known and recorded.

25.5.4 The Asian Tail

Often the averaging is confined to only a part of the life of the option. For example, if the averaging of the underlying is only over the final part of the option's life it is referred to as an **Asian tail**. Such a contract would reduce the exposure of the option to sudden moves in the underlying just before the payoff is received. A feature like this is also common in pension awards.

25.6 EARLY EXERCISE

There is not much to be said about early exercise that has not already been said elsewhere in this book. The only point to mention is that the details of the payoff on early exercise have to be well defined. The payoff at expiry depends on the value of the average up to expiry, which will, of course, not be known until expiry. Typically, on early exercise it is the average to date that is used. For example, in an American arithmetic average strike put the early payoff would be

$$\max\left(\frac{1}{t}\int_0^t S(\tau)\,d\tau - S, 0\right).$$

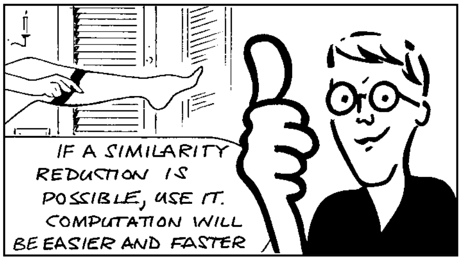
25.7 ASIAN OPTIONS IN HIGHER DIMENSIONS

We are not restricted to an average of a single underlying. The **anteater option**,[1] so called for obvious reasons, has a payoff defined in terms of the average of the ratio of two underlying S_1 and S_2:

$$I = \int_0^t \frac{S_1}{S_2}\,d\tau.$$

This contract will be in four dimensions, S_1, S_2, I and t.

25.8 SIMILARITY REDUCTIONS



As long as the stochastic differential equation or updating rule for the path-dependent quantity only contains references to S, t and the path-dependent quantity itself then the value of the option depends on three variables. Unless we are very lucky, the value of the option must be calculated numerically. Some options have a particular structure that permits a reduction in the dimensionality of the problem by use of a similarity variable. I will illustrate the idea with an example. The dimensionality of the continuously sampled arithmetic average strike option can be reduced from three to two.

The payoff for the call option is

$$\max\left(S - \frac{1}{T}\int_0^T S(\tau)\,d\tau,\ 0\right).$$

[1] Gunner Wilkins' favorite contract.

We can write the running payoff for the call option as

$$I \max\left(R - \frac{1}{t}, 0\right),$$

where

$$I = \int_0^t S(\tau)\, d\tau$$

and

$$R = \frac{S}{\int_0^t S(\tau)\, d\tau}. \tag{25.1}$$

The payoff at expiry may then be written as

$$I \max\left(R - \frac{1}{T}, 0\right).$$

In view of the form of the payoff function, it seems plausible that the option value takes the form

$$V(S, R, t) = I W(R, t), \quad \text{with} \quad R = \frac{S}{I}.$$

We find that W satisfies

$$\frac{\partial W}{\partial t} + \tfrac{1}{2}\sigma^2 R^2 \frac{\partial^2 W}{\partial R^2} + R(r - R)\frac{\partial W}{\partial R} - (r - R)W \le 0. \tag{25.2}$$

If the option is European we have strict equality in (25.2). If it is American we may have inequality in (25.2) but the constraint

$$W(R, t) \ge \max\left(R - \frac{1}{t}, 0\right)$$

must be satisfied. Moreover, if the option price ever meets the early exercise payoff it must do so smoothly. That is, the function $W(R, t)$ and its first R derivative must be continuous everywhere.

For the European option we must impose boundary conditions at both $R = 0$ and as $R \to \infty$:

$$W(0, t) = 0,$$

and

$$W(R, t) \sim R \quad \text{as } R \to \infty.$$

The solution of the European problem can be written as an infinite sum of confluent hypergeometric functions. I do not give this exact solution because it is easier (and certainly a more flexible approach) to obtain values by applying numerical methods directly to the partial differential equation.

In Figure 25.8 we see W against R at three months before expiry and with three months' averaging completed; $\sigma = 0.4$ and $r = 0.1$.

In the case of an American option, we have to solve the partial differential inequality (25.2) subject to the constraint, the final condition and the boundary conditions. We cannot do this analytically and we must find the solution numerically.

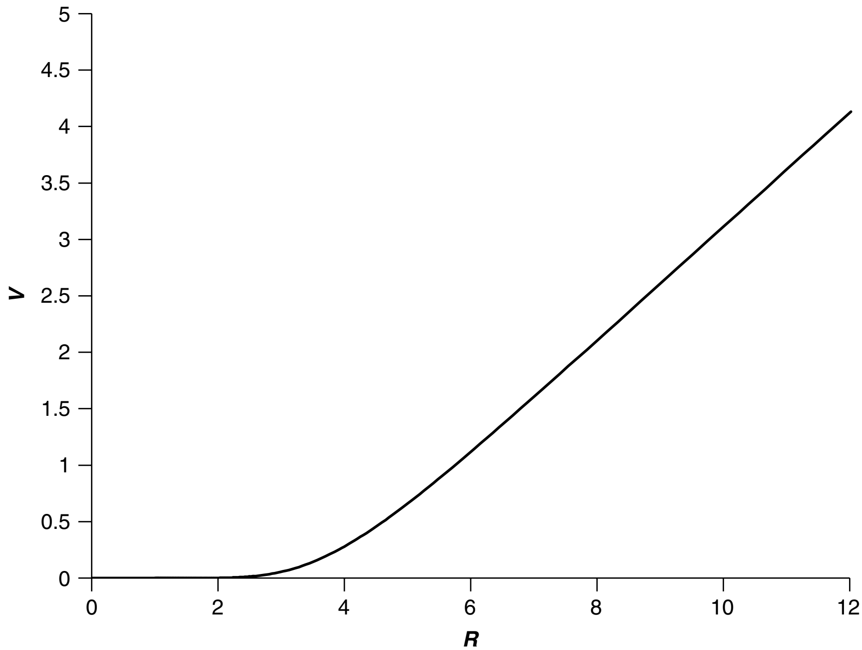


Figure 25.8 The European average strike call option; similarity variable *W* versus similarity variable *R* with $\sigma = 0.4$ and $r = 0.1$ at three months before expiry; there has already been three months' averaging.

25.8.1 Put-call Parity for the European Average Strike

The payoff at expiry for a portfolio of one European average strike call held long and one put held short is

$$I \max(R - 1/T, 0) - I \max(1/T - R, 0).$$

Whether R is greater or less than T at expiry, this payoff is simply

$$S - \frac{I}{T}.$$

The value of this portfolio is identical to one consisting of one asset and a financial product whose payoff is

$$-\frac{I}{T}.$$

In order to value this product find a solution of the average strike equation of the form

$$W(R, t) = b(t) + a(t)R \tag{25.3}$$

and with $a(T) = 0$ and $b(T) = -1/T$; such a solution would have the required payoff of $-I/T$. Substituting (25.3) into (25.2) and satisfying the boundary conditions, we find that

$$a(t) = -\frac{1}{rT}\left(1 - e^{-r(T-t)}\right), \quad b(t) = -\frac{1}{T}e^{-r(T-t)}.$$

We conclude that

$$V_C - V_P = S - \frac{S}{rT}\left(1 - e^{-r(T-t)}\right) - \frac{1}{T}e^{-r(T-t)}\int_0^t S(\tau)\,d\tau,$$

where V_C and V_P are the values of the European arithmetic average strike call and put. This is put-call parity for the European average strike option.

25.9 CLOSED-FORM SOLUTIONS AND APPROXIMATIONS

25.9.1 Kemna and Vorst (1990)

For geometric averaging there exist closed-form solutions because the geometric average of lognormal random variables is still lognormal.

These closed-form expressions are very like vanilla call and put formulae with two exceptions.

- The volatility must be replaced with $\sigma/\sqrt{3}$.
- The dividend yield must be replaced with $D + \sigma^2/6$.

25.9.2 Turnbull and Wakeman (1991)

There are no simple closed-form expressions for the values of arithmetically averaged Asians. The closed-form expressions due to Turnbull and Wakeman are only approximations. They have continuous sampling of the average. The formulae are not very accurate when sampling is discrete and/or volatility is large. The similar approach by Ed Levy also suffers from the same problems.

25.9.3 Curran (1992)

Curran has derived simple approximations based on the 'geometrical conditioning approach.' His results are more accurate than those previously mentioned.

25.9.4 Thompson (2000)

Thompson has derived upper and lower bounds for various Asian options in terms of double integrals. These expressions give quite a tight, and therefore very useful, range.

25.10 TERM-STRUCTURE EFFECTS

Is it more important to have an accurate model/approximation or to get the term structure of volatility correct?

First of all, how can we allow for time-varying volatility? The explanation for the square root of three rule for geometric options comes from considering how volatility at each instant is accumulated in the Asian option. To get an effective volatility for a vanilla option from time-varying local volatility you simply add up variances and average.

To price a vanilla option when volatility is time varying use an effective volatility of

$$\overline{\sigma}^2 = \frac{1}{T}\int_0^T \sigma(t)^2\,dt.$$

The effective volatility for a geometric Asian is also an average of local volatilities *but now weighted with time*:

$$\overline{\sigma}_G^2 = \frac{1}{T}\int_0^T \sigma(t)^2 \left(\frac{T-t}{T}\right)^2 dt.$$

(When $\sigma(t)$ is constant this gives the square root of three rule seen above.)

Which is more important, allowing for time-varying σ or having a good approximation and assuming constant volatility?

25.10.1 Some results

The following results (Table 25.1) are from Haug, Haug & Margrabe (2003). Price a six-month, fixed-strike arithmetic Asian option, with weekly samplings of the average. $S = 100$, $r = 0.05$, $\sigma = 0.2$ is base case. Upward sloping has implied volatility increasing by 0.5% per week with six-month volatility being 20%, downward sloping decreases by 0.25% per week to 20% at six months.

Table 25.1 FGA = Flat term-structure, Geometrical Average; FMC = Flat Monte Carlo; FL = Flat with Levy approximation; UGA = Upward-sloping term structure, Geometrical Average; UMA = Upward-sloping Monte Carlo; UL = Upward-sloping with Levy approximation, DGA = Downward-sloping term structure, Geometrical Average; DMC = Downward-sloping Monte Carlo; DL = Downward-sloping with Levy approximation.

Strike	FGA	FMC	FL	UGA	UMC	UL	DGA	DMC	DL
80	19.37	19.51	19.52	19.36	19.50	19.51	19.41	19.54	19.54
90	10.01	10.12	10.14	9.69	9.81	9.83	10.28	10.38	10.40
100	3.193	3.28	3.27	2.20	2.28	2.28	3.73	3.80	3.80
110	0.53	0.57	0.55	0.12	0.15	0.13	0.85	0.88	0.87
120	0.04	0.06	0.05	0.002	0.002	0.002	0.12	0.14	0.13

25.11 SOME FORMULAE

There are very few nice formulae for the values of Asian options. The most well known are for average rate calls and puts when the average is a continuously sampled, geometrical average.

The geometric average rate call

This option has payoff

$$\max(A - E, 0),$$

where A is the continuously sampled geometric average. This option has a Black–Scholes value of

$$e^{-r(T-t)}\left(G\exp\left(\frac{(r-D-\sigma^2/2)(T-t)^2}{2T}+\frac{\sigma^2(T-t)^3}{6T^2}\right)N(d_1)-EN(d_2)\right)$$

where

$$I=\int_0^t \log(S(\tau))\,d\tau,$$

$$G=e^{I/T}S^{(T-t)/T},$$

$$d_1=\frac{T\log(G/E)+(r-D-\sigma^2/2)(T-t)^2/2+\sigma^2(T-t)^3/3T}{\sigma\sqrt{(T-t)^3/3}}$$

and

$$d_2=\frac{T\log(G/E)+(r-D-\sigma^2/2)(T-t)^2/2}{\sigma\sqrt{(T-t)^3/3}}.$$

The geometric average of a lognormal random walk is itself lognormally distributed, but with a reduced volatility.

The geometric average rate put

This option has payoff

$$\max(E-A,0),$$

where A is the continuously sampled geometric average. This option has a Black–Scholes value of

$$e^{-r(T-t)}\left(EN(-d_2)-G\exp\left(\frac{(r-D-\sigma^2/2)(T-t)^2}{2T}+\frac{\sigma^2(T-t)^3}{6T^2}\right)N(d_1)\right).$$

25.12 SUMMARY

I applied the general theory of Chapter 24 to the problem of pricing Asian options, options with a payoff depending on an average. The partial differential equation approach is very powerful for these types of options, and hard to beat if your option has a similarity reduction or is American style. The approach can be generalized much further. We will see this in Chapter 28.

FURTHER READING

- Some exact solutions can be found in Boyle (1991) and Angus (1999).
- Bergman (1985) and Ingersoll (1987) present the partial differential equation formulation of some average strike options and demonstrate the similarity reduction.
- For another, rather abstract, method for valuing Asian options see Geman & Yor (1993).

- The application of the numerical Monte Carlo method is described by Kemna & Vorst (1990). They also derive some exact formulae.
- More examples of the methods described here can be found in Dewynne & Wilmott (1995b, c).
- For an approximate valuation of arithmetic Asian options see Levy (1990) who replaces the density function for the distribution of the average by a lognormal function. See also Curran (1992), Turnbull & Wakeman (1991) and Thompson (2000).
- Further detailed analysis of Asian options can be found in Lewis (2002).
- An interesting Asian option is considered by Krekel (2003).

CHAPTER 26
lookback options

In this Chapter...

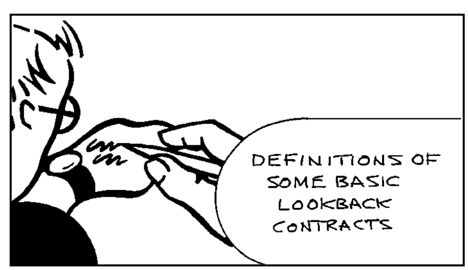
- the features that make up a lookback option
- how to put the lookback option into the Black–Scholes framework for both the continuous- and discrete-sampling cases

26.1 INTRODUCTION

The dream contract has to be one that pays the difference between the highest and the lowest asset prices realized by an asset over some period. Any speculator is trying to achieve such a trade. The contract that pays this is an example of a **lookback option**, an option that pays off some function of the realized maximum and/or minimum of the underlying asset over some prescribed period. Since lookback options have such an extreme payoff they tend to be expensive.

We can price these contracts in the Black–Scholes environment quite easily, theoretically. There are two cases to consider: Whether the maximum/minimum is measured continuously or discretely. Here I will always talk about the 'maximum' of the asset. The treatment of the 'minimum' should be obvious.

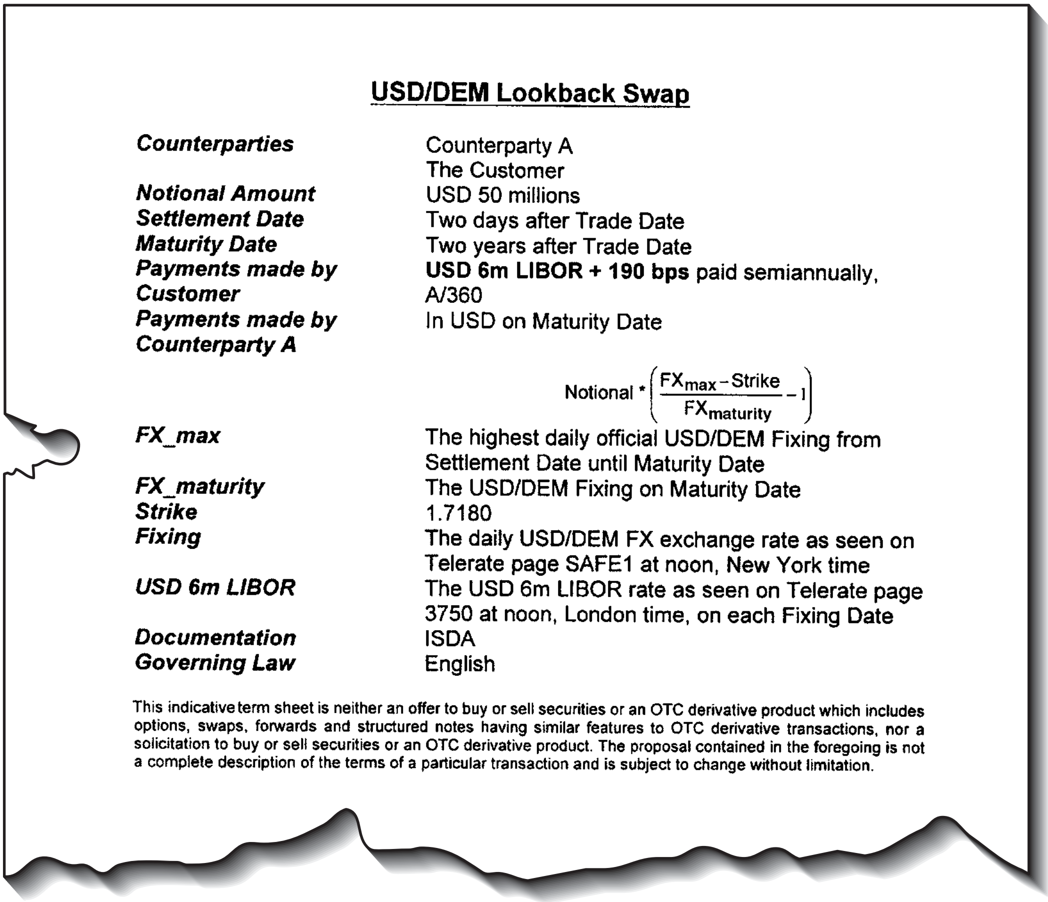
26.2 TYPES OF PAYOFF



For the basic lookback contracts, the payoff comes in two varieties, like the Asian option. These are the *rate* and the *strike* option, also called the **fixed strike** and the **floating strike** respectively. These have payoffs that are the same as vanilla options except that in the strike option the vanilla exercise price is replaced by the maximum. In the rate option it is the asset value in the vanilla option that is replaced by the maximum.

26.3 CONTINUOUS MEASUREMENT OF THE MAXIMUM

In Figure 26.1 is shown the term sheet for a foreign exchange lookback swap. One counterparty pays a floating interest rate every six months, while the other counterparty pays on maturity



USD/DEM Lookback Swap

Counterparties	Counterparty A The Customer
Notional Amount	USD 50 millions
Settlement Date	Two days after Trade Date
Maturity Date	Two years after Trade Date
Payments made by Customer	**USD 6m LIBOR + 190 bps** paid semiannually, A/360
Payments made by Counterparty A	In USD on Maturity Date $\text{Notional} * \left(\frac{FX_{max} - \text{Strike}}{FX_{maturity}} - 1 \right)$
FX_max	The highest daily official USD/DEM Fixing from Settlement Date until Maturity Date
FX_maturity	The USD/DEM Fixing on Maturity Date
Strike	1.7180
Fixing	The daily USD/DEM FX exchange rate as seen on Telerate page SAFE1 at noon, New York time
USD 6m LIBOR	The USD 6m LIBOR rate as seen on Telerate page 3750 at noon, London time, on each Fixing Date
Documentation	ISDA
Governing Law	English

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 26.1 Term sheet for USD/DEM lookback swap.

a linear function of the maximum realized level of the exchange rate. The floating interest payments of six-month LIBOR can be represented as a much simpler cashflow, as explained in Chapter 14. Then this contract becomes a very straightforward lookback option.

Introduce the new variable M as the realized maximum of the asset from the start of the sampling period $t = 0$, say, until the current time t:

$$M = \max_{0 \le \tau \le t} S(\tau).$$

In Figure 26.2 is shown a realization of an asset price path and the continuously sampled maximum. An obvious point about this plot, but one that is worth mentioning, is that the asset price is always below the maximum. (This will not be the case when we come to examine the discretely sampled case.) The value of our lookback option is a function of three variables, $V(S, M, t)$ but now we have the restriction

$$0 \le S \le M.$$

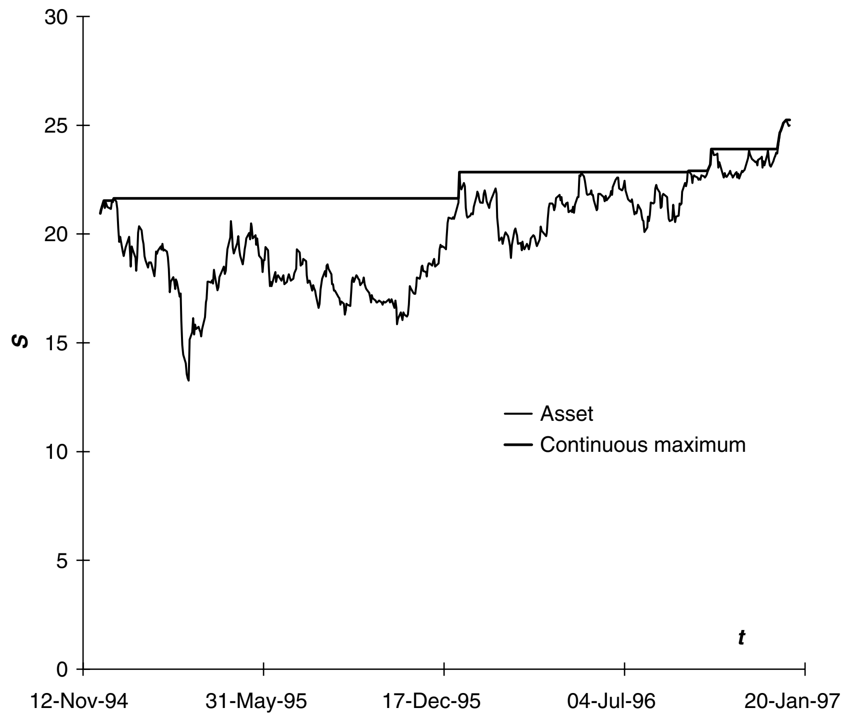


Figure 26.2 An asset price path and the continuously sampled maximum.

This observation will also lead us to the correct partial differential equation for the option's value, and the boundary conditions. We derive the equation in a heuristic fashion that can be made rigorous.

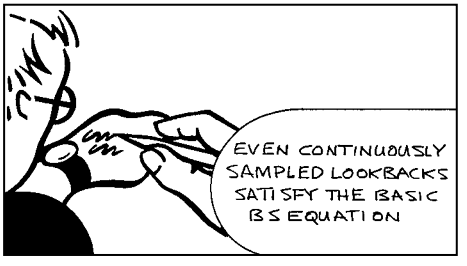
From Figure 26.2 we can see that for most of the time the asset price is below the maximum. But there are times when they coincide. When $S < M$ the maximum cannot change and so the variable M satisfies the stochastic differential equation

$$dM = 0.$$

Hold that thought. While $0 \leq S < M$ the governing equation must be Black–Scholes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

with M as a 'parameter' and only for $S < M$.



The behavior of the option value when $S = M$ tells us the *boundary condition* to apply there. The boundary condition is

$$\frac{\partial V}{\partial M} = 0 \quad \text{on} \quad S = M.$$

The reason for this boundary condition is that the option value is insensitive to the level of the maximum when the asset price is *at* the maximum. This is because the probability of the present maximum still being the maximum at expiry is zero. The rigorous derivation of this boundary condition is rather messy; see the original paper by Goldman, Sosin & Gatto (1979) for the details.

Finally, we must impose a condition at expiry to reflect the payoff. As an example, consider the lookback rate call option. This has a payoff given by

$$\max(M - E, 0).$$

The lookback strike put has a payoff given by

$$\max(M - S, 0).$$

26.4 DISCRETE MEASUREMENT OF THE MAXIMUM

The discretely sampled maximum is shown in Figure 26.3. Not only can the asset price go above the maximum, but in this figure we see that the maximum has very rarely been increased. Discrete sampling, as well as being more practical than continuous sampling, is used to decrease the value of a contract.

When the maximum is measured at discrete times we must first define the updating rule, from which follows the jump condition to apply across the sampling dates.

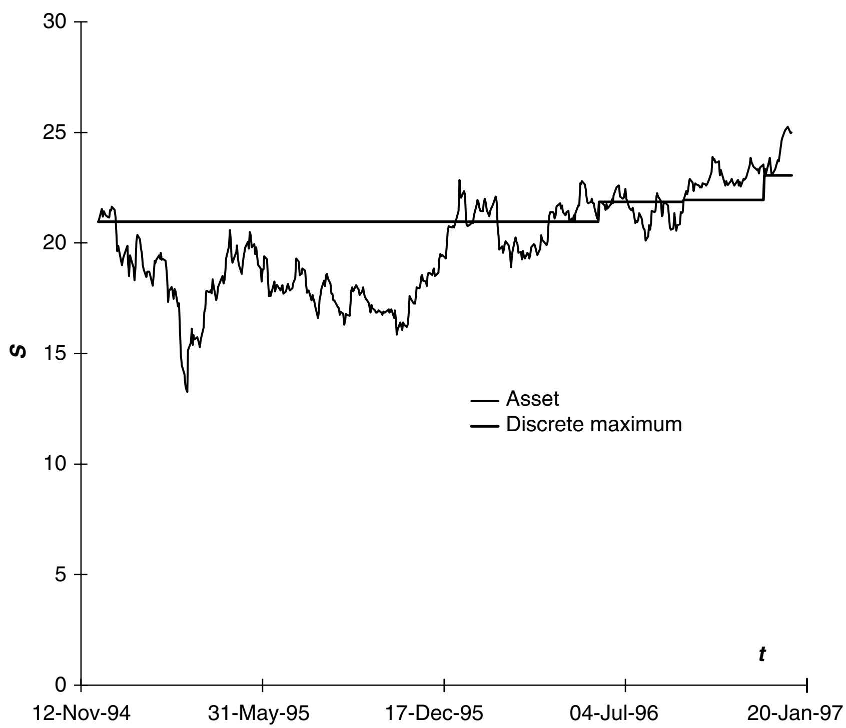


Figure 26.3 An asset price path and the discretely sampled maximum.

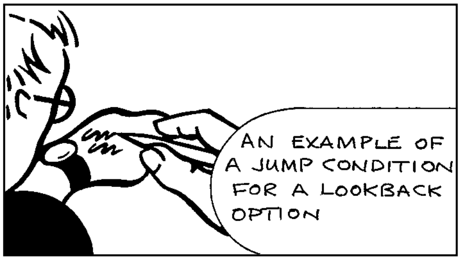
If the maximum is measured at times t_i then the updating rule is simply

$$M_i = \max(S(t_i), M_{i-1}).$$

The jump condition is then simply

$$V(S, M, t_i^-) = V(S, \max(S, M), t_i^+).$$

Note that the Black–Scholes equation is to be solved for all S; it is no longer constrained to be less than the maximum.



26.5 SIMILARITY REDUCTION

The general lookback option with a payoff depending on one path-dependent quantity is a three-dimensional problem. The three dimensions are asset price, the maximum and time. The numerical solution of this problem is more time consuming than a two-dimensional problem. However, there are some special, and important, cases when the dimensionality of the problem can be reduced.

This reduction relies on some symmetry properties in the equation and is not something that can be applied to all, or, indeed, many, lookback contracts. It is certainly possible if the payoff takes the form

$$M^\alpha P\left(\frac{S}{M}\right). \tag{26.1}$$

For example, this is true for the lookback strike put which has payoff

$$\max(M - S, 0) = M \max\left(1 - \frac{S}{M}, 0\right).$$

Generally, if the payoff takes the form (26.1), then the substitution

$$\xi = \frac{S}{M}$$

leads to a problem for $W(\xi, t)$ where

$$V(S, M, t) = M^\alpha W(\xi, t)$$

where W satisfies the governing equation

$$\frac{\partial W}{\partial t} + \tfrac{1}{2}\sigma^2\xi^2\frac{\partial^2 W}{\partial \xi^2} + r\xi\frac{\partial W}{\partial \xi} - rW = 0,$$

the final condition

$$W(\xi, T) = P(\xi)$$

and the boundary condition

$$\frac{\partial W}{\partial \xi} - \alpha W = 0 \quad \text{on} \quad \xi = 1.$$

26.6 SOME FORMULAE

Floating strike lookback call

The continuously sampled version of this option has a payoff

$$\max(S - M, 0) = S - M,$$

where M is the realized minimum of the asset price. In the Black–Scholes world the value is

$$Se^{-D(T-t)}N(d_1) - Me^{-r(T-t)}N(d_2) + Se^{-r(T-t)}\frac{\sigma^2}{2(r-D)}$$
$$\times\left(\left(\frac{S}{M}\right)^{-(2(r-D)/\sigma^2)} N\left(-d_1 + \frac{2(r-D)\sqrt{T-t}}{\sigma}\right) - e^{(r-D)(T-t)}N(-d_1)\right),$$

where

$$d_1 = \frac{\log(S/M) + (r - D + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

Floating strike lookback put

The continuously sampled version of this option has a payoff

$$\max(M - S, 0) = M - S,$$

where M is the realized maximum of the asset price. The value is

$$Me^{-r(T-t)}N(-d_2) - Se^{-D(T-t)}N(-d_1) + Se^{-r(T-t)}\frac{\sigma^2}{2(r-D)}$$
$$\times\left(-\left(\frac{S}{M}\right)^{-(2(r-D)/\sigma^2)} N\left(d_1 - \frac{2(r-D)\sqrt{T-t}}{\sigma}\right) + e^{(r-D)(T-t)}N(d_1)\right),$$

where

$$d_1 = \frac{\log(S/M) + (r - D + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

Fixed strike lookback call

This option has a payoff given by

$$\max(M - E, 0)$$

where M is the realized maximum. For $E > M$ the fair value is

$$Se^{-D(T-t)}N(d_1) - Ee^{-r(T-t)}N(d_2) + Se^{-r(T-t)}\frac{\sigma^2}{2(r-D)}$$
$$\times\left(-\left(\frac{S}{E}\right)^{-(2(r-D)/\sigma^2)}N\left(d_1 - \frac{2(r-D)\sqrt{T-t}}{\sigma}\right) + e^{(r-D)(T-t)}N(d_1)\right),$$

where

$$d_1 = \frac{\log(S/E) + (r - D + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

When $E < M$ the value is

$$(M-E)e^{-r(T-t)} + Se^{-D(T-t)}N(d_1) - Me^{-r(T-t)}N(d_2) + Se^{-r(T-t)}\frac{\sigma^2}{2(r-D)}$$
$$\times\left(-\left(\frac{S}{M}\right)^{-(2(r-D)/\sigma^2)}N\left(d_1 - \frac{2(r-D)\sqrt{T-t}}{\sigma}\right) + e^{(r-D)(T-t)}N(d_1)\right),$$

where

$$d_1 = \frac{\log(S/M) + (r - D + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

Fixed strike lookback put

This option has a payoff given by

$$\max(E - M, 0)$$

where M is the realized minimum. For $E < M$ the fair value is

$$Ee^{-r(T-t)}N(-d_2) - Se^{-D(T-t)}N(-d_1) + Se^{-r(T-t)}\frac{\sigma^2}{2(r-D)}$$
$$\times\left(\left(\frac{S}{E}\right)^{-(2(r-D)/\sigma^2)}N\left(-d_1 + \frac{2(r-D)\sqrt{T-t}}{\sigma}\right) - e^{(r-D)(T-t)}N(-d_1)\right),$$

where

$$d_1 = \frac{\log(S/E) + (r - D + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T - t}.$$

When $E > M$ the value is

$$(E - M)e^{-r(T-t)} - Se^{-D(T-t)}N(-d_1) + Me^{-r(T-t)}N(-d_2) + Se^{-r(T-t)}\frac{\sigma^2}{2(r - D)}$$
$$\times \left(\left(\frac{S}{M} \right)^{-(2(r-D)/\sigma^2)} N\left(-d_1 + \frac{2(r - D)\sqrt{T - t}}{\sigma} \right) - e^{(r-D)(T-t)}N(-d_1) \right),$$

where

$$d_1 = \frac{\log(S/M) + (r - D + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T - t}.$$

26.7 SUMMARY

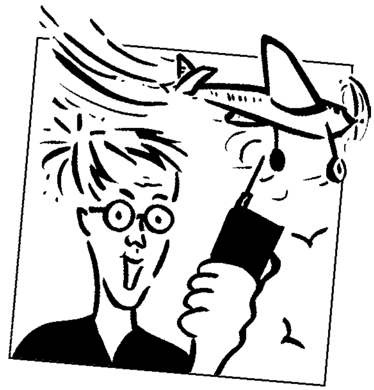
Lookback options, and lookback features generally, are seen in many types of contract. They are quite common in fixed-income products where an interest payment may depend on the maximum level that rates have reached over some previous period. The same partial differential equation framework that we have seen in the equity world, carries over in principle to the more complicated stochastic interest rate world.

FURTHER READING

- See Goldman, Sosin & Gatto (1979) for the first academic description of lookback options. They show how to derive the crucial boundary condition rigorously.
- Conze & Viswanathan (1991) give the derivation of formulae for several types of lookback option.
- Babbs (1992) puts the lookback option into a binomial setting.
- See Dewynne & Wilmott (1994b) for a derivation of the governing equation and boundary conditions.
- Heynen & Kat (1995) discuss the discrete and partial monitoring of the maximum.
- Two contracts that are related to lookback options are the stop-loss option, described by Fitt, Dewynne & Wilmott (1994), and the Russian option, see Duffie & Harrison (1992).
- The distribution of the maximum in all sorts of models is considered in Lewis (2004c).

CHAPTER 27

derivatives and stochastic control



In this Chapter...

- an intro to the subject of stochastic control
- valuing a rather complex exotic

27.1 INTRODUCTION

Some options give the holder an element of control during the life of the contract. The most common of these is the American option which the holder can exercise whenever he wants and doesn't have to wait until expiry. Some contracts give the holder even greater flexibility and require him to make many and complex decisions.

In this chapter I want to describe an option that sits outside the framework we have developed so far. It is not too difficult to analyze but it introduces some new ideas, and in particular it is a gentle introduction to the subject of **stochastic control**.

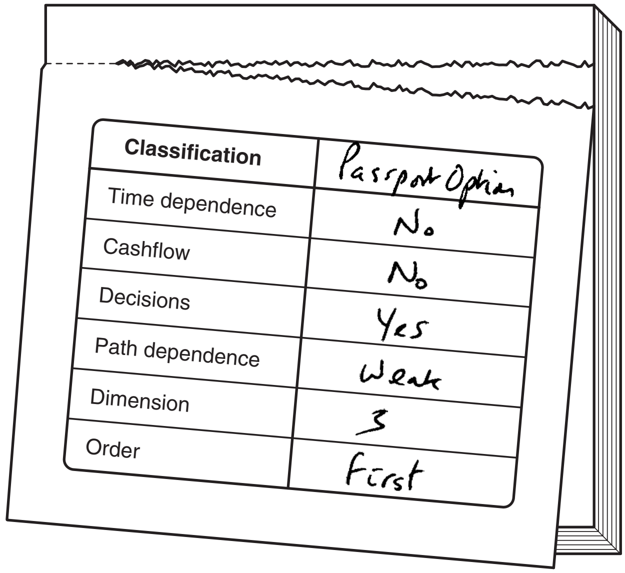
27.2 PERFECT TRADER AND PASSPORT OPTIONS

Suppose that you invest in a particular stock, keeping track of its movements and buying or selling according to your view of its future direction. The amount of money that you accumulate due to trading in this stock is called the **trading account**. If you are a good trader or lucky then the amount in the account will grow; if you are a bad trader or unlucky then the amount will be negative. How much would you pay to be insured against losing money over a given time horizon? A **perfect trader** or **passport option** is a call option on the trading account, giving the holder the amount in his account at the end of the horizon if it is positive, or zero if it is negative.

To value this contract we must introduce a new variable π which is the value of the trading account, meaning the value of the stocks held together with any cash accumulated.

We also introduce q as the amount of stock held at time t. This is the *control variable*. We will be deciding, as part of the valuation process, what this q should be at each time.

Since we hold q of the stock our total stock position is qS, and since our total worth is π we must be holding $\pi - qS$ in cash. Our stock position will change by $q\,dS$ while our cash



Classification option table for Passport Option.

position earns interest, an amount $r(\pi - qS)\,dt$.[1] Therefore the stochastic differential equation for π is

$$d\pi = r(\pi - qS)\,dt + q\,dS. \tag{27.1}$$

The quantity q is of our choosing, will vary as time and the stock price change, and is called the **strategy**; it will be a function of S, π and t. I will restrict the size of the position in the stock by insisting that $|q| \le 1$. Equation (27.1) contains a deterministic and a random term. The first term says that there is growth in the cash holding, $\pi - qS$, due to the addition of interest at a rate r, and the second term is due to the change in value of the stock holding.

The contract will pay off an amount

$$\max(\pi, 0)$$

at time T. This will be the final condition for our option value $V(S, \pi, t)$. Note that the option value is a function of three variables.

Now let us hedge this option:

$$\Pi = V - \Delta S.$$

We find that

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + q\sigma^2 S^2 \frac{\partial^2 V}{\partial S\,\partial \pi} + \tfrac{1}{2}q^2\sigma^2 S^2 \frac{\partial^2 V}{\partial \pi^2}\right) dt + \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial \pi} d\pi - \Delta\,dS.$$

Since $d\pi$ contains a dS term the correct hedge ratio is

$$\Delta = \frac{\partial V}{\partial S} + q\frac{\partial V}{\partial \pi}.$$

[1] I have assumed that the cash position earns interest. This may not be the case in practice so you will have to examine the relevant term sheet closely. What changes need to be made to the following if interest is *not* earned?

From the no-arbitrage principle follows the pricing equation

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+q\sigma^2S^2\frac{\partial^2 V}{\partial S\,\partial\pi}+\tfrac{1}{2}q^2\sigma^2S^2\frac{\partial^2 V}{\partial\pi^2}+rS\frac{\partial V}{\partial S}+r\pi\frac{\partial V}{\partial\pi}-rV=0. \tag{27.2}$$

This is not a diffusion equation in two space-like variables because S and π are perfectly correlated; the equation really has one space-like and two time-like variables.

We come to the stochastic control part of the problem in choosing q. If we are selling this contract then we should assume that the holder acts optimally, making the contract's value as high as possible. That doesn't mean the holder will follow such a strategy since he will have other priorities, he will have a view on the market and will not be hedging. The highest value for the contract occurs when q is chosen to maximize the q terms in (27.2):

$$\max_{|q|\le 1}\left(q\sigma^2S^2\frac{\partial^2 V}{\partial S\,\partial\pi}+\tfrac{1}{2}q^2\sigma^2S^2\frac{\partial^2 V}{\partial\pi^2}\right).$$

This is the only term containing q.

27.2.1 Similarity Solution

If the payoff is simply

$$V(S,\pi,T)=\max(\pi,0)$$

then we can find a similarity solution of the form

$$V(S,\pi,t)=SH(\xi,t),\quad \xi=\frac{\pi}{S}.$$

In this case Equation (27.2) becomes

$$\frac{\partial H}{\partial t}+\tfrac{1}{2}\sigma^2(\xi-q)^2\frac{\partial^2 H}{\partial\xi^2}=0, \tag{27.3}$$

the payoff is

$$H(\xi,T)=\max(\xi,0)$$

and the optimal strategy is

$$\max_{|q|\le 1}\left((\xi-q)^2\frac{\partial^2 H}{\partial\xi^2}\right).$$

Assuming that $\partial^2H/\partial\xi^2>0$ (and this can be verified *a posteriori*) because of the nature of the equation and its final condition, the optimal strategy is

$$q=\begin{cases}-1 & \text{when } \xi>0\\ \ \ 1 & \text{when } \xi<0.\end{cases}$$

For a more general payoff the strategy would depend on the sign of $\partial^2H/\partial\xi^2$.

USD/DEM 'Perfect Trader' Option

Notional Amount	USD 25,000,000+
Option Maturity	Three months from Trade Date
Allowed Position	Long or short up to Notional Amount
Transaction Frequency	Up to two times daily
Settlement Amount	Max(0,sum total in DEM of the gains + losses on each of the trades)
Upfront Premium	3.35% of Notional Amount

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 27.1 Term sheet for a perfect trader option.

The option value, assuming the optimal strategy, thus satisfies

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2(|\xi| + 1)^2 \frac{\partial^2 H}{\partial \xi^2} = 0.$$

The term sheet for a perfect trader option is shown in Figure 27.1.

The holder of the option is allowed to make a series of hypothetical trades on the USD/DEM exchange rate. Two trades are allowed per day with a maximum allowed position long and short. The holder then receives the positive part of his transactions. If his trades result in a negative net balance this is written off.

27.3 LIMITING THE NUMBER OF TRADES

The passport option described above allows the trader to buy and sell the underlying asset as often as he wishes. How valuable is the right to trade at will? Does it make much difference if the number of, or time between, trades is restricted? In practice, the trader can only make a finite number of purchases or sales.

Let's start by examining the case of limited number of trades.

Introduce the notation $V^{n+}(S, \pi, t)$ and $V^{n-}(S, \pi, t)$ to mean the value of the passport option when there are still n trades allowed and the $+/-$ refers to whether the trader is currently long or short the (maximum permitted quantity of the) underlying.

Given what you have read so far in this book you should be able to derive the governing equations without going through every single detail. So let me point out the relevant steps. First of all, the equations for $V^{n+}(S, \pi, t)$ and $V^{n-}(S, \pi, t)$ will be like Equation (27.2) except that q will be $+1$ for the V^{n+} equation and -1 for the V^{n-} equation. Second, because we are dealing with a contract in which there is a decision to flip from long to short or vice versa, the optimality that we assume in this strategy is very much like the optimality in the American option, and there we know that we are not dealing with an equation as such but an inequality.

So the two functions satisfy the inequalities

$$\frac{\partial V^{n+}}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{n+}}{\partial S^2}+\sigma^2S^2\frac{\partial^2 V^{n+}}{\partial S\,\partial \pi}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{n+}}{\partial \pi^2}+rS\frac{\partial V^{n+}}{\partial S}+r\pi\frac{\partial V^{n+}}{\partial \pi}-rV^{n+}\leq 0$$

and

$$\frac{\partial V^{n-}}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{n-}}{\partial S^2}-\sigma^2S^2\frac{\partial^2 V^{n-}}{\partial S\,\partial \pi}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{n-}}{\partial \pi^2}+rS\frac{\partial V^{n-}}{\partial S}+r\pi\frac{\partial V^{n-}}{\partial \pi}-rV^{n-}\leq 0.$$

Observe the sign difference between these.

The final condition is

$$V^{0\pm}(S,\pi,T)=\max(\pi,0).$$

A trade is optimal when the value with n trades left (and currently long/short) is the same as with $n-1$ trades left and the opposite position (short/long):

$$V^{n+}\geq V^{(n-1)-}\quad\text{and}\quad V^{n-}\geq V^{(n-1)+}.$$

This completes the formulation of the problem.

A simple modification is to include a fixed penalty P to be paid on each trade. This is modeled by

$$V^{n+}\geq V^{(n-1)-}+P\quad\text{and}\quad V^{n-}\geq V^{(n-1)+}+P.$$

27.4 LIMITING THE TIME BETWEEN TRADES

Instead of restricting the number of trades we could restrict the minimum time interval between trades. If a trade has been made then another is not allowed until a time ω has passed. We must introduce a 'clock' that keeps track of the time since the last trade. This clock is reset to zero as soon as a trade is made. The passport option value is now given by $V^+(S,\pi,t,\tau)$ and $V^-(S,\pi,t,\tau)$ where τ is the time on the clock.

These two functions satisfy the inequalities

$$\frac{\partial V^{+}}{\partial t}+\frac{\partial V^{+}}{\partial \tau}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{+}}{\partial S^2}+\sigma^2S^2\frac{\partial^2 V^{+}}{\partial S\,\partial \pi}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{+}}{\partial \pi^2}+rS\frac{\partial V^{+}}{\partial S}+r\pi\frac{\partial V^{+}}{\partial \pi}-rV^{+}\leq 0$$

and

$$\frac{\partial V^{-}}{\partial t}+\frac{\partial V^{-}}{\partial \tau}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{-}}{\partial S^2}-\sigma^2S^2\frac{\partial^2 V^{-}}{\partial S\,\partial \pi}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V^{-}}{\partial \pi^2}+rS\frac{\partial V^{-}}{\partial S}+r\pi\frac{\partial V^{-}}{\partial \pi}-rV^{-}\leq 0.$$

The final condition is

$$V^{\pm}(S,\pi,T,\tau)=\max(\pi,0).$$

A trade is optimal when the value of the option is the same as with the opposite position in the underlying and with a trade permitted (i.e. $\tau = \omega$):

$$V^+(S, \pi, t, \omega) \geq V^-(S, \pi, t, 0) \quad \text{and} \quad V^-(S, \pi, t, \omega) \geq V^+(S, \pi, t, 0).$$

In the absence of any penalties, there is a similarity solution for the passport option price of the form

$$V = S\Phi\left(\frac{\pi}{S}, t\right),$$

for some function Φ. Φ is a function also of the remaining variables. For the option with limited trades it is a function of n and for the time-limited problem it is a function of τ.

Example

Volatility of the underlying asset is 20%, six months to expiry. The plain passport option value is $0.0589 \times S$. In Figure 27.2 is shown the effect on the price of limiting the number of trades.

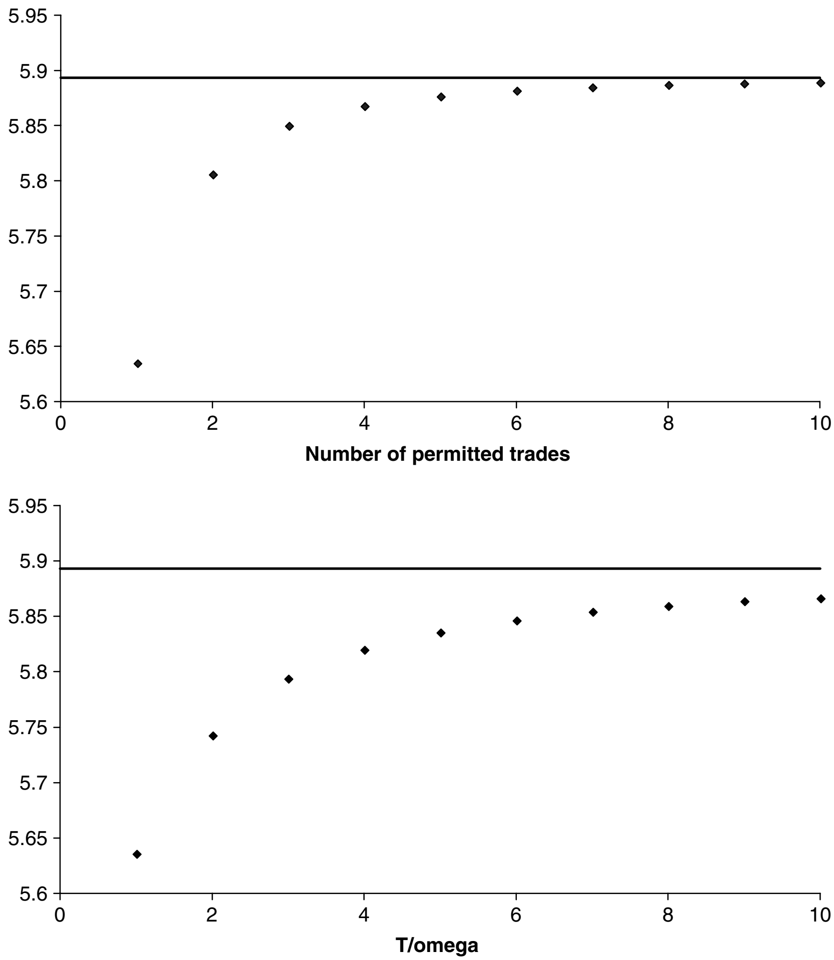
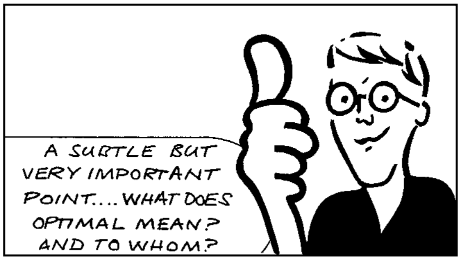


Figure 27.2 Values of different types of passport option.

In the upper plot is shown the option value (when $S = 100$) as a function of the number of permitted trades. The solid line is the option value with unlimited number of trades, the plain passport option. Note that even with only three or four trades allowed the option value is close to the unlimited trades case. The lower plot shows the option value against T/ω when the time between trades is restricted. The structure of the price is similar to the restricted number of trades case.

Allowing a trader to trade more frequently has little effect on the theoretical price of the contract. It does make hedging of the contract slightly harder since the option writer will be hedging at least as often as the owner trades. The more frequent the hedging the greater the effect of transaction costs.

27.5 NON-OPTIMAL TRADING AND THE BENEFITS TO THE WRITER



In this section I want to discuss issues to do with what is the best trading strategy for the option holder.

For the sake of argument, let's assume that we are in a Black–Scholes world. It's reasonable to also assume that the holder is not delta hedging. If he is then he's following the same strategy as the option writer. This would be a pointless strategy unless the holder thinks he has bought the option very cheaply. So let's suppose that he has another strategy in mind; maybe he sees himself as a hot trader with insight into the market's direction. Is he right to trade in a 'non-optimal' fashion?

We've used 'optimal' to mean something very specific; the strategy is optimal if it gives the option the highest value to someone who is delta hedging. So what is optimal to a delta hedger will generally not be optimal to our punter. Yet that is his concern. If he wants to take a chance, then I hope he gets a better return. If he trades well he would expect to make a greater return than the risk-free rate by taking more risk.

How does the option writer feel about this? He valued the contract on the assumption that the trader would trade optimally, with the hedger's definition of optimal. But the trader is following some other strategy. By definition, the 'value' of the contract to the holder is less than he paid for it *if he is following a non-optimal strategy.* In other words, the seller of the option is pleased that the holder is acting in a non-optimal fashion; he will probably make more money that way. Note that I am not saying that the trader is wrong to trade according to his own priorities, nor am I saying that the passport option is incorrectly valued. We'll discuss this again in Chapter 63.

27.6 SUMMARY

This new breed of derivative is particularly exciting. It is a challenge from the pricing and hedging points of view and especially because complex and continuous decisions must be made by the holder/trader. As with most quantitative finance, the math is quite straightforward once you've seen how it's done.

FURTHER READING

- Hyer, Lipton-Lifschitz & Pugachevsky (1997) describe the passport option.
- Ahn, Penaud & Wilmott (1998) and Penaud, Wilmott & Ahn (1998) discuss many, many extensions to the passport option concept.

CHAPTER 28

miscellaneous exotics

In this Chapter...

- contract specifications for many more exotic derivatives
- more 'tricks of the trade' for valuing exotic contracts in the partial differential equation framework

28.1 INTRODUCTION

The universe of exotic derivatives is large, and becoming larger all the time. I have tried to bring together, and even classify, many of these contracts. For example, a whole chapter was devoted entirely to Asian options, options depending on an average of the realized asset path. Nevertheless, because of the complexity of the instruments available and the increase in their number, the classification exercise can become a labor of Sisyphus. In this chapter I give up on this exercise and introduce a miscellany of exotics, with the aim of expanding the techniques available to the reader for pricing new contracts. Typically, I introduce the pricing and hedging concepts in an equity framework, but by the end of the book the astute reader will appreciate their applicability to other worlds, such as fixed income.

28.2 FORWARD-START OPTIONS

As its name suggests, a **forward-start** is an option that comes into being some time in the future. Let us consider an example: A forward-start call option is bought now, at time $t = 0$, but with a strike price that is not known until time T_1, when the strike is set at the asset price on that date, say. The option expires later at time T. There are two ways to solve this problem in a Black–Scholes world, the simple and the complicated. We begin with the former.

The simple way to price this contract is to ask what happens at time T_1. At that time we get an at-the-money option with a time $T - T_1$ left to expiry. If the stock price at time T_1 is S_1 then the value of the contract is simply the Black–Scholes value with $S = S_1$, $t = T - T_1$, $E = S_1$ and with given values for r and σ. For a call option this value, as a function of S_1, is

$$S_1 N(d_1) - S_1 e^{-r(T-T_1)} N(d_2),$$

where

$$d_1 = \frac{r + \frac{1}{2}\sigma^2}{\sigma}\sqrt{T - T_1}$$

and

$$d_2 = \frac{r - \frac{1}{2}\sigma^2}{\sigma}\sqrt{T - T_1}.$$

The value is proportional to S. Thus, at time T_1 we will hold an asset worth

$$S_1 f(T - T_1).$$

Since this is a constant multiplied by the asset price at time T_1 the value today must be

$$Sf(T - T_1)$$

where S is today's asset price.

The other way of valuing this option, within our general path-dependent framework, is to introduce a new state variable $\mathcal{S}$ which is defined for $t \geq T_1$ as being the asset price at time T_1,

$$\mathcal{S} = S(T_1). \tag{28.1}$$

For times before that, we set $\mathcal{S} = 0$, although it does not actually matter what it is. The result of this pricing method is, as we now see, identical to the above simple method but the technique of introducing a new variable has a very wide applicability.

The option has a value that depends on three variables: $V(S, \mathcal{S}, t)$. This function satisfies the Black–Scholes equation in S and t since $\mathcal{S}$ is not stochastic and is constant after the date T_1. At expiry we have

$$V(S, \mathcal{S}, T) = \max(S - \mathcal{S}, 0).$$

At the start date, T_1, the strike price is set to the current asset price, this is Equation (28.1). The jump condition across T_1 is simply

$$V(S, \mathcal{S}, T_1^-) = V(S, S, T_1^+).$$

And that's all there is to it.

For this, the simplest of forward-start options, we can take the analysis considerably further. We can either observe that the option value after time T_1 is that of a vanilla call, and therefore we have a formula for it as above, or we can use the similarity variable, $\xi = S/\mathcal{S}$ to transform the problem to

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2\xi^2\frac{\partial^2 H}{\partial \xi^2} + r\xi\frac{\partial H}{\partial \xi} - rH = 0$$

with

$$H(\xi, T) = \max(\xi - 1, 0)$$

and where $V(S, \mathcal{S}, t) = \mathcal{S}H(\xi, t)$.

Across time T_1 we have

$$V(S, \mathcal{S}, T_1^-) = V(S, S, T_1^+) = SH(1, T_1^+).$$

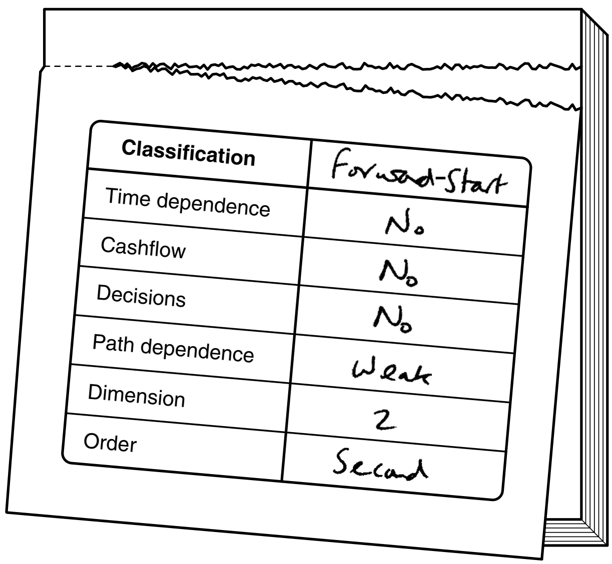
If we use this as the final condition (at time T_1) for the value for the option up to time T_1 then we see that for such times the option value is simply proportional to S. The unique solution is therefore

$$V(S, \mathcal{S}, t) = SH(1, T_1) \quad \text{for} \quad t < T_1.$$

Of course, $H(1, T_1)$ is just the value of an at-the-money call option with a strike of 1 at a time $T - T_1$ before expiry. This takes us back to the result of the simple approach.

I have stressed the path-dependent approach, although it is unnecessarily complicated for the simple forward-start option, because of its applicability to many other contracts.

Forward-start options can also come in a fixed notional variety. I will leave the pricing of these as an exercise.



Classification option table for Forward-Start.

28.3 SHOUT OPTIONS

A **shout** call option is a vanilla call option but with the extra feature that the holder can at any time reset the strike price of the option to the current level of the asset (if it is higher than the original strike). There is simultaneously a payment, usually of the difference between the old and the new strike prices. The action of resetting is called 'shouting.'

Since there is clearly an element of optimization in the matter of shouting, one would expect to see a free boundary problem occur quite naturally as with American options.

To value this contract introduce the two functions: $V_a(S, X, t)$ and $V_b(S, X, t)$. The former is the value of the option after shouting and the latter, before. S is the underlying asset value, X the strike level. Since the variable X is updated discretely, the relevant equation to solve is the basic Black–Scholes equation. The final conditions are

$$V_a(S, X, T) = V_b(S, X, T) = \max(S - X, 0).$$

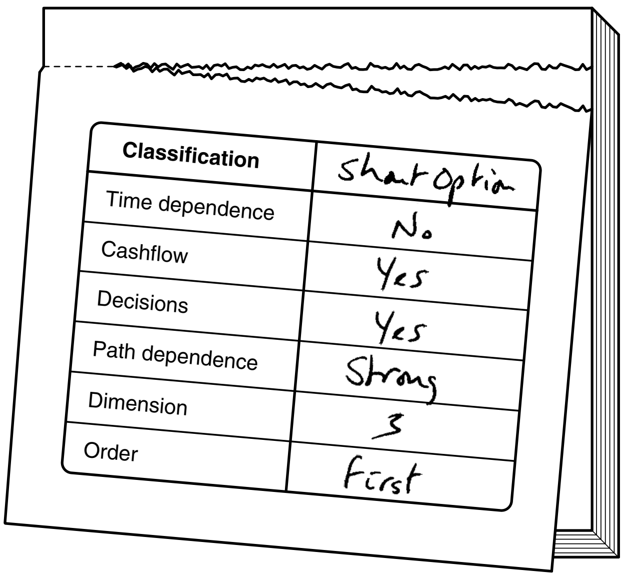
The function $V_b(S, X, t)$ must satisfy the constraint

$$V_b(S, X, t) \geq V_a(S, \max(S, X), t) - R(S, X),$$

with gradient continuity. Here $R(S, X)$ is the amount of money that must be paid on shouting. This represents the optimization of the shouting policy; when the two sides of this expression are equal it is optimal to shout.

This problem must be solved numerically; depending on the form of R there may be a similarity reduction to two dimensions. The option value is then $V(S_0, X_0, t_0)$ where the subscripts denote the initial values of the variables.

The definition of this simple shout option can be easily extended to allow for other rules about how the strike is reset, what the payment is on shouting, and to allow for multiple shouts.



Classification option table for Shout Option.

28.4 CAPPED LOOKBACKS AND ASIANS

In **capped lookbacks** and **capped Asians** there is some limit or guarantee placed on the size of the maximum, minimum or average. A typical example of a capped Asian would have the path-dependent quantity being the average of the lesser of the underlying asset and some other prescribed level. This is represented by

$$A = \frac{I}{t}$$

with

$$I = \int_0^t \min(S, S_u)\, d\tau.$$

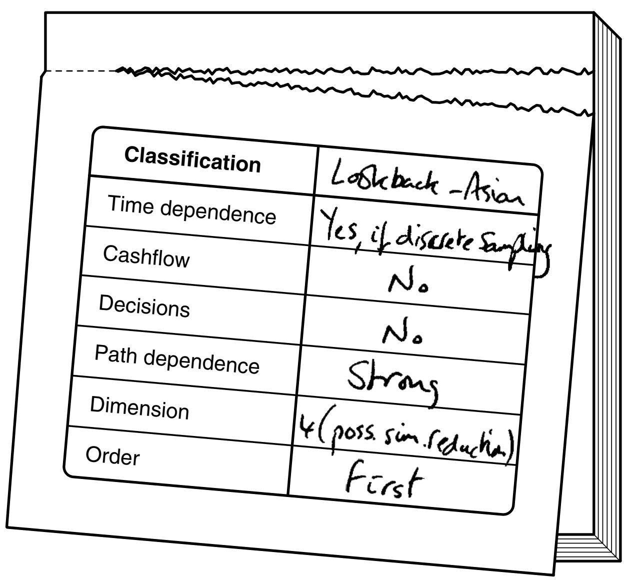
The stochastic differential equation for I from which follows the governing partial differential equation is

$$dI = \min(S, S_u)\,dt.$$

28.5 COMBINING PATH-DEPENDENT QUANTITIES: THE LOOKBACK-ASIAN ETC.

We have seen in Chapters 24–26 how to value options whose payoff, and therefore value, depends on various path-dependent quantities. There is no reason why we cannot price a contract that depends on more than one path-dependent quantity. Often, all that this requires is the use of one state-variable for each quantity.

As an example, let us consider the pricing of an option that we could call a **lookback-Asian**. By this, we mean a contract that depends on both a maximum (or minimum) and an average. But what do we mean by the 'maximum,' is it the realized maximum of the underlying asset or, perhaps, the maximum of the average? Clearly, there are a great many possible meanings for such a contract. We consider three of them here, although the reader can doubtless think of many more (and should).



Classification option table for Lookback-Asian.

28.5.1 The Maximum of the Asset and the Average of the Asset

The simplest example, and the one closest to the problems we have encountered so far, is that of a contract depending on both the realized maximum of the asset and the realized average of the asset. Suppose that the average is arithmetic and that both path-dependent quantities are sampled discretely, and on the same dates. These assumptions can easily be generalized.

First of all, we observe that the option value is a function of *four* variables, S, t, M, the maximum, and A, the average. The variables M and A are measured discretely according to the definitions

$$M_i = \max(S(t_1), S(t_2), ..., S(t_i))$$

and

$$A_i = \frac{1}{i}\sum_{j=1}^{i} S(t_j).$$

Now we need to find the jump condition across a sampling date. This follows directly from the updating rule across sampling dates. The updating rule across a sampling date for the maximum is, as we have seen,

$$M_i = \max(M_{i-1}, S(t_i)).$$

The updating rule for the average is

$$A_i = \frac{(i-1)A_{i-1} + S(t_i)}{i}.$$

Thus the jump condition across a sampling date for this type of lookback-Asian is therefore given by

$$V(S, M, A, t_i^-) = V\left(S, \max(M, S), \frac{(i-1)A + S}{i}, t_i^+\right).$$

This together with the Black–Scholes equation and a suitable final condition, is the full specification of this lookback-Asian.

28.5.2 The Average of the Asset and the Maximum of the Average

In this contract, the payoff depends on the average of the underlying asset and on the maximum *of that average*. We still have an option value that is a function of the four variables but now the updating rules are

$$A_i = \frac{(i-1)A_{i-1} + S(t_i)}{i},$$

and

$$M_i = \max(M_{i-1}, A_i).$$

Observe how we have taken the average first and then used that new average in the definition of the maximum. Again, there is plenty of scope for generalization.

Thus the jump condition across a sampling date for this type of lookback-Asian is given by

$$V(S, M, A, t_i^-) = V\left(S, \max\left(M, \frac{(i-1)A + S}{i}\right), \frac{(i-1)A + S}{i}, t_i^+\right).$$

28.5.3 The Maximum of the Asset and the Average of the Maximum

In this final example, we just swap the role of maximum and average in the previous case. We take the maximum of the asset first and then the average of that maximum, thus we have the following updating rules

$$M_i = \max(M_{i-1}, S(t_i))$$

and

$$A_i = \frac{(i-1)A_{i-1} + M_i}{i},$$

and consequently we have the jump condition

$$V(S, M, A, t_i^-) = V\left(S, \max(M, S), \frac{(i-1)A + \max(M, S)}{i}, t_i^+\right).$$

Finally, let us comment that for many of the obvious and natural payoffs, there is a similarity reduction, so that we need only solve a three-dimensional problem. I leave this issue to the reader to follow up.

28.6 THE VOLATILITY OPTION

A particularly interesting path-dependent quantity that can be the payoff for an exotic option is the realized historical volatility. By this, I mean a statistical quantity such as

$$\sqrt{\frac{1}{\delta t}\frac{1}{M-1}\sum_{j=1}^{M}\left(\log(S(t_j))/S(t_{j-1})\right)^2},$$

where δt is the time interval between samples of the asset price. (Note that I have not taken off the drift of the asset. This becomes increasingly less significant as we let δt tend to zero.) The data points used in this expression are shown in Figure 28.1 and a term sheet is shown in Figure 28.2.

The reader might ask: 'Won't this quantity simply be the volatility that we put into the model?' If so, what is the point of having a model for the historical volatility at all? The answer is that either we do not take δt sufficiently small for the above quantity to be necessarily close to the input volatility, or we assume a more complicated model for the volatility (such as stochastic or uncertain volatility, see Chapters 51 and 52, or assume an implied volatility surface, see Chapter 50).

Let us begin by valuing a contract with the above payoff in a Black–Scholes, constant volatility world, and then briefly discuss improvements.

The trick is to introduce *two* path-dependent quantities, the running volatility and the last sampled asset price:

$$I_i = \sqrt{\frac{1}{\delta t(i-1)}\sum_{j=1}^{i}\left(\log(S(t_j))/S(t_{j-1})\right)^2};$$

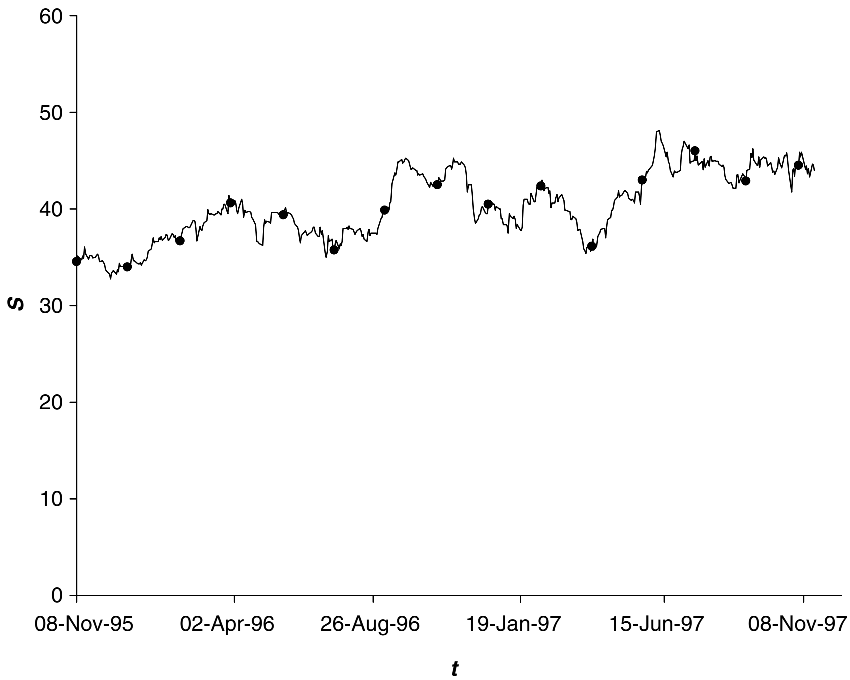


Figure 28.1 A schematic representation of the calculation of the historical volatility.

$$\mathcal{S}_i = S(t_{i-1}).$$

The option value is a function of four variables: $V(S, \mathcal{S}, I, t)$. The updating rules at time t_i for the two path-dependent quantities are

$$\mathcal{S}_i = S(t_{i-1})$$

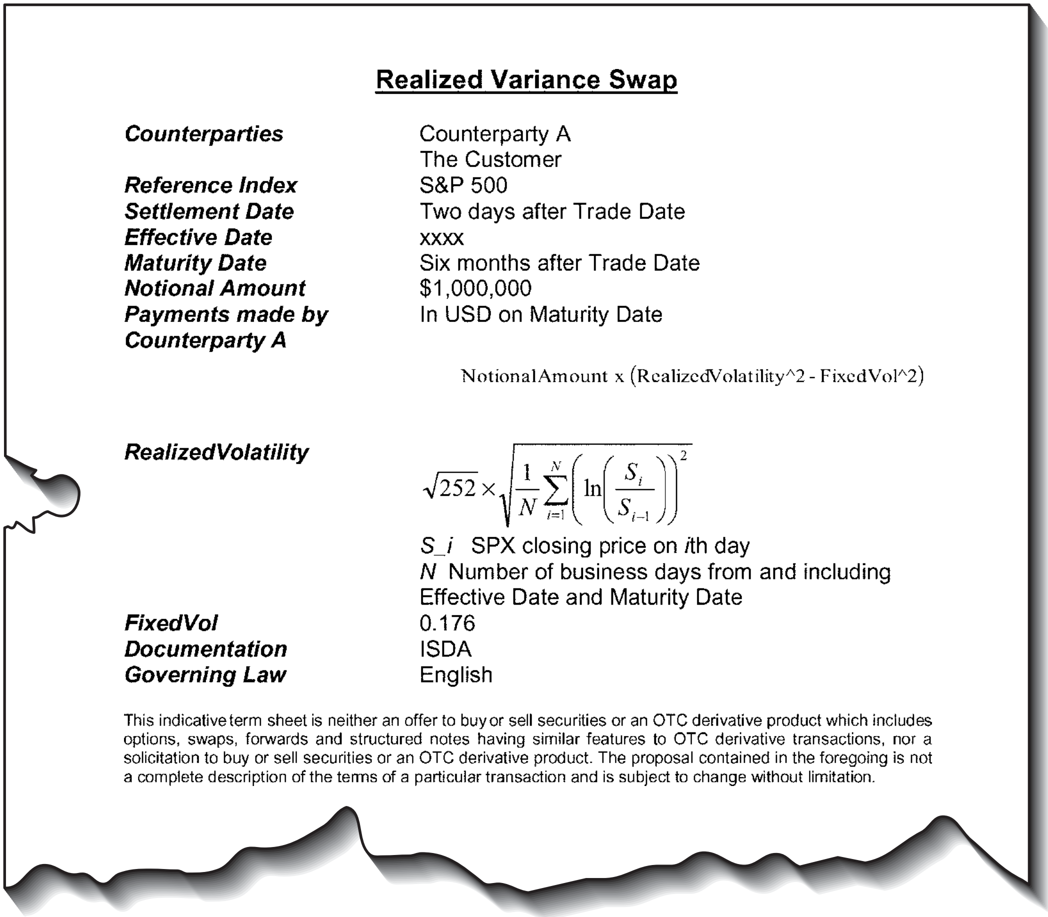
and

$$\begin{aligned} I_i &= \sqrt{\frac{1}{\delta t\,(i-1)} \sum_{j=1}^{i} \left(\log(S(t_j))/S(t_{j-1})\right)^2} \\ &= \sqrt{\frac{i-2}{i-1} I_{i-1}^2 + \frac{1}{\delta t (i-1)} \left(\log(S(t_i)) - \log(\mathcal{S}_i)\right)^2}. \end{aligned}$$

We can see from the second updating rule why we had to keep track of the old sampled asset price: it is used in the updating rule for the running volatility.

The jump condition across a sampling date is therefore

$$V(S, \mathcal{S}, I, t_i^-) = V\left(S, S, \sqrt{\frac{i-2}{i-1} I^2 + \frac{1}{\delta t (i-1)} \left(\log(S) - \log(\mathcal{S})\right)^2}, t_i^+\right).$$



Realized Variance Swap

Counterparties	Counterparty A The Customer
Reference Index	S&P 500
Settlement Date	Two days after Trade Date
Effective Date	xxxx
Maturity Date	Six months after Trade Date
Notional Amount	$1,000,000
Payments made by Counterparty A	In USD on Maturity Date
	NotionalAmount x (RealizedVolatility^2 - FixedVol^2)
RealizedVolatility	$\sqrt{252} \times \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\ln\left(\frac{S_i}{S_{i-1}}\right)\right)^2}$ *S_i* SPX closing price on *i*th day *N* Number of business days from and including Effective Date and Maturity Date
FixedVol	0.176
Documentation	ISDA
Governing Law	English

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

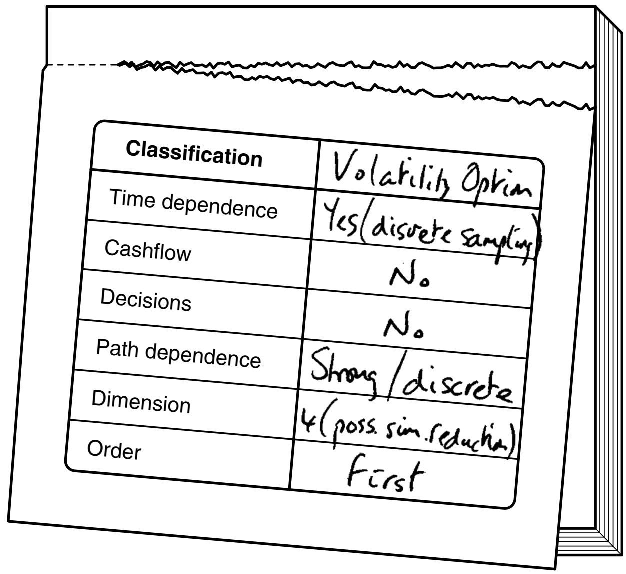
Figure 28.2 Term sheet for a typical variance swap.

If the option pays off the realized volatility at expiry, T, then

$$V(S, \mathcal{S}, I, T) = I.$$

The dimensionality of this problem can be reduced to three by the use of the similarity variable $S/\mathcal{S}$.

If we do not believe in constant volatility then we could introduce a stochastic volatility model. This does not change the specification of our model in any way other than to introduce a new variable σ which satisfies some stochastic differential equation. The problem to solve is then in five dimensions, becoming four with the use of a similarity reduction, with the same path-dependent quantities and the same updating rules and jump condition. The high (four) dimensionality makes this problem computationally intensive and it may well be a candidate for a Monte Carlo simulation (see Chapter 80).



Classification option table for Volatility Option.

28.6.1 The Continuous-time Limit

We can get to the continuous-sampling limit for this problem by letting $\delta t \to 0$. The analysis is rather messy and is more easily derived directly, as we'll now see.

For the problem to make sense we must allow the volatility to be non-deterministic. For example, we could model it via a stochastic differential equation. However, the simplest model that is internally consistent is to just allow the volatility to be a function of the asset price and time $\sigma(S, t)$. Introduce the new state variable I defined by

$$I = \sqrt{\frac{1}{t}\int_0^t \sigma(S, t)^2 \, dt}.$$

This is a continuous-time version of the above discrete definition for I since

$$\left(\log(S(t_j)/S(t_{j-1}))\right)^2 \approx \sigma(S_j)^2 \, dX^2.$$

From this we have

$$dI = \frac{(\sigma(S, t)^2 - I^2)}{2tI} \, dt$$

from which we get

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma(S, t)^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \frac{(\sigma(S, t)^2 - I^2)}{2tI}\frac{\partial V}{\partial I} = 0.$$

The option value is a function of three variables $V(S, I, t)$ and we are on familiar territory.

28.6.2 Hedging Variance Swaps with Vanilla Options

It is possible to construct a portfolio that has a constant vega, regardless of the position of the underlying. Conceivably, such a portfolio may also act as a hedge against variance swaps. This requires holding a continuous distribution of vanilla options. Let's see how this may be achieved.

What is the value of a portfolio made up of $f(E)$ call options with strike E and the relevant expiry? If this is a continuous distribution of options, which is what we want (although this is impossible in practice), the answer is

$$\int_0^\infty f(E)\, V(S, E; \sigma)\, dE.$$

V is just the Black–Scholes call formula, and σ is the implied volatility.

The vega for an individual option is

$$S\sqrt{T-t}\; e^{-D(T-t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}$$

with

$$d_1 = \frac{\log(S/E) + (r - D + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

Let's introduce $x = S/E$, so that the vega can be written in terms of

$$d_1(x) = \frac{\log(x) + (r - D + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

The portfolio vega is then

$$\int_0^\infty f(E) S\sqrt{T-t}\; e^{-D(T-t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1(S/E)^2}{2}}\, dE.$$

Changing the integration variable $E = \xi S$ gives the following expression for the portfolio vega

$$\int_0^\infty f(\xi S) S^2 \sqrt{T-t} e^{-D(T-t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1(1/\xi)^2}{2}}\, d\xi.$$

The details of this expression don't matter, all we need to know is that it is of the form

$$\int_0^\infty f(\xi S) S^2 \ldots\, d\xi$$

where the $\ldots$ are independent of S. To make this, and therefore the vega, independent of the underlying requires

$$\frac{\partial}{\partial S}\left(f(\xi S) S^2\right) = 0.$$

In other words,

$$\xi S^2 f'(\xi S) + 2S f(\xi S) = S(E f'(E) + 2 f(E)) = 0$$

where $'$ means differentiate with respect to the argument. The solution of this is just

$$f(E) = \frac{k}{E^2}$$

for an arbitrary constant k.

This is a very famous result that variance swaps can be hedged with vanilla options, using the 'one over strike squared' rule.

28.7 CORRELATION SWAP

In Figure 28.3 is a term sheet for a multi-asset contract, similar to a volatility or variance swap except that it pays off the realized correlation. Carr & Madan (1999) show how to replicate the payoff for such contracts by statically hedging with vanilla contracts. The idea is similar to that for variance swaps described above.

28.7.1 Dispersion Trading

A popular, and sometimes profitable, trading strategy, is **dispersion trading**. This is the exploitation of inefficiencies between the pricing of baskets of options on individual stocks and the option on the basket of stocks, index options, for example. This is in essence a play on correlation. In a world of constant and known parameters we have the relationship between the volatility of an index and the volatilities and correlations of its constituents:

$$\sigma_I^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} W_i W_j \rho_{ij} \sigma_i \sigma_j,$$

where the Ws are the weights of each component.[1]

Being long dispersion means that you buy the individual options and sell the index options. You are hoping that every now and then you'll get some nice big, stock specific moves which will make you money on your long volatility single stock positions without affecting the index as a whole.

28.8 LADDERS

The **ladder option** is a lookback option that is discretely sampled, but this time discretely sampled in asset price rather than time. Thus the option receives a payoff that is a function of the highest asset price achieved out of a given set. For example, the ladder is set at multiples of \$5: ... , 50, 55, 60, 65, ... If during the life of the contract the asset reached a maximum of 58, then the maximum registered would be 55. Such an option would clearly be cheaper than the continuous version.

This contract can be decomposed into a series of barrier-type options triggered at each of the rungs of the ladder. Alternatively, using the framework of Chapter 26, we simply have a payoff that is a step function of the maximum, M (see Figure 28.4).

[1] This is not quite correct since the sum of lognormal assets is not lognormal. But it is close.

1-year Nikkei225/S&P 500 Correlation Swap

Trade Date	XXXX
Expiration Date	First Index Date + 249 Days
Correlation Buyer	X
Correlation Seller	Counterparty
Strike	41%
Notional Amount	USD 2,000,000
Underlying Indices	Nikkei 225 Index (NKY) S&P 500 Index (SPX)
Settlement	Cash Settlement is Applicable
Cash Settlement Amount	Notional Amount x [Realized Correlation – Strike] x 100 If positive the Buyer will receive this amount. If negative then the Seller will receive the absolute amount

Realized Correlation

The realized 250-day correlation between the 5-day log Returns of the two Underlying Indices calculated as follows. At the Expiration Date, the closing levels of the previous 250 days when both Underlying Indices traded are considered. The 5-day log Returns for the two Underlying Indices are calculated according to the formula

$\mathrm{Log_e}$[Index Closing Level @ day(i)] –
$\mathrm{Log_e}$[Index Closing Level @ day(i-5)]

Where i=6 to 250 and i=250 corresponds to Expiration Date.

The Realized Correlation is then

$$\frac{245\left[\sum_{6}^{250} X(i)Y(i)\right]-\left[\sum_{6}^{250} X(i)\right]\left[\sum_{6}^{250} Y(i)\right]}{\left[\sqrt{245\sum_{6}^{250} X(i)^2-\left(\sum_{6}^{250} X(i)\right)^2}\right]\left[\sqrt{245\sum_{6}^{250} Y(i)^2-\left(\sum_{6}^{250} Y(i)\right)^2}\right]}$$

Where X(i) and Y(i) are the 5-day log Returns for the two Underlying Indices.

Documentation ISDA Master & Confirmation using the 1996 Equity Derivatives Definitions

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 28.3 Term sheet for a correlation swap.

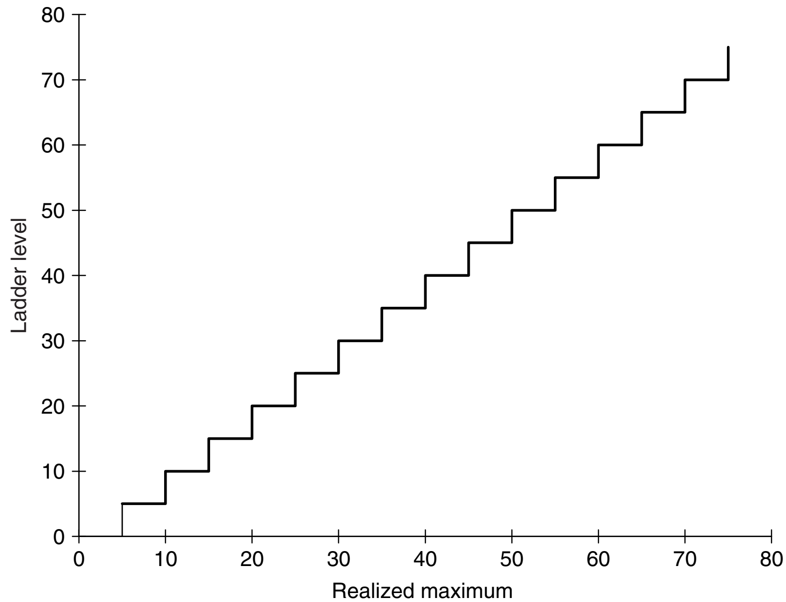


Figure 28.4 The payoff for a ladder option as a function of the realized maximum.

28.9 PARISIAN OPTIONS

Parisian options are barrier options for which the barrier feature (knock-in or knock-out) is only triggered after the underlying has spent a certain prescribed time beyond the barrier. The effect of this more rigorous triggering criterion is to 'smooth' the option value (and delta and gamma) near the barrier to make hedging somewhat easier. It also makes manipulation of the triggering, by manipulation of the underlying asset, much harder. In the classical Parisian contract the 'clock' measuring the time outside the barrier is reset when the asset returns to within the barrier. In the **Parasian** contract the clock is not reset. We only consider the former here, the latter is a simple modification.

In Figure 28.5 is shown a representation of a Parisian contract. The bottom curve is the (scaled) time that the stock price has been above the barrier level. Once it has reached ten days (here scaled to one) the barrier is triggered.

Parisian options are clearly strongly path-dependent, but in a way that is perfectly manageable in the differential equation framework. We do not need all the details of the path taken; the only extra information we need to know is the value of the variable τ, defined as the length of time the asset has been beyond the barrier:

$$\tau = t - \sup\left\{t' \leq t | S(t') \leq S_u\right\}$$

for an up barrier at S_u, and there is a similar expression for a down barrier. The stochastic differential equation for τ is given by

$$d\tau = \begin{cases} dt & S > S_u \\ -\tau^- & S = S_u \\ 0 & S < S_u \end{cases}$$

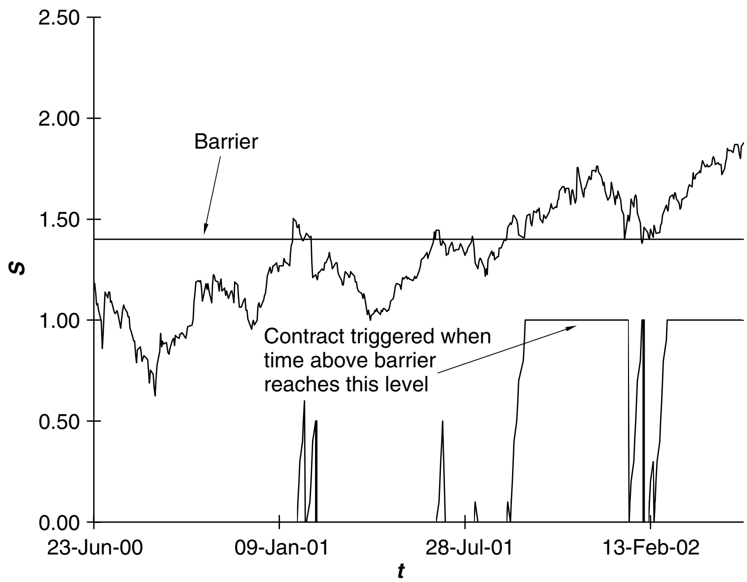


Figure 28.5 Representation of a Parisian contract.

where τ^- is the value of τ before it jumps to zero on resetting. We use this equation to derive the partial differential equation for the option value. Notice how, outside the barrier, real time t and the trigger time τ are increasing at the same rate. The barrier is triggered when τ reaches the value ω.

We must solve for V as a function of three variables S, t and τ, $V(S, t, \tau)$. We solve for V in two regions: inside the barrier and outside. Inside the barrier the clock is switched off and we solve

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

In this region there is no τ dependence (this is not true for Parasians). Outside the barrier the clock is ticking and we have

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial \tau} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

At $S = S_u$, where τ is reset, we must impose continuity of the option value:

$$V(S_u, t, \tau) = V(S_u, t, 0).$$

Now we come to the exact specification of the payoff in the event of triggering (or the event of not-triggering). If the barrier has not been triggered by expiry T then the option has the payoff $F(S, \tau)$. If the barrier has been triggered before expiry the option pays off $G(S)$ at expiry. For example, an up-and-in Parisian put would have $F = 0$ and $G = \max(E - S, 0)$. An up-and-out call would have $F = \max(S - E, 0)$ and $G = 0$. In this framework ins and outs are treated the same. The boundary conditions are applied as follows:

$$V(S, T, \omega) = G(S)$$

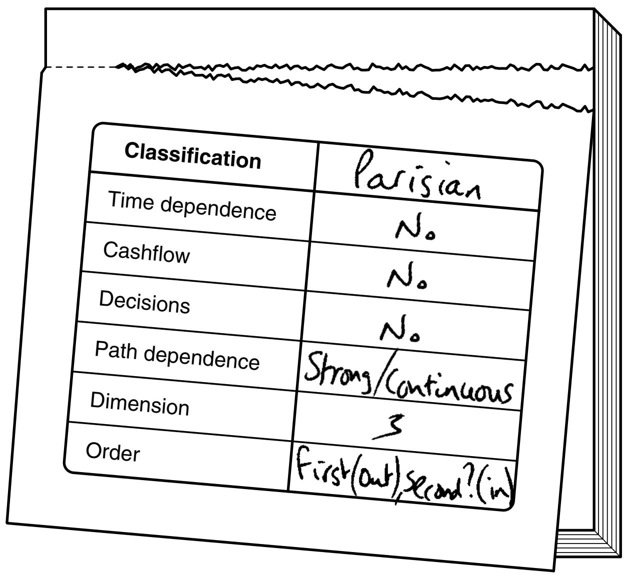
and

$$V(S, T, \tau) = F(S, \tau).$$

American-style Parisians (Henry Miller options?) have the additional constraint that

$$V(S, t, \tau) \geq A(S, t, \tau),$$

with continuity of the delta, where the function A defines the payoff in the event of early exercise.



Classification option table for Parisian Option.

28.9.1 Examples

First consider the case of a Parisian, European, down-and-in put on an asset with no dividends, an expiration time of $T = 0.25$ years, a volatility of $\sigma = 0.2$, an interest rate of $r = 0.08$, strike $E = 10$, barrier $S = 8$, and barrier trigger time $\omega = 0.05$. Figure 28.6 depicts the option value V versus price S and time t. As expected, the option is worthless for values of time less than the barrier time ω. However for time greater than the barrier time the function appears quite smooth. This is validated by examining the hedge value versus price and time, as in Figure 28.7. The diffusion, prior to the barrier time, acts to smooth the data such that the hedge ratio remains reasonably manageable, compared to a traditional knock-out barrier.

Next consider the more sophisticated example of a Parisian, American, up-and-out call with dividend rate $D = 0.04$, an expiration time of $T = 0.25$ years, a volatility of $\sigma = 0.2$, an interest rate of $r = 0.08$, strike $E = 8$, barrier $S = 10$, and barrier time $\omega = 0.05$. The resulting plots of option value V and hedge ratio of this more complicated example are shown in Figures 28.8 and 28.9, respectively.

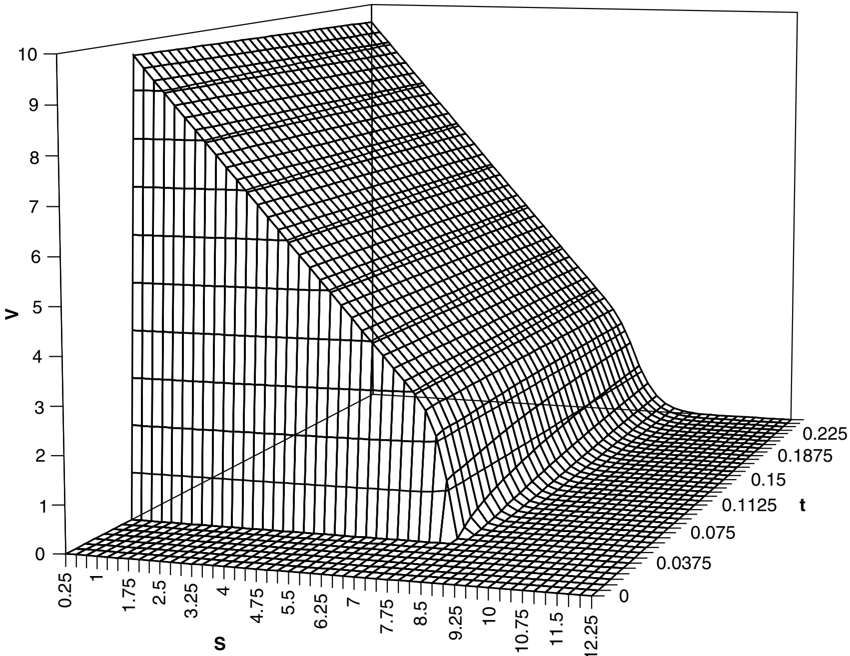


Figure 28.6 Euro Parisian value, see text for details.

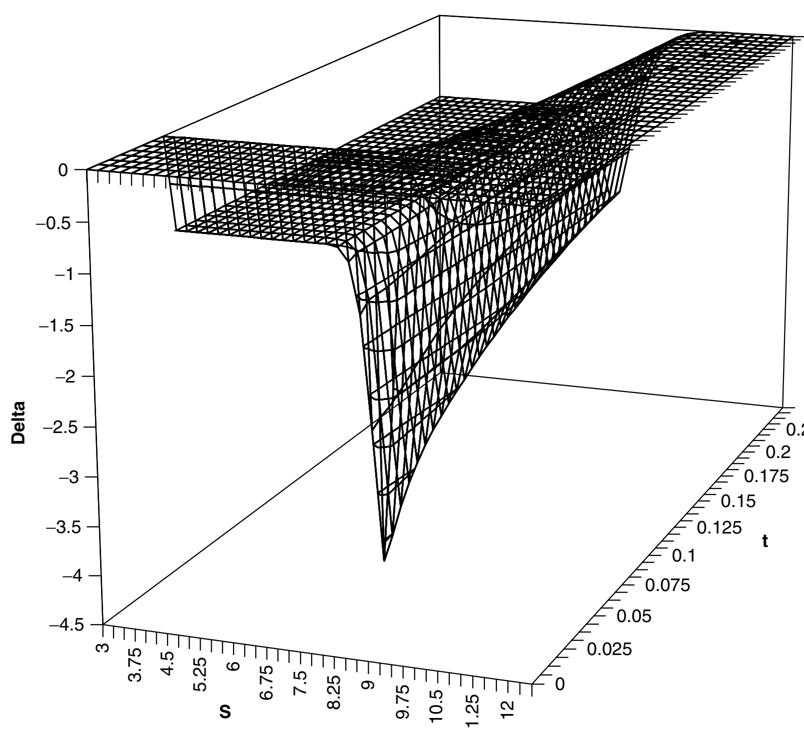


Figure 28.7 Euro Parisian delta, see text for details.

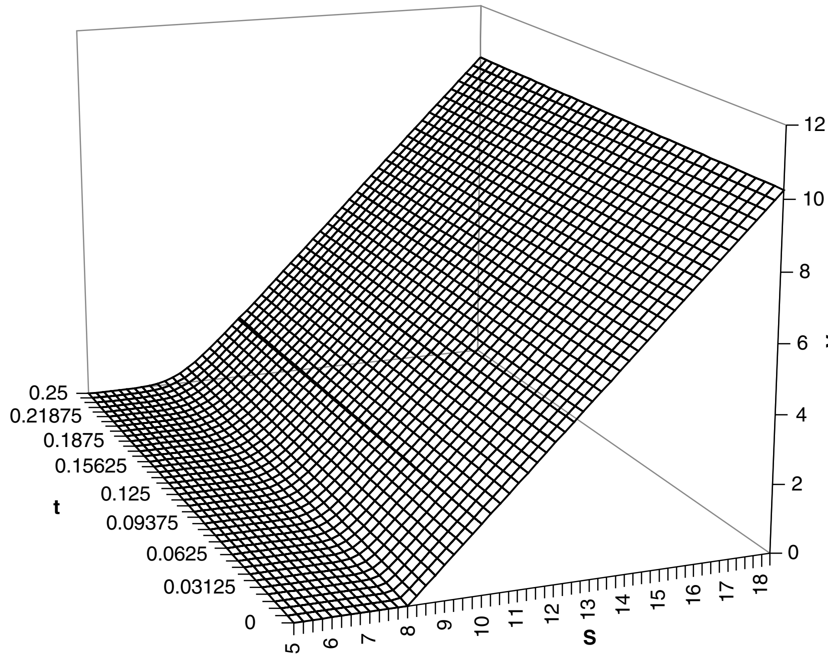


Figure 28.8 American Parisian value, see text for details.

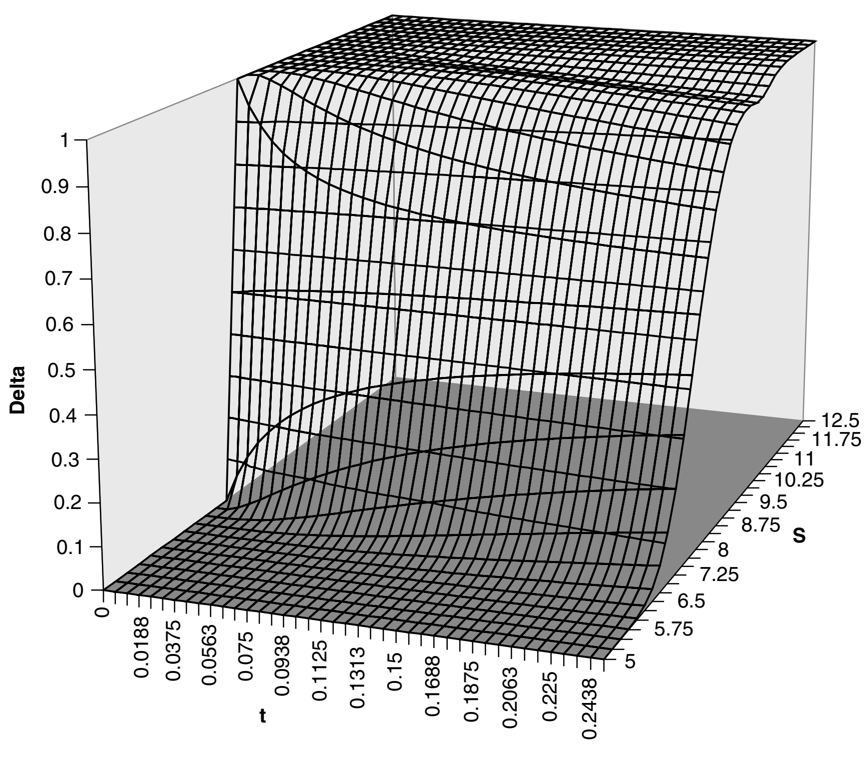


Figure 28.9 American Parisian delta, see text for details.

28.10 YET MORE EXOTICS

Balloon option

The quantity of option bought will increase if certain conditions are met, such as barriers being triggered.

Break/cancelable forward

A forward contract, usually FX, where the holder can terminate the contract at certain times if they so wish.

Contingent premium option

A premium, or a further premium, is payable if certain conditions are met.

Coupe option

A periodic option in which the strike gets reset to the worst of the underlying and the previous strike. Similar to a cliquet option (Chapter 56), but cheaper.

Extendible option/swap

An option/swap whose expiration can be extended by the holder.

Hawai'ian option

These are a cross between Asian and American, as you'd expect.

Himalayan option

These are multi-asset options in which the best performing stock is thrown out of the basket at specified sampling dates, leaving just one asset in at the end on which the payoff is based. There are many other, similar, **mountain range options**.

HYPER option

High Yielding Performance Enhancing Reversible options are like American options but which you can exercise over and over again. On each exercise the option flips from call to put or vice versa. These can be priced by introducing a price function when in the call state and another when in the put state. The Black–Scholes partial differential equation is solved for each of these, subject to certain optimality constraints.

28.11 SUMMARY

I hope that after reading the chapters in this part of the book the reader will feel confident to find partial differential equation formulations for many other types of derivative contract. In Part Three, on interest rates and products, it is assumed that the reader can transfer ideas from the lognormal asset world to the more complicated world of fixed income.

FURTHER READING

- See Dewynne & Wilmott (1994c) for a mixed bag of exotics and their modeling.
- Chesney *et al*. (1997) price Parisians via Laplace transforms.
- See Haber, Schönbucher & Wilmott (1997) for a more detailed description of the partial differential equation approach to pricing Parisian options. A fully-functional standalone Parisian option pricer may be downloaded from `www.wilmott.com`.
- See Whaley (1993) for a description of his implied volatility index and options on it.
- See Demeterfi *et al*. (1999) for a vast amount of information on volatility contracts.
- See `http://my.dreamwiz.com/stoneq/products` for a comprehensive list of exotics.

CHAPTER 29
equity and FX term sheets

In this Chapter...

- some exotics
- some analysis
- some code

29.1 INTRODUCTION

In this chapter we look at a few term sheets in detail. We start with a very simple, not even exotic, contract and build up. We look at the purpose of the exotics, how to price them, and in a few cases provide some code. The code will make most sense after you have read about numerical methods. As mentioned earlier, although these are real term sheets many of them are incomplete. I received them while they were works in progress and so you will often see bits missing (e.g. [ ], or X, or ? will represent those quantities that hadn't been decided upon at the time).

29.2 CONTINGENT PREMIUM PUT

Our first term sheet is a special kind of put option on the SP500 index (see Figure 29.1). Read the specifications in the term sheet carefully before looking at the explanation.

The underlying for this contract is an equity index, the SP500, so we will probably want to work with the lognormal random walk assumption. The payoff is

$$\text{\#Contracts} * \max[0, \text{S\&Pstrike} - \text{S\&Pfinal}].$$

In symbols this is just

$$q \max(E - S, 0)$$

where q defines the quantity and E the strike, and both of these are in terms of the level of the SP500 index when the contract is initiated. Let's assume that $q = 1$ and that E is given. Because of the 95% in the contract this option starts out as an out-of-the-money put.

Over-the-counter Option linked to the S&P500 Index

Option Type	European put option, with contingent premium feature
Option Seller	XXXX
Option Buyer	[dealing name to be advised]
Notional Amount	USD 20MM
Trade Date	[]
Expiration Date	[]
Underlying Index	S&P500
Settlement	Cash settlement
Cash Settlement Date	5 business days after the Expiration Date
Cash Settlement Amount	Calculated as per the following formula: #Contracts * max[0, S&Pstrike – S&Pfinal] where #Contracts = Notional Amount / S&Pinitial This is the same as a conventional put option: **S&Pstrike** will be equal **to 95% of the closing price on the Trade Date** **S&Pfinal** will be the level of the Underlying Index at the valuation time on the Expiration Date **S&Pinitial** is the level of the Underlying Index at the time of execution
Initial Premium Amount	[2%] of Notional Amount
Initial Premium Payment Date	5 business days after Trade Date
Additional Premium Amounts	[1.43%] of Notional Amount per Trigger Level
Additional Premium Payment Dates	The Additional Premium Amounts shall be due only if the Underlying Index at any time from and including the Trade Date and to and including the Expiration Date is equal to or greater than any of the Trigger Levels.
Trigger Levels	103%, 106% and 109% of **S&P500initial**
Documentation	ISDA
Governing law	New York

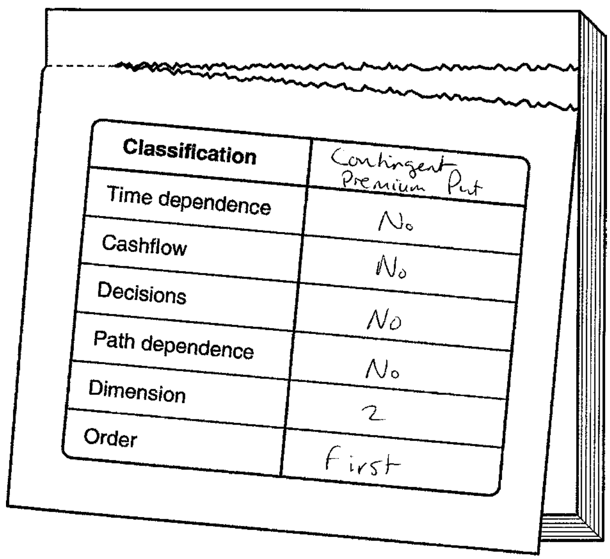
This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.1 Contingent premium put term sheet.

So far this is just a vanilla put. The novel feature is that the option holder is obliged to pay extra premiums if the SP500 index should rise to certain trigger levels. Again, these trigger levels are defined in terms of the level of the SP500 index when the contract is initiated.

This contract isn't really all that exciting because even the novel feature is easily valued. If a contract is truly exotic then it can't be decomposed into simpler instruments. This one can. You can treat each of the contingent premiums as short positions (for the option holder) in simple American binary calls, also known as one-touch options, three of them, each with a different strike. So you would value this contract as a long put and short three different one-touch calls. In the Black–Scholes world there are formulae for all of these.

A person would buy this contract because they pay less upfront for it than a vanilla put. The extra premia only have to be paid if the stock price rises above the trigger levels.



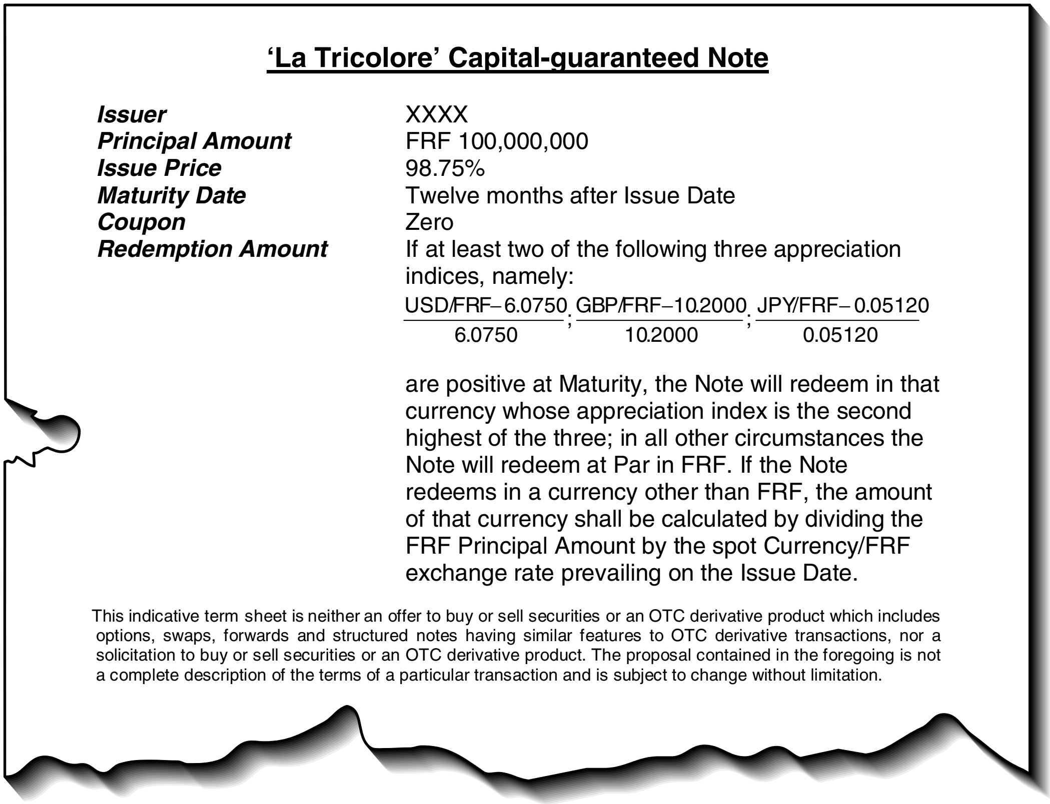
Classification option table for Contingent Premium Put.

29.3 BASKET OPTIONS

29.3.1 Simple Basket Option

In Figure 29.2 we see a quite straightforward basket option, or option on several different underlyings. The payoff is the second highest performing out of three exchange rates. There are only the three underlyings so when you do the actual calculations you can choose either Monte Carlo simulations or finite-difference methods. Since the underlyings are all exchange rates you will probably want to use lognormal random walks for each, and they will be correlated with each other. In this special case there is the closed-form formula for the value, in terms of a three-dimensional integral, so all you need to do is the integration.[1] This can be done very easily, and quickly, by the methods described in Chapter 81. Don't forget that, because the underlyings are exchange rates, you will need to incorporate an adjustment to the drift rates to allow for foreign interest. Finally, there will be three correlations to input; be wary of their stability.

[1] This wouldn't be true if early exercise was allowed.

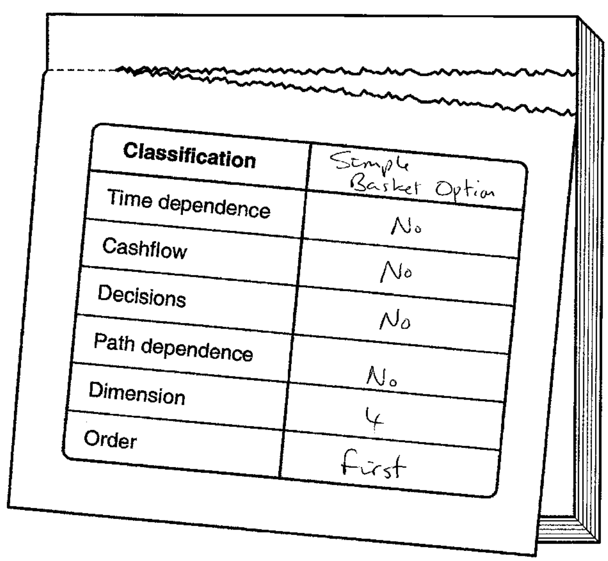


'La Tricolore' Capital-guaranteed Note

Issuer	XXXX
Principal Amount	FRF 100,000,000
Issue Price	98.75%
Maturity Date	Twelve months after Issue Date
Coupon	Zero
Redemption Amount	If at least two of the following three appreciation indices, namely: $\frac{\text{USD/FRF}-6.0750}{6.0750};\ \frac{\text{GBP/FRF}-10.2000}{10.2000};\ \frac{\text{JPY/FRF}-0.05120}{0.05120}$ are positive at Maturity, the Note will redeem in that currency whose appreciation index is the second highest of the three; in all other circumstances the Note will redeem at Par in FRF. If the Note redeems in a currency other than FRF, the amount of that currency shall be calculated by dividing the FRF Principal Amount by the spot Currency/FRF exchange rate prevailing on the Issue Date.

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.2 Basket option term sheet, three underlyings.



Classification	Simple Basket Option
Time dependence	No
Cashflow	No
Decisions	No
Path dependence	No
Dimension	4
Order	First

Classification option table for Simple Basket Option.

29.3.2 Basket Option with Averaging Over Time

The next term sheet, shown in Figure 29.3, is another basket option. However, this time there are 20 underlyings. There is also the added complexity that several different currencies are involved, since the stocks are from different countries. This makes this contract a quanto option. The payoff is also defined in terms of the average of the asset prices at specified sampling dates, giving this contract an Asian feature. Finally, there are also the 'swap' payments, but these can easily be stripped out from the valuation procedure and treated separately.

International Pharmaceutical Basket Equity Swap

Indicative terms	
Trade Date	[ ]
Initial Valuation Date	[ ]
Effective Date	[ ]
Final Valuation Date	26th September 2002
Averaging Dates	The monthly anniversaries of the Initial Valuation Date commencing 26th March 2002 and up to and including the Expiration Date
Notional Amount	US$25,000,000
Counterparty floating amounts (US$ LIBOR)	
Floating Rate Payer	[ ]
Floating Rate Index	USD-LIBOR
Designated Maturity	Three months
Spread	Minus 0.25%
Day Count Fraction	Actual/360
Floating Rate Payment Dates	Each quarterly anniversary of the Effective Date
Initial Floating Rate Index	[ ]
The Bank Fixed and Floating Amounts (Fee, Equity Option)	
Fixed Amount Payer	XXXX
Fixed Amount	1.30% of Notional Amount
Fixed Amount Payment Date	Effective Date
Basket	A basket comprising 20 stocks and constructed as described in attached Appendix
Initial Basket Level	Will be set at 100 on the Initial Valuation Date
Floating Equity Amount Payer	XXXX

Figure 29.3 Basket option term sheet, 20 underlyings.

Floating Equity Amount	Will be calculated according to the performance of the basket of stocks in the following way:

$$\text{Notional Amount} * \max\left[0, \left(\frac{\text{BASKET}_{\text{average}} - 100}{100}\right)\right]$$

where

$$\text{BASKET}_{\text{average}} = 100 * \sum_{20\,\text{stocks}} \left(\text{Weight} * \frac{P_{\text{average}}}{P_{\text{initial}}}\right)$$

And for each stock the Weight is given in the Appendix
P_initial is the local currency price of each stock on the Initial Valuation Date
P_average is the arithmetic average of the local currency price of each stock on each of the Averaging Dates

Floating Equity Amount Payment Date	Termination Date

Appendix
Each of the following stocks are equally weighted (5%):
Astra (Sweden), Glaxo Wellcome (UK), Smithkline Beecham (UK), Zeneca Group (UK), Novartis (Switzerland), Roche Holding Genus (Switzerland), Sanofi (France), Synthelabo (France), Bayer (Germany), Abbott Labs (US), Bristol Myers Squibb (US), American Home Products (US), Amgen (US), Eli Lilly (US), Medtronic (US), Merck (US), Pfizer (US), Schering-Plough (US), Sankyo (Japan), Takeda Chemical (Japan).

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.3 (*continued*)

How many dimensions is this contract? One per stock, exchange rate and then another 20 for the averages, and time, of course.

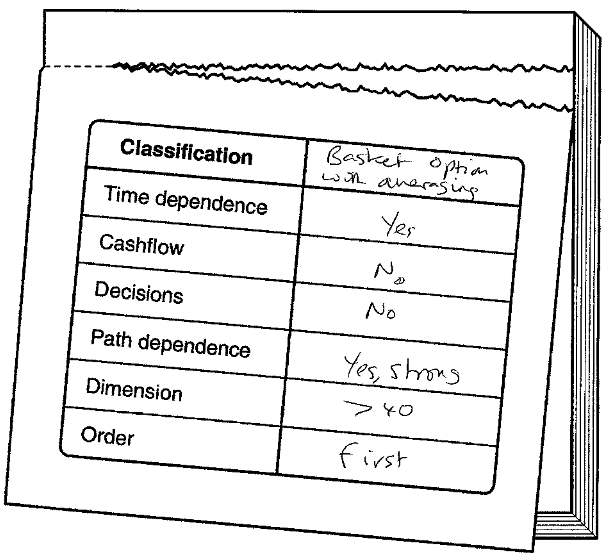
Undoubtedly you would want to price this contract by performing the risk-neutral simulations of the asset paths. You will need to do the full paths because of the averaging feature in the contract. There is no way that you would contemplate a finite-difference solution of this contract.

With so many underlyings, both stock prices and exchange rates, there will be a large number of parameters to be input into lognormal models. With N factors you will have N volatilities and $N(N-1)/2$ correlations. We know that the correlations in particular will be quite unstable. So GIGO. Maybe there is some rationalization that could be done first to decrease the number of parameters to more manageable and more stable numbers.

The term sheet specifies two sorts of averaging: pathwise and across the stocks. Averaging is a smoothing operation which in itself will make the option value less sensitive to the model. Contrast this with an option whose payoff involves the *differences* between many underlyings. With such a contract you really have to model the dissimilarities between underlyings, and this is a much harder task than modeling average behavior. To make this more concrete think of how to value options with the following two payoffs,

$$\max(S_1 + S_2 + \cdots + S_N - E, 0) \quad \text{versus} \quad \max(S_1 - E_1, S_2, E_2, \cdots, S_N - E_N, 0).$$

The latter is far, far more sensitive to each stock than the former. In fact the former is just like an option on an index. Inspired by the difference between these two payoffs we would expect that our pharmaceutical basket option is not all that bad. One simplifying approximation that we could make, in the interest of robustness, would be to have one 'asset/basket' per currency. This would dramatically reduce the number of input parameters. So, we would replace all the US pharmaceuticals in this contract with a single representative 'index.'



Classification option table for Basket Option with Averaging.

29.4 KNOCKOUT OPTIONS

29.4.1 Double Knockout

In Figure 29.4 is the term sheet for a quite basic double knockout option. The underlying is an exchange rate, so watch out for the foreign interest rate in the governing model. Without any doubt you would price this using a finite-difference method. It is very low-dimensional so you would not benefit in any way from using a simulation. Simulations would be slower and less flexible.

USD/MXN Double Knock-Out Note

Principal Amount	USD 10,000,000
Issuer	XXXX
Maturity	6 months from Trade Date
Issue Price	100%
Coupon	If the USD/MXN spot exchange rate trades above the Upper Barrier or below the Lower Barrier at any time during the term of the Note: Zero Otherwise: $400\% \times \max\left(0, \frac{8.2500-\text{FX}}{\text{FX}}\right)$ where FX is the USD/MXN spot exchange rate at Maturity
Redemption Amount	100%
Upper Barrier Level	8.2500
Lower Barrier Level	7.4500

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.4 Double knockout term sheet.

The payoff for this option is

$$\max(8.25\ S - 1, 0)$$

where I have used S as 1/FX. So 1 Mexican peso = S US dollars. The holder of this contract wants the exchange rate to remain within the range defined by the barriers, but would prefer S to be closer to 1/7.45 at expiration. We must solve

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - r_f) S \frac{\partial V}{\partial S} - rV = 0$$

where r_f is the Mexican interest rate. There will be the boundary conditions

$$V(7.45, t) = V(8.45, t) = 0$$

with the final condition

$$V(S, T) = \max(8.25\ S - 1, 0).$$

Figure 29.5 shows results for this contract with above contract specifications and volatility of 5%, domestic interest rate of 5% and foreign interest rate of 8%.

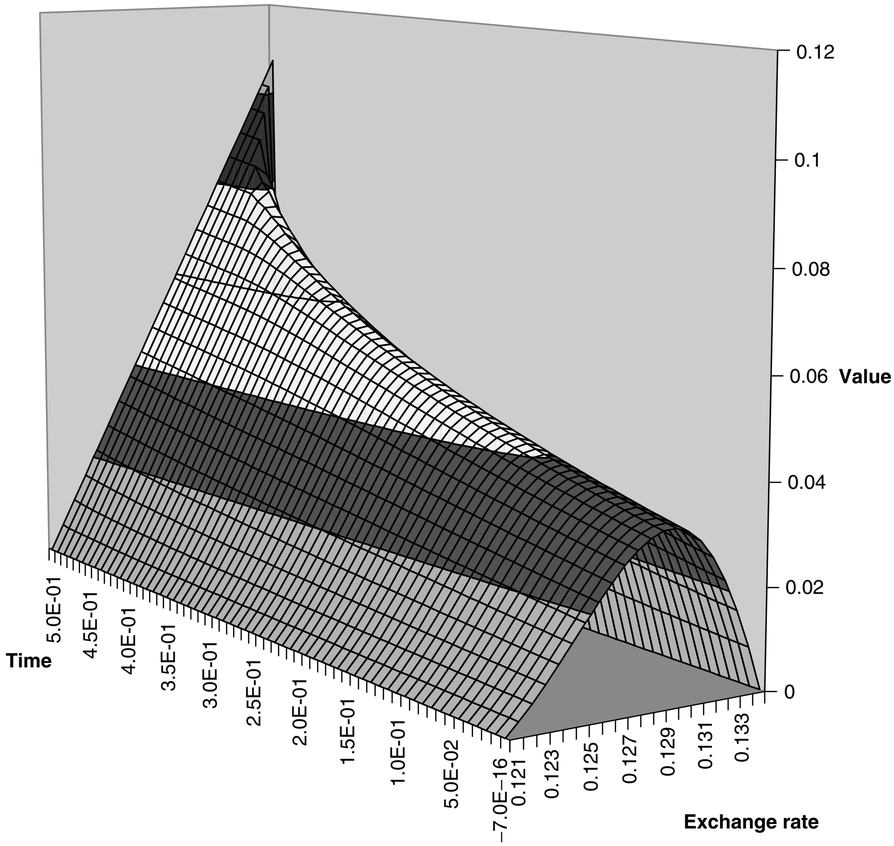
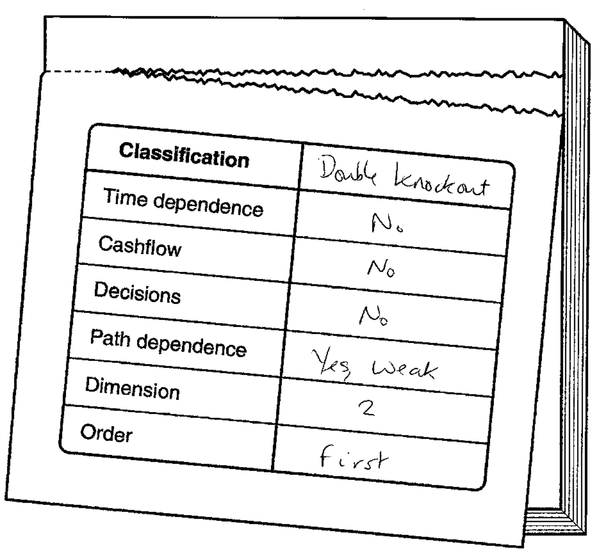


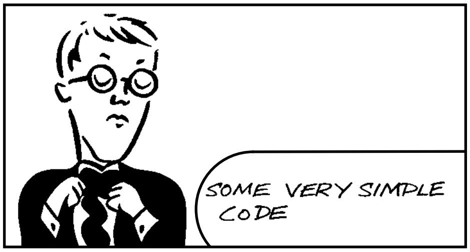
Figure 29.5 Double knockout results.



Classification	Double knockout
Time dependence	No
Cashflow	No
Decisions	No
Path dependence	Yes, weak
Dimension	2
Order	First

Classification option table for Double Knockout.

Below is some very simple explicit finite-difference code for this. However, please read the chapters on numerical methods before you implement it.



```
Function DoubleKO(Vol, Int_Rate, For_Int_Rate, _
            Lower_Barrier, Upper_Barrier, Strike, _
            Expiration, Qty, NAS)

ReDim S(0 To NAS) As Double

dS = (Upper_Barrier - Lower_Barrier) / NAS
dt = 0.9 * dS * dS / Vol / Vol / Upper_Barrier _
            / Upper_Barrier

NTS = Int(Expiration / dt) + 1
dt = Expiration / NTS

ReDim V(0 To NAS, 0 To NTS)

For i = 0 To NAS
S(i) = i * dS + Lower_Barrier
V(i, 0) = Qty * Application.Max(S(i) - Strike, 0)
Next i

For j = 1 To NTS
For i = 1 To NAS - 1
Delta = (V(i + 1, j - 1) - V(i - 1, j - 1)) / 2 / dS
Gamma = (V(i + 1, j - 1) - 2 * V(i, j - 1) _
            + V(i - 1, j - 1)) / dS / dS
Theta = -0.5 * Vol * Vol * S(i) * S(i) * Gamma - _
            (Int_Rate - For_Int_Rate) * S(i) * Delta _
            + Int_Rate * V(i, j - 1)
V(i, j) = V(i, j - 1) - dt * Theta
Next i
V(0, j) = 0
V(NAS, j) = 0
Next j

DoubleKO = V

End Function
```

29.4.2 Instalment Knockout

The next contract is our first in this chapter in which the holder has to make decisions. This is a straightforward up-and-out put option with the twist that the contract is paid for in instalments during its life rather than all upfront (see Figure 29.6). And those instalments are optional, meaning that if the holder so wishes he doesn't have to pay, but of course he then loses the option.

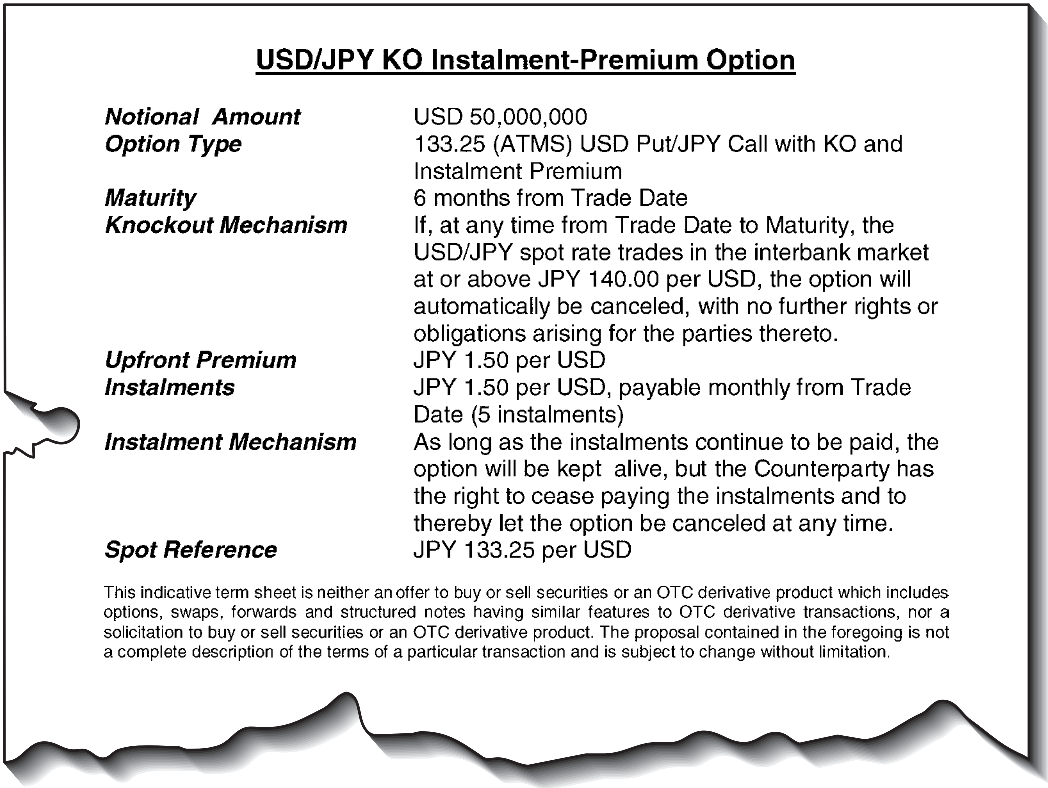
This contract is best valued by a finite-difference solution of the Black–Scholes partial differential equation. The final condition is that of a put payoff

$$V(S,T) = \max(E - S, 0)$$

and there is the usual knockout boundary condition

$$V(S_u, t) = 0,$$

where S_u is the position of the barrier.

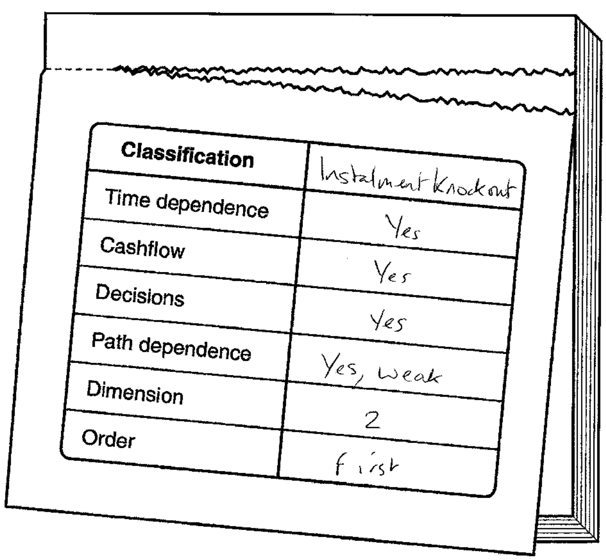


USD/JPY KO Instalment-Premium Option

Notional Amount	USD 50,000,000
Option Type	133.25 (ATMS) USD Put/JPY Call with KO and Instalment Premium
Maturity	6 months from Trade Date
Knockout Mechanism	If, at any time from Trade Date to Maturity, the USD/JPY spot rate trades in the interbank market at or above JPY 140.00 per USD, the option will automatically be canceled, with no further rights or obligations arising for the parties thereto.
Upfront Premium	JPY 1.50 per USD
Instalments	JPY 1.50 per USD, payable monthly from Trade Date (5 instalments)
Instalment Mechanism	As long as the instalments continue to be paid, the option will be kept alive, but the Counterparty has the right to cease paying the instalments and to thereby let the option be canceled at any time.
Spot Reference	JPY 133.25 per USD

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.6 Instalment knockout term sheet.



Classification	Instalment Knockout
Time dependence	Yes
Cashflow	Yes
Decisions	Yes
Path dependence	Yes, weak
Dimension	2
Order	first

Classification option table for Instalment Knockout.

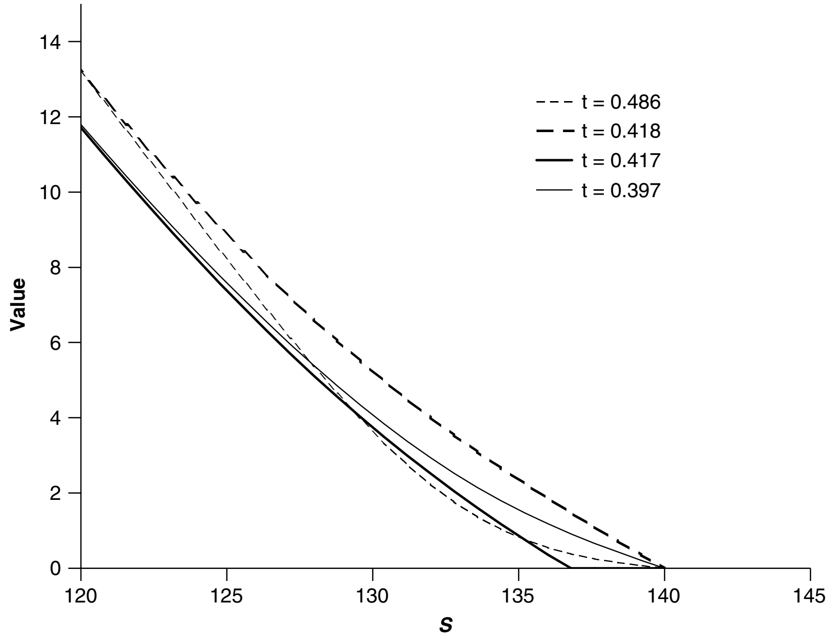


Figure 29.7 Instalment knockout evolution before and after payment.

Imagine solving this backwards in time from expiration until a time that is just *after* the last instalment is due. What you have solved so far is a simple up-and-out put option. Now think what happens as you go further backwards in time, back to just *before* the last instalment is due. We know that across this date a payment of 1.50 may be made. If that payment is made then the contract value before the payment and the contract value after the payment must differ by 1.50. Think of having a car on which you have to pay instalments; you'd rather have a car on which there are no instalments left to pay than one on which there is still one payment left. So, the contract value before the last instalment is less than the value just after:

$$V(S, t_i^-) = V(S, t_i^+) - 1.50.$$

Although, this is not quite true with our contract because the payment of each instalment is optional. It is quite possible that the contract value after the instalment has been paid is just 1.00. If that is the case you would have paid 1.50 for something worth 1.00. And that doesn't make economic sense. In that case you would prefer to throw the option away, and not pay the instalment. The general idea is that you would never pay the instalment if it means that the contract value before the payment goes negative:

$$V(S, t_i^-) = \max\left(V(S, t_i^+) - 1.50, 0\right).$$

The results of this can be seen in Figure 29.7. Look at the two bold lines in this figure. The dashed line is at time 0.418, just after the last instalment, and the solid bold line is at 0.417, just before the last instalment. See how there is a difference of 1.50 between these where both are above the axis. The solid line is 1.50 below the dashed one, except that it has been cut off at the axis. It would not be optimal to pay the 1.50 if at time 0.417 the asset were above approximately 137, it would be like paying \$150 for a car which would then only be worth \$100.

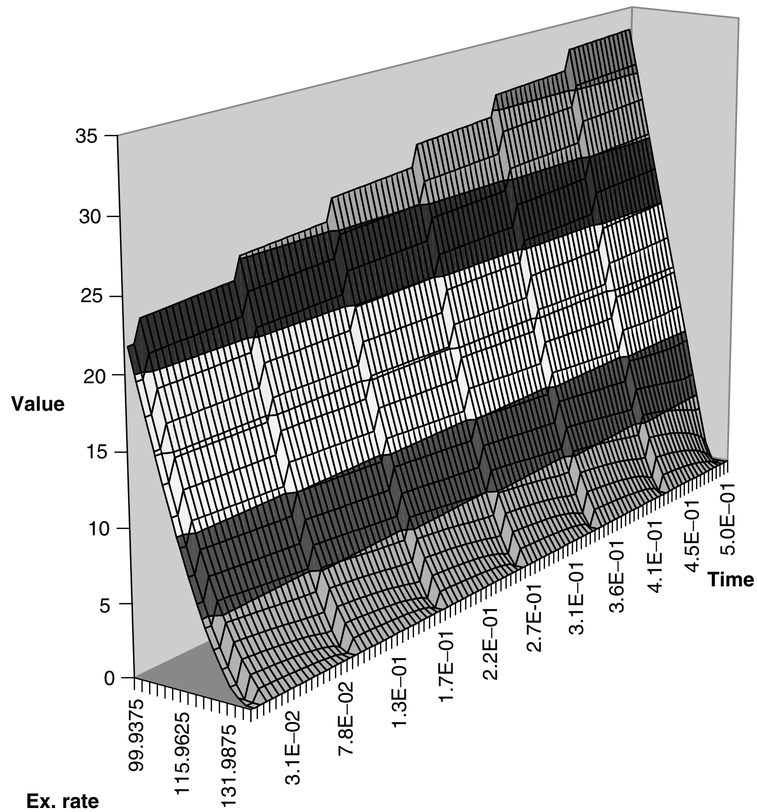
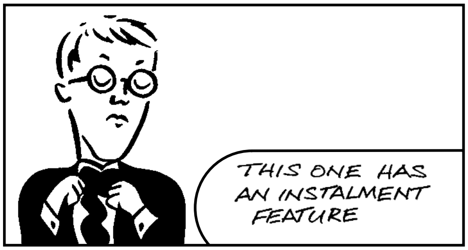


Figure 29.8 Instalment knockout solution.

You can see how this works going backwards in time in Figure 29.8. This is a three-dimensional plot of the contract value against stock price and time. You can easily see the regular jumps in contract value. And below is the finite-difference code.

```
Function InstalmentKO(Vol, Int_Rate, For_Int_Rate, _
            Upper_Barrier, Strike, Expiration, Paymnt, Freq, NAS)

ReDim S(0 To NAS) As Double

Lower_Limit = 3 * Strike / 4
dS = (Upper_Barrier - Lower_Limit) / NAS
dt = 0.9 * dS * dS / Vol / Vol / Upper_Barrier / Upper_Barrier

NTS = Int(Expiration / dt) + 1
dt = Expiration / NTS

ReDim V(0 To NAS, 0 To NTS)

For i = 0 To NAS
S(i) = i * dS + Lower_Limit
V(i, 0) = Application.Max(Strike - S(i), 0)
Next i
```



```
For j = 1 To NTS
For i = 1 To NAS - 1
Delta = (V(i + 1, j - 1) - V(i - 1, j - 1)) / 2 / dS
Gamma = (V(i + 1, j - 1) - 2 * V(i, j - 1) _
         + V(i - 1, j - 1)) / dS / dS
Theta = -0.5 * Vol * Vol * S(i) * S(i) * Gamma - _
         (Int_Rate - For_Int_Rate) * S(i) * Delta _
         + Int_Rate * V(i, j - 1)
V(i, j) = V(i, j - 1) - dt * Theta
Next i
V(0, j) = 2 * V(1, j) - V(2, j)
V(NAS, j) = 0

' Test for payment date
If Int((j + 1) * dt / Freq) - Int(j * dt / Freq) _
         > 0 Then
For i = 0 To NAS
V(i, j) = Application.Max(V(i, j) - Paymnt, 0)
Next i
End If

Next j

InstalmentKO = V

End Function
```

29.5 RANGE NOTES

The next contract is an example of a range note, see the term sheet in Figure 29.9. This contract pays off at expiration according to how long an exchange rate has been within a specified range.

Usually these contracts pay an amount *proportional* to the fraction of time the exchange rate has been within the range. A typical contract would pay

$$\frac{\text{Number of days in range}}{\text{Number of days in period}}.$$

Such a contract would be only weakly path-dependent. It could be valued in several, equivalent ways: a) Add a source term to the Black–Scholes equation representing the cash accumulation (be careful with this, you may need to do some PVing if the money isn't received until expiration); b) Treat as an integral (infinite number) of binary options (this is a rare example of being able to decompose a contract into other contracts);[2] c) Introduce a new state variable. The last of these is what we must do with the specific contract we have here.

The part of the contract saying 'subject to a minimum of zero,' while looking quite innocuous, actually makes this contract strongly path-dependent. To do it correctly we must introduce a new state variable, a clock, just like with the Parisian option.

$$\tau = \int_0^t f(S)\,dt,$$

where

$$f(S) = +1 \quad \text{for} \quad S_l \le S \le S_u, \quad \text{and} \quad -1 \quad \text{otherwise.}$$

[2] As ever, I am not usually keen on 'special cases' such as this because a small change to a contract could make the decomposition into vanillas invalid.

6 Month In-Out Range Accrual Option on MXN/USD FX Rate

Settlement Date	One week from Trade Date
Maturity Date	6 months from Trade Date
Option Premium	USD 50,000+
Option Type	In MINUS Out Range Accrual on MXN/USD FX rate
Option Payment Date	2 business days after Maturity Date
Option Payout	USD 125,000 * Index
Where Index	$\dfrac{\text{FX daily In MINUS FX daily Out}}{\text{Total Business Days}}$ (subject to a minimum of zero)
FX daily In	The number of business days Spot MXN/USD Exchange Rate is within Range
FX daily Out	The number of business days Spot MXN/USD Exchange Rate is outside Range
Range	MXN/USD 7.7200-8.1300
Spot MXN/USD Exchange Rate	Official spot exchange rate as determined by the Bank of Mexico as appearing on Reuters page "BNMX" at approximately 3:00 p.m. New York time.
Current Spot MXN/USD	7.7800

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.9 In-out range accrual term sheet.

The clock is shown for a realized path in Figure 29.10. In this example the range is defined by 90 and 110.

With this definition for τ the governing partial differential equation is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + f(S)\frac{\partial V}{\partial \tau} = 0$$

with final condition

$$V(S, T, \tau) = \max(\tau, 0).$$

29.5.1 A Really Simple Range Note

A contract that pays off

$$\frac{\text{Number of days in range}}{\text{Number of days in period}}.$$

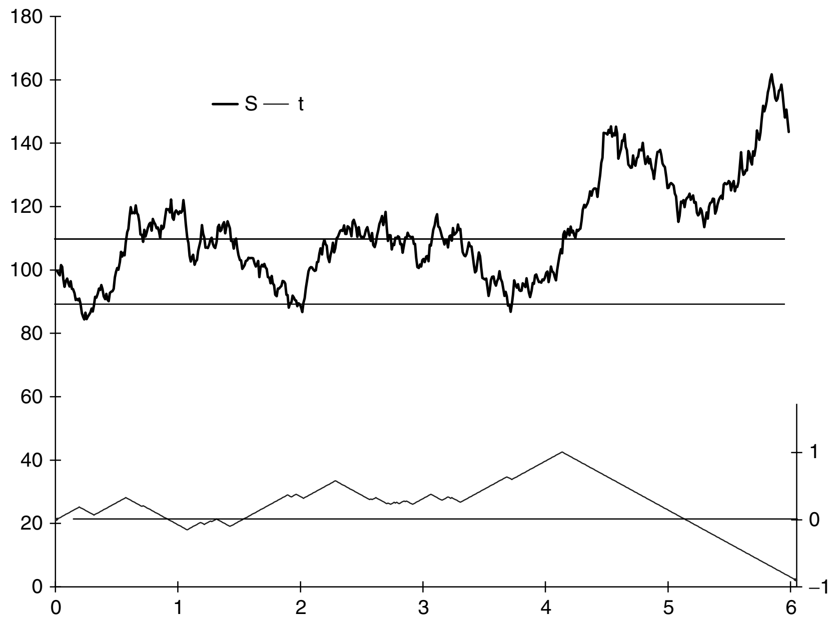
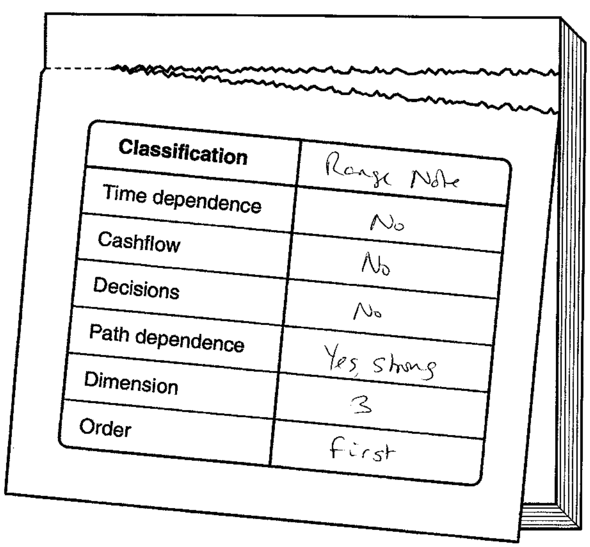


Figure 29.10 The clock in action.



Classification option table for Range Note.

without the 'subject to a minimum of zero' in the above example is much easier to price, because it is only weakly path-dependent (see code, below, and picture in Figure 29.11).[3] In this contract there is a credit of dt accumulated for every period dy that the asset is in the

[3] You can decompose this contract exactly into a series of digitals. I don't do this here just in case we want to use some of the non-linear pricing models we'll be seeing later, for which such decompositions are obviously not possible.

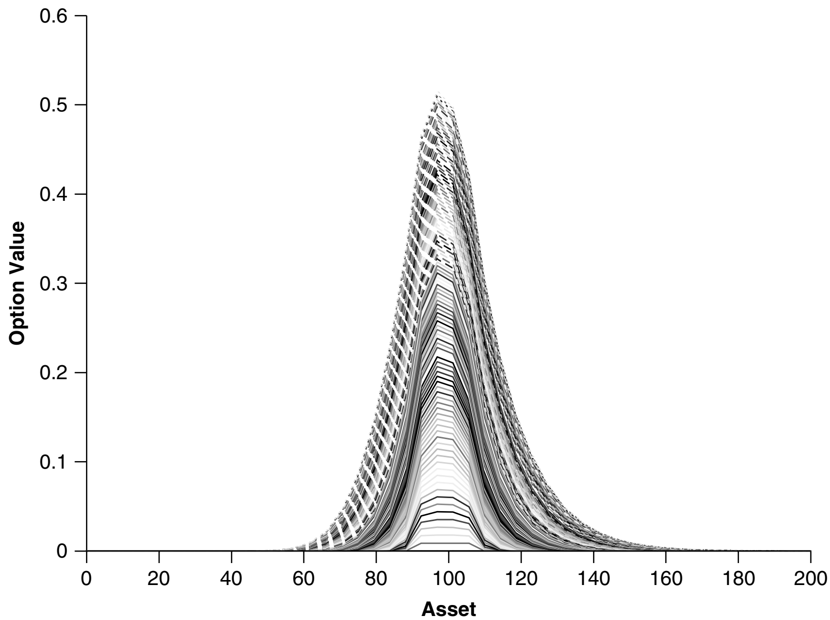
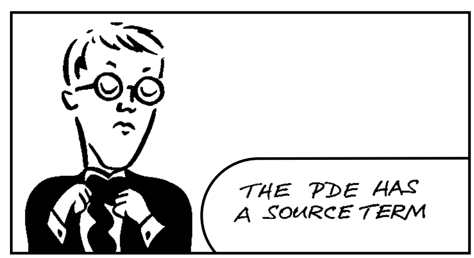


Figure 29.11 Results for a simple range note, evolution through time. Upper level is 110, lower is 90, expiration is one year, $r = 0.05$ and $\sigma = 0.2$.

range `LowerLevel` to `UpperLevel`, but this is only paid out at expiration (so you'll see a present valuing coefficient in the 'source term').

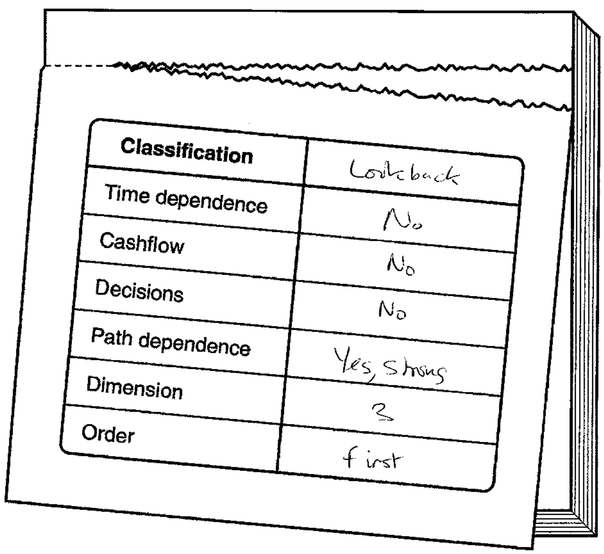



```
Function RangeNote(LowerLevel, UpperLevel, Expn, _
            Vol, Int_Rate, NAS)

dS = 2 * UpperLevel / NAS
dt = 0.9 / NAS / NAS / Vol / Vol
NTS = Int(Expn / dt) + 1
dt = Expn / NTS

ReDim S(0 To NAS) As Double
ReDim V(0 To NAS, 0 To NTS)

For i = 0 To NAS
S(i) = i * dS
V(i, 0) = 0
Next i

For j = 1 To NTS
For i = 1 To NAS - 1
Delta = (V(i + 1, j - 1) - V(i - 1, j - 1)) / 2 / dS
Gamma = (V(i + 1, j - 1) - 2 * V(i, j - 1) + V(i - 1, j - 1)) / dS / dS
Theta = -0.5 * Vol * Vol * S(i) * S(i) * Gamma - _
            Int_Rate * S(i) * Delta + Int_Rate * V(i, j - 1) _
            - F(S(i), LowerLevel, UpperLevel) * Exp(-Int_Rate * j * dt)
            ' The source term, valued at expiration
V(i, j) = V(i, j - 1) - dt * Theta
Next i
```

```
V(0, j) = V(0, j - 1) * (1 - Int_Rate * dt)
V(NAS, j) = 2 * V(NAS - 1, j) - V(NAS - 2, j)
Next j

RangeNote = V

End Function


Function F(S, LowerLevel, UpperLevel) As Double
If S > LowerLevel And S < UpperLevel Then F = 1
End Function
```

29.6 LOOKBACKS

The term sheet in Figure 29.12 is of a very vanilla lookback option, so see Chapter 26. There are just two things to note. First, the lookback is continuously sampled: this is slightly unusual. Second, there is a swap element to it, but that can easily be stripped out and treated separately.



Classification option table for Lookback.

Figure 29.13 shows a second lookback with discrete sampling.

The value of this lookback option will be a function of the stock price, S, the realized maximum, M and t, $V(S, M, t)$. It satisfies the Black–Scholes equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

with final condition representing the payoff:

$$V(S, M, T) = \max(M - S, 0).$$

If the maximum is resampled at times t_i then we have the jump condition

$$V(S, M, t_i^-) = V(S, \max(S, M), t_i^+).$$

USD/DEM Lookback Swap

Counterparties	Counterparty A The Customer
Notional Amount	USD 50 millions
Settlement Date	Two days after Trade Date
Maturity Date	Two years after Trade Date
Payments made by Customer	USD 6m LIBOR + 190 bps paid semiannually, A/360
Payments made by Counterparty A	In USD on Maturity Date $\text{Notional} * \left(\frac{FX_{max} - \text{Strike}}{FX_{maturity}} - 1 \right)$
FX_max	The highest daily official USD/DEM Fixing from Settlement Date until Maturity Date
FX_maturity	The USD/DEM Fixing on Maturity Date
Strike	1.7180
Fixing	The daily USD/DEM FX exchange rate as seen on Telerate page SAFE1 at noon, New York time
USD 6m LIBOR	The USD 6m LIBOR rate as seen on Telerate page 3750 at noon, London time, on each Fixing Date
Documentation	ISDA
Governing Law	English

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.12 Lookback term sheet.

This problem has a similarity solution (provided volatility doesn't have any S dependence) of the form

$$V(S, M, t) = M\ H(\xi, t)$$

where

$$\xi = \frac{S}{M}.$$

The governing equation is now

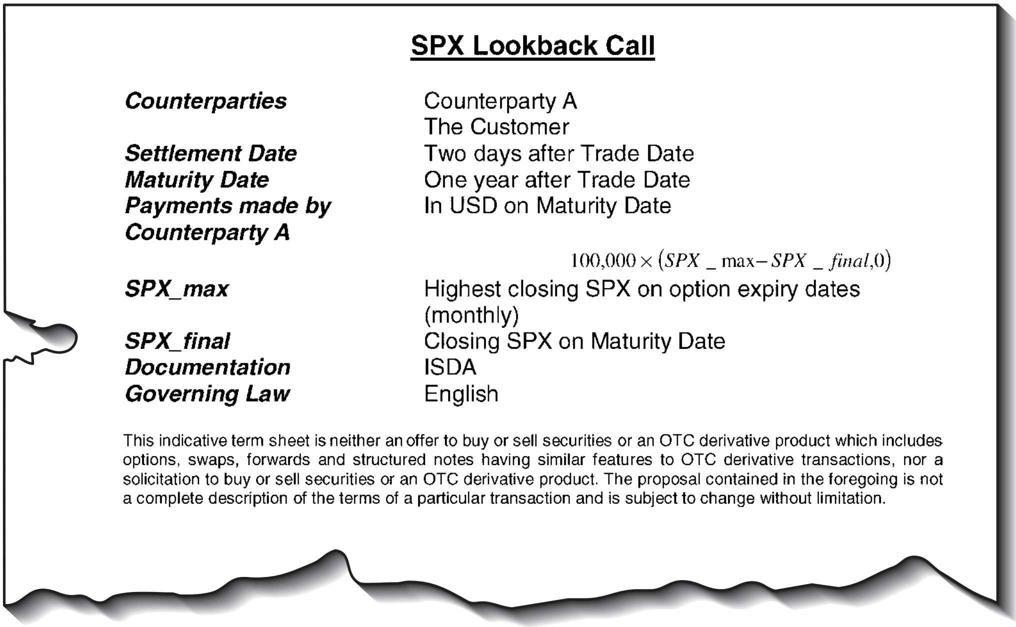
$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2 \xi^2 \frac{\partial^2 H}{\partial \xi^2} + r\xi \frac{\partial H}{\partial \xi} - rH = 0,$$

with final condition

$$H(\xi, T) = \max(1 - \xi, 0).$$

The jump condition becomes

$$H(\xi, t_i^-) = \max(\xi, 1)\ H\left(\frac{\xi}{\max(\xi, 1)}, t_i^+ \right).$$



SPX Lookback Call

Counterparties	Counterparty A The Customer
Settlement Date	Two days after Trade Date
Maturity Date	One year after Trade Date
Payments made by Counterparty A	In USD on Maturity Date $100{,}000 \times (SPX\_\max - SPX\_final, 0)$
SPX_max	Highest closing SPX on option expiry dates (monthly)
SPX_final	Closing SPX on Maturity Date
Documentation	ISDA
Governing Law	English

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.13 Lookback term sheet, discrete sampling.

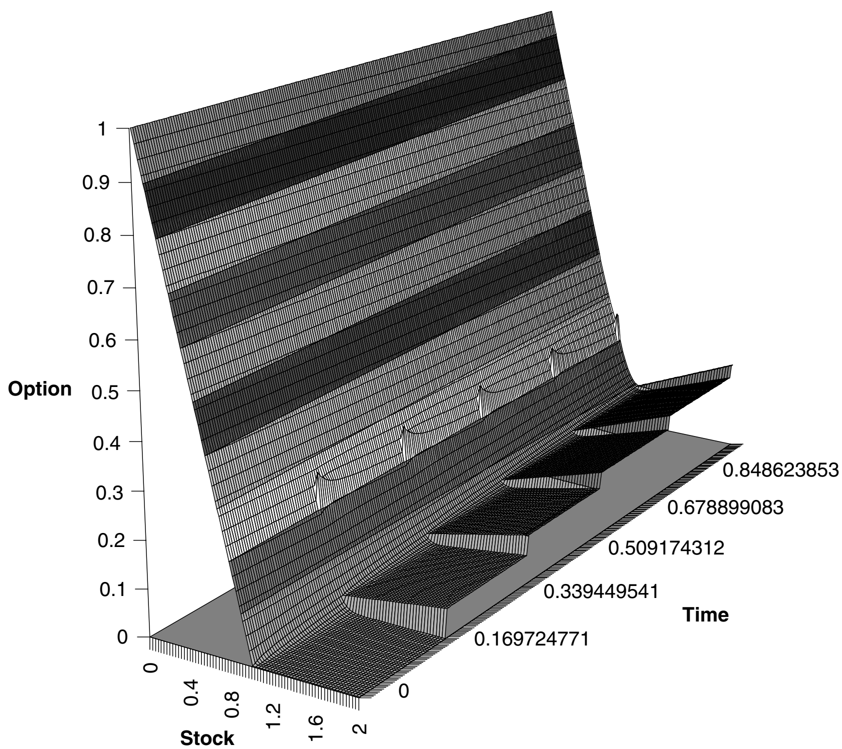
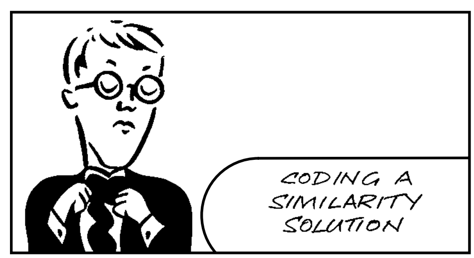


Figure 29.14 Lookback results.

Figure 29.14 shows results for this contract and the code is given below.

```
Function Lookback(Vol, IntRate, Expiration, FixPer, NXS)

dX = 2 / NXS
dt = 0.9 / NXS ^ 2 / Vol ^ 2
NTS = Int(Expiration / dt) + 1
dt = Expiration / NTS

ReDim X(0 To NXS)
ReDim H(0 To NXS, 0 To NTS)

For i = 0 To NXS
X(i) = i * dX
H(i, 0) = Application.Max(1 - X(i), 0)
Next i

For k = 1 To NTS
For i = 1 To NXS - 1
Delta = (H(i + 1, k - 1) - H(i - 1, k - 1)) / 2 / dX
Gamma = (H(i + 1, k - 1) - 2 * H(i, k - 1) _
            + H(i - 1, k - 1)) / dX / dX
Theta = -0.5 * Vol ^ 2 * X(i) ^ 2 * Gamma - _
            IntRate * X(i) * Delta + IntRate _
            * H(i, k - 1)
H(i, k) = H(i, k - 1) - Theta * dt
Next i
H(0, k) = H(0, k - 1) * (1 - IntRate * dt)
H(NXS, k) = 2 * H(NXS - 1, k) - H(NXS - 2, k)

If Int(k * dt / FixPer) <> Int((k + 1) * dt _
            / FixPer) Then
For i = 0 To NXS
M = Application.Max(X(i), 1)
X_after = X(i) / M
i_after = Int(X_after / dX)
Frac = (i_after * dX - X_after) / dX
H(i, k) = M * ((1 - Frac) * H(i_after, k) + Frac _
            * H(i_after + 1, k))
Next i
End If

Next k

Lookback = H

End Function
```



29.7 CLIQUET OPTION

Cliquet options are quite popular in the world of equity derivatives. These contracts, illustrated by the term sheet in Figure 29.15, are appealing to the investor because of their protection against downside risk, coupled with significant upside potential. Capping the maximum, as in this globally floored, locally capped example, ensures that the payoff is never too extreme and therefore that the value of the contract is not too outrageous.

Five-year Minimum Coupon Cliquet on AEX Index

Option Buyer	XXXX
Option Seller	YYYY
Notional Amount	EUR 25mio
Trade Date	20 December 2000
Start Date	31 January 2001
Maturity Date	Start Date + Five years
Option Seller Pays at Maturity	Notional * $\max\left(\sum_{i=1}^{5}\max\left(0,\min\left(\text{Cap},\frac{S_i-S_{i-1}}{S_{i-1}}\right)\right),16.25\%\right)$
Index	AEX Index
Cap	8%
Option Premium	?
Index Levels	S_i = Closing Level of Index on Start Date + *i* years

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.15 Cliquet term sheet.

From the point of view of the sell side, aiming to minimize market risk by delta hedging, their main exposure is to volatility risk. However, the contract is very subtle in its dependence on the assumed model for volatility. Because of this, I will devote an entire chapter, Chapter 56, to their analysis. Here I will just develop the very basics of the pricing, in a nice simple constant volatility world.

29.7.1 Path Dependency, Constant Volatility

We will be working in the classical lognormal framework for the underlying

$$dS = \mu S\,dt + \sigma S\,dX.$$

Assuming for the moment that volatility is constant, or at most a deterministic function of stock price S and time t, we can approach the pricing from the two most common directions: Monte Carlo simulation and partial differential equations. A brief glance at the term sheet shows that there are none of the nasties such as early exercise, convertibility or other decision processes that make Monte Carlo difficult to implement.

Monte Carlo

Monte Carlo pricing requires a simulation of the risk-neutral random walk for S, the calculation of the payoff for many, tens of thousands, say, of paths, and the present valuing of the resulting average payoff. This can be speeded up by many of the now common techniques. Calculation of the greeks is slightly more time consuming but still straightforward.

PDE

To derive a partial differential equation which one then solves via, for example, finite-difference methods, requires one to work out the amount of path dependency in the option and to count the number of dimensions. This is not difficult.

In all non-trivial problems we always have the two given dimensions, S and t. In order to be able to keep track, before expiry, of the progress of the possible option payoff we also need the following two new 'state variables'

$$S' \quad \text{and} \quad Q,$$

where

$$S' = \text{the value of } S \text{ at the previous fixing} = S_i$$

and

$$Q = \text{the sum to date of the bit inside the max function} =$$

$$\sum_{j=1}^{i} \max\left(0, \min\left(\text{Cap}, \frac{S_j - S_{j-1}}{S_{j-1}}\right)\right).$$

Here I am using the index i to denote the fixing just prior to the current time, t. This is all made clear in Figure 29.16.

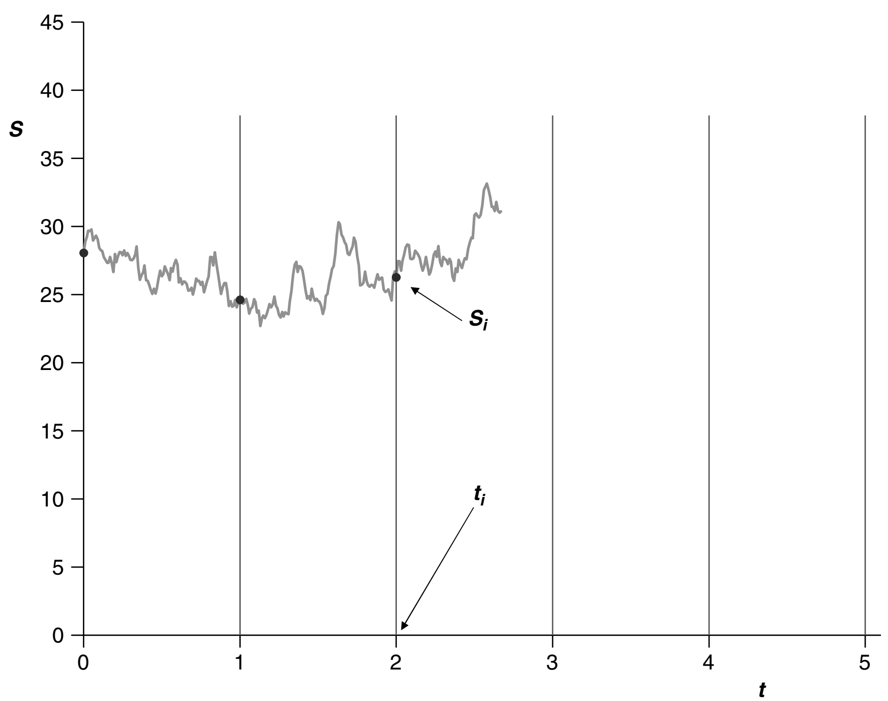


Figure 29.16 Calculating the payoff for a cliquet option.

Since S' and Q are only updated discretely, at each fixing date, the pricing problem for $V(S,t,S',Q)$ becomes

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV=0$$

where r is the risk-free interest rate. In other words, the vanilla Black–Scholes equation. The twist is that V is a function of four variables, and must further satisfy the jump condition across the fixing date

$$V(S,t_i^-,S',Q)=V\left(S,t_i^+,S,Q+\min\left(E_1,\frac{S-S'}{S'}\right)\right)$$

and the final condition

$$V(S,T,S',Q)=\max(Q,E_2).$$

Here E_1 is the local cap and E_2 the global floor. (More general payoff structures can readily be imagined.)

Being a four-dimensional problem it is a toss up as to whether a Monte Carlo or a finite-difference solution is going to be the faster. However, the structure of the payoff, and the assumption of lognormality, mean that a similarity reduction is possible, taking the problem down to only three dimensions and thus comfortably within the domain of usefulness of finite-difference methods. The similarity variable is

$$\xi=\frac{S}{S'}.$$

The option value is now a function of ξ, t and Q. The governing equation for $V(\xi,t,Q)$ (loose notation, but the most clear) is

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2\xi^2\frac{\partial^2 V}{\partial \xi^2}+r\xi\frac{\partial V}{\partial \xi}-rV=0.$$

USD/DEM 'Perfect Trader' Option

Notional Amount	USD 25,000,000+
Option Maturity	Three months from Trade Date
Allowed Position	Long or short up to Notional Amount
Transaction Frequency	Up to two times daily
Settlement Amount	Max(0,sum total in DEM of the gains + losses on each of the trades)
Upfront Premium	3.35% of Notional Amount

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 29.17 Passport option term sheet.

For Discussion Purposes Only

BUND/BTP Future Linked 18 months "Dual Passport" Note

Aggregate Principal Amount	DEM 50,000,000
Trade Date	[ ] January 1998
Issue Date	20 January 1998
Settlement Date	20 January 1998
Maturity Date	20 July 1999
Issue Price	100%
Redemption Amount	102.5% PLUS Notional Income Where "Notional Income" is the aggregate sum of the notional profit or loss from each executed transaction.

Formula

$$\max\left(\left[\sum_{i=1}^{N1} U_{i-1}\times(\text{Price}_i-\text{Price}_{i-1})\times 25\times 100\right]+\left[\sum_{j=1}^{N2} V_{j-1}\times(\text{Price}_j-\text{Price}_{j-1})\times 20.000\times 100\right]/\text{FX}\right)$$

Where	N1 is the total number of BUND transactions during the life of the option N2 is the total number of BTP transactions during the life of the option U is the BUND position V is the BTP position FX is equal to the prevailing DEM/ITL exchange rate at Option Maturity Price means the reference price for each transaction
Coupon	Zero
Option Maturity Date	6 months and 2 business days from the Issue Date
Position	The noteholder may have a non-zero position in EITHER the BUND future Index Units or the BTP future Index Units and may switch between these. The noteholder may not hold positions in both of these simultaneously. The maximum BUND position is plus or minus 190 Index Units. The maximum BTP position is plus or minus 200 Index Units.
Transaction Frequency	The maximum number of transactions per day is four.
Index Units	Index Units are either LIFFE BUND futures or LIFFE BTP futures contracts. The current reference month will be Mar 98 until this becomes deliverable, at which point the noteholder must close out the position in this contract and enter transactions based on the Jun 98 contract. Once the Jun 98 contract becomes deliverable the noteholder must close out the position in this contract and enter transactions based on the Sep 98 contract.

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

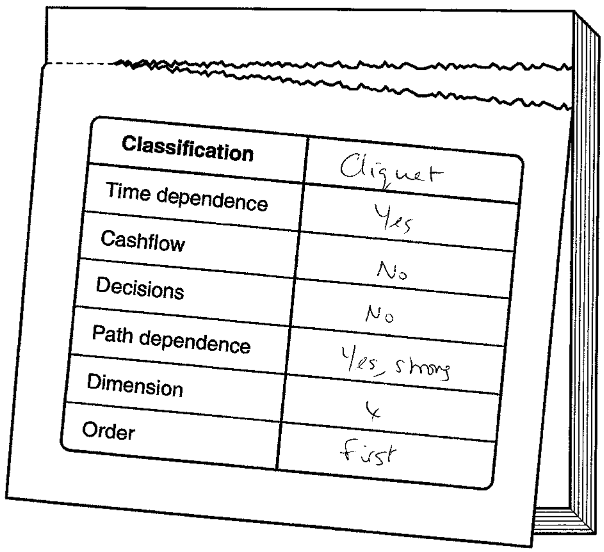
Figure 29.18 Dual passport option term sheet.

The jump condition becomes

$$V(\xi, t_i^-, Q) = V\left(1, t_i^+, Q + \min\left(E_1, \xi - 1\right)\right)$$

and the final condition is

$$V(\xi, T, Q) = \max(Q, E_2).$$



Classification option table for Cliquet.

29.8 PASSPORT OPTIONS

The first passport option, shown in Figure 29.17, is very standard. There is a restriction on the number of trades per day but this will not affect the price too much. Usually the holder of the passport option has to pay bid-offer spread on the underlying asset when he trades, so that will not affect the price of the option.

The next passport option, in Figure 29.18, is more complicated. The holder is insured against trading losses in *two* underlyings. The twist in this contract is that the trader can only hold a position in one of the underlyings at each time, not both simultaneously.

29.9 DECOMPOSITION OF EXOTICS INTO VANILLAS

If a contract can be decomposed into simpler, vanilla products, then that's what you should do for pricing and hedging. If it can't be decomposed it can be quite dangerous to try and *approximately* decompose it. The cliquet option above is a perfect example. The individual building blocks look like simple call spreads but the subtlety in the optionality, the floor for example, can seriously mess this up. That's why we've devoted a whole chapter, Chapter 56, to a close study of the cliquet option and how its value depends on the volatility model.

Given the large, and increasing, number of exotic structures around, I no longer find it helpful to think of exotics in terms of simpler vanillas. That's what the classification is for; to help you think about new products in mathematical terms.

PART THREE
fixed-income modeling and derivatives

This part of the book introduces stochastic models for interest rates, and derives models for many products.

The mathematics of this part of the book is mostly no different from what you have seen already. And again we apply the ideas of hedging and no arbitrage. However, the pricing and hedging of fixed-income products is technically more complicated than the pricing and hedging of equity instruments. One of the reasons for this is that the variable that we will be modeling, the short-term interest rate, is not itself a traded quantity. A consequence of this is that we must introduce the confusing quantity, the market price of interest rate risk.

Chapter 30: One-factor Interest Rate Modeling Interest rates are no more predictable than stock prices. I therefore show how to model them as stochastic variables. We concentrate on modeling the short-term interest rate and see how the prices of other instruments satisfy a parabolic partial differential equation. I discuss the properties of several well-known models.

Chapter 31: Yield Curve Fitting We would like an output of our simple model to be the yield curve as it is today in the market. I show how to choose parameters in the model so as to ensure that this is the case. This approach is not easy to justify, as I explain.

Chapter 32: Interest Rate Derivatives I go deeper into the pricing of interest-rate derivatives, relating the pricing equations to those from the equity world.

Chapter 33: Convertible Bonds Convertible bonds have a value that depends on a stock price. Because of their long lifespan we cannot assume that interest rates are known, and this leads us to model the value of CBs via a two-factor model with stochastic asset and stochastic interest rates.

Chapter 34: Mortgage-backed Securities Mortgage-backed securities are simply many mortgages lumped together and sold on as a financial product. I say 'simply' but the valuation of these products can be quite complicated. The main issue in the valuation is how to treat the possible early repayment of mortgage loans by individual homeowners.

Chapter 35: Multi-factor Interest Rate Modeling If we have only the one factor, the short-term interest rate, determining the behavior of the whole yield curve then we will get unnatural results. For example, the one-factor theory says that we can hedge a ten-year instrument with a one-year bond, which is clearly not true. To better model reality we discuss multi-factor modeling, where now there are several sources of randomness and hence more complicated relationships between rates of different maturities.

Chapter 36: Empirical Behavior of the Spot Interest Rate The popular models for the short-term interest rate make various assumptions about the behaviors of the volatility and drift of the rate. In this chapter I show how to examine the data to decide which, if any, of the models are close to reality.

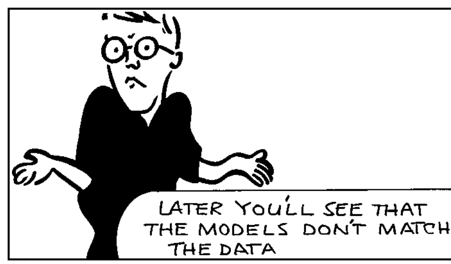
Chapter 37: The Heath, Jarrow & Morton and Brace, Gatarek & Musiela Models Instead of modeling a short-term rate and then finding the full yield curve, Heath, Jarrow & Morton model the whole forward rate curve in one go. There are many advantages in this method, but also a few disadvantages. The Brace, Gatarek & Musiela is a discrete-compounding, and therefore more realistic, version of this idea.

Chapter 38: Fixed-income Term Sheets In this chapter you will find a selection of term sheets, analysis and Visual Basic code.

CHAPTER 30
one-factor interest rate modeling

In this Chapter...

- stochastic models for interest rates
- how to derive the bond pricing equation for many fixed-income products
- the structure of many popular interest rate models



30.1 INTRODUCTION

Until now I have assumed that interest rates are either constant or a known function of time. This may be a reasonable assumption for short-dated equity contracts. But for longer-dated contracts the interest rate must be more accurately modeled. This is not an easy task. In this chapter I introduce the ideas behind modeling interest rates using a single source of randomness. This is **one-factor interest rate modeling**. The model will allow the short-term interest rate, the spot rate, to follow a random walk. This model leads to a parabolic partial differential equation for the prices of bonds and other interest rate derivative products.

The 'spot rate' that we will be modeling is a very loosely-defined quantity, meant to represent the yield on a bond of infinitesimal maturity. In practice one should take this rate to be the yield on a liquid finite-maturity bond, say one of one month. Bonds with one *day* to expiry do exist but their price is not necessarily a guide to other short-term rates. I will continue to be vague about the precise definition of the spot interest rate. We could argue that if we are pricing a complex product that is highly model-dependent then the *exact* definition of the independent variable will be relatively unimportant compared with the choice of model.

30.2 STOCHASTIC INTEREST RATES

Since we cannot realistically forecast the future course of an interest rate, it is natural to model it as a random variable. We are going to model the behavior of r, the interest rate received by the shortest possible deposit. From this we will see the development of a model for all other rates. The interest rate for the shortest possible deposit is commonly called the **spot interest rate**.

Figure 30.1 shows the time series of a one-month US interest rate. We will often use the one-month rate as a proxy for the spot rate.

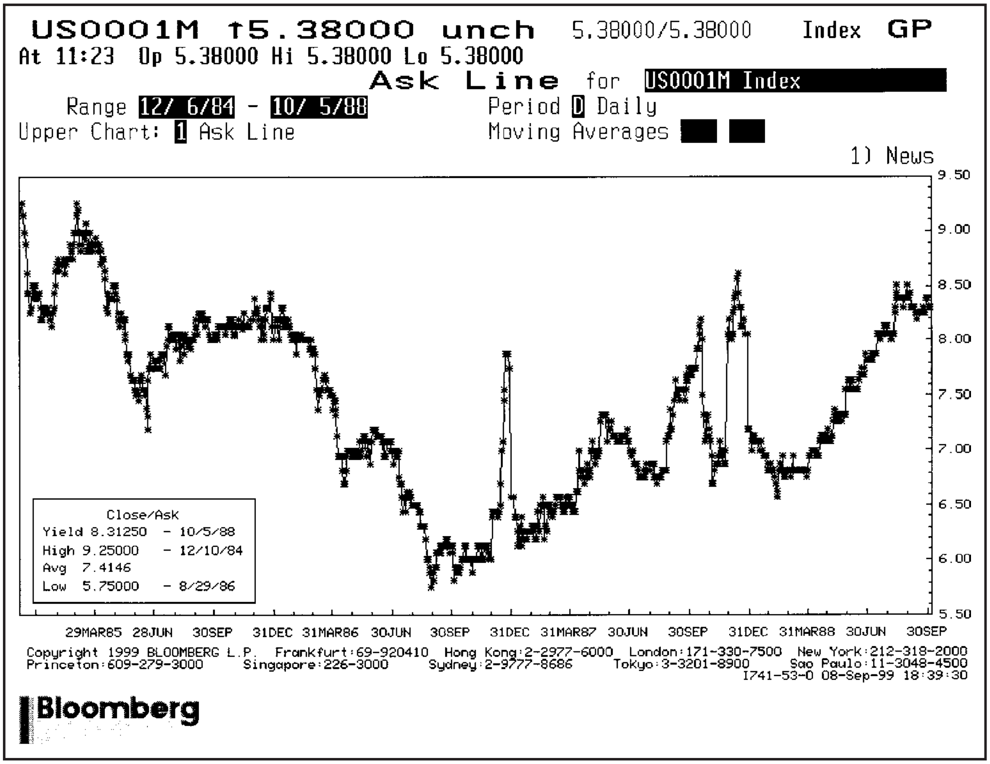


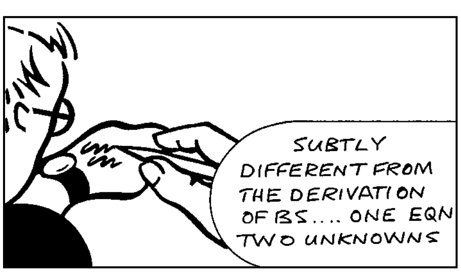
Figure 30.1 One-month interest rate time series. Source: Bloomberg L.P.

Earlier I proposed a model for the asset price as a stochastic differential equation, the lognormal random walk. Now let us suppose that the interest rate r is governed by another stochastic differential equation of the form

$$dr = u(r, t)\, dt + w(r, t)\, dX. \tag{30.1}$$

The functional forms of $u(r, t)$ and $w(r, t)$ determine the behavior of the spot rate r. For the present I will not specify any particular choices for these functions.

We use this random walk to derive a partial differential equation for the price of a bond using similar arguments to those in the derivation of the Black–Scholes equation. Later I describe functional forms for u and w that have become popular with practitioners.



30.3 THE BOND PRICING EQUATION FOR THE GENERAL MODEL

When interest rates are stochastic a bond has a price of the form $V(r, t; T)$. The reader should think for the moment in terms of simple bonds, but the governing equation will be far more general and may be used to price many other contracts. That's why I'm using the notation V rather than our earlier Z, for zero-coupon bonds.

Pricing a bond presents new technical problems, and is in a sense harder than pricing an option since *there is no underlying asset with which to hedge*. We are therefore not modeling a *traded* asset; the traded asset (the bond, say) is a derivative of our independent variable r. The only way to construct a hedged portfolio is by hedging one bond with a bond of a different maturity. We set up a portfolio containing two bonds with different maturities T_1 and T_2. The bond with maturity T_1 has price $V_1(r, t; T_1)$ and the bond with maturity T_2 has price $V_2(r, t; T_2)$. We hold one of the former and a number $-\Delta$ of the latter. We have

$$\Pi = V_1 - \Delta V_2. \tag{30.2}$$

The change in this portfolio in a time dt is given by

$$d\Pi = \frac{\partial V_1}{\partial t}dt + \frac{\partial V_1}{\partial r}dr + \tfrac{1}{2}w^2\frac{\partial^2 V_1}{\partial r^2}dt - \Delta\left(\frac{\partial V_2}{\partial t}dt + \frac{\partial V_2}{\partial r}dr + \tfrac{1}{2}w^2\frac{\partial^2 V_2}{\partial r^2}dt\right), \tag{30.3}$$

where we have applied Itô's lemma to functions of r and t. Which of these terms are random? Once you've identified them you'll see that the choice

$$\Delta = \frac{\partial V_1}{\partial r}\bigg/\frac{\partial V_2}{\partial r}$$

eliminates all randomness in $d\Pi$. This is because it makes the coefficient of dr zero. We then have

$$\begin{aligned} d\Pi &= \left(\frac{\partial V_1}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V_1}{\partial r^2} - \left(\frac{\partial V_1}{\partial r}\bigg/\frac{\partial V_2}{\partial r}\right)\left(\frac{\partial V_2}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V_2}{\partial r^2}\right)\right)dt \\ &= r\,\Pi\,dt = r\left(V_1 - \left(\frac{\partial V_1}{\partial r}\bigg/\frac{\partial V_2}{\partial r}\right)V_2\right)dt, \end{aligned}$$

where we have used arbitrage arguments to set the return on the portfolio equal to the risk-free rate. This risk-free rate is just the spot rate.

Collecting all V_1 terms on the left-hand side and all V_2 terms on the right-hand side we find that[1]

$$\frac{\dfrac{\partial V_1}{\partial t} + \tfrac{1}{2}w^2\dfrac{\partial^2 V_1}{\partial r^2} - rV_1}{\dfrac{\partial V_1}{\partial r}} = \frac{\dfrac{\partial V_2}{\partial t} + \tfrac{1}{2}w^2\dfrac{\partial^2 V_2}{\partial r^2} - rV_2}{\dfrac{\partial V_2}{\partial r}}.$$

At this point the distinction between the equity and interest-rate worlds starts to become apparent. This is *one* equation in *two* unknowns.[2] Fortunately, the left-hand side is a function of T_1 but not T_2 and the right-hand side is a function of T_2 but not T_1. The only way for this to be possible is for both sides to be independent of the maturity date. Dropping the subscript from V, we have

$$\frac{\dfrac{\partial V}{\partial t} + \tfrac{1}{2}w^2\dfrac{\partial^2 V}{\partial r^2} - rV}{\dfrac{\partial V}{\partial r}} = a(r, t)$$

[1] You will see parallels with the later stochastic volatility model.

[2] Whenever we model something as stochastic that is not traded we end up with too few equations for the number of unknowns. We get around that sticky problem by introducing a market price of risk.

for some function $a(r, t)$. I shall find it convenient to write

$$a(r, t) = w(r, t)\lambda(r, t) - u(r, t);$$

for a given $u(r, t)$ and non-zero $w(r, t)$, this is always possible. The function $\lambda(r, t)$ is as yet unspecified.

The bond pricing equation is therefore

$$\boxed{\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV = 0.} \tag{30.4}$$

To find a unique solution of (30.4) we must impose one final and two boundary conditions. The final condition corresponds to the payoff on maturity and so for a zero-coupon bond

$$V(r, T; T) = 1.$$

Boundary conditions depend on the form of $u(r, t)$ and $w(r, t)$ and are discussed later for a special model.

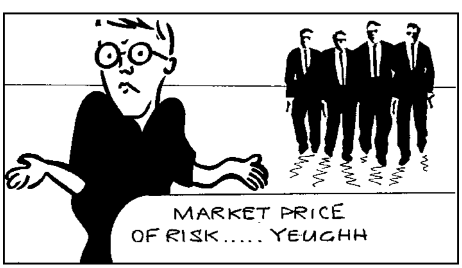
It is easy to incorporate coupon payments into the model. If an amount $K(r, t)\,dt$ is received in a period dt then

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV + K(r, t) = 0.$$

When this coupon is paid discretely, arbitrage considerations lead to jump condition

$$V(r, t_c^-; T) = V(r, t_c^+; T) + K(r, t_c),$$

where a coupon of $K(r, t_c)$ is received at time t_c.



30.4 WHAT IS THE MARKET PRICE OF RISK?

I now give an interpretation of the function $\lambda(r, t)$. Imagine that you hold an unhedged position in one bond with maturity date T. In a time-step dt this bond changes in value by

$$dV = w\frac{\partial V}{\partial r}\,dX + \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + u\frac{\partial V}{\partial r}\right)dt.$$

From (30.4) this may be written as

$$dV = w\frac{\partial V}{\partial r}\,dX + \left(w\lambda\frac{\partial V}{\partial r} + rV\right)dt,$$

or

$$dV - rV\,dt = w\frac{\partial V}{\partial r}(dX + \lambda\,dt). \tag{30.5}$$

The right-hand side of this expression contains two terms: a deterministic term in dt and a random term in dX. The presence of dX in (30.5) shows that this is not a riskless portfolio. The deterministic term may be interpreted as the excess return above the risk-free rate for accepting a certain level of risk. In return for taking the extra risk the portfolio profits by an extra $\lambda\, dt$ per unit of extra risk, dX. The function λ is therefore called the **market price of risk**.

30.5 INTERPRETING THE MARKET PRICE OF RISK, AND RISK NEUTRALITY

The bond pricing equation (30.4) contains references to the functions $u - \lambda w$ and w. The former is the coefficient of the first-order derivative with respect to the spot rate, and the latter appears in the coefficient of the diffusive, second-order derivative. The four terms in the equation represent, in order as written, time decay, diffusion, drift and discounting. The equation is similar to the backward equation for a probability density function, see Chapter 10, except for the final discounting term. As such we can interpret the solution of the bond pricing equation as the expected present value of all cashflows. Suppose that we get a 'Payoff' at time T, then the value of that contract today would be

$$E\left[e^{-\int_t^T r(\tau)\, d\tau}\text{Payoff}\right].$$

Notice that the present value (exponential) term goes inside the expectation since it is also random when interest rates are random.

As with equity options, this expectation is not with respect to the *real* random variable, but instead with respect to the *risk-neutral* variable. There is this difference because the drift term in the equation is not the drift of the real spot rate u, but the drift of another rate, called the **risk-neutral spot rate**. This rate has a drift of $u - \lambda w$. When pricing interest rate derivatives (including bonds of finite maturity) it is important to model, and price, using the risk-neutral rate. This rate satisfies

$$dr = (u - \lambda w)\, dt + w\, dX.$$

We need the new market-price-of-risk term because our modeled variable, r, is not traded.

If we set λ to zero then any results we find are applicable to the real world. If, for example, we want to find the distribution of the spot interest rate at some time in the future then we would solve a Fokker–Planck equation with the real, and not the risk-neutral, drift.

Because we can't observe the function λ, except possibly via the whole yield curve (see Chapter 31), I tend to think of it as a great big carpet under which we can brush all kinds of nasty, inconvenient things.

30.6 TRACTABLE MODELS AND SOLUTIONS OF THE BOND PRICING EQUATION

We have built up the bond pricing equation for an arbitrary model. That is, we have not specified the risk-neutral drift, $u - \lambda w$, or the volatility, w. How can we choose these functions to give us a good model? First of all, a simple lognormal random walk would *not* be suitable for r, since it would predict exponentially rising or falling rates. This rules out the equity price model as an interest rate model. So we must think more carefully about how to choose the drift and volatility.

Modeling interest rates is far, far harder than modeling stock prices because we have no economic clues on which to hang our model. When we set out to model equity prices we observed that the actual level of the stock price is immaterial, only the returns matter. This made the lognormal random walk model an obvious choice. With interest rates the level *does* matter; a 5% interest rate and a 500% interest rate are clearly totally different beasts.

One thing we can do, although I don't approve, is to choose a model that makes further analysis easy.

Let us examine some choices for the risk-neutral drift and volatility that lead to tractable models, that is, models for which the solution of the bond pricing equation for zero-coupon bonds can be found analytically. We will discuss these models and see what properties we like or dislike.

For example, assume that $u - \lambda w$ and w take the form

$$u(r,t) - \lambda(r,t)w(r,t) = \eta(t) - \gamma(t)r, \tag{30.6}$$

$$w(r,t) = \sqrt{\alpha(t)r + \beta(t)}. \tag{30.7}$$

Note that we are describing a model for the risk-neutral spot rate. I will allow the functions α, β, γ, η and λ that appear in (30.7) and (30.6) to be functions of time. By suitably restricting these time-dependent functions, we can ensure that the random walk (30.1) for r has the following nice properties:

- **Positive interest rates**: Except for a few pathological cases, such as Switzerland in the 1960s, interest rates are positive. With the above model the spot rate can be bounded below by a positive number if $\alpha(t) > 0$ and $\beta \le 0$. The lower bound is $-\beta/\alpha$. (In the special case $\alpha(t) = 0$ we must take $\beta(t) \ge 0$.) Note that r can still go to infinity, but with probability zero.
- **Mean reversion**: Examining the drift term, we see that for large r the (risk-neutral) interest rate will tend to decrease towards the mean, which may be a function of time. When the rate is small it will move up on average.

We also want the lower bound to be non-attainable; we don't want the spot interest rate to get forever stuck at the lower bound or have to impose further conditions to say how fast the spot rate moves away from this value. This requirement means that

$$\eta(t) \ge -\beta(t)\gamma(t)/\alpha(t) + \alpha(t)/2,$$

and it is discussed further below.

With the model (30.7) and (30.6) the boundary conditions for (30.4), for a zero-coupon bond, are, first, that

$$V(r,t;T) \to 0 \quad \text{as} \quad r \to \infty,$$

and, second, that on $r=-\beta/\alpha$, V remains finite. When r is bounded below by $-\beta/\alpha$, a local analysis of the partial differential equation can be carried out near $r=-\beta/\alpha$. Briefly, balancing the terms

$$\tfrac{1}{2}(\alpha r+\beta)\frac{\partial^2 V}{\partial r^2} \quad \text{and} \quad (\eta-\gamma r)\frac{\partial V}{\partial r}$$

shows that finiteness of V at $r=-\beta/\alpha$ is a sufficient boundary condition only if $\eta \geq -\beta\gamma/\alpha + \alpha/2$.

I chose u and w in the stochastic differential equation for r to take the special functional forms (30.7) and (30.6) for a very special reason. With these choices the solution of (30.4) for the zero-coupon bond is of the simple form

$$Z(r,t;T)=e^{A(t;T)-rB(t;T)}. \tag{30.8}$$

We are going to be looking at zero-coupon bonds specifically for a while, hence the change of our notation from V, meaning many interest rate products, to the very specific Z for zero coupon bonds.

The model with all of α, β, γ and η non-zero is the most general stochastic differential equation for r which leads to a solution of (30.4) of the form (30.8). This is easily shown.

Substitute (30.8) into the bond pricing equation (30.4). This gives

$$\frac{\partial A}{\partial t}-r\frac{\partial B}{\partial t}+\tfrac{1}{2}w^2B^2-(u-\lambda w)B-r=0. \tag{30.9}$$

Some of these terms are functions of t and T (i.e. A and B) and others are functions of r and t (i.e. u and w). Differentiating (30.9) with respect to r gives

$$-\frac{\partial B}{\partial t}+\tfrac{1}{2}B^2\frac{\partial}{\partial r}(w^2)-B\frac{\partial}{\partial r}(u-\lambda w)-1=0.$$

Differentiate again with respect to r, and after dividing through by B, you get

$$\tfrac{1}{2}B\frac{\partial^2}{\partial r^2}(w^2)-\frac{\partial^2}{\partial r^2}(u-\lambda w)=0.$$

In this, only B is a function of T, therefore we must have

$$\frac{\partial^2}{\partial r^2}(w^2)=0 \tag{30.10}$$

and

$$\frac{\partial^2}{\partial r^2}(u-\lambda w)=0. \tag{30.11}$$

From this follow (30.7) and (30.6).

The substitution of (30.7) and (30.6) into (30.9) yields the following equations for A and B:

$$\frac{\partial A}{\partial t}=\eta(t)B-\tfrac{1}{2}\beta(t)B^2 \tag{30.12}$$

and

$$\frac{\partial B}{\partial t} = \tfrac{1}{2}\alpha(t)B^2 + \gamma(t)B - 1. \tag{30.13}$$

In order to satisfy the final data that $Z(r, T; T) = 1$ we must have

$$A(T; T) = 0 \quad \text{and} \quad B(T; T) = 0.$$

30.7 SOLUTION FOR CONSTANT PARAMETERS

The solution for arbitrary α, β, γ and η is found by integrating the two ordinary differential equations (30.12) and (30.13). Generally speaking, though, when these parameters are time-dependent this integration cannot be done explicitly. But in some special cases this integration *can* be done explicitly.

The simplest case is when α, β, γ and η are all constant. In this case we have

$$\frac{dB}{dt} = \tfrac{1}{2}\alpha B^2 + \gamma B - 1.$$

This can be integrated when written in the form

$$\int_0^B \frac{1}{(B' - a)(B' + b)} dB' = \tfrac{1}{2}\alpha \int_T^t dt,$$

where

$$b, a = \frac{\pm\gamma + \sqrt{\gamma^2 + 2\alpha}}{\alpha}.$$

This incorporates the final condition at $t = T$. The result is that

$$B(t; T) = \frac{2(e^{\psi_1(T-t)} - 1)}{(\gamma + \psi_1)(e^{\psi_1(T-t)} - 1) + 2\psi_1}, \tag{30.14}$$

where

$$\psi_1 = \sqrt{\gamma^2 + 2\alpha} \quad \text{and} \quad \psi_2 = \frac{\eta - a\beta/2}{a + b}.$$

The equation for A is

$$\frac{dA}{dt} = \eta B - \tfrac{1}{2}\beta B^2.$$

Dividing this by the ordinary differential equation for B gives

$$\frac{dA}{dB} = \frac{\eta B - \frac{1}{2}\beta B^2}{\frac{1}{2}\alpha B^2 + \gamma B - 1}.$$

This can be integrated to give

$$\frac{\alpha}{2}A = a\psi_2 \log(a - B) + (\psi_2 + \tfrac{1}{2}\beta)b \log((B + b)/b) - \tfrac{1}{2}B\beta - a\psi_2 \log a, \tag{30.15}$$

which has incorporated the final condition.

When all four of the parameters are constant it is obvious that both A and B are functions of only the one variable $\tau = T - t$, and not t and T individually; this would not necessarily be the case if any of the parameters were time-dependent.

A wide variety of yield curves can be predicted by the model. As $\tau \to \infty$,

$$B \to \frac{2}{\gamma + \psi_1}$$

and the yield curve Y has long-term behavior given by

$$Y \to \frac{2}{(\gamma + \psi_1)^2}\left(\eta(\gamma + \psi_1) - \beta\right).$$

Thus for constant and fixed parameters the model leads to a fixed long-term interest rate, independent of the spot rate.

The probability density function, $P(r, t)$, for the risk-neutral spot rate satisfies

$$\frac{\partial P}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial r^2}(w^2 P) - \frac{\partial}{\partial r}((u - \lambda w)P).$$

In the long term this settles down to a distribution, $P_\infty(r)$, that is independent of the *initial* value of the rate. This distribution satisfies the ordinary differential equation

$$\frac{1}{2}\frac{d^2}{dr^2}(w^2 P_\infty) = \frac{d}{dr}((u - \lambda w)P_\infty).$$

The solution of this for the general affine model with constant parameters is

$$P_\infty(r) = \frac{\left(\frac{2\gamma}{\alpha}\right)^k}{\Gamma(k)}\left(r + \frac{\beta}{\alpha}\right)^{k-1} e^{-\frac{2\gamma}{\alpha}(r + \frac{\beta}{\alpha})} \tag{30.16}$$

where

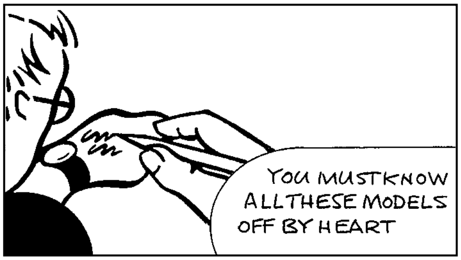
$$k = \frac{2\eta}{\alpha} + \frac{2\beta\gamma}{\alpha^2}$$

and $\Gamma(\cdot)$ is the gamma function. The boundary $r = -\beta/\alpha$ is non-attainable if $k > 1$. The mean of the steady-state distribution is

$$\frac{\alpha k}{2\gamma} - \frac{\beta}{\alpha}.$$

30.8 NAMED MODELS

There are many interest rate models, which are associated with the names of their inventors. The stochastic differential equation (30.1) for the risk-neutral interest rate process, with risk-neutral drift and volatility given by (30.6) and (30.7), incorporates the models of Vasicek, Cox, Ingersoll & Ross, Ho & Lee, and Hull & White.



30.8.1 Vasicek

The Vasicek model takes the form of (30.6) and (30.7) but with $\alpha = 0$, $\beta > 0$ and with all other parameters independent of time:

$$dr = (\eta - \gamma r)\,dt + \beta^{1/2} dX.$$

This model is so 'tractable' that there are explicit formulae for many interest rate derivatives. The value of a zero-coupon bond is given by

$$e^{A(t;T) - rB(t;T)}$$

where

$$B = \frac{1}{\gamma}\left(1 - e^{-\gamma(T-t)}\right)$$

and

$$A = \frac{1}{\gamma^2}(B(t;T) - T + t)(\eta\gamma - \tfrac{1}{2}\beta) - \frac{\beta B(t;T)^2}{4\gamma}.$$

The model is mean reverting to a constant level, which is a good property, but interest rates can easily go negative, which is a very bad property.

In Figure 30.2 are shown three types of yield curves predicted by the Vasicek model, each using different parameters. (It is quite difficult to get the humped yield curve with reasonable numbers.)

The steady-state probability density function for the Vasicek model is a degenerate case of (30.16), since $\alpha = 0$. We find that

$$P_\infty(r) = \sqrt{\frac{\gamma}{\beta\pi}}\, e^{-\frac{\gamma}{\beta}\left(r - \frac{\eta}{\gamma}\right)^2}.$$

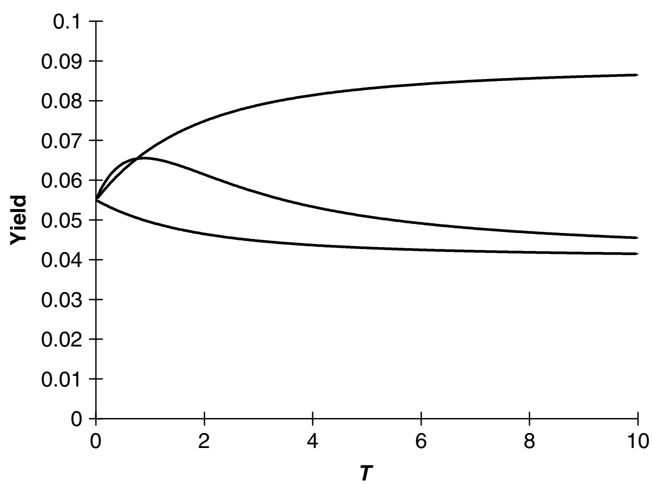


Figure 30.2 Three types of yield curve given by the Vasicek model.

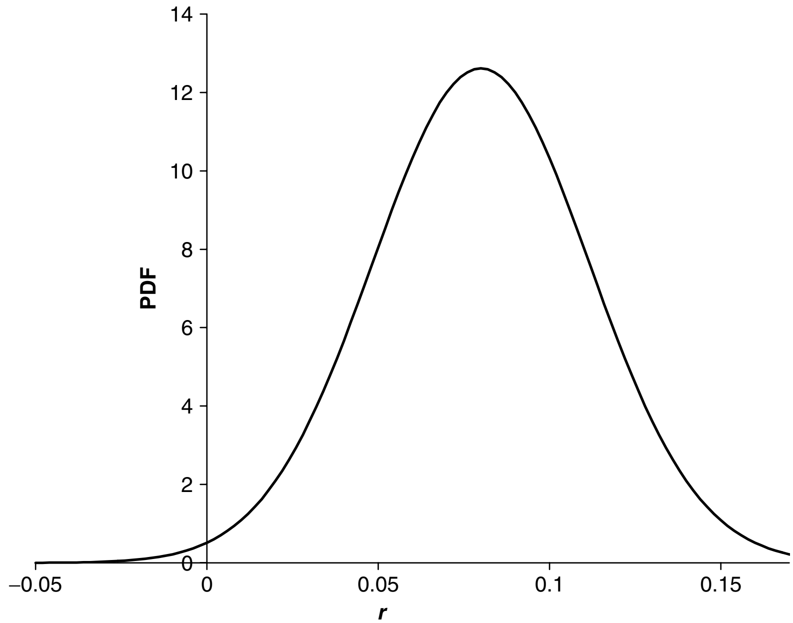


Figure 30.3 The steady-state probability density function for the risk-neutral spot rate in the Vasicek model.

This is plotted in Figure 30.3. Thus, in the long run, the spot rate is Normally distributed in the Vasicek model. The mean of this distribution is

$$\frac{\eta}{\gamma}.$$

(The parameters in the figure have been deliberately chosen to give an alarming probability of a negative interest rate. For reasonable parameters the probability of negative rates is not that worrying, but then with reasonable parameters it's hard to get realistic looking yield curves.)

30.8.2 Cox, Ingersoll & Ross

The CIR model takes (30.6) and (30.7) as the interest rate model but with $\beta = 0$, and again no time dependence in the parameters:

$$dr = (\eta - \gamma r)\, dt + \sqrt{\alpha r}\, dX.$$

The spot rate is mean reverting and if $\eta > \alpha/2$ the spot rate stays positive. There are some explicit solutions for interest rate derivatives, although typically involving integrals of the non-central chi-squared distribution. The value of a zero-coupon bond is

$$e^{A(t;T)-rB(t;T)}$$

where A and B are given by (30.15) and (30.14) with $\beta = 0$. The resulting expression is not much simpler than in the non-zero β case.

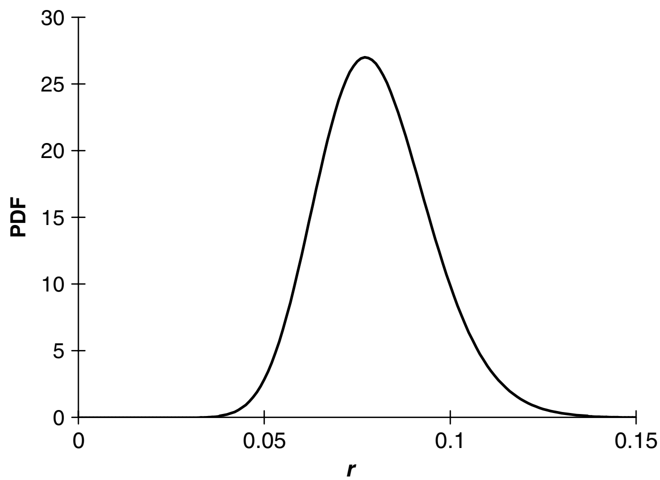


Figure 30.4 The steady-state probability density function for the risk-neutral spot rate in the CIR model.

The steady-state probability density function for the spot rate is a special case of (30.16). A plot of this function is shown in Figure 30.4. The mean of the steady-state distribution is again

$$\frac{\eta}{\gamma}.$$

In Figure 30.5 are simulations of the Vasicek and CIR models using the same random numbers. The parameters have been chosen to give similar mean and standard deviations for the two processes.

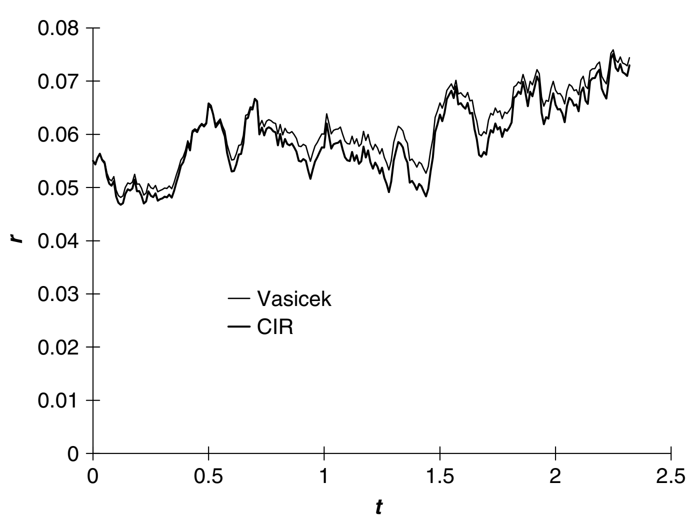


Figure 30.5 A simulation of the Vasicek and CIR models using the same random numbers.

30.8.3 Ho & Lee

Ho & Lee have $\alpha = \gamma = 0$, $\beta > 0$ and constant but η can be a function of time:

$$dr = \eta(t)\, dt + \beta^{1/2} dX.$$

The value of zero-coupon bonds is given by

$$e^{A(t;T) - rB(t;T)}$$

where

$$B = T - t$$

and

$$A = -\int_t^T \eta(s)(T-s)ds + \tfrac{1}{6}\beta(T-t)^3.$$

This model was the first 'no-arbitrage model' of the term structure of interest rates. By this is meant that the careful choice of the function $\eta(t)$ will result in theoretical zero-coupon bonds prices, output by the model, which are the same as market prices. This technique is also called **yield curve fitting**. This careful choice is

$$\eta(t) = -\frac{\partial^2}{\partial t^2} \log Z_M(t^*; t) + \beta(t - t^*)$$

where today is time $t = t^*$. In this $Z_M(t^*; T)$ is the market price today of zero-coupon bonds with maturity T. Clearly this assumes that there are bonds of all maturities and that the prices are twice differentiable with respect to the maturity. We will see why this should give the 'correct' prices later. This analytically tractable model also yields simple explicit formulae for bond options. The business of 'yield curve fitting' is the subject of Chapter 31.

30.8.4 Hull & White

Hull & White have extended both the Vasicek and the CIR models to incorporate time-dependent parameters. This time dependence again allows the yield curve (and even a volatility structure) to be fitted. We will explore this model in greater depth later.

30.9 EQUITY AND FX FORWARDS AND FUTURES WHEN RATES ARE STOCHASTIC

Recall from Chapter 5 that forward prices and futures prices are the same if rates are constant? How does this change, if at all, when rates are stochastic? We must repeat the analysis of that chapter but now with

$$dS = \mu S\, dt + \sigma S\, dX_1$$

and

$$dr = u(r,t)\, dt + w(r,t)\, dX_2.$$

We are in the world of correlated random walks, as described in Chapter 11. The correlation coefficient is ρ.

30.9.1 Forward Contracts

$V(S, r, t)$ will be the value of the forward contract at any time during its life on the underlying asset S, and maturing at time T. As in Chapter 5, I'll assume that the delivery price is known and then find the forward contract's value.

Set up the portfolio of one long forward contract and short Δ of the underlying asset, and Δ_1 of a risk-free bond:

$$\Pi = V(S, t) - \Delta\ S - \Delta_1\ Z.$$

I won't go through all the details because the conclusion is the obvious one:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S w \frac{\partial^2 V}{\partial S \partial r} + \tfrac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2} + rS \frac{\partial V}{\partial S} + (u - \lambda w)\frac{\partial V}{\partial r} - rV = 0.$$

The final condition for the equation is simply the difference between the asset price S and the fixed delivery price $\overline{S}$. So

$$V(S, r, T) = S - \overline{S}.$$

The solution of the equation with this final condition is

$$V(S, r, t) = S - \overline{S}Z.$$

At this point Z is not just any old risk-free bond, it is a zero-coupon bond having the same maturity as the forward contract. This is the forward contract's value during its life.

Remember that the delivery price is set initially to $t = t_0$ as the price that gives the forward contract zero value. If the underlying asset is S_0 at t_0 then

$$0 = S_0 - \overline{S}Z$$

or

$$\overline{S} = \frac{S_0}{Z}.$$

The quoted forward price is therefore

$$\text{Forward price } = \frac{S}{Z}.$$

Remember that Z satisfies

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2} w^2 \frac{\partial^2 Z}{\partial r^2} + (u - \lambda w)\frac{\partial Z}{\partial r} - rZ = 0$$

with

$$Z(r, T) = 1.$$

30.9.2 Futures contracts

Use $F(S, r, t)$ to denote the futures price.

Set up a portfolio of one long futures contract and short Δ of the underlying, and Δ_1 of a risk-free bond:

$$\Pi = -\Delta\ S - \Delta_1\ Z.$$

(Remember that the futures contract has no value.)

$$d\Pi = dF - \Delta\, dS - \Delta_1\, dZ.$$

Following the usual routine we get

$$\frac{\partial F}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} + \rho\sigma S w \frac{\partial^2 F}{\partial S \partial r} + \tfrac{1}{2} w^2 \frac{\partial^2 F}{\partial r^2} + r S \frac{\partial F}{\partial S} + (u - \lambda w)\frac{\partial F}{\partial r} = 0.$$

The final condition is

$$F(S, r, T) = S.$$

Let's write the solution of this as

$$F(S, r, t) = \frac{S}{p(r, t)}.$$

Why? Two reasons. First, a similarity solution is to be expected; the price should be proportional to the asset price. Second, I want to make a comparison between the futures price and the forward price. The latter is

$$\frac{S}{Z}.$$

So it's natural to ask, how similar are Z and p?

It turns out that p satisfies

$$\frac{\partial p}{\partial t} + \tfrac{1}{2} w^2 \frac{\partial^2 p}{\partial r^2} + (u - \lambda w)\frac{\partial p}{\partial r} - rp \underline{- w^2 \frac{\left(\frac{\partial p}{\partial r}\right)^2}{q} + \rho\sigma\beta \frac{\partial p}{\partial r}} = 0. \tag{30.17}$$

(Just plug the similarity form into the equation to see this.)

The final condition is

$$p(r, T) = 1.$$

The differences between the p and Z equations are in the underlined terms in Equation (30.17).

30.9.3 The Convexity Adjustment

There is clearly a difference between the prices of forwards and futures when interest rates are stochastic. From Equation (30.17) you can see that the difference depends on the volatility of the spot interest rate, the volatility of the underlying and the correlation between them. Provided that $\rho \geq 0$ the futures price is always greater than the equivalent forward price. Should the correlation be zero then the volatility of the stock is irrelevant. If the interest rate volatility is zero then rates are deterministic and forward and futures prices are the same.

Since the difference in price between forwards and futures depends on the spot rate volatility, market practitioners tend to think in terms of **convexity adjustments** to get from one to the other. Clearly, the convexity adjustment will depend on the precise nature of the model. For the popular models described above, the p Equation (30.17) still has simple solutions.

30.10 SUMMARY

In this chapter I introduced the idea of a random interest rate. The interest rate that we modeled was the 'spot rate,' a short-term interest rate. Several popular spot rate models were described. These models were chosen because simple forms of the coefficients make the solution of the basic bond pricing equation straightforward analytically.

From a model for this spot rate we can derive the whole yield curve. This is certainly unrealistic and in later chapters we will see how to make the model of more practical use.

FURTHER READING

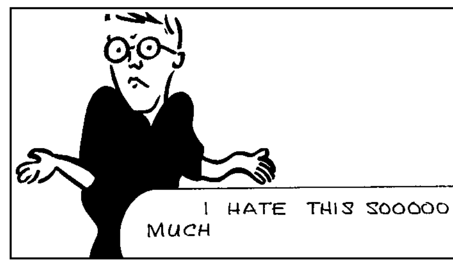
- See the original interest rate models by Vasicek (1977), Dothan (1978), Cox, Ingersoll & Ross (1985), Ho & Lee (1986) and Black, Derman & Toy (1990).
- For details of the general affine model see the papers by Pearson & Sun (1989), Duffie (1992), Klugman (1992) and Klugman & Wilmott (1994).
- The comprehensive book by Rebonato (1996) describes all of the popular interest rate models in detail.

CHAPTER 31

yield curve fitting

In this Chapter...

- how to choose time-dependent parameters in one-factor models so that today's yield curve is an output of the model
- the advantages and disadvantages of yield curve fitting



31.1 INTRODUCTION

One-factor models for the spot rate build up an entire yield curve from a knowledge of the spot rate and the parameters in the model. In using a one-factor model we have to decide how to choose the parameters and whether to believe the output of the model. If we choose parameters using historical time series data then one of the outputs of the model will be a theoretical yield curve. Unless we are very, very lucky this theoretical curve will not be the same as the market yield curve. Which do we believe? Do we believe the theoretical yield curve or do we believe the prices trading in the market? You have to be very brave to ignore the market prices for such liquid instruments as bonds and swaps. Even if you are pricing very complex products you must still hedge with simpler, more liquid, traded contracts for which you would like to get the price right.

Because of this need to price liquid instruments correctly the idea of **yield curve fitting** or **calibration** has become popular. When one-factor models are used in practice they are almost always fitted. This means that one or more of the parameters in the model is allowed to depend on time. This functional dependence on time is then carefully chosen to make an output of the model, the price of zero-coupon bonds, exactly match the market prices for these instruments. Yield curve fitting is the subject of this chapter.

31.2 HO & LEE

The Ho & Lee spot interest rate model is the simplest that can be used to fit the yield curve. It will be useful to examine this model in detail to see one way in which fitting is done in practice.

In the Ho & Lee model the process for the risk-neutral spot rate is

$$dr = \eta(t)\,dt + c\,dX.$$

The standard deviation of the spot rate process, c, is constant, the drift rate η is time-dependent.

In this model the solution of the bond pricing equation for a zero-coupon bond is simply

$$Z(r, t; T) = e^{A(t;T)-r(T-t)},$$

where

$$A(t; T) = -\int_t^T \eta(s)(T-s)\,ds + \tfrac{1}{6}c^2(T-t)^3.$$

If we know $\eta(t)$ then the above gives us the theoretical value of zero-coupon bonds of all maturities. Now turn this relationship around and ask the question: 'What functional form must we choose for $\eta(t)$ to make the theoretical value of the discount rates for all maturities equal to the market values?' Call this special choice for η, $\eta^*(t)$. The yield curve is to be fitted today, $t = t^*$, when the spot interest rate is r^* and the discount factors in the market are $Z_M(t^*; T)$. To match the market and theoretical bond prices, we must solve

$$Z_M(t^*; T) = e^{A(t^*;T)-r^*(T-t^*)}.$$

Taking logarithms of this and rearranging slightly we get

$$\int_{t^*}^T \eta^*(s)(T-s)\,ds = -\log(Z_M(t^*; T)) - r^*(T-t^*) + \tfrac{1}{6}c^2(T-t^*)^3. \tag{31.1}$$

Observe that I am carrying around in the notation today's date t^*. This is a constant but I want to emphasize that we are doing the calibration to *today's* yield curve. If we calibrate again tomorrow, the market yield curve will have changed.

Differentiate (31.1) twice with respect to T to get

$$\eta^*(t) = c^2(t-t^*) - \frac{\partial^2}{\partial t^2}\log(Z_M(t^*; t)).$$

With this choice for the time-dependent parameter $\eta(t)$ the theoretical and actual market prices of zero-coupon bonds are the same. It also follows that

$$A(t; T) = \log\left(\frac{Z_M(t^*; T)}{Z_M(t^*; t)}\right) - (T-t)\frac{\partial}{\partial t}\log(Z_M(t^*; t)) - \tfrac{1}{2}c^2(t-t^*)(T-t)^2.$$

31.3 THE EXTENDED VASICEK MODEL OF HULL & WHITE

The Ho & Lee model isn't the only one that can be calibrated, it's just the easiest. Most one-factor models have the potential for fitting, but the more tractable the model the easier the fitting. If the model is not at all tractable, having no nice explicit zero-coupon bond price formula, then we can always resort to numerical methods.

The next easiest model to fit is the Vasicek model. The Vasicek model has the following stochastic differential equation for the risk-neutral spot rate

$$dr = (\eta - \gamma r)\,dt + c\,dX.$$

Hull & White extend this to include a time-dependent parameter

$$dr = (\eta(t) - \gamma r)\, dt + c\, dX.$$

Assuming that γ and c have been estimated statistically, say, we choose $\eta = \eta^*(t)$ at time t^* so that our theoretical and the market prices of bonds coincide.

Under this risk-neutral process the value of zero-coupon bonds

$$Z(r, t; T) = e^{A(t;T) - rB(t;T)},$$

where

$$A(t; T) = -\int_t^T \eta^*(s) B(s; T)\, ds + \frac{c^2}{2\gamma^2}\left(T - t + \frac{2}{\gamma}e^{-\gamma(T-t)} - \frac{1}{2\gamma}e^{-2\gamma(T-t)} - \frac{3}{2\gamma}\right).$$

and

$$B(t; T) = \frac{1}{\gamma}\left(1 - e^{-\gamma(T-t)}\right).$$

To fit the yield curve at time t^* we must make $\eta^*(t)$ satisfy

$$\begin{aligned} A(t^*; T) &= -\int_{t^*}^T \eta^*(s) B(s; T)\, ds + \frac{c^2}{2\gamma^2}\left(T - t^* + \frac{2}{\gamma}e^{-\gamma(T-t^*)} - \frac{1}{2\gamma}e^{-2\gamma(T-t^*)} - \frac{3}{2\gamma}\right) \\ &= \log(Z_M(t^*; T)) + r^* B(t^*, T). \end{aligned} \tag{31.2}$$

This is an integral equation for $\eta^*(t)$ if we are given all of the other parameters and functions, such as the market prices of bonds, $Z_M(t^*; T)$.

Although (31.2) may be solved by Laplace transform methods, it is particularly easy to solve by differentiating the equation twice with respect to T. This gives

$$\eta^*(t) = -\frac{\partial^2}{\partial t^2}\log(Z_M(t^*; t)) - \gamma\frac{\partial}{\partial t}\log(Z_M(t^*; t)) + \frac{c^2}{2\gamma}\left(1 - e^{-2\gamma(t-t^*)}\right). \tag{31.3}$$

From this expression we can now find the function $A(t; T)$,

$$\begin{aligned} A(t; T) = \log\left(\frac{Z_M(t^*; T)}{Z_M(t^*; t)}\right) - B(t; T)\frac{\partial}{\partial t}\log(Z_M(t^*; t)) \\ - \frac{c^2}{4\gamma^3}\left(e^{-\gamma(T-t^*)} - e^{-\gamma(t-t^*)}\right)^2\left(e^{2\gamma(t-t^*)} - 1\right). \end{aligned}$$

31.4 YIELD-CURVE FITTING: FOR AND AGAINST

31.4.1 For

The building blocks of the bond pricing equation are delta hedging and no arbitrage. If we are to use a one-factor model correctly then we must abide by the delta-hedging assumptions. We must buy and sell instruments to remain delta neutral. The buying and selling of instruments must be done at the market prices. We *cannot* buy and sell at a theoretical price. But we are not modeling the bond prices directly; we model the spot rate and bond prices are then derivatives of the spot rate. This means that there is a real likelihood that our output bond prices will differ markedly from the market prices. This is useless if we are to hedge with these bonds. The model thus collapses and cannot be used for pricing other instruments, unless we can find a way to generate the correct prices for our hedging instruments from the model; this is yield curve fitting.

Once we have fitted the prices of traded products we then dynamically or statically hedge with these products. The idea is that even if the model is wrong so that we lose money on the contract we are pricing, we should then make that money back on the hedging instruments. Exactly this idea is discussed in Chapter 60.

31.4.2 Against

If the market prices of simple bonds were correctly given by a model, such as Ho & Lee or Hull & White, fitted at time t^* then, when we come back a week later, $t^* +$ one week, say, to refit the function $\eta^*(t)$, we would find that this function *had not changed* in the meantime. This *never* happens in practice. We find that the function η^* has changed out of all recognition. What does this mean? Clearly the model is wrong.[1]

By simply looking for a Taylor series solution of the bond-pricing equation for short times to expiry, we can relate the value of the risk-adjusted drift rate at the short end to the slope and curvature of the market yield curve. This is done as follows. Look for a solution of (30.4) of the form

$$Z(r, t; T) \sim 1 + a(r)(T-t) + b(r)(T-t)^2 + c(r)(T-t)^3 + \cdots.$$

Substitute this into the bond pricing equation:

$$\begin{aligned} &-a - 2b(T-t) - 3c(T-t)^2 + \tfrac{1}{2}\left(w^2 - 2(T-t)w\frac{\partial w}{\partial t}\right)\left((T-t)\frac{d^2a}{dr^2} + (T-t)^2\frac{d^2b}{dr^2}\right) \\ &\quad + \left((u-\lambda w) - (T-t)\frac{\partial(u-\lambda w)}{\partial t}\right)(T-t)\left(\frac{da}{dr} + (T-t)^2\frac{db}{dr}\right) \\ &\quad - r\left(1 + a(T-t) + c(T-t)^2\right) + \cdots = 0. \end{aligned}$$

Note how I have expanded the drift and volatility terms about $t = T$; in the above these are evaluated at r and T. By equating powers of $(T-t)$ we find that

$$a(r) = -r, \quad b(r) = \tfrac{1}{2}r^2 - \tfrac{1}{2}(u-\lambda w)$$

and

$$\begin{aligned} c(r) = \tfrac{1}{12}w^2\frac{\partial^2}{\partial r^2}(r^2 - r(u-\lambda w)) - \tfrac{1}{6}(u-\lambda w)\frac{\partial}{\partial r}(r^2 - r(u-\lambda w)) \\ - \tfrac{1}{3}\frac{\partial}{\partial t}(u-\lambda w) + \tfrac{1}{6}r^2(r - (u-\lambda w)). \end{aligned}$$

In all of these $u - \lambda w$ and w are evaluated at r and T.

[1] This doesn't mean that it isn't useful, or profitable. This is a much more subtle point.

From the Taylor series expression for Z we find that the yield to maturity is given by

$$-\frac{\log(Z(r,t;T))}{T-t} \sim -a + \left(\tfrac{1}{2}a^2 - b\right)(T-t) + \left(ab - c - \tfrac{1}{3}a^3\right)(T-t)^2 + \cdots$$

for short times to maturity.

The yield curve takes the value $-a(r) = r$ at maturity, obviously. The slope of the yield curve is

$$\tfrac{1}{2}a^2 - b = \tfrac{1}{2}(u - \lambda w),$$

i.e. one half of the risk-neutral drift. The curvature of the yield curve at the short end is proportional to

$$ab - c - \tfrac{1}{3}a^3,$$

which contains a term that is the derivative of the risk-neutral drift with respect to time via c. Let me stress the key points of this analysis. The slope of the yield curve at the short end depends on the risk-neutral drift, and vice versa. The curvature of the yield curve at the short end depends on the time derivative of the risk-neutral drift, and vice versa.

If we choose time-dependent parameters within the risk-adjusted drift rate such that the market prices are fitted at time t^* then we have

$$Z(r^*, t^*; T) = Z_M(t^*; T)$$

which is one equation for the time-dependent parameters.

Thus, for Ho & Lee, for example, the value of the function $\eta^*(t)$ at the short end, $t = t^*$, depends on the slope of the market yield curve. Moreover, the slope of $\eta^*(t)$ depends on the *curvature* of the yield curve at the short end. Results such as these are typical for all fitted models. These, seemingly harmless results, are actually quite profound.

It is common for the slope of the yield curve to be quite large and positive, the difference between very short and not quite so short rates is large. But then for longer maturities typically the yield curve flattens out. This means that the yield curve has a large negative curvature. If one performs the fitting procedure as outlined here for the Ho & Lee or extended Vasicek models, one typically finds the following:

- The value of $\eta^*(t)$ at $t = t^*$ is very large. This is because the yield-curve slope at the short end is often large.
- The slope of $\eta^*(t)$ at $t = t^*$ is large and negative. This is because the curvature of the yield curve is often large and negative.

A typical plot of $\eta^*(t)$ versus t is shown in Figure 31.1. This shows the high value for the fitted function and the large negative slope.[2] So far, so good. Maybe this is correct, maybe this is really what the fitted parameter should look like. But what happens when we come back in a few months to look at how our fitted parameter is doing? If the model is correct then we would find that the fitted curve looked like the bold part of the curve in the figure. The previous data should have just dropped off the end; the rest of the curve should remain unchanged.

[2] The strange oscillation of the function η^* beyond the short end is usually little more than numerical errors.

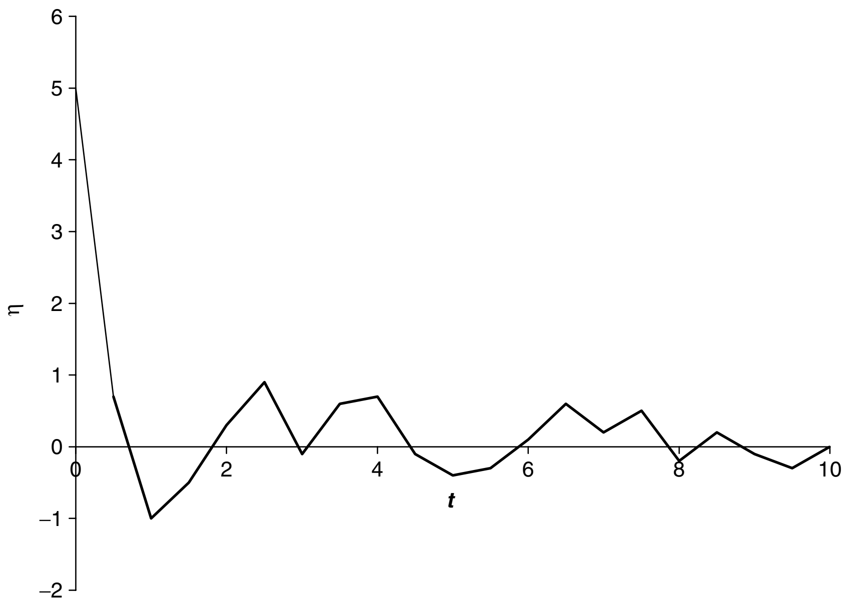


Figure 31.1 Typical fitted function $\eta^*(t)$.

We would then see a corresponding dramatic flattening of the yield curve. Does this in fact happen? No. The situation looks more like that in Figure 31.2, which is really just a translation of the curve in time. Again we see the high value at the short end, the large negative slope and the oscillations. The recalibrated function in Figure 31.2 looks nothing like the bold line in Figure 31.1. This is because the yield curve has not changed that much in the meantime. It still has the high slope and curvature. In fact, we don't even have to wait for a few months for the deviation to be significant; it becomes apparent in weeks or even days.

We can conclude from this that yield curve fitting is an inconsistent and dangerous business. The results presented here are by no means restricted to the models I have named; *no* one-factor model will capture the high slope and curvature that is usual for yield curves; they 'may' give reasonable results when the yield curve is fairly flat. We will discuss these criticisms of yield curve modeling in Chapter 36. In fact, very few models can accommodate the large slope and curvature that is common for yield curves. Exceptions include some Heath, Jarrow & Morton models (HJM, see Chapter 37) and a non-probabilistic model described in Chapter 68.

31.5 OTHER MODELS

Other models for the short-term interest rate have been proposed. One of the most popular (but one for which there are no explicit solutions) is the **Black, Derman & Toy** (**BDT**) model where the risk-neutral spot interest rate satisfies

$$d(\log r) = \left(\theta(t) - \frac{\sigma'(t)}{\sigma(t)}\log r\right) dt + \sigma(t)\, dX.$$

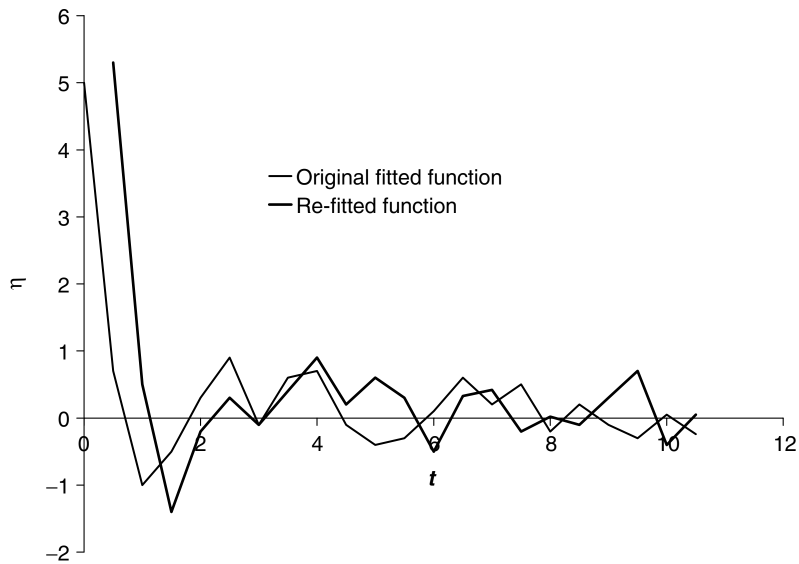


Figure 31.2 Typical *re*-fitted function $\eta^*(t)$, a short time later.

The two functions of time σ and θ allow both zero-coupon bonds and their volatilities to be matched. An even more general model is the **Black & Karasinski** model

$$d \log r = (\theta(t) - a(t) \log r)\, dt + \sigma(t)\, dX.$$

These models are popular because fitting can be done quite simply by a numerical scheme.

Any criticisms of yield curve fitting in general, of course, apply to these models. On the other hand, as we shall see in Chapter 36, the dependence of the volatility of r on r itself is, for these models, not dissimilar to that found in data for US interest rates.

31.6 SUMMARY

I have outlined why the yield curve is fitted, and how it is fitted in some simple models. From a practical perspective it is hard to argue against calibration; you cannot hedge with something if your theoretical price is very different from its traded price. But from a modeling and empirical perspective it is hard to argue in its favor; the data show how inconsistent the concept is. This is always going to be a problem with one-factor Brownian motion models, unless yield curves suddenly decide not to be so steep. There is not a great deal that can be done theoretically. If you are concerned with consistency then avoid one-factor models, or you will end up painting yourself into a corner.

On the other hand, people seem to make money using these models and I guess that is the correct test of a model. Unless you are speculating with an interest rate derivative, you will have to delta hedge and therefore have to calibrate. Practitioners go much further than I have shown here; they fit as many market prices and properties as they can. Put in another time-dependent parameter and you can fit interest rate volatilities of different maturities, yet another parameter and you can fit the market prices of caps. By fitting more and more data, are you digging a deeper and deeper grave or are you improving and refining the accuracy of your model?

As an aside, suppose we are not interested in hedging but want to speculate with some fixed-income instruments. It is common knowledge that the yield curve is a poor predictor of real future interest rates. In this case it could be unnecessary or even dangerous to fit the yield curve. In this situation one could 'value' the instrument using the *real* spot rate process. This would give a 'value' for the instrument that was the expected present value of all cashflows under the *real* random walk. To do this one needs a model for the real drift u. There are ways of doing this, discussed in Chapter 36, that give satisfactory results and they don't need any fitting. This subject of valuing under speculation is the subject of Chapter 59.

FURTHER READING

- A more sophisticated choice of time-dependent parameters is described by Hull & White (1990).
- Klugman & Wilmott (1994) consider the fitting of the general affine model.
- Baker (1977) gives details of the numerical solution of integral equations.
- See Black, Derman & Toy (1990) for details of their popular model.
- Rebonato (1996) discusses calibration in depth for many popular models.
- See Derman (2004) for the story behind the BDT model.

CHAPTER 32
interest rate derivatives

In this Chapter...

- common fixed-income contracts such as bond options, caps and floors
- how to price interest rate products in the consistent partial differential equation framework
- how to price contracts the market way
- path dependency in interest rate products, such as the index amortizing rate swap

32.1 INTRODUCTION

In the first part of this book I derived a theory for pricing and hedging many different types of options on equities, currencies and commodities. In Chapter 30 I presented the theory for zero-coupon bonds, boldly saying that the model may be applied to other contracts.

In the equity options world we have seen different degrees of complexity. The simple contracts have no path dependency. These include the vanilla calls and puts and contracts having different final conditions such as binaries or straddles. At the next stage of complexity we find the weakly path-dependent contracts such as American options or barriers for which, technically speaking, the path taken by the underlying is important, yet for which we only need to solve in two dimensions (S and t). Finally, we have seen strongly path-dependent contracts such as Asians or lookbacks for which we must introduce a new state variable to keep track of the key features of the path of the underlying. Many of these ideas are mirrored in the theory of interest rate derivatives.

In this chapter we delve deeper into the subject of fixed-income contracts by considering interest rate derivatives such as bond options, caps and floors, swaptions, captions and floortions, and more complicated and path-dependent contracts such as the index amortizing rate swap.

32.2 CALLABLE BONDS

As a gentle introduction to more complex fixed-income products, consider the **callable bond**. This is a simple coupon-bearing bond, but one that the issuer may call back on specified dates for a specified amount. The amount for which it may be called back may be time-dependent.

This feature reduces the value of the bond; if rates are low, so that the bond value is high, the issuer will call the bond back.

The callable bond satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV = 0,$$

with

$$V(r, T) = 1,$$

and

$$V(r, t_c^-) = V(r, t_c^+) + K_c,$$

across coupon dates. If the bond can be called back for an amount $C(t)$ then we have the constraint on the bond's value

$$V(r, t) \leq C(t),$$

together with continuity of $\partial V/\partial r$.

32.3 BOND OPTIONS

The stochastic model for the spot rate presented in Chapter 30 allows us to value contingent claims such as bond options. A **bond option** is identical to an equity option except that the underlying asset is a bond. Both European and American versions exist.

As a simple example, we derive the differential equation satisfied by a call option, with exercise price E and expiry date T, on a zero-coupon bond with maturity date $T_B \geq T$. Before finding the value of the option to buy a bond we must find the value of the bond itself.

Let us write $Z(r, t; T_B)$ for the value of the bond. Thus, Z satisfies

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2} + (u - \lambda w)\frac{\partial Z}{\partial r} - rZ = 0 \qquad (32.1)$$

with

$$Z(r, T_B; T_B) = 1$$

and suitable boundary conditions. Now write $V(r, t)$ for the value of the call option on this bond. Since V also depends on the random variable r, it too must satisfy equation (32.1). The only difference is that the final value for the option is

$$V(r, T) = \max(Z(r, t; T_B) - E, 0).$$

This payoff is shown in Figure 32.1.

Figure 32.2 shows the Bloomberg option calculator for bond options. In this case the model used is Black, Derman & Toy.

32.3.1 Market Practice

The above is all well and good, but suffers from the problem that any inaccuracy in the model is magnified by the process of solving once for the bond and then again for the bond option.

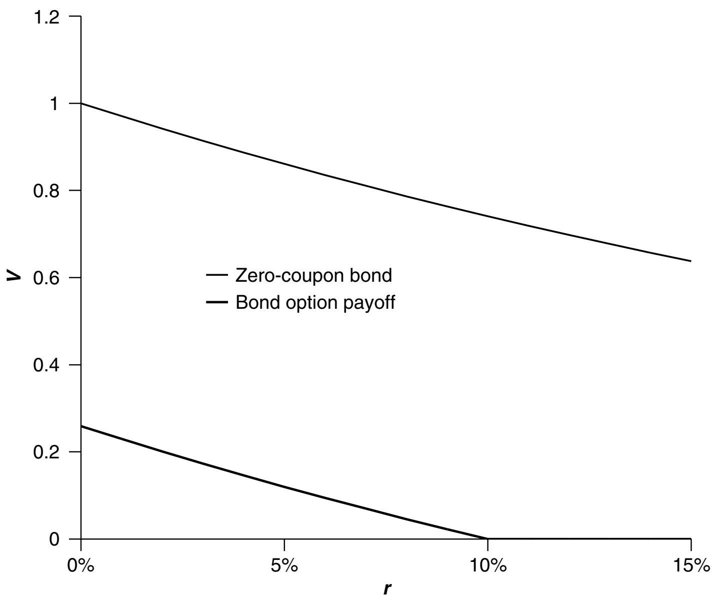


Figure 32.1 Zero-coupon bond price as a function of spot, and the payoff for a call option on the bond.

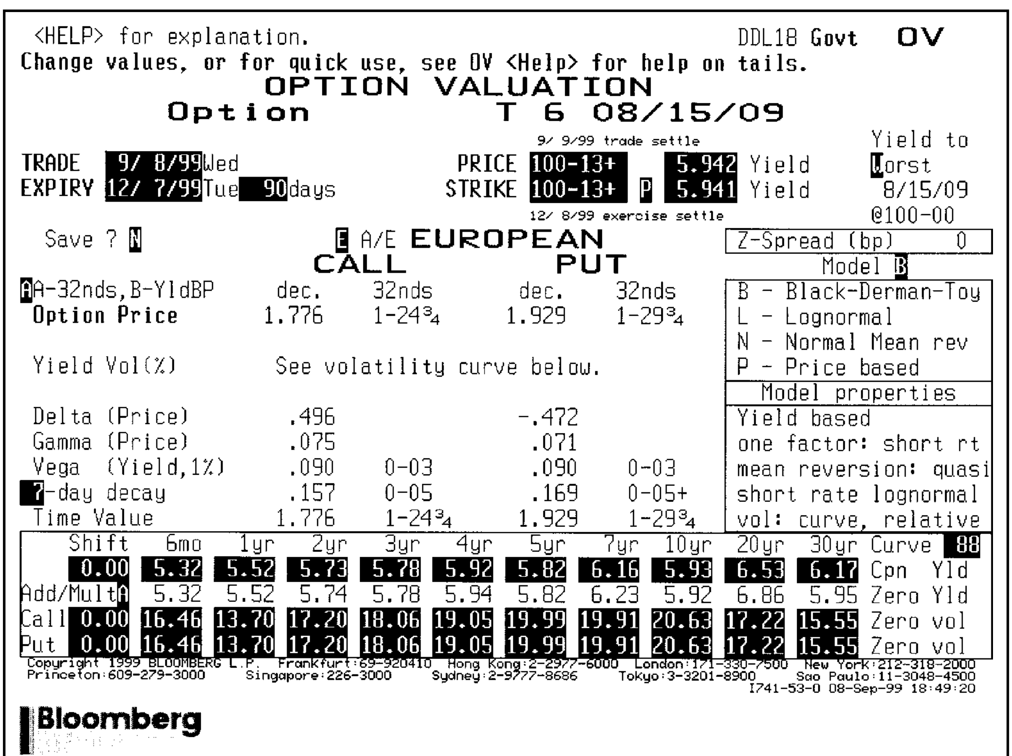


Figure 32.2 Bond option valuation. Source: Bloomberg L.P.

This makes the contract second order, see Chapter 22. When the time comes to exercise the option the amount you receive will, for a call, be the difference between the *actual* bond price and the exercise price, not the difference between the *theoretical* bond price and the exercise price. So the model had better be correct for the bond price. Of course, this model can never be correct, and so we must treat the pricing of bond options with care. Practitioners tend to use an approach that is internally inconsistent but which is less likely to be very wrong. They use the Black–Scholes equity option pricing equation and formulae assuming that the underlying is the bond. That is, they assume that the bond behaves like a lognormal asset. This requires them to estimate a volatility for the bond, either measured statistically or implied from another contract, and an interest rate for the lifetime of the bond option. This will be a good model provided that the expiry of the bond option is much shorter than the maturity of the underlying bond. Over short time periods, well away from maturity, the bond does behave stochastically, with a measurable volatility.

The price of a European bond call option in this model is

$$e^{-r(T-t)}\left(FN(d_1') - EN(d_2')\right),$$

and the put has value

$$e^{-r(T-t)}\left(EN(-d_2') + FN(-d_1')\right),$$

where F is the forward price of the bond at expiry of the option and

$$d_1' = \frac{\log(F/X) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2' = d_1' - \sigma\sqrt{T-t}.$$

This model should not be used when the life of the option is comparable to the maturity of the bond, because then there is an appreciable **pull to par**, that is, the value of the bond at maturity is the principal plus last coupon; the bond cannot behave lognormally close to maturity because we know where it must end up. This contrasts greatly with the behavior of an equity for which there is no date in the future on which we know its value for certain. This pull to par is shown in Figure 32.3.

Another approach used in practice is to model the yield to maturity of the underlying bond. The usual assumption is that this yield follows a lognormal random walk. By modeling the yield and then calculating the bond price based on this yield, we get a bond that behaves well close to its maturity; the pull to par is incorporated.

There is one technical point about the definition of the bond option concerning the meaning of 'price.' One must be careful to use whichever of the clean or dirty price is correct for the option in question. This amounts to knowing whether or not accrued interest should be included in the payoff, see Chapter 13.

32.4 CAPS AND FLOORS

A **cap** is a contract that guarantees to its holder that otherwise floating rates will not exceed a specified amount; the variable rate is thus capped.

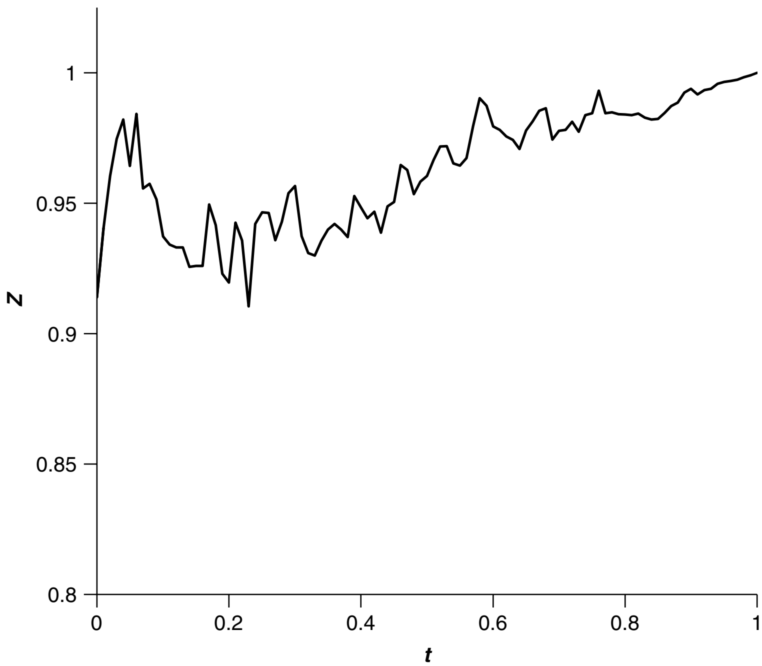


Figure 32.3 The pull to par for a zero-coupon bond.

A typical cap contract involves the payment at times t_i, each quarter, say, of a variable interest on a principal with the cashflow taking the form

$$\max(r_L - r_c, 0),$$

multiplied by the principal (and by the tenor of the interest payments, 0.5 if the payments are semi-annual, for example). Here r_L is the basic floating rate, for example three-month LIBOR if the payments are made quarterly, and r_c is the fixed cap rate. These payments continue for the lifetime of the cap. The rate r_L to be paid at time t_i is set at time t_{i-1}. Each of the individual cashflows is called a **caplet**; a cap is thus the sum of many caplets.

The cashflow of a caplet is shown in Figure 32.4.

If we assume that the actual floating rate is the spot rate i.e. $r_L \approx r$ (and this approximation may not be important) then a single caplet may be priced by solving

$$\frac{\partial V}{\partial t} + \tfrac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV = 0, \tag{32.2}$$

with

$$V(r, T) = \max(r - r_c, 0).$$

Mathematically, this is similar to a call option on the floating rate r.

Figure 32.5 shows the Bloomberg calculator for caps.

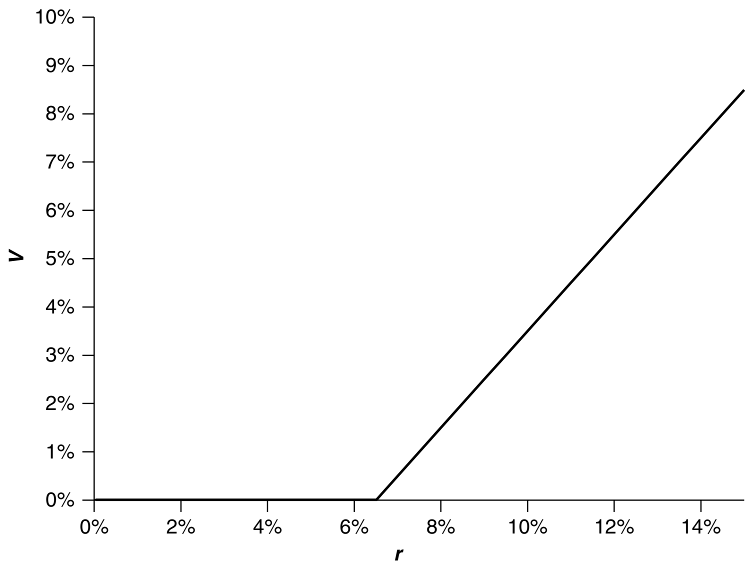


Figure 32.4 The dependence of the payment of a caplet on LIBOR.

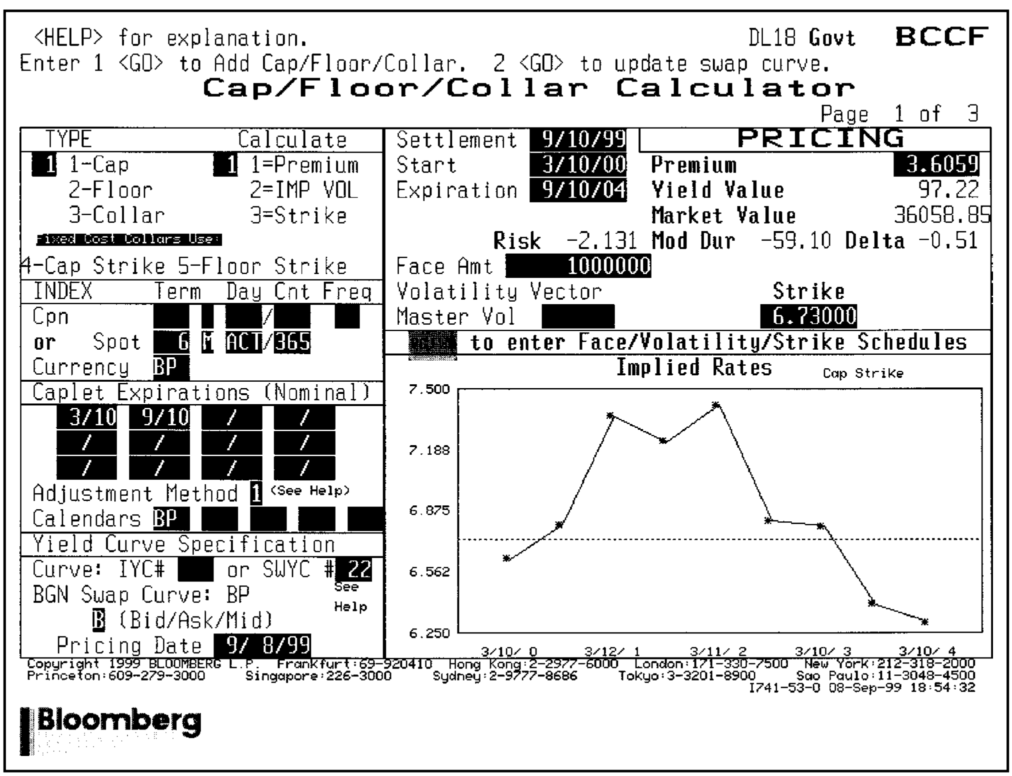


Figure 32.5 Cap/floor/collar calculator. Source: Bloomberg L.P.

A **floor** is similar to a cap except that the floor ensures that the interest rate is bounded below, by r_f. A floor is made up of a sum of floorlets, each of which has a cashflow of

$$\max(r_f - r_L, 0).$$

We can approximate r_L by r again, in which case the floorlet satisfies the bond pricing equation but with

$$V(r, T) = \max(r_f - r, 0).$$

A floorlet is thus a put on the spot rate.

32.4.1 Cap/floor Parity

A portfolio of a long caplet and a short floorlet (with $r_c = r_f$) has the cashflow

$$\max(r_L - r_c, 0) - \max(r_c - r_L, 0) = r_L - r_c.$$

This is the same cashflow as one payment of a swap. Thus there is the model-independent no-arbitrage relationship

$$\text{cap} = \text{floor} + \text{swap}.$$

32.4.2 The Relationship Between a Caplet and a Bond Option

A caplet has the following cashflow

$$\max(r_L - r_c, 0).$$

This is received at time t_i but the rate r_L is set at t_{i-1}. This cashflow is *exactly* the same as the cashflow

$$\frac{1}{1 + r_L\tau}\max(r_L - r_c, 0)$$

received at time t_{i-1}, after all, that is the definition of r_L. Here τ is the time period between interest payments. We can rewrite this cashflow as

$$\max\left(1 - \frac{1 + r_c\tau}{1 + r_L\tau}, 0\right).$$

But

$$\frac{1 + r_c\tau}{1 + r_L\tau}$$

is the price at time t_{i-1} of a bond paying $1 + r_c\tau$ at time t_i. We can conclude that a caplet is equivalent to a put option expiring at time t_{i-1} on a bond maturing at time t_i.

32.4.3 Market Practice

Again, because the Black–Scholes formulae are so simple to use, it is common to use them to price caps and floors. This is done as follows. Each individual caplet (or floorlet) is priced as a call (or put) on a lognormally distributed interest rate. The inputs for this model are the volatility of the interest rate, the strike price r_c (or r_f), the time to the cashflow $t_i - t$, and *two* interest rates. One interest rate takes the place of the stock price and will be the current forward rate applying between times t_{i-1} and t_i. The other interest rate, used for discounting to the present, is the yield on a bond maturing at time t_i. For a caplet the relevant formula is

$$e^{-r^*(t_i-t)}\left(F(t,t_{i-1},t_i)N(d_1') - r_c N(d_2')\right).$$

Here $F(t,t_{i-1},t_i)$ is the forward rate today between t_{i-1} and t_i, r^* is the yield to maturity for a maturity of $t_i - t$,

$$d_1' = \frac{\log(F/r_c) + \frac{1}{2}\sigma^2 t_{i-1}}{\sigma\sqrt{t_{i-1}}} \quad \text{and} \quad d_2' = d_1' - \sigma\sqrt{t_{i-1}}.$$

σ is the volatility of the $(t_i - t_{i-1})$ interest rate. The floorlet value is

$$e^{-r^*(t_i-t)}\left(-F(t,t_{i-1},t_i)N(-d_1') + r_c N(-d_2')\right).$$

32.4.4 Collars

A **collar** places both an upper and a lower bound on the interest payments. It can be valued as a long cap and a short floor.

32.4.5 Step-up Swaps, Caps and Floors

Step-up swaps etc. have swap (cap etc.) rates that vary with time in a prescribed manner.

32.5 RANGE NOTES

The **range note** pays interest on a notional principal for every day that an interest rate lies between prescribed lower and upper bounds. Let us assume that the relevant interest rate is our spot rate r. In this case we must solve

$$\frac{\partial V}{\partial t} + \frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV + \mathcal{I}(r) = 0,$$

with

$$V(r,T) = 0,$$

where

$$\mathcal{I}(r) = r \quad \text{if} \quad r_l < r < r_u \quad \text{and is zero otherwise.}$$

This is only an approximation to the correct value since in practice the relevant interest rate will have a finite (not infinitesimal) maturity.

32.6 SWAPTIONS, CAPTIONS AND FLOORTIONS

A swaption has a strike rate, r_E, that is the fixed rate that will be swapped against floating if the option is exercised. In a call swaption or **payer swaption** the buyer has the right to become the fixed rate payer; in a put swaption or receiver swaption the buyer has the right to become the payer of the floating leg.

Captions and **floortions** are options on caps and floors respectively. These contracts can be put into the partial differential equation framework with little difficulty. However, these contracts are second order, meaning that their value depends on another, more basic, contract, see Chapter 22. Although the partial differential equation approach is possible, and certainly consistent across instruments, it is likely to be time consuming computationally and prone to serious mispricings because of the high order of the contracts.

32.6.1 Market Practice

With some squeezing the Black–Scholes formulae can be used to value European swaptions. Perhaps this is not entirely consistent, but it is certainly easier than solving a partial differential equation.

The underlying is assumed to be the fixed leg of a par swap with maturity T_S, call this r_f. It is assumed to follow a lognormal random walk with a measurable volatility. If at time T the par swap rate is greater than the strike rate r_E then the option is in the money. At this time the value of the swaption is

$$\max(r_f - r_E, 0) \times \text{present value of all future cashflows.}$$

It is important that we are 'modeling' the par rate because the par rate measures the rate at which the present value of the floating legs is equal to the present value of the fixed legs. Thus in this expression we only need worry about the excess of the par rate over the strike rate. This expression looks like a typical call option payoff; all we need to price the swaption in the Black–Scholes framework are the volatility of the par rate, the times to exercise and maturity and the correct discount factors. The payer swaption formula in this framework is

$$\frac{1}{F}e^{-r(T-t)}\left(1-\left(1+\tfrac{1}{2}F\right)^{-2(T_S-T)}\right)\left(FN(d_1')-EN(d_2')\right)$$

and the receiver swaption formula is

$$\frac{1}{F}e^{-r(T-t)}\left(1-\left(1+\tfrac{1}{2}F\right)^{-2(T_S-T)}\right)\left(EN(-d_2')-FN(-d_1')\right)$$

where F is the forward rate of the swap, T_S is the maturity of the swap and

$$d_1' = \frac{\log(F/E)+\frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2' = d_1' - \sigma\sqrt{T-t}.$$

These formulae assume that interest payments in the swap are exchanged every six months.

Figure 32.6 shows the Bloomberg swaption valuation page. This uses the Black model for pricing.

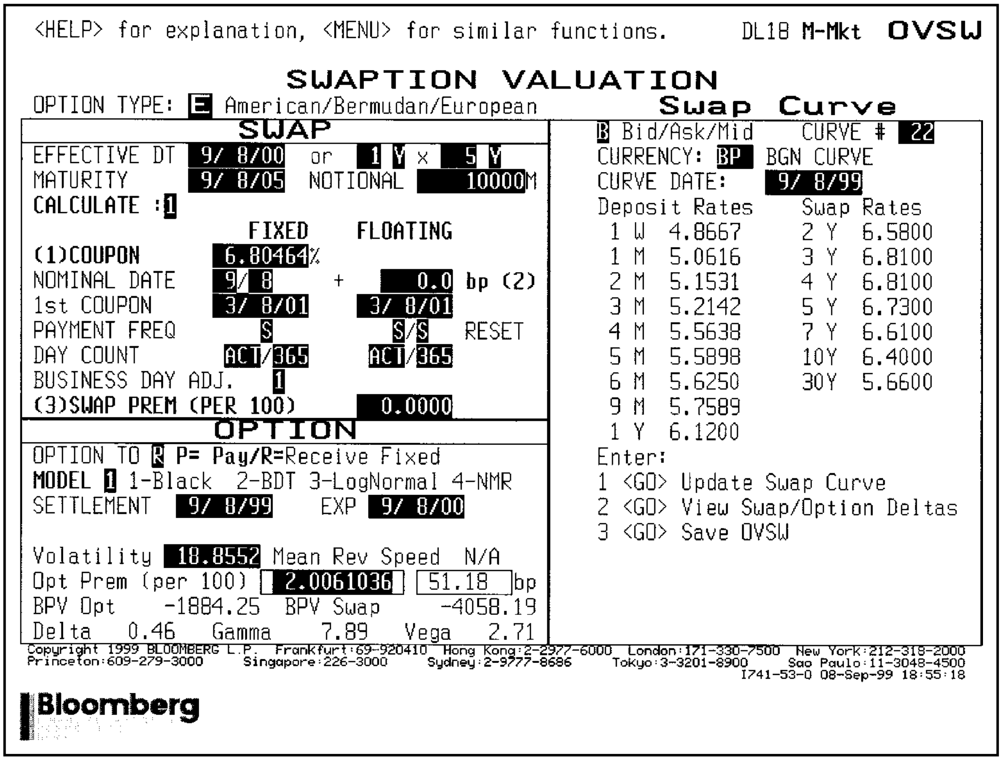


Figure 32.6 Swaption valuation. Source: Bloomberg L.P.

32.7 SPREAD OPTIONS

Spread options have a payoff that depends on the difference between two interest rates. In the simplest case the two rates both come from the same yield curve, more generally the spread could be between rates on different but related curves (yield curve versus swap curve, LIBOR versus Treasury bills), risky and riskless curves (credit derivatives, see Chapter 41) or rates in different currencies.

Can we price this contract in the framework we have built up? No. The contract crucially depends on the tilting of the yield curve. In our one-factor world all rates are correlated and there is little room for random behavior in the spread. We will have to wait until Chapter 35 before we can price such a contract in a consistent framework.

Another method for pricing this contract is to squeeze it into the Black–Scholes-type framework. This amounts to modeling the spread directly as a lognormal (or Normal) variable and choosing suitable rates for discounting. This latter method is the market practice and although intellectually less satisfying it is also less prone to major errors.

32.8 INDEX AMORTIZING RATE SWAPS

A swap is an agreement between two parties to exchange interest payments on some principal; usually one payment is at a fixed rate and the other at a floating rate. In an index amortizing

Table 32.1 Typical amortizing schedule.

Spot rate (%)	Principal reduction (%)
less than 3	100
4	60
5	40
6	20
8	10
over 12	0

rate (IAR) swap the amount of this principal decreases, or **amortizes**, according to the behavior of an 'index' over the life of the swap; typically, that index is a short-term interest rate. The easiest way to understand such a swap is by example, which I keep simple.

Example

Suppose that the principal begins at $10,000,000 with interest payments being at 5% from one party to the other, and $r\%$, the spot interest rate, in the other direction. These payments are to be made quarterly.[1] At each quarter, there is an exchange of $r - 5\%$ of the principal. However, at each quarter the principal may also amortize according to the level of the spot rate at that time. In the Table 32.1 we see a typical amortizing schedule.

We read this table as follows. First, on a reset date, each quarter, there is an exchange of interest payments on the principal as it then stands. What happens next depends on the level of the spot rate. If the spot interest rate (or whatever index the amortization schedule depends on) is below 3% on the date of the exchange of payments then the principal on which future interest is paid is then amortized 100%; in other words, this new level of the principal is zero and thus no further payments are made. If the spot rate is 4% then the amortization is 60%, i.e. the principal falls to just 40% of its level before this reset date. If the spot rate is 8% then the principal amortizes by just 10%. If the rate is over 12% there is no amortization at all and the principal remains at the same level. For levels of the rate between the values in the first column of the table the amount of amortization is a linear interpolation. This interpolation is shown in Figure 32.7 and the function of r that it represents I call $g(r)$.

So, although the principal begins at $10,000,000, it can change after each quarter. This feature makes the index amortizing rate swap path dependent.

The party receiving the fixed rate payments will suffer if rates rise because he will pay out a rising floating rate while the principal does not decrease. If rates fall the principal amortizes and so his lower floating rate payments are unfortunately on a lower principal. Again, he suffers. Thus the receiver of the fixed rate wants rates to remain steady and is said to be selling volatility.

In Figure 32.8 is shown the term sheet for a USD IAR swap. In this contract there is an exchange every six months of a fixed rate and six-month LIBOR. This is a vanilla IAR swap with no extra features and can be priced in the way described above. Terms in square bracket would be set at the time that the deal was made.

We will look at the details of the pricing of these contracts in Chapter 38.

[1] In which case, r would, in practice, be a three-month rate and not the spot rate. The IAR swap is so path-dependent that this difference may not be of major importance.

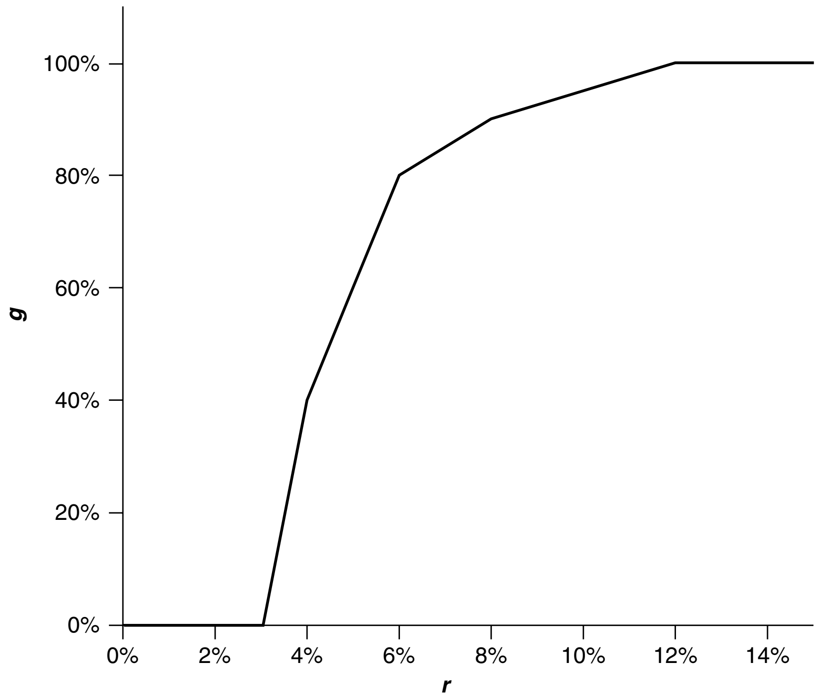
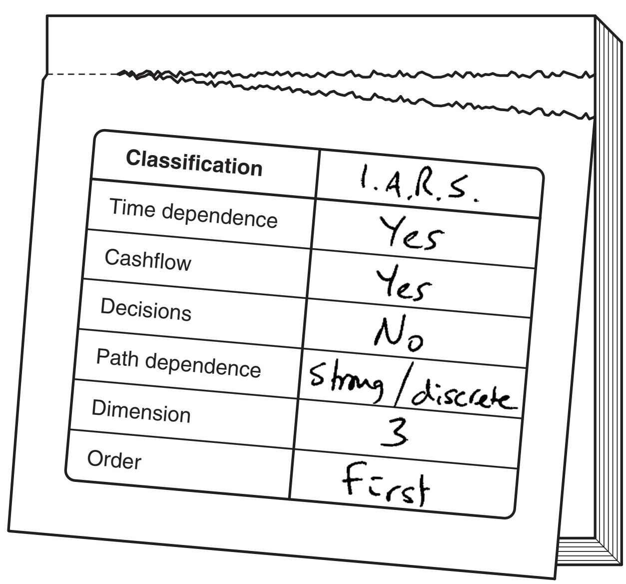


Figure 32.7 A typical amortizing schedule.



Classification option table for IARs.

USD Index Amortizing Swap

Counterparties	XXXX The Customer
Notional Amount	USD 50 millions, subject Amortization Schedule
Settlement Date	Two days after Trade Date
Maximum Maturity Date	Five years after Trade Date
Early Maturity Date	On any Fixing Date leading to a Notional Amount equal to 0
Payments made by Customer	**USD 6m LIBOR** paid semiannually, A/360
Payments made by XXXX	**In USD X% p.a.** paid semiannually, 30/360
Index Rate	USD 6m LIBOR
Base Rate	[ ]%
Amortization Schedule (after 1st coupon period)	(see schedule below)

USD 6m LIBOR – Base Rate	Amortization
−3%	–[ ]%
−2%	–[ ]%
−1%	–[ ]%
0	–[ ]%
1%	0%
2%	0%

NB If the observed difference falls between two entries of this schedule, the amortization amount is interpolated

Fixing Dates	2 business days before each coupon period
USD 6m LIBOR	The USD 6m LIBOR rate as seen on Telerate page 3750 at noon, London time, on each Fixing Date
Documentation	ISDA
Governing Law	English

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 32.8 Term sheet for a USD Index Amortizing Swap.

32.8.1 Other Features in the Index Amortizing Rate Swap

Lockout period

There is often a 'lockout' period, usually at the start of the contract's life, during which there are no reductions in the principal. During this period the interest payments are like those of a vanilla swap. Mathematically, we can model this feature by allowing the amortizing schedule, previously $g(r)$, to be time-dependent: $g(r, i)$. In this case the amount of amortizing depends on the reset date, t_i, as well as the spot interest rate. Such a model can be used for far more sophisticated structures than the simple lockout period. The similarity structure is retained.

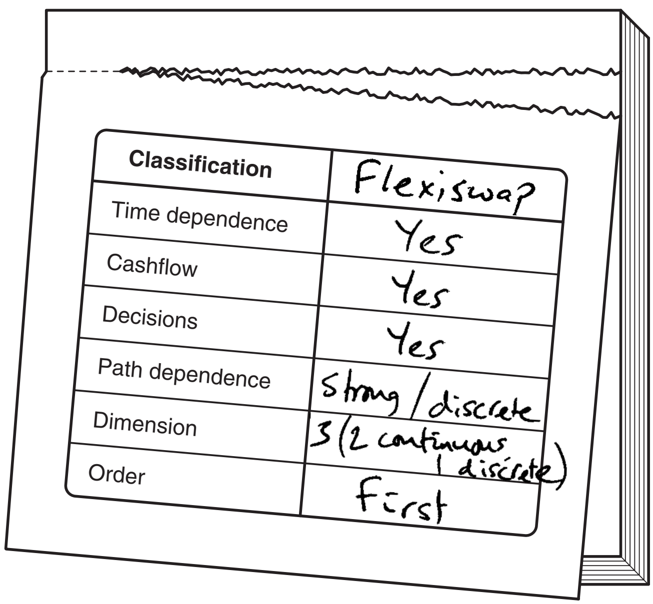
Clean up

Some contracts have that if the principal ever falls to a specified percentage of the original principal then it is reduced all the way to zero. Such a feature, as well as barriers, eliminates the possibility of similarity reductions.

32.9 CONTRACTS WITH EMBEDDED DECISIONS

The following contract is interesting because it requires the holder to make a complex series of decisions on whether to accept or reject cashflows. The contract is path dependent.

This contract, called a **flexiswap**, is a swap with M cashflows of floating minus fixed during its life. The catch is that the holder must choose to accept exactly $m \leq M$ of these cashflows. At each cashflow date the holder must say whether they want the cashflow or not, they cannot later change their mind. When they have taken m of them they can take no more.



Classification option table for Flexiswap.

Before I show how to price this contract, a few words about the representation of the cashflows. If this were a vanilla swap then we would rewrite the floating legs of the swap in terms of zero-coupon bonds and the whole contract could then be priced off the prevailing discount curve. We would certainly not solve any partial differential equations. However, since the present problem is so path dependent it is probably quite safe to make the usual assumption that each cashflow takes the form

$$r - r_f,$$

where r is the spot interest rate.

The trick to pricing this contract is to introduce $m+1$ functions of r and t: $V(r,t,i)$ with $i=0,\dots,m$. The index i represents how many cashflows are still left to be taken. For example, when $i=0$ all the cashflows have been taken up already. Each function will satisfy whichever one-factor interest rate model you fancy.

Since all of its cashflows have been taken up, the function $V(r,t,0)$ is exactly zero. But what about the other functions? Introducing t_j as the available dates on which there are cashflows, with $j=1,\dots,M$, we have two things to ensure. First of all, if we have i cashflows still left to choose when there are only i cashflows left then we must accept all of them. Second, if we have a genuine choice then we must make it optimal.

The first point is guaranteed by the following conditions at cashflow dates:

$$V(r,t_{M-i},i+1)=V(r,t_{M-i},i)+r-r_f \quad \text{for} \quad i=0,\dots,m-1.$$

These equations mean that the last i cashflows are accepted if there are i left to choose.

The optimality of the remaining, earlier, choices is ensured by an American-style constraint,

$$V(r,t_j,i+1)\geq V(r,t_j,i)+r-r_f.$$

Whichever interest rate model you choose, the time taken to solve for the value of this contract will be approximately $m+1$ times the time taken to solve for a single, non-path-dependent contract of the same maturity.

There also exist **flexicaps** and **flexifloors**.

32.10 WHEN THE INTEREST RATE IS *NOT* THE SPOT RATE

We have seen in the interest rate chapters the almost interchangeable use of spot interest rate, one-month rate, three-month rate and even six-month rate. Invariably, the contract in question would specify which interest rate is to be used. If cashflows occur each quarter then the contract almost always specifies that the interest rate is a three-month rate. We have even exploited this fact in Chapter 14 to decompose a swap into a basket of zero-coupon bonds. At other times, I have, in a cavalier fashion replaced a real rate with the impossible-to-observe spot rate. The reason for this is that it is r, the spot rate, that we are modeling. It is nice to have cashflows written in terms of the directly modeled variable.

The following is a rough guide to when to use the theoretically useful spot rate, and when to stay with the correct rate (even though that is usually much, much harder to implement consistently).

Use the spot rate to price a contract if all of the following conditions apply:

- The real rates quoted in the contract details are six months or less.
- The contract has cashflows that are highly non linear in the rate or have early exercise or are path dependent.
- There are no more-liquid instruments available that have cashflows that match our contract

To be on the safe side you should also statically hedge with available instruments to reduce all cashflows and then price the residual cashflows. This will reduce as much as possible any model errors. This is the real problem with making what might appear to be minor approximations: You might inadvertently introduce arbitrage.

32.10.1 The Relationship Between the Spot Interest Rate and Other Rates

The relationship between the spot rate of infinitesimal maturity and rates of finite maturity is through the zero-coupon bond pricing equation. For a general spot-rate model this relationship is complicated, unless the rates are sufficiently short.

For short maturities we can solve the zero-coupon bond pricing equation (30.4) by a Taylor series expansion. We did this at length in Chapter 31 but I will remind you of the result here. Substituting

$$Z(r, t; T) = 1 + a(r)(T - t) + \tfrac{1}{2}b(r)(T - t)^2 + \cdots$$

into (30.4) we find that

$$a(r) = -r$$

and

$$b(r) = r^2 - (u - \lambda w).$$

From this we find that the yield curve is given, for small maturities, by

$$-\frac{\log Z}{T - t} \sim r + \tfrac{1}{2}(u - \lambda w)(T - t) + \cdots \quad \text{as} \quad t \to T. \tag{32.3}$$

As part of our spot rate model we know the risk-adjusted drift $u - \lambda w$, as a function of r and t. Therefore the above gives us an approximate relationship between any short maturity interest rate and the spot rate. The use of this is demonstrated in the following example.

A caplet has a cashflow of

$$\max(r_L - r_c, 0),$$

where r_L is three-month LIBOR, say. We can write this approximately as

$$\max\left(r + \tfrac{1}{8}(u - \lambda w) - r_c, 0\right).$$

The $\frac{1}{8}$ comes from $\frac{1}{2}$ multiplied by the maturity of the three-month rate measured in years. If we believe our one-factor model then we must believe this to be a better representation of the actual cashflow than that given by simply replacing r_L with r.

32.11 SOME EXAMPLES

The term sheet in Figure 32.9 shows details of a Sterling/Deutschmark deconvergence swap. This contract allows the counterparty to express the view that rates between Germany and the UK will widen. Pricing this contract requires models for both UK and German interest rates and the correlation between them.

Figure 32.10 shows a one-year USD fixed rate note with redemption linked to World Bank bonds. The interesting point about this contract is that the issuer of the bond gets to choose whether to redeem at par or to redeem using a choice of three World bank bonds. Obviously the issuer chooses whichever will be cheapest to deliver at the time of redemption. Hence this is an example of a **cheapest-to-deliver** bond.

Sterling/Deutschmark Deconvergence Swap

Start Date	10th March 1999
Maturity Date	10th March 2003
Counterparty 1 payments	
Floating rate	3m DEM LIBOR plus SPREAD
SPREAD	2.35%
Counterparty 2 payments	
Floating rate	3m GBP LIBOR

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 32.9 Term sheet for a Sterling/Deutschmark deconvergence swap.

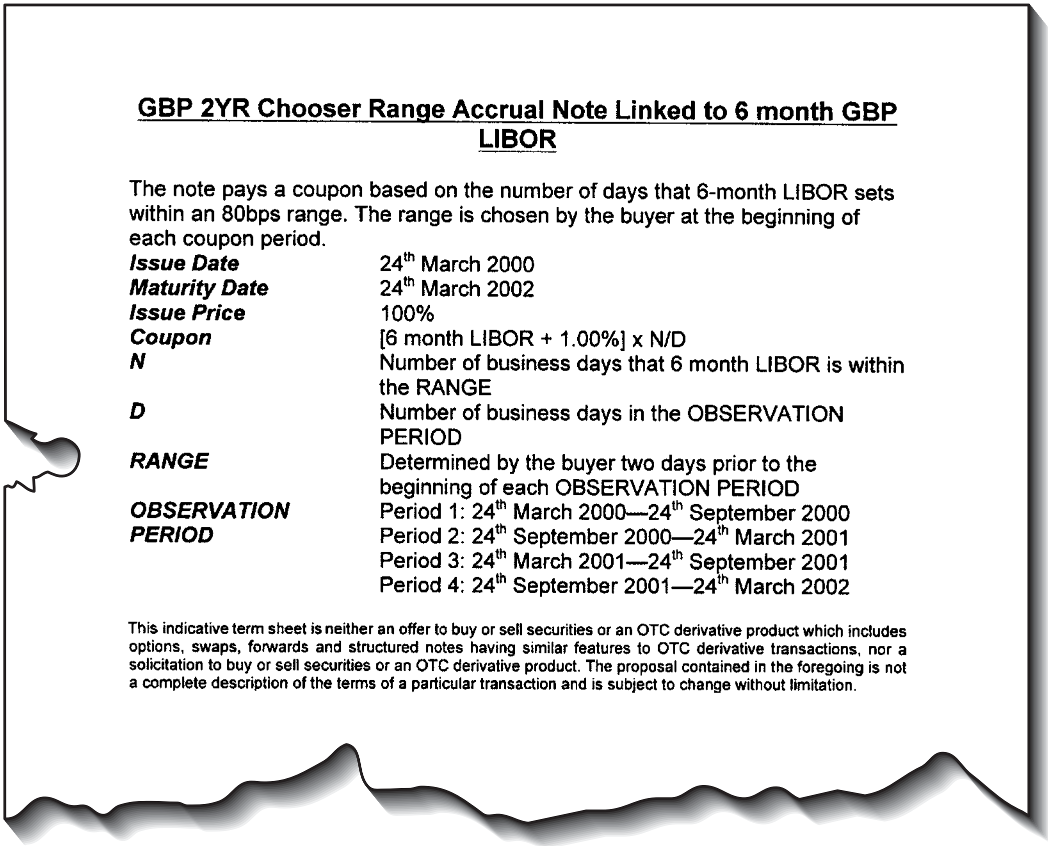
USD 1YR Fixed Rate Note with Redemption linked to World Bank Zero-Coupon Bonds

The note either redeems at par or the investor is delivered World Bank zero-coupon bonds maturing on either 1st March 2009, 15th April 2009 or 15th July 2009. The choice of redemption is at the Issuer's option. If World Bank bonds are delivered, this will be at effective prices of 57.24% for the March bond, 56.85% for the April bond and 56.05% for the July bond.

Issue Date	15 April 2002
Maturity Date	15 April 2003
Issue price	100%
Coupon	7.00%

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 32.10 Term sheet for a USD fixed rate note.



GBP 2YR Chooser Range Accrual Note Linked to 6 month GBP LIBOR

The note pays a coupon based on the number of days that 6-month LIBOR sets within an 80bps range. The range is chosen by the buyer at the beginning of each coupon period.

Issue Date	24th March 2000
Maturity Date	24th March 2002
Issue Price	100%
Coupon	[6 month LIBOR + 1.00%] x N/D
N	Number of business days that 6 month LIBOR is within the RANGE
D	Number of business days in the OBSERVATION PERIOD
RANGE	Determined by the buyer two days prior to the beginning of each OBSERVATION PERIOD
OBSERVATION PERIOD	Period 1: 24th March 2000—24th September 2000 Period 2: 24th September 2000—24th March 2001 Period 3: 24th March 2001—24th September 2001 Period 4: 24th September 2001—24th March 2002

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 32.11 Term sheet for a chooser range accrual note.

The term sheet in Figure 32.11 is of a GBP two-year chooser range accrual note linked to GBP LIBOR. The contract pays a daily coupon equivalent to an annual six-month LIBOR + 100bps. But this is only paid while LIBOR is within an 80 bps range. The novel feature about this range note is that the holder chooses the 80 bps range at the start of each coupon period.

32.12 MORE INTEREST RATE DERIVATIVES

The following require a stochastic interest rate model for their pricing; they are model-dependent.

- **Accordion swap:** A swap (fixed-income instrument, see Chapter 14) whose maturity can be lengthened or shortened at the wish of the holder.
- **Barrier cap/floor:** An interest rate cap or floor with a barrier feature.
- **Basis swap:** A swap in which both legs are floating, of different maturities or currencies, say.

- **Bermudan swaptions:** Bermudan swaptions are like vanilla swaptions in that they give the holder the right to pay (payer swaption) or receive (receiver swaption) the fixed leg of a swap. The Bermudan characteristic allows the holder to exercise into this at specified dates.
- **Bounded cap or floor:** An interest rate cap or floor whose total payout is bounded.
- **Callable swap:** A swap which may be called back by the fixed rate payer.
- **Constant Maturity Swap:** A swap in which one leg is itself a swap rate of constant tenor (rather than the more standard LIBOR rate).
- **LIBOR-in-arrears swap:** LIBOR-in-arrears swap is an interest rate swap in which the floating payment is made at the same time that it is set. In the plain vanilla swap the rate is fixed prior to the payment, so that a swap with six-monthly payments of six-month LIBOR has the floating rate set six months before it is paid. This subtle difference means that the LIBOR-in-arrears swap cannot be decomposed into bonds and the pricing is not model-independent. The analysis that leads to (32.3) shows that there will be a slight difference between the vanilla swap and the LIBOR-in-arrears swap. Since the difference depends on the slope of the forward curve, the LIBOR-in-arrears swap is often thought of as a play on the steepening or flattening of the yield curve.
- **Moving average cap/floor:** An interest rate cap/floor with payout determined by an average interest rate over a period.
- **Putable swap:** A swap which may be called back by the floating rate payer.
- **Ratchets and one-way floaters:** Ratchets and one-way floaters are floating rate notes where the amount of the periodic payments is reset, usually in a monotonically increasing (or decreasing) manner. The amount of the reset will depend on a specified floating interest rate.
- **Reflex cap/floor:** As a cap or floor but with payments depending on a trigger being reached.
- **Reverse floater:** A floating rate note with coupon that rises as the underlying rate falls and vice versa.
- **Rolling cap/floor:** A cap or floor in which the out-of-the-money portion of each payment is carried forward into the next period.
- **Triggers:** Triggers are just like barrier options in that payments are received until (or after) a specified financial asset trades above or below a specified level. For example, the trigger swap is like a plain vanilla swap of fixed and floating until the reference LIBOR rate fixes above/below a specified rate. You can imagine that they come in in and out, up and down varieties.

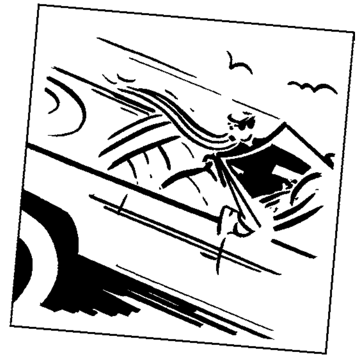
32.13 SUMMARY

There are a vast number of contracts in the fixed-income world. It is an impossible task to describe and model any but a small quantity of these. In this chapter I have tried to show two of the possible approaches to the modeling in a few special cases. These two approaches to the modeling are the consistent way via a partial differential equation or the practitioner way via the Black–Scholes equity model and formulae. The former is nice because it can be made consistent across all instruments, but is dangerous to use for liquid, and high-order contracts. Save this technique for the more complex, illiquid and path-dependent contracts. The alternative approach is, as everyone admits, a fudge, requiring a contract to be squeezed and bent until it looks like a call or a put on something vaguely lognormal. Although completely inconsistent across instruments it is far less likely to lead to serious mispricings.

The reader is encouraged to find out more about the pricing of products in these two distinct ways. Better still, the reader should model new contracts for himself as well.

FURTHER READING

- Black (1976) models the value of bond options assuming the bond is a lognormal asset.
- See Hull & White (1996) for more examples of pricing index amortizing rate swaps.
- Everything by Jamshidian on the pricing of interest rate derivatives is popular with practitioners. See the bibliography for some references.
- The best technical book on interest rate derivatives, their pricing and hedging, is by Rebonato (1996).
- See `http://my.dreamwiz.com/stoneq/products` for a comprehensive list of interest rate derivatives.
- Hagan (2002) explains how to express the volatility risk of exotics in terms of their natural vanilla hedging instruments.
- For an explanation of convexity corrections for several instruments see Hagan (2003).
- See Jäckel (2003) for cap pricing.
- See Berrahoui (2004) for CMS pricing.
- Jäckel & Kawai (2005) present formulae for the prices of interest rate futures contracts allowing for volatility skew.



CHAPTER 33
convertible bonds

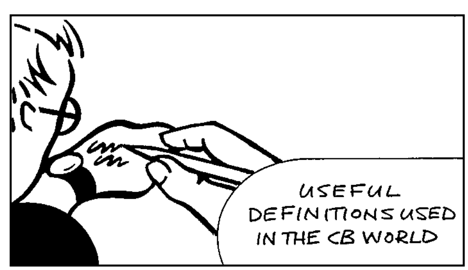
In this Chapter...

- the basic Convertible Bond (CB)
- the purpose of convertible bonds
- market conventions for the pricing and analysis of CBs
- how to price CBs in a one-factor setting
- how to price CBs in a two-factor setting
- features that can make CBs path-dependent

33.1 INTRODUCTION

The conversion feature of convertible bonds makes these contracts similar mathematically to American options. They are also particularly interesting because of their dependence on a stock price and on interest rates. I begin this chapter with some definitions used by the market in the analysis of convertible bonds. I then explain the contract terms of these bonds in a one-factor, stochastic asset price setting. Finally, I show how to price convertibles in a two-factor world of stochastic asset and stochastic interest rate.

33.2 CONVERTIBLE BOND BASICS



The **convertible bond** or **CB** on a stock pays specified coupons with return of the principal at maturity, *unless* at some previous time the owner has converted the bond into the underlying asset. A convertible bond thus has the characteristics of an ordinary bond but with the extra feature that the bond may, at a time of the holder's choosing, be exchanged for a specified asset. This exchange is called **conversion**.

The value of a CB is clearly bounded below by both

- its **conversion value**, which is the amount received if the bond is converted immediately (regardless of whether this is optimal)

$$\text{conversion value} = \text{market price of stock} \times \text{conversion ratio}.$$

- its value as a corporate bond, with a final principal and coupons during its life. This is called its **straight value**.

The latter point shows how there are credit risk issues in the pricing of CBs. These credit risk issues are very important but we will not discuss them in any detail here. We assume that there is no risk of default. Credit risk is discussed in depth in Chapter 40.

33.2.1 What are CBs for?

Companies issue bonds for the purpose of raising capital. Corporate bonds come in many, many varieties. At one end there are the simple vanilla bonds, with fixed coupon and repayment of the principal at maturity. Most corporate bonds have a call provision so that the issuing company can call them back prior to maturity. They would do this if interest rates were to fall sufficiently below the then prevailing market rates. Some bonds have put provisions so that the holder can sell them back to the issuer prior to maturity. Because of the possibility of default corporate bonds are ranked according to precedence in receiving assets in the event of bankruptcy.

Convertible bonds are just another type of corporate bond. From the issuer's perspective they have advantages over more vanilla bonds in that should the company be successful, their share price will rise, the holder will convert the bond and the bond issuer will issue new shares in the company. Issuing new shares may be considered a better option than the continual payment of interest. The bond holder buys the CB to gain some exposure to the company with some downside protection. Of course, we are in the world of commercial paper and there will always be the risk of default.

Most convertibles are issued by smaller companies. Often the coupon is low. The small company has low interest rate obligations in the short term. In the long term, if they are successful, they will hand over shares to happy investors. On the other hand, CBs can be issued by more speculative firms who would then be tempted to take risks with the capital so raised. What is their downside?

The convertible bond has characteristics that make it sometimes behave like stock, and sometimes like a bond. The conversion feature of convertible bonds also makes these contracts similar to American options. The question of when to exercise an American option is very similar to the question of when to convert a convertible bond. It is this 'optionality' that adds value to the convertible bond. CBs are different from options, however, in that on conversion, new shares are issued.

The CB is an example of a **hybrid** instrument since it has features of both equity and debt.

33.2.2 The Issuers of CBs

- CBs are issued by corporations.
- These companies are often not of the highest quality, in terms of credit risk.

By selling bonds which can be later converted to equity the issuer can get away with a lower coupon than might otherwise be expected.

33.2.3 Why Issue a Convertible?

Corporations in need of capital have many choices available to them. These choices have two common themes: Issue equity; Issue debt.

- **Issue equity:** Dilutes earnings per share but has low initial financing costs.
- **Issue debt:** Does not dilute earnings per share but may have high initial financing costs.

The other possibility is to issue a hybrid instrument with both of these features: The Convertible Bond.

- The bond can be sold with a lower coupon than a plain bond with the same maturity and price. (Or equivalently, it can be sold for a higher price with the same coupon.)
- It does not dilute earnings per share, until the bond is converted and new shares are issued.
- If the bond is converted, the principal does not have to be repaid.

The typical issuer of CBs has high cash requirements, perhaps with a very rapid use of cash and low credit quality. They may be startups. In the US only 30% of the CB market is Investment Grade (but rising). In Europe and Japan this figure is 85%.

Examples

Technology/Media/Telecommunications (TMT) and Biotechnology companies, high burn rate (use of cash), high volatility, high risk of default. (High volatility increases price of CBs, high risk of default lowers the price.)

33.2.4 Why Buy a Convertible?

Convertible bonds can be very attractive investments, for the following reasons.

- Upside participation with downside protection. (Unless there is a default.)
- Coupon is typically greater than the dividend yield of the equity.
- Some investors may be barred from participating directly in the equity market. The debt nature of the instrument may make it appealing.
- Legally rank, being debt securities, above equity in case of default.

33.2.5 Some Statistics

CBs are no longer bought by the private investor; hedge funds have dominated this market for some time now.

- Globally, capitalization of CBs is $500 billion.
- 400 hedge funds focused on CB arbitrage (out of 7000 hedge funds in total).
- CB arbitrage is the third best (since 1993) hedge fund strategy after Equity Market Neutral and Event Driven, with a Sharpe ratio of 1.04. (N.B. S&P500 Sharpe ratio was 1.18 over same period.)
- Hedge funds hold 70% of CBs.

That was then, this is now. CB arbitrage has not been a profitable strategy for a while. So many hedge funds have jumped onto this bandwagon that the wheels have dropped off. Arbitrage possibilities have evaporated. Anyway . . .

33.3 MARKET PRACTICE

As we have just seen, the value of a CB is at least its conversion value. The bond component of the CB pushes the value up. We can thus calculate the **market conversion price**. This is the amount that the CB holder effectively pays for a stock if he exercises the option to convert immediately:

$$\text{market conversion price} = \frac{\text{market price of CB}}{\text{conversion ratio}} \geq \text{price of underlying.}$$

The purchaser of the CB pays a premium over the market price of the underlying stock. This is measured by the **market conversion premium ratio**:

$$\text{market conversion premium ratio} \\ = \frac{\text{market conversion price} - \text{current market price of the CB}}{\text{market price of underlying stock}}.$$

By holding the underlying stock you will receive any dividends; by holding the CB you get the coupons on the bond but not the dividends on the stock. A measure of the benefit of holding the bond is the **favorable income differential**, measured by

$$\text{favorable income differential} \\ = \frac{\text{coupon interest from bond} - (\text{conversion ratio} \times \text{dividend per share})}{\text{conversion ratio}}.$$

After calculating this we can estimate the **premium payback period** i.e. how long it will take to recover the premium over the stock price:

$$\text{premium payback period} = \frac{\text{market conversion price} - \text{current market price of the CB}}{\text{favorable income differential}}.$$

This does not, of course, allow for the time value of money.

For small values of the underlying stock the possibility of conversion is remote and the CB trades very much like a vanilla bond. If the stock price rises high enough the CB will be converted and so begins to act like the stock.

Figure 33.1 shows the details of a particular convertible bond. Figure 33.2 shows the calculation of the price.

33.4 CONVERTS AS OPTIONS

The 'payoff' for a convertible is similar to that for a vanilla call option. In some special cases the convertible can be decomposed into a pure bond and a vanilla call on the stock. For this to be valid the CB cannot be callable or putable. To make things simple let's also assume that there are no dividends on the stock. There are three steps to pricing CBs as if they are options:

1. Calculate the value of the straight bond component.

```
                                                        DL18 Corp   ODE5
Enter 10 <GO> for News, 11 <GO> for Involved Parties
                       SECURITY  DISPLAY                      PAGE    1/ 3
SOL MELIA EUROPE SOLSM 1 09/15/04    99.3750/99.8750   (3.38/3.28) BGN  @18:00
CONV TO          66.6489 SHRS(PER    1000.0)SOL   (MADR) €12.33  (  0.06)DP 100%
CONVERTIBLE UNTIL  9/ 8/ 4                       ISSUER INFORMATION
 SECURITY INFORMATION                            NAME    SOL MELIA EUROPE, BV
 CPN FREQ        ANNUAL           NOT SEASONED   TYPE    INDUSTRIAL
 CPN TYPE        FIXED             IDENTIFICATION #'s    REDEMPTION INFO
 MTY/REFUND TYP  CONVERTIBLE       ISIN     XS0101633310 MATURITY DT   9/15/04
 CALC TYP (   1)STREET CONVENTION  MLNUM FC287           REFUNDING DT
 DAY COUNT( 1) ACT/ACT                                   NEXT CALL DT
 MARKET ISS      EURO NON-DOLLAR   COMMON    010163331   WORKOUT DT    9/15/04
 COUNTRY/CURR    ES  /EUR                                RISK FACTOR   4.8215
 COLLATERAL TYP  COMPANY GUARNT         ISSUANCE INFO    REDMP VAL    112.0200
 AMT ISSUED           200,000(M)   ANNOUNCE DT    8/31/99      RATINGS
 AMT OUTSTAND         200,000(M)   1ST SETTLE DT  9/15/99  MOODY   NA
 MIN PC/INC       1,000/     1,000 1ST CPN DT     9/15/00  S & P   NA
 PAR AMT               1,000.00    INT ACCRUE DT  9/15/99  COMP    NR
 LEADMGR-BOOKS WDR-sole            PRICE @ ISSUE  100      DCR     NA
 APPLIED       LUXEMBOURG                                  FI      NA

NOTES  NO PROSPECTUS
CALL @ACC VAL FROM 9/29/02 ONLY IF CM STK @LST=€19.5052 FOR 20/30 T/D. PRX/SHR=
€15.004. INIT CVR PREM=21%. CVTS FROM 10/25/99..GTD BY SOL MELIA. SR, UNSEC'D.
Copyright 1999 BLOOMBERG L.P.  Frankfurt:69-920410  Hong Kong:2-2977-6000  London:171-330-7500  New York:212-318-2000
Princeton:609-279-3000   Singapore:226-3000   Sydney:2-9777-8686   Tokyo:3-3201-8900   Sao Paulo:11-3048-4500
                                                              I741-53-0 08-Sep-99 18:59:09
Bloomberg
```

Figure 33.1 CB info. Source: Bloomberg L.P.

2. Calculate the strike of the vanilla call.
3. Use Black–Scholes to calculate the value of the option component.

Example

- Five-year CB, no call or put, principal \$1000
- 2% coupon, paid in two semi-annual instalments
- Conversion ratio of 20
- Stock at \$47
- Five-year continuously compounded interest rate 5% (equivalent to 5.06% semi-annual)
- Stock volatility 22%

Step 1: Straight bond component The present value (at a rate of 5% continuously compounded) per annum of the \$1000 is \$786.59. The present value of all the coupons adds up to \$79.59.

The value of the CB as a straight bond is therefore \$866.18.

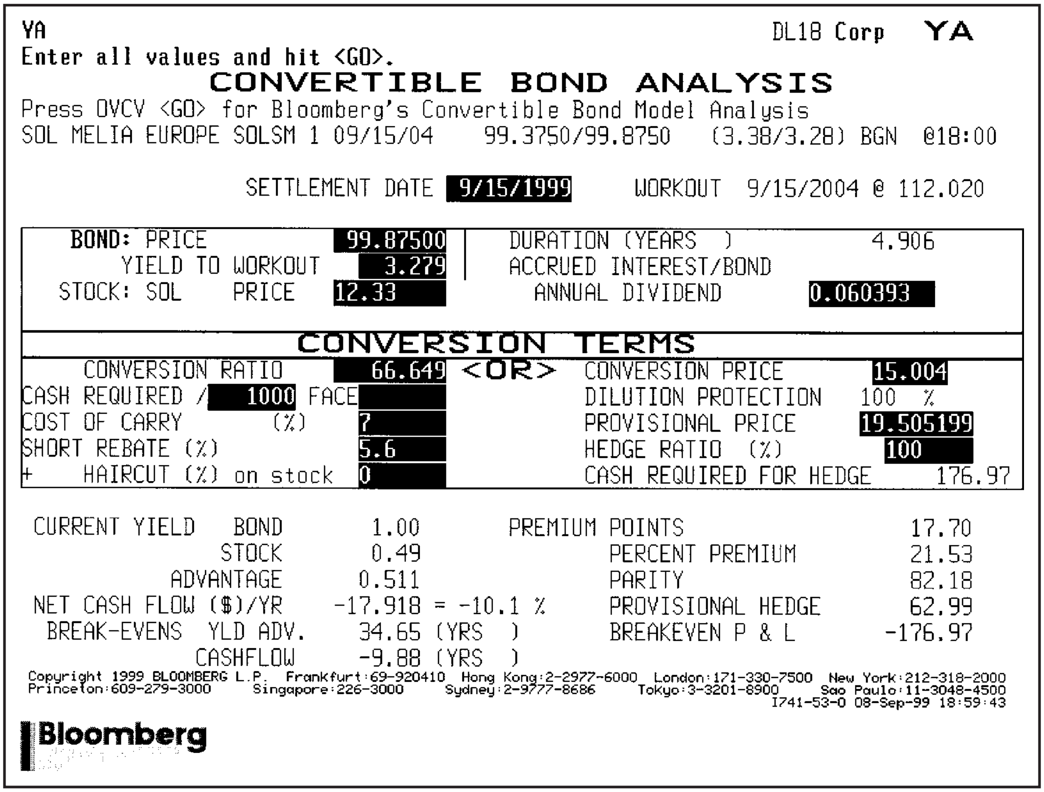


Figure 33.2 CB calculator. Source: Bloomberg L.P.

Step 2: The strike The effective strike is simply the principal divided by the conversion ratio:

$$\frac{1000}{20} = 50.$$

Step 3: The option What is the value of a call option with a strike of \$50 expiring in five years when the underlying is \$47, the volatility is 22% and the risk-free rate is 5%?

Plug these numbers into a Black–Scholes calculator and you will get \$12.97 per share, or \$259.40 for the 20 shares.

Therefore the value of the CB is the sum of the straight bond and option components:

$$866.18 + 259.40 = 1125.58.$$

Note that the delta of the option is 0.74 so the CB is behaving more like a stock than a bond.

That technique for pricing is fine as far as it goes, but it leaves a lot to be desired when we start to introduce features that are commonly seen in convertibles. So now let's build up a flexible, consistent general theory, but still in the same framework of Black and Scholes.

33.5 PRICING CBs WITH KNOWN INTEREST RATE

I will continue to use S to mean the underlying asset price, the maturity date is T and the CB can be converted into n of the underlying. To introduce the ideas behind pricing convertibles I will start by assuming that interest rates are deterministic for the life of the bond. Since the bond value depends on the price of that asset we have

$$V = V(S, t);$$

the contract value depends on an asset price and on the time to maturity. Repeating the Black–Scholes analysis, with a portfolio consisting of one convertible bond and $-\Delta$ assets, we find that the change in the value of the portfolio is

$$d\Pi = \frac{\partial V}{\partial t}\,dt + \frac{\partial V}{\partial S}dS + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\,dt - \Delta\,dS.$$

As before, choose

$$\Delta = \frac{\partial V}{\partial S}$$

to eliminate risk from this portfolio.[1]

The return on this risk-free portfolio is at most that from a bank deposit and so

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (rS - D(S,t))\,\frac{\partial V}{\partial S} - rV \le 0. \qquad (33.1)$$

This inequality is the basic Black–Scholes inequality. Scaling the principal to \$1, the final condition is

$$V(S, T) = 1.$$

Coupons are paid discretely every quarter or half year and so we have the jump condition across each coupon date

$$V(S, t_c^-) = V(S, t_c^+) + K,$$

where K is the amount of the discrete coupon paid on date t_c.

Since the bond may be converted into n assets we have the constraint

$$V \ge nS.$$

In addition to this constraint, we require the continuity of V and $\partial V/\partial S$.

The early conversion feature makes the convertible bond similar to an American option problem; mathematically, we have another free boundary problem. It is interesting to note that

[1] With so many convertibles in the hands of hedge funds, and with these hedge funds delta hedging we find a rather interesting phenomenon. The funds must sell stock in order to be delta neutral, and must sell more stock as the share price rises. Stock rises, sell stock. Conversely, if the share price falls they buy it back. Stock falls, buy stock. The end result is that volatility tends to be dampened down. See Chapter 61 for the maths.

the final data do not satisfy the pricing constraint. Thus, although the value *at* maturity may be \$1 the value *just before* is

$$\max(nS, 1).$$

In the absence of credit risk issues, boundary conditions are

$$V(S, t) \sim nS \quad \text{as} \quad S \to \infty$$

and on $S = 0$ the bond value is the present value of the principal and all the coupons:

$$V(0, t) = e^{-r(T-t)} + \sum K e^{-r(t_c-t)}.$$

Here the sum is taken over all future coupons. As said earlier, when the asset price is high the bond behaves like the underlying, but for small asset price it behaves like a non-convertible bond.

In Figures 33.3, 33.4 and 33.5 are shown the values of a convertible bond with $n = 1$, $r = 0.06$, $\sigma = 0.25$ and with two years before maturity.

In Figure 33.3 there are no coupons paid on the bond and no dividend paid on the underlying. As with the American call on a stock paying no dividends it is not optimal to exercise/convert until expiry. The fine line is the stock price, into which the CB can convert. The CB value is always above this line.

In Figure 33.4 there is a continuous dividend yield of 5%. In this case there is a free boundary, marked on the figure: for sufficiently large S the bond should be converted. The bold line is the bond value, the fine lines are the stock and the bond price without the dividends.

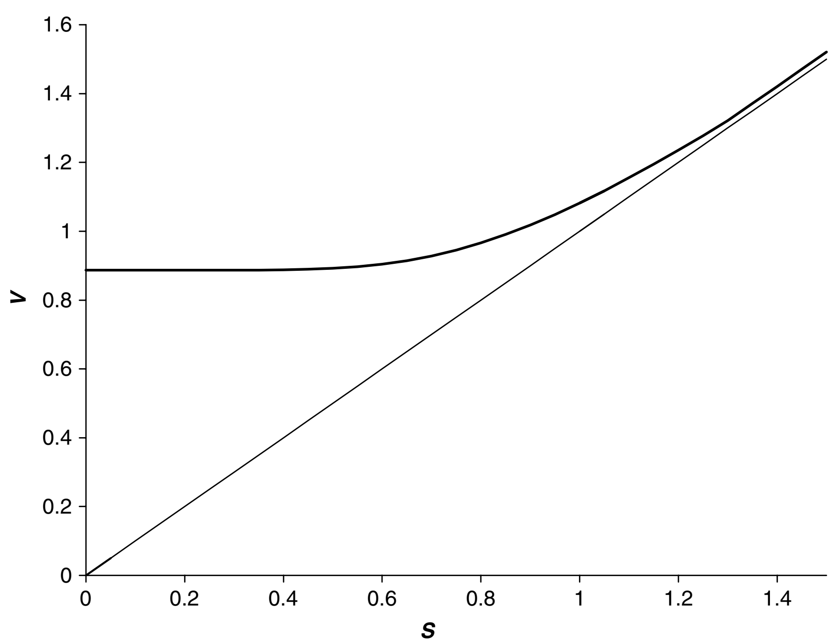


Figure 33.3 The value of a convertible bond with constant interest rate. Zero coupon and dividend. See text for details.

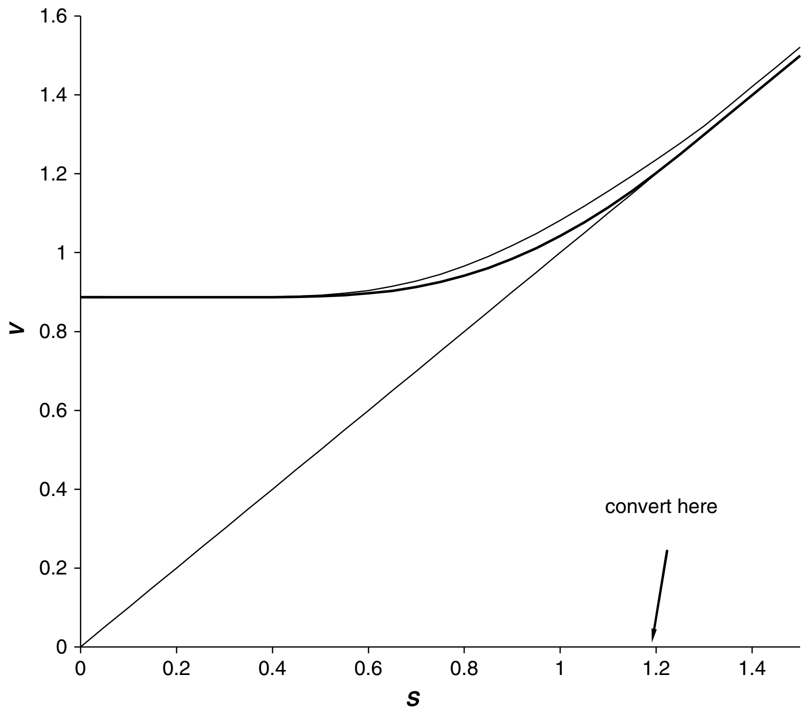


Figure 33.4 The value of a convertible bond with constant interest rate and a dividend paid on the underlying. See text for details.

In Figure 33.5 there is no dividend on the underlying but a coupon of 3% paid twice a year. The value of the bond is higher than in the previous two examples. Should this ever be converted before maturity? The fine lines are the stock and the bond value with no coupons for comparison.

The number of the underlying into which the bond may be converted, n, can be time-dependent. Sometimes the bond may only be converted during specified periods. This is called **intermittent conversion**. If this is the case then the constraint only needs to be satisfied during these times; at other times the contract is European.

33.5.1 Call and Put Features

The convertible bond permits the holder to exchange the bond for a certain number of the underlying asset at any time of their choosing. CBs often also have a **call feature** which gives the issuing company the right to purchase back the bond at any time (or during specified periods) for a specified amount. Sometimes this amount varies with time. The bond with a call feature is clearly worth less than the bond without. This is modeled exactly like US-style exercise again.

If the bond can be repurchased by the company for an amount M_C then elimination of arbitrage opportunities leads to

$$V(S, t) \le M_C.$$

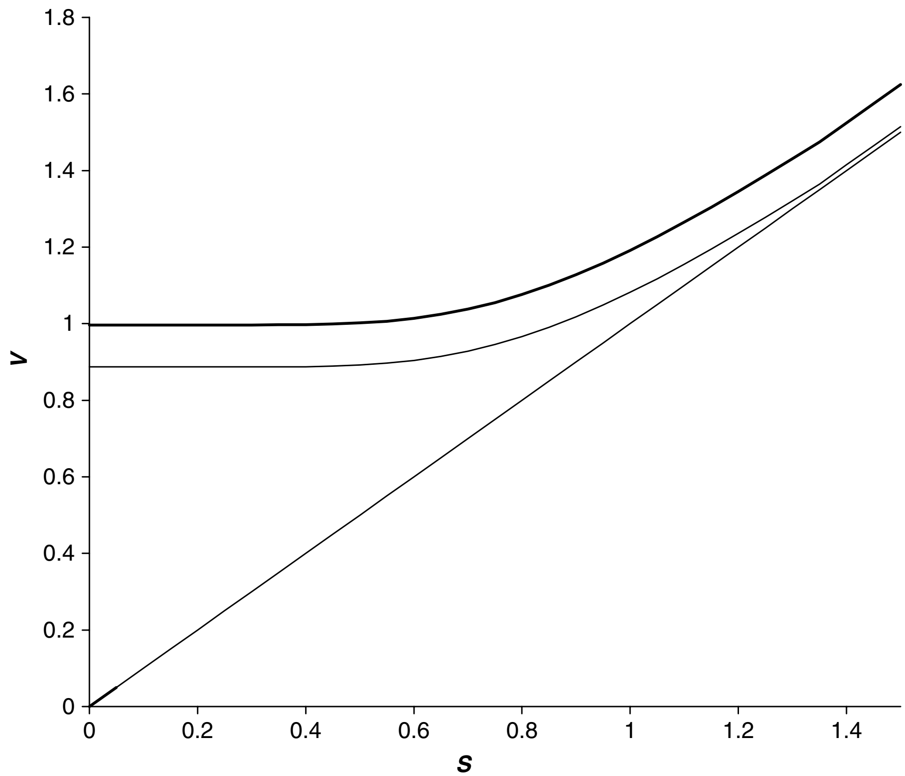


Figure 33.5 The value of a convertible bond with constant interest rate and coupons. See text for details.

The value M_C can be time dependent. Now we must solve a constrained problem in which our bond price is bounded below by nS and above by M_C. To eliminate arbitrage and to optimize the bond's value, V and $\partial V/\partial S$ must be continuous. As with the intermittent conversion feature it is also simple to incorporate the intermittent call feature, according to which the company can only repurchase the bond during certain time periods.

In Figure 33.6 is shown an example of a CB value against the underlying where the effect of a call feature is apparent. Here the bond can be called back at any time for 1.2. The fine line is the same bond without the call feature.

Some convertible bonds incorporate a **put feature**. This right permits the holder of the bond to return it to the issuing company for an amount M_P, say. The value M_P can be time-dependent. Now we must impose the constraint

$$V(S, t) \geq M_P.$$

This feature increases the value of the bond to the holder.

In Figure 33.7 is shown an example of a CB value against the underlying where the effect of the put feature is apparent. The bond is puttable for 0.95 at any time. The fine line is the bond without the put feature.

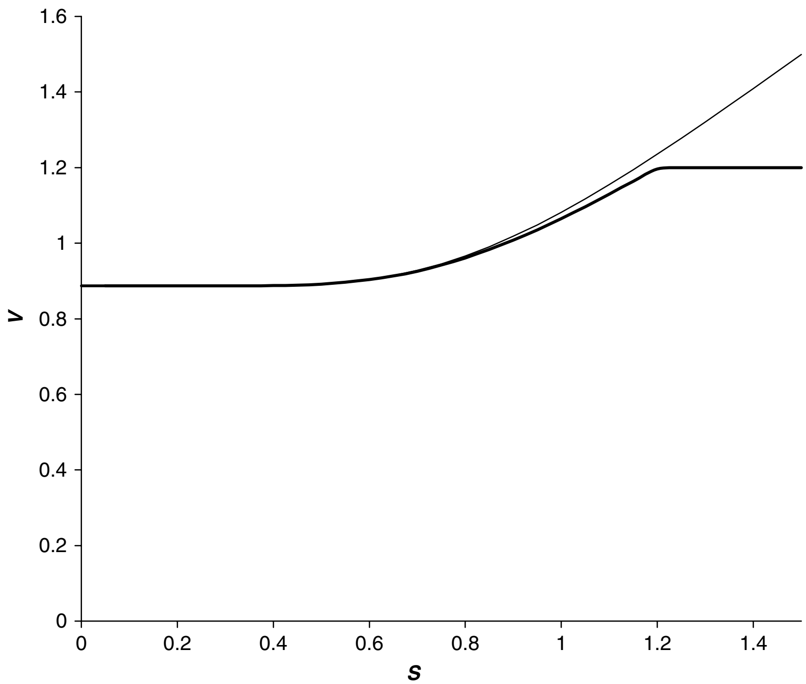


Figure 33.6 Value of a CB against the underlying when the bond is callable.

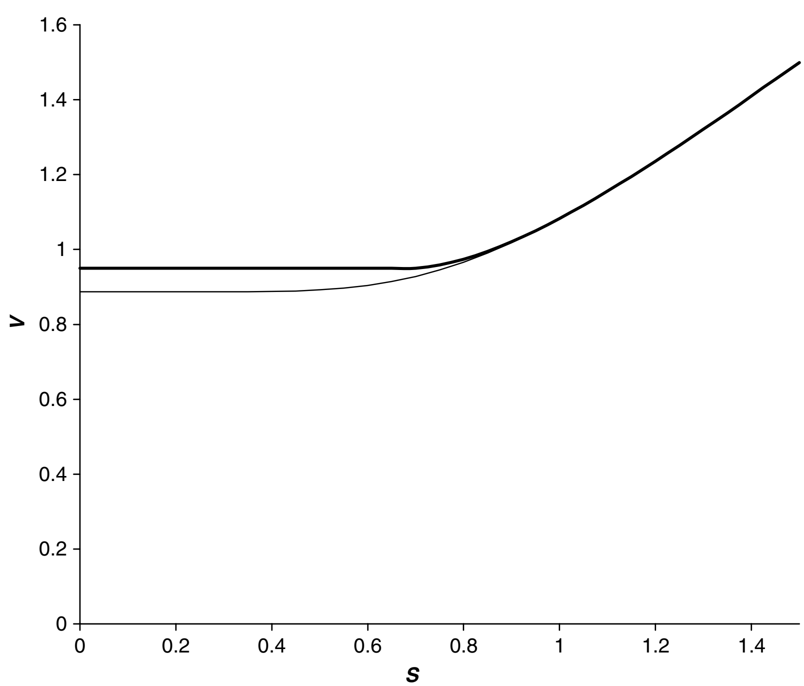


Figure 33.7 Value of a CB against the underlying when the bond is putable.

33.6 TWO-FACTOR MODELING: CONVERTIBLE BONDS WITH STOCHASTIC INTEREST RATE

The lifespan of a typical convertible is much longer than that for a traded option. It is therefore safer to price CBs using a stochastic interest rate model. When interest rates are stochastic, the convertible bond has a value of the form

$$V = V(S, r, t).$$

Before, r was just a parameter, now it is an independent variable.

We continue to assume that the asset price is governed by the lognormal model

$$dS = \mu S\, dt + \sigma S\, dX_1, \tag{33.2}$$

and the interest rate by

$$dr = u(r, t)\, dt + w(r, t)\, dX_2. \tag{33.3}$$

Eventually we are going to be finding solutions numerically, so we allow u and w to be any functions of r and t. Observe that in (33.2) and (33.3) there are two Wiener processes. This is because S and r are governed by two different random variables; this is a **two-factor model**. dX_1 and dX_2 are both still drawn from Normal distributions with zero mean and variance dt, but they are not the same random variable. They may, however, be correlated and we assume that

$$E[dX_1\, dX_2] = \rho\, dt,$$

with $-1 \leq \rho(r, S, t) \leq 1$. The subject of correlated random walks was discussed in Chapter 11. There we saw other examples of **multi-factor model**, in which there are two (or more) sources of risk and hence two independent variables in addition to t. The theory of functions of several random variables was covered in Chapter 11; recall that Itô's lemma can be applied to functions of two or more random variables. The usual Taylor series expansion together with a few rules of thumb results in the correct expression for the small change in any function of both S and r. These rules of thumb are

- $dX_1^2 = dt$;
- $dX_2^2 = dt$;
- $dX_1\, dX_2 = \rho\, dt$.

The equation for dV is

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial r}dr + \frac{1}{2}\left(\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + 2\rho\sigma S w \frac{\partial^2 V}{\partial S \partial r} + w^2 \frac{\partial^2 V}{\partial r^2}\right) dt.$$

Now we come to the pricing of the convertible bond. Construct a portfolio consisting of the convertible bond with maturity T_1, $-\Delta_2$ zero-coupon bonds with maturity date T_2 and $-\Delta_1$ of the underlying asset. We are therefore going to hedge both the interest rate risk and the underlying asset risk. Thus

$$\Pi = V - \Delta_2 Z - \Delta_1 S.$$

The analysis is much as before; the choice

$$\Delta_2 = \frac{\partial V}{\partial r} \bigg/ \frac{\partial Z}{\partial r}$$

and

$$\Delta_1 = \frac{\partial V}{\partial S}$$

eliminates risk from the portfolio. Terms involving T_1 and T_2 may be grouped together separately to find that

$$\boxed{\begin{aligned}&\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S w \frac{\partial^2 V}{\partial S \partial r} + \tfrac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2} \\ &\quad + rS\frac{\partial V}{\partial S} + (u - \lambda w)\frac{\partial V}{\partial r} - rV = 0.\end{aligned}} \tag{33.4}$$

where again $\lambda(r, S, t)$ is the market price of interest rate risk. This is exactly the same market price of risk as for ordinary bonds with no asset dependence and so we would expect it not to be a function of S, only of r and t.

This is the convertible bond pricing equation. There are two special cases of this equation that we have seen before:

- When $u = 0 = w$ we have constant interest rate r; Equation (33.4) collapses to the Black–Scholes equation.

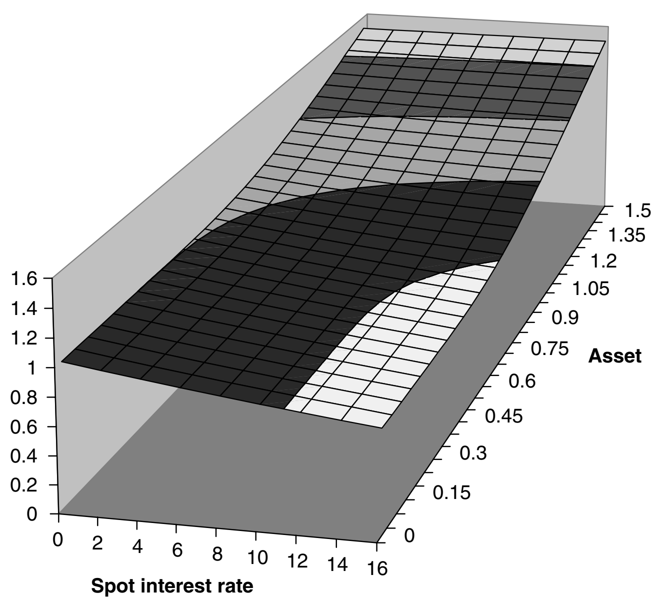


Figure 33.8 The value of a CB with stochastic asset and interest rates.

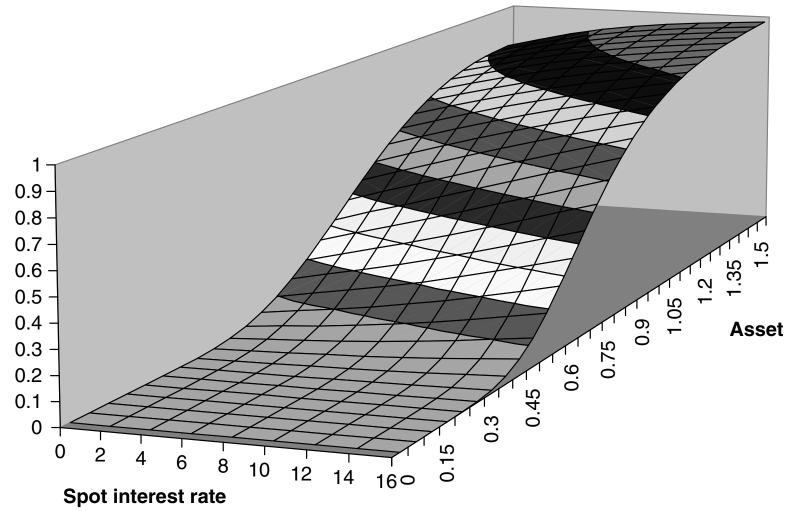


Figure 33.9 $\partial V/\partial S$ for a CB with stochastic asset and interest rates.

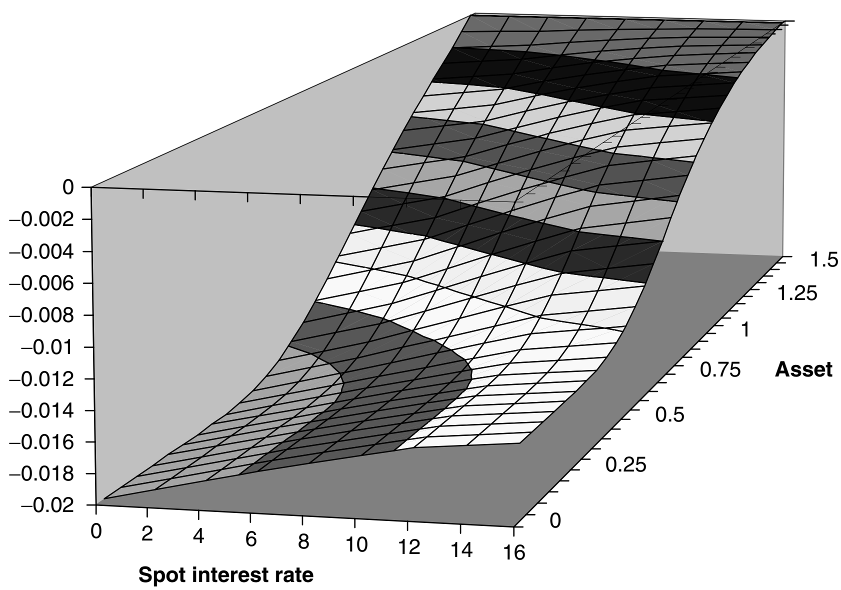


Figure 33.10 $\partial V/\partial r$ for a CB with stochastic asset and interest rates.

- When there is no dependence on an asset $\partial/\partial S = 0$ we return to the basic bond pricing equation.

Dividends and coupons are incorporated in the manner discussed in Chapter 8 and earlier in this chapter. For discrete dividends and discrete coupons we have the usual jump conditions.

The condition at maturity and constraints are exactly as before; there is one constraint for each of the convertibility feature, the call feature and the put feature.

In Figure 33.8 is shown the value of a CB when the underlying asset is lognormal and interest rates evolve according to the Vasicek model fitted to a flat 7% yield curve. In Figures 33.9 and 33.10 are shown the two hedge ratios $\partial V/\partial S$ and $\partial V/\partial r$.

33.7 A SPECIAL MODEL

In some circumstances, and for a very narrow choice of interest rate model, we can find a similarity reduction. For example, if we use the Vasicek model for the risk-neutral interest rate

$$dr = (\eta - \gamma r)\,dt + \beta^{1/2} dX_2$$

then we can look for convertible bond prices of the form

$$V(S, r, t) = g(r, t) H\left(\frac{S}{g(r, t)}, t\right).$$

Skipping some of the details, the function $g(r, t)$ must be the value of a zero-coupon bond with the same maturity as the CB:

$$g(r, t) = Z(r, t; T) = e^{A(t;T) - rB(t;T)}$$

where

$$B(t; T) = \frac{1}{\gamma}(1 - e^{-\gamma(T-t)})$$

and

$$A(t; T) = \frac{1}{\gamma^2}\left((B(t; T) - T + t)(\eta\gamma - \tfrac{1}{2}\beta)\right) - \frac{\beta B(t; T)^2}{4\gamma}.$$

The function $H(\xi, t)$ then satisfies

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\xi^2\left(\sigma^2 + 2B(t; T)\rho\beta^{1/2}\sigma + B(t; T)^2\beta\right)\frac{\partial^2 H}{\partial \xi^2} = 0.$$

This problem must be solved subject to

$$H(\xi, T) = \max(n\xi, 1)$$

and

$$H(\xi, t) \geq n\xi.$$

Unfortunately call and put features do not fit into this similarity formulation. This is because the constraint

$$V(S, t) \leq M_C$$

becomes

$$H(\xi, t) \leq \frac{M_C}{Z(r, t; T)}$$

the right-hand side of which cannot be written in terms of just ξ and t.

33.8 PATH DEPENDENCE IN CONVERTIBLE BONDS

The above-mentioned convertible bond is the simplest of its kind; they can be far more complex. One source of complexity is, as ever, path dependency. A typical path-dependent bond would be the following.

The bond pays \$1 at maturity, time $t = T$. Before maturity, it may be converted, at any time, for n of the underlying. Initially, n is set to some constant n_b. At time T_0 the conversion ratio n is set to some function of the underlying *at that time*, $n_a(S(T_0))$. Restricting our attention to a deterministic interest rate, this three-dimensional problem (for which there is no similarity solution) satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \le 0.$$

Here $V(S, \mathcal{S}, t)$ is the bond value with $\mathcal{S}$ being the value of S at time T_0. Dividends and coupons are to be added to this equation as necessary. The CB value satisfies the final condition

$$V(S, \mathcal{S}, T) = 1,$$

the constraint

$$V(S, \mathcal{S}, t) \ge n(S, \mathcal{S}),$$

where

$$n(S, \mathcal{S}) = \begin{cases} n_b & \text{for } t \le T_0 \\ n_a(\mathcal{S}) & \text{for } t > T_0. \end{cases}$$

and the jump condition

$$V(S, \mathcal{S}, T_0^-) = V(S, S, T_0^+).$$

This problem is three-dimensional, it has independent variables S, $\mathcal{S}$ and t. We could introduce stochastic interest rates with little extra theoretical work (other than choosing the interest rate model) but the computing time required for the resulting four-dimensional problem might make this infeasible.

33.9 DILUTION

In reality, the conversion of the bond into the underlying stock requires the company to issue n new shares in the company. This contrasts with options for which exercise leaves the number of shares unchanged.

I'm going to subtly redefine S as follows. If N is the number of shares before conversion then the total worth of the company is $NS - V$ before conversion. The $-V$ in this is due to the company's obligations with respect to the CB. This means that the share price is actually

$$\frac{NS - V}{N}$$

and not S. The constraint that the CB value must be greater than the share price is

$$V \ge n\frac{NS - V}{N}$$

which can be rewritten as

$$V \geq \frac{N}{n+N} nS. \tag{33.5}$$

We must also have

$$V \leq S \tag{33.6}$$

and

$$V(S, T) = 1.$$

Constraint (33.5) bounds the bond price below by its value on conversion, which is lower than it was previously. Constraint (33.6) allows the company to declare bankruptcy if the bond becomes too valuable. The factor $N/(n+N)$ is known as the **dilution**. In the limit $n/N \to 0$ we return to $V \geq nS$.

33.10 CREDIT RISK ISSUES

The risk of default in CBs is very important. If the issuing company goes bankrupt, say, you will not receive any coupons and nor will the ability to convert into the stock have much value. We have said that in the absence of default issues, the CB trades like a bond when the

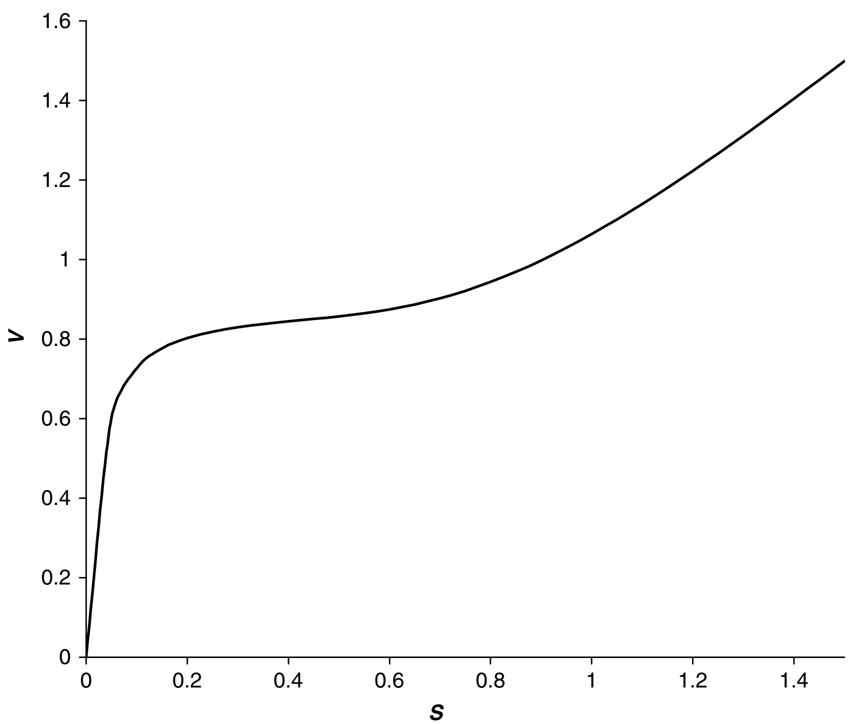


Figure 33.11 The value of a CB allowing for risk of default.

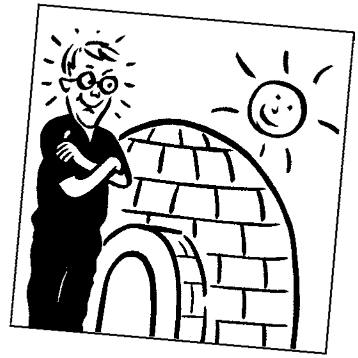
stock price is low. This is true, except that the bond behaves like a risky bond, one that may default. Furthermore, if the stock price is very low it is usually indicative of a none-too-healthy company. We can expect, and this is seen in the markets, that the price of a CB versus the stock looks rather like the plot in Figure 33.11. Here the CB price goes to zero as $S \to 0$. We will see the model that led to this picture in Chapter 40.

33.11 SUMMARY

Convertible bonds are a very important type of contract, playing a major role in the financing of companies. From a pricing and hedging perspective they are highly complex instruments. They have the early exercise feature of American options but in *three* guises—the option to convert, call and put—sometimes behaving like a bond and sometimes like a stock. They have long lifespans, five years, say, meaning that the assumption of constant interest rates is not valid. In more complicated cases they can be path-dependent. Finally, they are not without the risk of default. Put all these together and you have quite a sophisticated contract.

FURTHER READING

- For details of the effect of the issue of new shares on the value of convertible bonds see Brennan & Schwartz (1977), Cox & Rubinstein (1985) and Gemmill (1992).
- See Fabozzi (1996) and Nyborg (1996) for more information about the market practice of valuing CBs.
- For state-of-the-art modeling of converts look up research by Ayache *et al*. (2002, for example).
- Cross-currency convertibles are considered by Ouachani & Zhang (2004).



CHAPTER 34

mortgage-backed securities

In this Chapter...

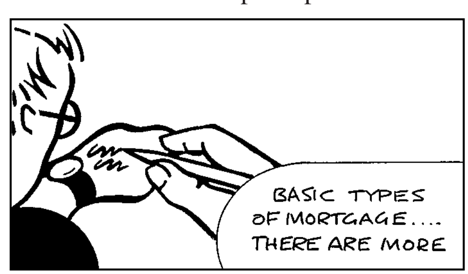
- types of mortgages
- what makes mortgage-backed securities special
- modeling prepayment

34.1 INTRODUCTION

Most homeowners know about mortgages. And mortgage-backed securities are simply many mortgages lumped together for onward sale. The most interesting aspect of mortgage-backed securities is that you must model the behavior of homeowners, whether or not they pay off their mortgages early, refinance etc. Because there are so many individual mortgages in each security it is only necessary to model the *average* behavior of the homeowners.

34.2 INDIVIDUAL MORTGAGES

A **mortgage** is a loan, usually to private individuals, to help them with the purchase of a home. They come in a variety of forms. The rate of interest could be fixed, or floating, or fixed for a while and then left to float. Sometimes the monthly payment to the lender covers both interest and gradual repayment of the principal. Sometimes only the interest is covered and the principal must be repaid at the end of the loan via some other form of funding. The lifetime of most mortgages is long, typically 20, 25 or 30 years.



The size of the loan that an individual can borrow will depend on their income and estimated credit-worthiness. The loan itself is backed by the property. There is risk to the lender above normal interest risk because of the possibility of default combined with the possibility of falling property values.

- **Fixed rate mortgages** have a rate of interest that is fixed for the life of the mortgage. Monthly payments remain at the same level and include payment of interest and a gradual repayment of principal. These are the most common form of mortgage in the US.

- **Floating rate mortgages** have interest payments linked to a variable interest rate. This variable rate is set by the mortgagee, the lender, and varies discretely and is loosely associated with the interest rates of the currency. These are the most common mortgages in the UK.

34.2.1 Monthly Payments in the Fixed Rate Mortgage

Since the monthly payments are constant, they represent both payment of interest and gradual repayment of principal. To start with the interest forms the bulk of the payment later in the life of the mortgage it is the repayment of principal that makes up the major part of each payment. If the interest rate is r_M and the amount of the loan is scaled to one and each monthly payment is x then

$$1 = \sum_{i=1}^{12\,N} \frac{x}{\left(1 + \frac{r_M}{12}\right)^i},$$

where N is the number of years over which the loan is paid off, the mortgage maturity. The monthly payment can be calculated from

$$x = \frac{r_M/12}{1 - (1 + r_M/12)^{-12\,N}}.$$

The remaining balance, when there are M payments left, is then given by

$$P = \sum_{i=1}^{M} \frac{x}{\left(1 + \frac{r_M}{12}\right)^i}.$$

These calculations use a discrete-time interest rate. If we were to use a continuous-time version then the sums would become integrals.

For more details of how mortgages work play a game of Monopoly.

34.2.2 Prepayment

Although mortgages are very much like other fixed-income securities they do have one novel feature; the borrower may decide to pay back the outstanding balance before the maturity of the loan. This prepayment is rather like the callable feature in a more usual bond. The difference between the two situations is important. We've seen how to model the callability of a bond in Chapter 32. There the bond callability was treated very much like the early exercise of an American option, with the call being made at an optimal time. With mortgages the time of prepayment depends very much on the individual's circumstances and his personal definition of 'optimal.' (There is more on what this might mean for the lender in Chapter 63.)

The reasons for prepayment could be any of the following:

- The owner of the property comes into some money, he is risk averse and pays off the mortgage early (perhaps mortgage rates are higher than bank interest rates).
- The owner decides to move house and pays off the mortgage with the proceeds from the sale.

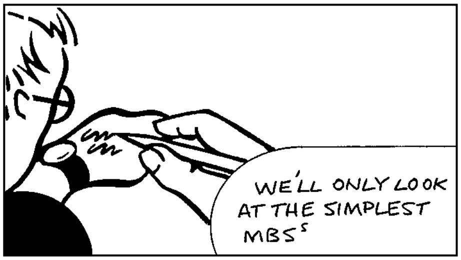
- The house falls down in an earthquake and the insurance goes to the lender.
- The householder defaults, and the insurance pays off the loan.
- Interest rates fall and the owner finds a better deal from another lender; this is known as refinancing.

Some of these are more important than the others, but most of them are outside the normal pricing assumption that the owner pays off when interest rates change suitably. Only the last one in the list approaches this 'rational' explanation, and even then there is a great deal of inertia stopping people from bothering with refinancing, or even knowing that the possibility exists.

We'll look at the stats of prepayment and how to model it later on.

34.3 MORTGAGE-BACKED SECURITIES

Mortgage-backed securities (**MBS**) are created by pooling together many individual mortgages. Investors then buy a piece of this pool and in return get the sum of all the interest and principal payments. These MBSs can often then be bought and sold through a secondary market.



By buying into this pool of mortgages the investor gets a stake in the housing-loan market, but with less of a prepayment risk. Most prepayers do not act 'rationally' on an individual basis, but when there are a lot of them the 'average' prepayment can be considered. This is rather like diversification. We'll see how the pricing is done via a prepayment function that represents the average behavior of the borrower.

Collateralized mortgage obligations (**CMO**) are securities based on MBSs but in which there has been further pooling and/or splitting so as to create securities with different maturities for example. A typical CMO might receive interest and principal only over a certain future time frame.

MBSs can be stripped into principal and interest components. **Principal only** (**PO**) MBSs receive only the principal payments and become worth more as prepayment increases. **Interest only** (**IO**) MBSs receive only the interest payments. The latter can be very risky since high levels of prepayment mean much fewer interest payments.

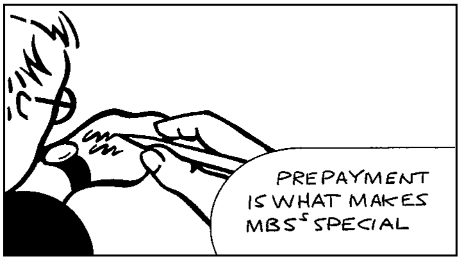
34.3.1 The Issuers

In the US most MBSs are issued by the following organizations:

- Government National Mortgage Association (GNMA or, colloquially, Ginnie Mae)
- Federal National Mortgage Association (FNMA or Fannie Mae)
- Federal Home Loan Mortgage Corporation (FHLMC or Freddie Mac)

These are all US agencies or government sponsored.

Others, such as investment banks and house builders, also issue private-label MBSs.



34.4 MODELING PREPAYMENT

If prepayment is what distinguishes MBSs from other fixed-income securities and derivatives then we need a model for this prepayment.

We could go to the extreme of calculating the optimum time for prepayment as we did for American options. This would give the theoretical maximum value for the MBS. This is not the usual practice since the mortgagors (the homeowners) do not behave in this 'rational' fashion, and the resulting MBS value would be unrealistically high. Instead we model what the mortgagors actually do in practice. I've put the word rational in inverted commas because rationality is in the eyes of the beholder, as it were (see Chapter 63).[1]

In the next few sections we'll look at prepayment models. We'll assume that the underlying mortgages are all fixed rate.

34.4.1 The Statistics of Repayment

In Figure 34.1 is shown a contour map of the value of new GNMA 30-year loans in billions of US dollars by year of origination and by coupon. Periods of falling interest rates are accompanied by periods of falling coupon rates. Figure 34.2 shows US 30-year rates for the same period.

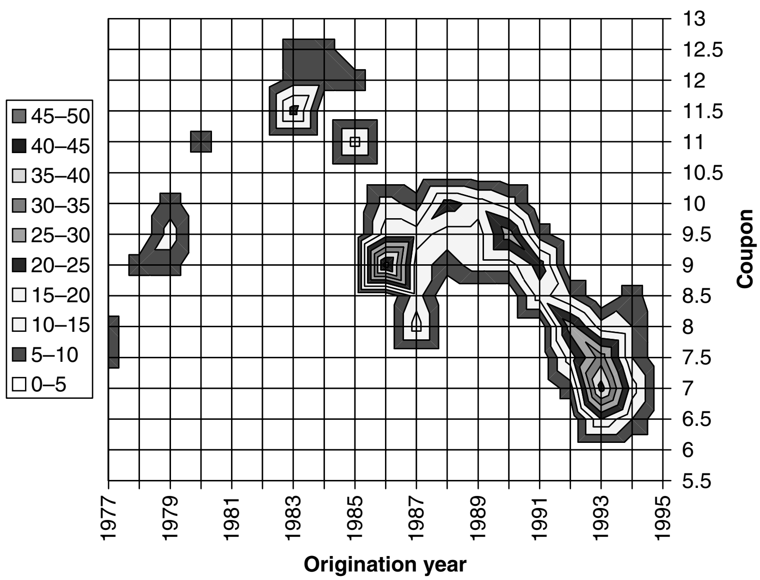


Figure 34.1 New Ginny Mae 30-year loans by year and coupon rate. Source: Davidson & Herskovitz (1996).

[1] Is there an arbitrage here? If so, it would involve the cooperation of millions of homeowners.

Figure 34.2 US 30-year interest rates.

Clearly, there is a wide spread of coupon rates. So, in any model of MBSs we are going to have to work in terms of averages of coupons. Thus we tend to think in terms of the **Weight-averaged Coupon** (**WAC**), where the weighting is by value of the loan.

Here are a couple of definitions used in prepayment models.

Single Monthly Mortality (**SMM**) is the amount prepaid in any month as a percentage of the expected mortgage balance:

$$\text{SMM} = \frac{\text{Scheduled balance} - \text{Actual balance}}{\text{Scheduled balance}}.$$

Conditional Prepayment rate (**CPR**) is an annualized version of the SMM:

$$\text{CPR} = 1 - (1 - \text{SMM})^{12}.$$

34.4.2 The PSA Model

The simplest model is the **Public Securities Association Model** (**PSA**) which takes no prepayment effects into account other than the age of the mortgage. The model assumes that prepayment starts at zero, for a mortgage just initiated, rising at 0.2% per month for the first 30 months and then stays constant at 6%. These numbers are in annualized, CPR, terms. See Figures 34.3 and 34.4 for the CPR and the percentage of remaining mortgages for the PSA model. How did I work out the data for Figure 34.4?

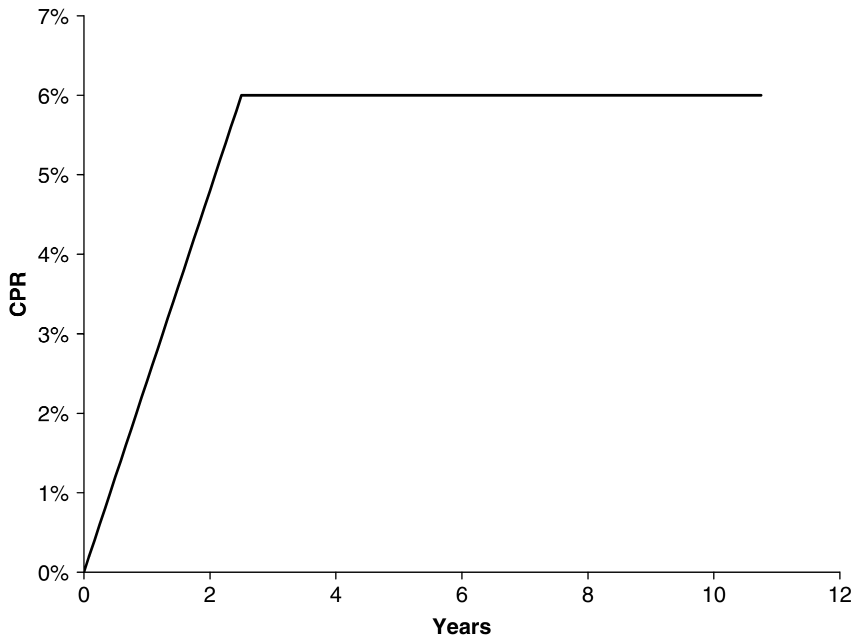


Figure 34.3 CPR in the PSA model.

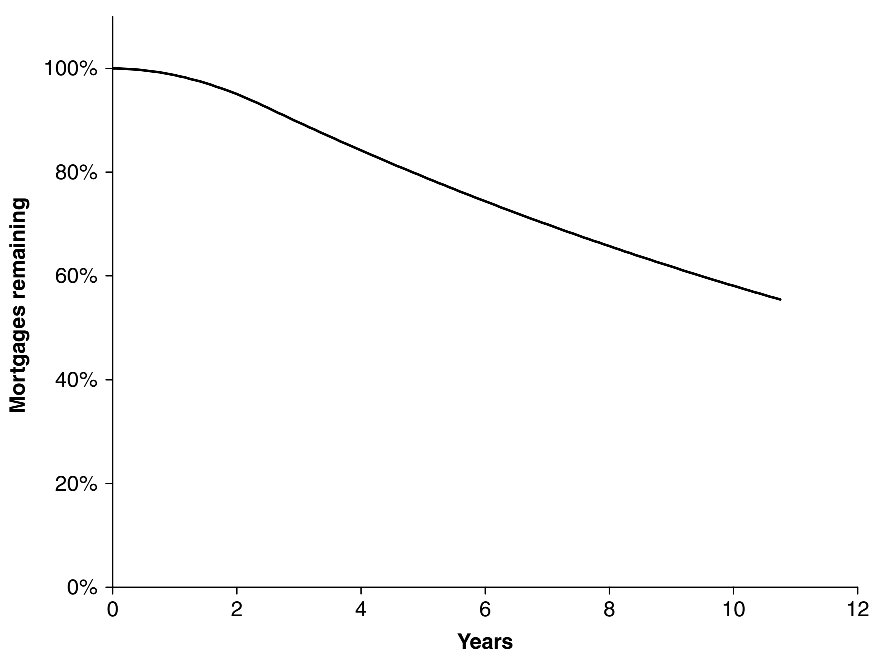


Figure 34.4 Percentage of remaining mortgages in the PSA model.

Using the PSA model is quite straightforward. Assuming that we can value (with whatever model) the value of all the cashflows associated with the MBS *without* prepayment it is simple enough to value the MBS *with* prepayment. After all, in the PSA model the prepayments are completely deterministic, depending only on the time since the initiation of the mortgages. Of course, we are really working in a world of expectations, and *real* expectations at that. There is no guarantee that the prepayment will occur as modeled, and nor can we hedge the risk that we are wrong.

34.4.3 More Realistic Models

The main drawback of the PSA model is that it does not adequately capture the effect of market rates on prepayment; prepayment increases as the then market rate decreases. This is simply because homeowners can refinance their mortgages at a lower interest rate; they behave rationally to some extent.

In Figure 34.5 is shown the CPR for 30-year GNMA loans, now as a function of the spread between the weight-average coupon and the prevailing coupon. This time no account has been taken of the maturity of the mortgages.

Presumably for negative spreads the prepayment is due to moving home, coming into some money etc. It is only when the spread goes above about 1% that the refinancing effect seems to take off.

Once we decide to make prepayment levels dependent on current mortgage rates we enter the world of path dependency.

As a proxy for the mortgage rate at which we could refinance our loan we could use the short-term interest rate, plus a spread. That spread is to cover the mortgagee against the credit

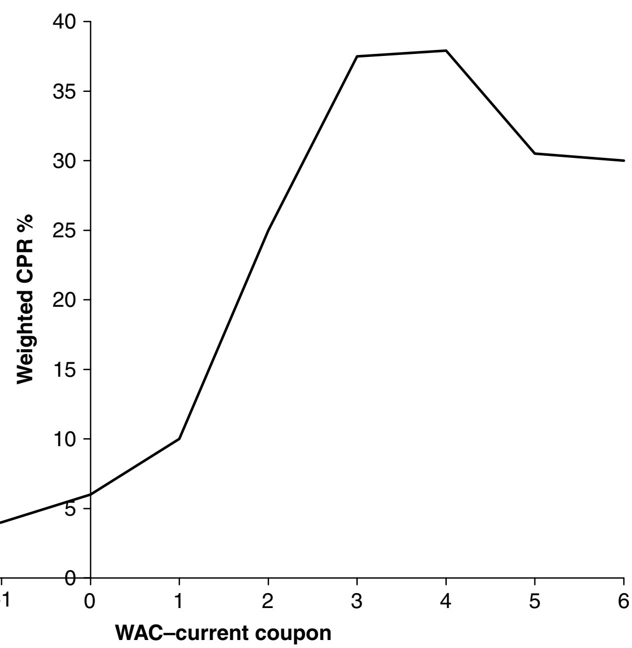


Figure 34.5 CPR versus spread between WAC and prevailing mortgage rate.

risk issues. Even though individual loans may be backed by the home as collateral, there is no guarantee that at the time of default house prices have not fallen, leaving the lender with a loss.

A possible model for prepayment would be to decompose the maturity effect and the interest effect into the form

$$\text{CPR} = a(t) f(r).$$

Here $a(t)$ captures the dependence of prepayment on time, the age of the mortgage. This is just like in the PSA model. The function $f(r)$ captures the dependence on the rate level, i.e. the prevailing mortgage rate.

There are many other factors affecting prepayment, such as seasonality and general economic conditions such as unemployment level. Interestingly from a modeling perspective, there is a time lag of approximately two or three months between a change in mortgage rate and prepayment rates.

34.5 VALUING MBSs

I think I've given enough clues in the above to enable the reader to start on the road to valuing MBSs. If you do want to incorporate the interest rate effect into prepayment then you will have a path-dependent problem on your hands. You could use the spot rate r and time t as the two obvious variables, and then the outstanding balance P as the path-dependent variable. We will also need the percentage of mortgages still left in the pool, Q. So we will have a value for our MBS of the form $V(r, P, Q, t)$.

Let's consider the simplest example in which the MBS consists of a pool of mortgages, all with the same age and maturity and the same fixed coupon. We'll also assume that the payments are constant throughout the life of the mortgage.

To make the analysis simpler, at least for me, we'll work in continuous time. Forgetting about prepayment for the moment, the equation for the balance in terms of the continuous payment x is

$$P = x \int_t^T e^{-r_M(\tau - t)} d\tau,$$

where T is the maturity date. The payment x on a mortgage starting at time $t = 0$ on a loan of \$1 is given by

$$x = \frac{1}{\int_0^T e^{-r_M \tau}\, d\tau}.$$

All things being equal, the balance will change according to

$$dP = (r_M P - x)\, dt.$$

This is completely deterministic in the absence of prepayment.

Now introduce the fraction of remaining mortgages, Q:

$$P = Qx \int_t^T e^{-r_M(\tau - t)}\, d\tau.$$

The dynamics of P are given by

$$dP = (r_M P - xQ)\,dt + \frac{P}{Q}dQ.$$

Suppose we have a prepayment (on average) of the form

$$a(t)f(r)\,dt$$

during a time step dt, a time t into the life of the mortgage and when the short-term interest rate is r. This is the fraction of the balance that is repaid and therefore the fraction of the pool that disappears. Thus

$$dQ = -a(t)f(r)Q\,dt.$$

We can now write

$$dP = (r_M P - xQ + a(t)f(r)P)\,dt.$$

This combines the expected change in balance due to the regular payments and the unexpected change due to early prepayment.

What is the governing equation for $V(r, P, Q, t)$? Using a risk-neutral spot rate model of

$$dr = (u - \lambda w)dt + w\,dX$$

we have

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} + (r_M P - xQ - a(t)f(r)P)\,\frac{\partial V}{\partial P}$$
$$- a(t)f(r)Q\frac{\partial V}{\partial Q} - rV + (a(t)f(r)P + xQ) = 0.$$

The final term is the cashflow due to the regular payments and the early prepayments. (Note that when r is constant and equal to r_M the solution is $V = P$, which makes sense.)

At maturity

$$V(r, P, Q, T) = 0.$$

There is a similarity reduction, to three variables, of the form

$$V(r, P, Q, t) = PH\left(r, \frac{P}{Q}, t\right).$$

How would the problem be changed to value IOs and POs?

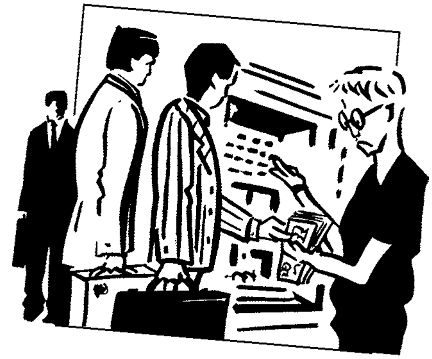
34.6 SUMMARY

There is much more to MBSs than I have described here. But I hope that you've seen how interesting they are from a modeling perspective. See the books mentioned below for further details about other contracts and subtleties of the modeling.

FURTHER READING

- See Arditti (1996) and Davidson & Herskovitz (1996) for more information on mortgage-backed securities.

CHAPTER 35
multi-factor interest rate modeling



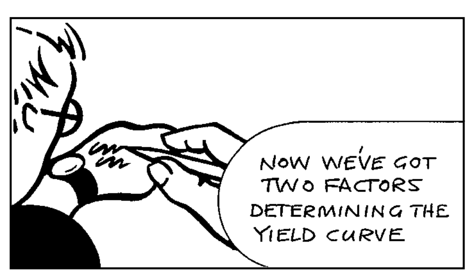
In this Chapter...

- the theoretical framework for multi-factor interest rate modeling
- why the 'long rate' is a good second factor to use
- popular two-factor models
- market prices of risk

35.1 INTRODUCTION

The simple stochastic spot interest rate models of Chapter 30 cannot hope to capture the rich yield curve structure found in practice: From a given spot rate at a given time they will predict the whole yield curve. Generally speaking, the one source of randomness, the spot rate, may be good at modeling the overall level of the yield curve but it will not necessarily model shifts in the yield curve that are substantially different at different maturities. For some instruments this may not be important. For example, for instruments that depend on the *level* of the yield curve it may be sufficient to have one source of randomness, i.e. one factor. More sophisticated products depend on the difference between yields of different maturities and for these products it is important to model the tilting of the yield curve. One way to do this is to invoke a second factor, a second source of randomness. This idea can be extended in principle to any number of factors.

In this chapter I describe the theory behind multi-factor interest rate modeling. I mention a few popular two-factor models explicitly. In these models the second factor is often, although not necessarily, a long-term interest rate.



35.2 THEORETICAL FRAMEWORK FOR TWO FACTORS

Assume that zero-coupon bonds (and all other simple interest rate instruments, for that matter) depend on two variables r, the spot interest rate, and another independent variable l. Thus, for example, a zero-coupon bond with maturity T has a price $Z(r, l, t; T)$. The variables satisfy

$$dr = u\,dt + w\,dX_1$$

and

$$dl = p\,dt + q\,dX_2.$$

All of u, w, p and q are allowed to be functions of r, l and t. The correlation coefficient ρ between dX_1 and dX_2 may also depend on r, l and t.

Remember, as yet I have deliberately not said what l is. It could be another interest rate, a long rate, say, or the yield curve slope at the short end, or the volatility of the spot rate, for example. We set up the framework in general and look at specific models later.

Since we have two sources of randomness now, in pricing one zero-coupon bond we must hedge with *two* others to eliminate the risk:

$$\Pi = Z(r, l, t; T) - \Delta_1 Z(r, l, t; T_1) - \Delta_2 Z(r, l, t; T_2).$$

The change in the value of this portfolio is given by

$$\begin{aligned}&(\mathcal{L}(Z) - \Delta_1\mathcal{L}(Z_1) - \Delta_2\mathcal{L}(Z_2))\,dt\\ &\quad+\left(\frac{\partial Z}{\partial r} - \Delta_1\frac{\partial Z_1}{\partial r} - \Delta_2\frac{\partial Z_2}{\partial r}\right)dr + \left(\frac{\partial Z}{\partial l} - \Delta_1\frac{\partial Z_1}{\partial l} - \Delta_2\frac{\partial Z_2}{\partial l}\right)dl,\end{aligned} \tag{35.1}$$

with the obvious notation for Z, Z_1 and Z_2. Here

$$\mathcal{L}(Z) = \frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2} + \rho wq\frac{\partial^2 Z}{\partial r\,\partial l} + \tfrac{1}{2}q^2\frac{\partial^2 Z}{\partial l^2}.$$

Now choose Δ_1 and Δ_2 to make the coefficients of dr and dl in (35.1) equal to zero. The corresponding portfolio is risk-free and should earn the risk-free rate of interest, r.

We thus have the three equations

$$\begin{aligned}\frac{\partial Z}{\partial r} - \Delta_1\frac{\partial Z_1}{\partial r} - \Delta_2\frac{\partial Z_2}{\partial r} &= 0,\\ \frac{\partial Z}{\partial l} - \Delta_1\frac{\partial Z_1}{\partial l} - \Delta_2\frac{\partial Z_2}{\partial l} &= 0\end{aligned}$$

and

$$\mathcal{L}'(Z) - \Delta_1\mathcal{L}'(Z_1) - \Delta_2\mathcal{L}'(Z_2) = 0$$

where

$$\mathcal{L}'(Z) = \mathcal{L}(Z) - rZ.$$

These are three simultaneous equations for Δ_1 and Δ_2. As such, this system is over-prescribed and for the equations to be consistent we require

$$\det(\mathbf{M}) = 0$$

where

$$\mathbf{M} = \begin{pmatrix} \mathcal{L}'(Z) & \mathcal{L}'(Z_1) & \mathcal{L}'(Z_2) \\ \partial Z/\partial r & \partial Z_1/\partial r & \partial Z_2/\partial r \\ \partial Z/\partial l & \partial Z_1/\partial l & \partial Z_2/\partial l \end{pmatrix}.$$

The first row of the matrix **M** is a linear combination of the second and third rows. We can therefore write

$$\mathcal{L}'(Z) = (\lambda_r w - u)\frac{\partial Z}{\partial r} + (\lambda_l q - p)\frac{\partial Z}{\partial l}$$

where the two functions $\lambda_r(r, l, t)$ and $\lambda_l(r, l, t)$ are the market prices of risk for r and l respectively, and are again independent of the maturity of any bond. In full, we have

$$\boxed{\begin{aligned} &\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2} + \rho w q\frac{\partial^2 Z}{\partial r\,\partial l} + \tfrac{1}{2}q^2\frac{\partial^2 Z}{\partial l^2} \\ &\quad + (u - \lambda_r w)\frac{\partial Z}{\partial r} + (p - \lambda_l q)\frac{\partial Z}{\partial l} - rZ = 0. \end{aligned}} \tag{35.2}$$

The model for interest rate derivatives is defined by the choices of w, q, ρ, and the risk-adjusted drift rates $u - \lambda_r w$ and $p - \lambda_l q$.

In Section 35.3 we look at popular two-factor models in more detail. First, however, we examine briefly the special case in which l is a long term interest rate.

Before I forget,

$$\Delta_1 = \frac{\partial Z/\partial r\ \partial Z_2/\partial l - \partial Z/\partial l\ \partial Z_2/\partial r}{\partial Z_1/\partial r\ \partial Z_2/\partial l - \partial Z_1/\partial l\ \partial Z_2/\partial r}$$

and

$$\Delta_2 = \frac{\partial Z/\partial r\ \partial Z_1/\partial l - \partial Z/\partial l\ \partial Z_1/\partial r}{\partial Z_1/\partial r\ \partial Z_2/\partial l - \partial Z_1/\partial l\ \partial Z_2/\partial r}.$$

35.2.1 Special Case: Modeling a Long-term Rate

We have not yet chosen what to model as l. There is one natural and special case that has certain advantages: A long-term interest rate. Specifically, here we examine the choice of l as the yield on a consol bond.

A simple **consol bond** is a fixed-coupon-bearing bond of infinite maturity. If C_o is the value of this bond when the coupon is \$1 p.a. then the yield is defined as

$$l = \frac{1}{C_o}.$$

This will be our definition of l. This choice for the variable l is very special because this consol bond is traded and thus its value must satisfy a pricing equation. The relevant equation is identical to (35.2) except there is an extra term:

$$\mathcal{L}'(C_o) + 1 = (\lambda_r w - u)\frac{\partial C_o}{\partial r} + (\lambda_l q - p)\frac{\partial C_o}{\partial l};$$

the extra term, 1, is due to the coupon payment.[1] Substituting $1/l$ into this equation we find

$$p - \lambda_l q = l^2 - rl + \frac{q^2}{l^2}.$$

In other words, because we are modeling a traded quantity, we find an expression for the market price of risk for that factor. Thus the bond-pricing equation with the consol bond yield as a factor is

$$\mathcal{L}'(Z) = (\lambda_r w - u)\frac{\partial Z}{\partial r} - \left(l^2 - rl + \frac{q^2}{l^2}\right)\frac{\partial Z}{\partial l},$$

or, in full,

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2} + \rho wq\frac{\partial^2 Z}{\partial r\,\partial l} + \tfrac{1}{2}q^2\frac{\partial^2 Z}{\partial l^2} + (u - \lambda_r w)\frac{\partial Z}{\partial r} + \left(l^2 - rl + \frac{q^2}{l^2}\right)\frac{\partial Z}{\partial l} - rZ = 0.$$

Observe that the coefficient of $\partial Z/\partial l$ no longer contains a market price of risk term. Compare this result with options on traded stocks. In that case, since the stock is traded we do not see a market price of risk for stocks; in fact, we find a simple expression for this market price of risk in terms of the real asset drift rate. Now contrast this with the spot interest rate term; the spot rate is not traded and so we see the market price for the spot rate appear in the equation.

Because we need to model one function fewer when we use the consol yield as a factor, it is an obvious and popular choice for the second factor in two-factor interest rate modeling.

Actually, all is not as rosy as I have just made out. The long rate and the short rate are more closely linked than the above suggests; it is *not* possible to model both a short rate and a long rate completely independently. For example, in the affine one-factor world the long rate is just a simple function of the short rate. And the same applies in the two-factor world. The internal consistency requirement amounts to a restriction on the volatility of the long rate. See Duffie, Ma & Yong (1994) for more details of the restrictions.

35.2.2 Special Case: Modeling the Spread Between the Long and the Short Rate

A version of the above is to model the spread between the long and the short rate as the new variable. Call this variable s. Because

$$s = \frac{1}{C_o} - r,$$

and C_o is traded, we find a simple expression for the market price of risk of the spread. Furthermore, there is evidence that, at least in the US, the short rate and the spread are uncorrelated. This eliminates another function from the modeling, leaving us with just three terms: Volatilities of short rate and spread, and the risk-adjusted spot-rate drift.

35.3 POPULAR MODELS

In this section I briefly describe some popular models. Most of these models are popular because the pricing equations (35.2) for these models have explicit solutions. In these models sometimes the second factor is the long rate and sometimes it is some other, usually unobservable, variable.

[1] Note that I have made the approximation that this coupon is paid continuously throughout the year. In practice it is paid discretely, but this has only a small effect on the following analysis.

Brennan and Schwartz (1982)

In the Brennan & Schwartz model the risk-adjusted spot rate satisfies

$$dr = (a_1 + b_1(l - r))\,dt + \sigma_1 r\,dX_1$$

and the long rate satisfies

$$dl = l(a_2 - b_2 r + c_2 l)\,dt + \sigma_2 l\,dX_2.$$

Brennan & Schwartz choose the parameters statistically. Because of the relatively complicated functional forms of the terms in these equations there are no simple solutions of the bond pricing equation. The random terms in these two stochastic differential equations are of the lognormal form, but the drift terms are more complicated than that, having some mean reversion character.

The main problems with the Brennan & Schwartz models are twofold. First, the long and short rates must satisfy certain internal consistency requirements. Second, as pointed out by Hogan (1993), the Brennan & Schwartz models can blow up in a finite time, meaning that rates can go to infinity. This is not a good property for an interest rate model.

General affine model

If r and l satisfy the following:

- the risk-adjusted drifts of both r and l are linear in r and l (but can have an arbitrary time dependence)
- the random terms for both r and l are both square roots of functions linear in r and l (but can have an arbitrary time dependence)
- the stochastic processes for r and l are uncorrelated

then the two-factor bond pricing equation (35.2) for a zero-coupon bond has a solution of the form

$$e^{A(t;T)-B(t;T)r-C(t;T)l}.$$

This result is a two-factor version of that found in Chapter 30 for a single factor.

The ordinary differential equations for A, B and C must in general be solved numerically.

Fong and Vasicek (1991)

Fong & Vasicek consider the following model for risk-adjusted variables:

$$dr = a(\overline{r} - r)\,dt + \sqrt{\xi}\,dX_1$$

and

$$d\xi = b(\overline{\xi} - \xi)\,dt + c\sqrt{\xi}\,dX_2.$$

Thus they model r, the risk-adjusted spot rate, and ξ the square root of the volatility of the spot rate. The latter cannot be observed, and this is an obvious weakness of the model. But it also makes it harder to show that the model is wrong. The simple linear mean reversion and the square roots in these equations result in explicit equations for simple interest rate products.

Longstaff and Schwartz (1992)

Longstaff & Schwartz consider the following model for risk-adjusted variables:

$$dx = a(\overline{x} - x)\,dt + \sqrt{x}\,dX_1$$

and

$$dy = b(\overline{y} - y)\,dt + \sqrt{y}\,dX_2,$$

where the spot interest rate is given by

$$r = cx + dy.$$

Again, the simple nature of the terms in these equations results in explicit equations for simple interest rate products.

Hull and White

The risk-neutral model

$$dr = (\eta(t) - u - \gamma r)\,dt + c\,dX_1$$

and

$$du = -au\,dt + b\,dX_2$$

is a two-factor version of the one-factor Hull & White. The spot rate r is mean reverting back to a level $(\eta(t) - u)/a$ that is time varying and also stochastic. The function $\eta(t)$ is used for fitting the initial yield curve. The solution for a zero-coupon bond is

$$e^{A(t;T)-B(t;T)r-C(t;T)}$$

where

$$\begin{aligned}
A(t;T) = &-\int_t^T \eta(s)B(s;T)\,ds + \frac{c^2}{2\gamma^2}\left(T - t + \frac{2}{\gamma}e^{-\gamma(T-t)} - \frac{1}{2\gamma}e^{-2\gamma(T-t)} - \frac{3}{2\gamma}\right)\\
&- \rho cb\left(-\frac{1}{\gamma^2 a}\left(-(T-t) - \frac{1}{\gamma}e^{-\gamma(T-t)} + \frac{1}{\gamma}\right)\right.\\
&+ \frac{1}{\gamma^2(a-\gamma)}\left(\frac{1}{\gamma}e^{-\gamma(T-t)} - \frac{1}{2\gamma} - \frac{1}{2\gamma}e^{-2\gamma(T-t)}\right)\\
&\left.- \frac{1}{a\gamma(a-\gamma)}\left(\frac{1}{a}e^{-a(T-t)} - \frac{1}{a} - \frac{1}{\gamma+a}e^{-(\gamma+a)(T-t)} + \frac{1}{\gamma+a}\right)\right)\\
&- \tfrac{1}{2}b^2\left(-\frac{T-t}{\gamma^2 a^2} + \frac{1}{2\gamma^3(a-\gamma)^2}\left(e^{-2\gamma(T-t)} - 1\right) + \frac{1}{2a^3(a-\gamma)^2}\left(e^{-2a(T-t)} - 1\right)\right.\\
&- \frac{2}{\gamma^3 a(a-\gamma)}\left(e^{-\gamma(T-t)} - 1\right) + \frac{2}{\gamma a^3(a-\gamma)}\left(e^{-a(T-t)} - 1\right)\\
&\left.- \frac{2}{a\gamma(a+\gamma)^2(a-\gamma)}\left(e^{-(\gamma+a)(T-t)} - 1\right)\right),
\end{aligned}$$

$$B(t;T) = \frac{1}{\gamma}(1 - e^{-\gamma(T-t)})$$

and

$$C(t;T) = -\frac{1}{\gamma a} + \frac{1}{\gamma(a-\gamma)}e^{-\gamma(T-t)} - \frac{1}{a(a-\gamma)}e^{-a(T-t)}.$$

The correlation between the two random walks is ρ, assumed constant.

35.4 THE MARKET PRICE OF RISK AS A RANDOM FACTOR

Whenever you model a quantity that isn't traded you will end up with a market price of risk for that quantity. So in the above we find that there are two market prices of risk if r and l aren't traded. Well, what if we make one of the random quantities which we model just the market price of risk for r? The effect is that we only have one 'arbitrary' market price of risk, the market price of market price of risk risk. (No that isn't a typo!) As will be seen in Chapter 36 there is plenty of evidence to suggest that the market price of spot rate is indeed very random, so this makes a lot of sense from a modeling perspective. Let's do the math.

Suppose that we have the two real random walks

$$dr = u\,dt + w\,dX_1$$

and

$$d\lambda = p\,dt + q\,dX_2,$$

where λ is the market price of r risk. The zero-coupon bond pricing equation is then

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2} + \rho wq\frac{\partial^2 Z}{\partial r\,\partial\lambda} + \tfrac{1}{2}q^2\frac{\partial^2 Z}{\partial\lambda^2} + (u-\lambda w)\frac{\partial Z}{\partial r} + (p-\lambda_\lambda q)\frac{\partial Z}{\partial\lambda} - rZ = 0.$$

The sole arbitrary function is λ_λ, the market price of market price of risk risk. I see having fewer arbitrary functions as being a positive feature of a model; those who like fudge factors will disagree.

As we saw in Chapter 31, and we shall see again in Chapter 36, there is a relationship between the market price of spot rate risk, λ, and the slope of the yield curve at the short end. So modeling r and λ is like modeling the behavior of the yield curve close to maturity.

35.5 THE PHASE PLANE IN THE ABSENCE OF RANDOMNESS

There is a very simple analysis that can be done on two-factor models to determine whether they are well behaved. This analysis is the examination of the dynamics in the *absence of randomness*. It requires the drift coefficients for the two factors to be independent of time. I'm also going to cheat a little in the following. Ideally we would transform the stochastic differential equations into equations with constant volatilities, but I'm going to omit that step here. It will make no difference to anything I say about the singularities, but may have an effect on my description of the global behavior of rates.

To demonstrate the technique, let us consider the Fong & Vasicek model. In the absence of randomness, this model becomes

$$dr = a(\overline{r} - r)\,dt$$

and

$$d\xi = b(\overline{\xi} - \xi)\,dt.$$

Dividing one by the other we get

$$\frac{d\xi}{dr} = \frac{b(\overline{\xi} - \xi)}{a(\overline{r} - r)}. \tag{35.3}$$

We can plot the solutions of this first-order ordinary differential equation in a **phase plane**, see Figure 35.1.

The idea behind phase planes can be described by examining this model. Each line in this plot represents a possible evolution of r and ξ. In this particular example there is a **singularity** at the point $(\overline{r}, \overline{\xi})$ where the numerator and denominator of (35.3) are both zero and so the slope $d\xi/dr$ is undefined. This singularity is in this case called a **node**. In the Fong & Vasicek model with a and b positive this singularity is **stable**, meaning that solutions that pass close

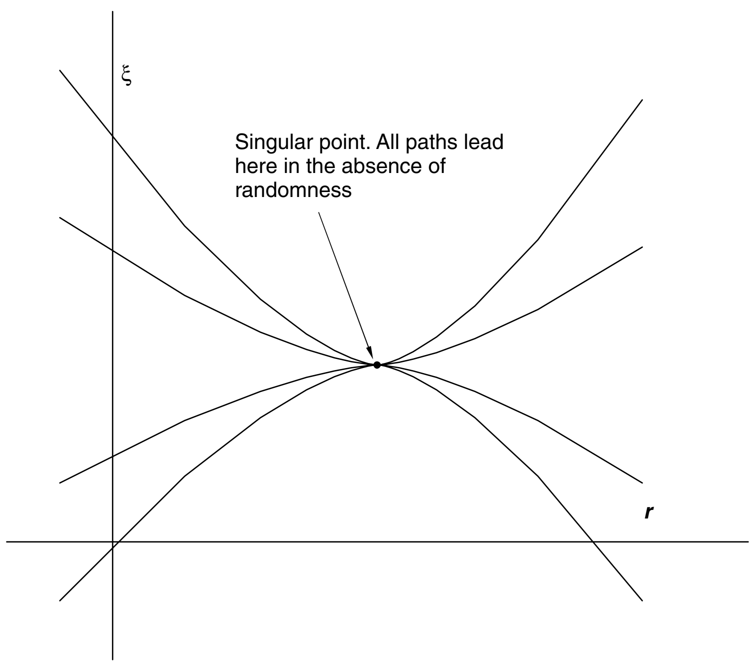


Figure 35.1 Phase plane for the Fong & Vasicek model (no randomness).

to the singular point are drawn into it. (In this case, *all* solutions are drawn to this point.) Some more types of possible singularities for other two-factor (non-random) models are shown in Figure 35.2; there are others. The two-factor Hull & White model has a nice deterministic phase plane.

When randomness is put back into the model the solution paths are no longer deterministic. A possible evolution of a single realization is shown in Figure 35.3 for the Fong & Vasicek model. What's wrong with this figure, and with Figure 35.2? The answer is that the phase plane only looks the way I have drawn it for points close to the singularity. Remember, I said that ideally one should transform the governing stochastic differential equations into equations with constant volatilities, but this I have not done. For the Fong & Vasicek model in particular, as

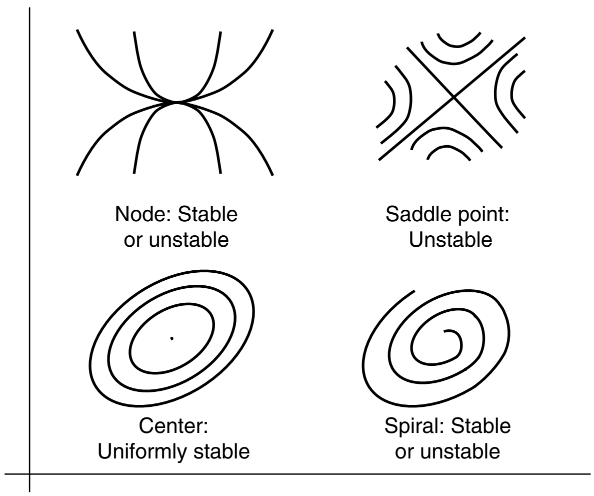


Figure 35.2 Possible phase-plane singularities.

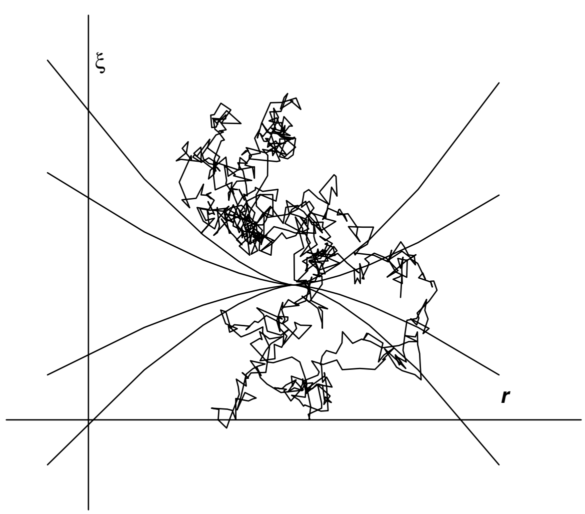


Figure 35.3 Typical realization of the Fong & Vasicek model, phase plane and time series.

long as c is not too big the variable ξ cannot go negative because of the square root in front of the random term.

We can use these phase planes to decide whether the properties of our model are realistic. A very useful example is the Brennan and Schwartz model. For certain parameter ranges the equations are so unstable that the long and short rates tend to infinity. Not a realistic property for an interest rate model.[2]

35.6 THE YIELD CURVE SWAP

One of the most important contracts for which you may need a two-factor model is the **yield curve swap**. This is a contract in which one counterparty pays a floating rate based on one part of the yield curve and the other counterparty pays a floating rate based on a different point on the curve. In Figure 35.4 is the term sheet for a yield curve swap on sterling. Every quarter one

Sterling Yield Curve Swap

Swap Notional	GBP [ ]
Swap Effective Date	4 November 2000
Swap Maturity Date	4 November 2002
Counterparty A payments	
Floating Coupon	GBP 2yr Swap (Mid) plus 0.16%, as fixed on Telerate page 42279 at 11:00am London time on each period start date
Payment Frequency	Quarterly
Payment Dates	4 February, May, August, November
Calculation Basis	Act/365
Counterparty B payments	
Floating Index	GBP 3 Month LIBOR flat
Payment Frequency	Quarterly
Payment Dates	4 February, May, August, November
Calculation Basis	Act/365

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 35.4 Term sheet for a sterling yield curve swap.

[2] Do not forget, of course, that in the examples given in this section, the phase plane is for the *risk-adjusted* spot interest rate, not the real rate.

side pays three-month GBP LIBOR and the other pays the two-year swap rate plus 16 basis points. A **basis point** is one hundredth of a percent.

There are two ways to approach the pricing of this contract. One is to model the spread between these two rates as a lognormal, or more probably, Normal variable. And the other is to use a two-factor interest rate model.

If we assume that the spread is a Normal variable then we can find its volatility, from historical data, say, or implied from another contract, and price the contract in an almost Black–Scholes world. You will also need to discount using some interest rates and the natural choice would be the yield curve as it stands at the time of pricing. This is a moderately robust way to price this contract.

If we are to use a stochastic interest rate model, why must it be two-factor? This is because the net cashflows depend on the spread between two rates. If you were to use a one-factor model then these two rates would be perfectly correlated, resulting in no volatility in the spread. The two-factor model is needed for there to be some volatility in the spread, and thus a non-deterministic outcome for each cashflow. The two interest payments are treated slightly differently. The three-month rate can be thought of as the spot interest rate but the two-year swap rate is more complicated; you will have to first model the swap rates using the model. Having modeled the swap rate, using the two-factor differential equation, you can then price the yield curve swap, using the same equation but with different cashflows. The contract can therefore be thought of as second order. Not only is this procedure messy and time consuming, but the result may depend quite strongly on your chosen model.

35.7 GENERAL MULTI-FACTOR THEORY

Suppose that the world of fixed income can be modeled by N factors x_i such that

$$dx_i = \mu_i(\mathbf{x}, t)\,dt + \sigma_i(\mathbf{x}, t)\,dX_i.$$

Here $\mathbf{x}$ denotes the vector with ith entry x_i. The correlation between the factors is $\rho_{ij}(\mathbf{x}, t)$, also a function of the N factors. With the market prices of risk denoted by $\lambda_i(\mathbf{x}, t)$ the pricing equation for non-path-dependent contracts is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{ij}\sigma_i\sigma_j\frac{\partial^2 V}{\partial x_i\,\partial x_j} + \sum_{i=1}^{N}(\mu_i - \lambda_i\sigma_i)\frac{\partial V}{\partial x_i} - r(\mathbf{x}, t)V = 0. \tag{35.4}$$

Notice how we need to model the spot interest rate $r(\mathbf{x}, t)$ as a function of the factors.

It's hard enough accurately modeling one factor, so you can imagine what a task it is when you have two or more factors.

35.7.1 Tractable Affine Models

Some simple functional forms for the volatilities and risk-neutral drift rates lead to tractable models in the sense that (35.4) is easily solved explicitly. The results for many factors are similar to those of Chapter 30 for a single factor.

There are two obvious examples: A multi-factor Vasicek and a multi-factor CIR (or a combination).

Multi-factor Vasicek

In the Vasicek version we have volatilities and correlations that are independent of the factors but may be arbitrary functions of time:

$$\sigma_i = c_i(t) \quad \text{and} \quad \rho_{ij} = d_{ij}(t).$$

The risk-neutral drifts must be linear in the factors:

$$\mu_i - \lambda_i \sigma_i = a_i(t) - \sum_{j=1}^{N} b_{ij}(t) x_j.$$

The spot rate must also be linear in the factors:

$$r = g_0(t) + \sum_{i=1}^{N} g_i(t) x_i.$$

Multi-factor CIR

In the CIR version we have volatilities which are the square root of linear in the N factors and the drift should be linear in the factors. The correlations must all be zero.

For both the Vasicek and CIR models we have solutions for zero-coupon bonds of the form

$$Z(\mathbf{x}, t; T) = \exp\left(f_0(t; T) + \sum_{i=1}^{N} f_i(t; T) x_i \right).$$

The functions $f_i(t; T)$ satisfy ordinary differential equations with independent variable t.

These models are particularly interesting because the yields, being defined as

$$Y(t; T) = -\frac{\log Z(\mathbf{x}, t; T)}{T - t}$$

are represented by

$$-\frac{1}{T - t}\left(f_0(t; T) + \sum_{i=1}^{N} f_i(t; T) x_i \right).$$

They are also linear in the factors.

The simplest example of this is the multi-factor Vasicek with constant parameters and no correlation between the factors, so that $d_{ij} = 0$. We must solve for the fs from

$$\dot{f}_0 + \tfrac{1}{2} \sum_{i=1}^{N} c_i^2 f_i^2 + \sum_{i=1}^{N} a_i f_i - g_0 = 0$$

and

$$\dot{f}_i - \sum_{k=1}^{N} b_{ki} f_k - g_i = 0 \quad \text{for all } i.$$

The dot over the fs means differentiation with respect to time. In this model, we can, without loss of generality, assume that $b_{ij} = 0$ unless $i = j$ and also that $g_0 = 0$. This simplifies these equations in such a way that it is easy to see that the solution for a zero-coupon bond is just like the product of one-factor Vasicek solutions.

35.8 SUMMARY

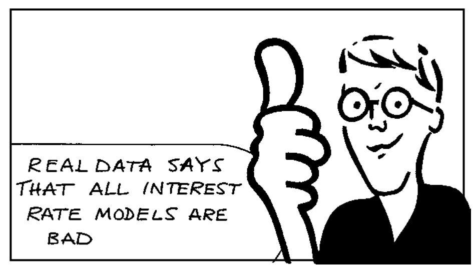
In this chapter we have seen the general theory for multi-factor interest rate models. Multi-factor models are better than one-factor in that they will allow for a richer yield curve structure. But again, as with the one-factor models, the theoretical yield curve will not be the same as the market yield curve unless the model has been calibrated. Even then, all of the problems that we saw in the one-factor fitted models are still seen here: The slope and curvature of the yield curve are in general too large to be consistently modeled by even a multi-factor model. Nevertheless, having many factors does allow different parts of the yield curve to be less than one hundred percent correlated and that has to be an improvement. Indeed, for any contract that pays off the spread between two different points on the yield curve a one-factor model cannot be used.

FURTHER READING

- See the original papers by Brennan & Schwartz (1982) and Longstaff & Schwartz (1992).
- See Duffie, Ma & Yong (1994) for an explanation of the relationship between the short and the long rate and why they can't be modeled independently.
- Hogan (1993) analyzes the Brennan and Schwartz models and shows that they can have some alarming features, such as rates going to infinity in a finite time.
- Rebonato (1996) discusses theoretical and practical issues concerning several two-factor models in some detail.
- Jordan & Smith (1977) is the classic reference text for phase-plane analysis.
- See Hagan *et al*. (2002) for the famous SABR model. This is a two-factor model, but the beauty lies in the use of asymptotic analysis to find relatively simple solutions.

CHAPTER 36
empirical behavior of the spot interest rate

In this Chapter...



- how to analyze short-term interest rates to determine the best model for the volatility and the real drift
- how to analyze the slope of the yield curve to get information about the market price of risk
- how to build up sensible one- or two-factor spot rate models

36.1 INTRODUCTION

Earlier I described the general framework for one-factor interest rate modeling and we have seen several popular models in detail. The one-factor models that we saw in Chapter 30 for the spot interest rate were all chosen for their nice properties; for most of them we were able to find simple closed-form solutions of the bond-pricing equation. We discussed the pros and cons of these models in that chapter.

In this chapter we will see how to deduce a model for the spot rate from data. The method that we use assumes that the model is time homogeneous, and ensures that the spot rate is well behaved. The principle is the same as that in Chapter 53, where we find a volatility model from volatility data. Here, though, we take the idea further, finding also the market price of risk from yield curve data. The downside to the resulting model is that we cannot find closed-form solutions for contract values; the risk-neutral drift and the volatility don't have a sufficiently nice structure.

I describe the analysis and modeling using empirical US spot interest rate data, making comparisons with the popular models. We use these data to determine which models seem to fit reality best and to derive a spot rate model that has certain desirable properties: It has a realistic volatility structure and a meaningful steady-state probability density function. By basing our modeling on real data we inevitably lose tractability when we come to solve the bond pricing equation. Our only concession to tractability is the assumption of time-independence of our random walk. Using this assumption, and by examining data over a long period of time, we hope to model, in a sense, the 'average' behavior of the spot rate. Such a model could, for example, be used to price long-dated instruments. I conclude with some suggestions for how to build up a two-factor model for interest rates.

36.2 POPULAR ONE-FACTOR SPOT-RATE MODELS

Because of the complexity of fixed-income instruments the most popular of the interest-rate models are chosen because they result in closed-form expressions for the prices of simple contracts. That is they are *tractable*. This reduces to a minimum the computer time required for bond calculations. Unfortunately, this also means that the models are not necessarily a good description of reality. Let us briefly re-examine some one-factor models.

Consider the classical one-factor models of Vasicek (1977), Cox, Ingersoll & Ross (1985), Ho & Lee (1986) and the extensions of Hull & White (1990). In these models the spot rate r satisfies the stochastic differential equation

$$dr = u(r,t)\,dt + w(r,t)dX. \tag{36.1}$$

Table 36.1 shows the structural forms adopted by the above authors.

Here $\lambda(r,t)$ denotes the market price of risk. The function $u - \lambda w$ is the risk-adjusted drift. The time-dependent coefficients in these models allow for the fitting of the yield curve and other interest-rate instruments as we have seen in Chapter 31.

If the spot rate equation is (36.1) then, as shown in Chapter 30, the zero-coupon bond pricing equation is

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2}w(r,t)^2\frac{\partial^2 Z}{\partial r^2} + (u(r,t) - \lambda(r,t)w(r,t))\frac{\partial Z}{\partial r} - rZ = 0, \tag{36.2}$$

with $Z(r,T;T) = 1$. It is no coincidence that most of the models in the table give zero-coupon bond prices of the form

$$Z(r,t;T) = e^{A(t,T) - rB(t,T)}. \tag{36.3}$$

As shown independently by numerous people (Duffie (1992), Klugman (1992), Klugman & Wilmott (1994), Ritchken (1994) and others), if the solution of (36.2) for the zero-coupon bond takes the form (36.3) then the most general forms for the coefficients must be as shown at the

Table 36.1 Popular one-factor spot interest rate models.

Model	$u(r,t) - \lambda(r,t)w(r,t)$	$w(r,t)$
Vasicek	$a - br$	c
CIR	$a - br$	$cr^{1/2}$
Ho & Lee	$a(t)$	c
HW I	$a(t) - b(t)r$	$c(t)$
HW II	$a(t) - b(t)r$	$c(t)r^{1/2}$
General affine	$a(t) - b(t)r$	$(c(t)r - d(t))^{1/2}$

bottom of Table 36.1; this general affine model contains four time-dependent parameters. We saw in Chapter 30 that this is the most general form which leads to explicit formulae.

Black, Derman & Toy (BDT), on the other hand, have not developed their model, given by

$$d(\log r) = \left(\theta(t) - \frac{w'(t)}{w(t)} \log r\right) dt + w(t)\, dX,$$

for reasons of tractability with respect to the solution of a partial differential equation. Their model is chosen to enable them to fit market data easily. They have a volatility structure (the coefficient of dX in the spot rate equation) that is proportional to the spot rate with a time-dependent coefficient. This coefficient and the other time-dependent coefficient in the risk-adjusted drift rate are chosen so that the BDT model correctly fits today's yield curve and interest rate options. This data fitting is, of course, very appealing, but, as we have seen, it must be treated with care.

36.3 IMPLIED MODELING: ONE FACTOR

Now let us examine an alternative approach to the modeling of interest rates. We build up our model in stages and try not to be too sidetracked by 'tractability.' Remember, we are assuming time-independent parameters.

In Figure 36.1 is shown the US one-month LIBOR rates, daily, for the period 1977–1994, and this is the data that we use in our analysis. The ideas that we introduce can be applied to any currency, but here we use US data for illustration.

There are three key stages:

1. By differencing spot rate time series data we determine the volatility dependence on the spot rate $w(r)$.

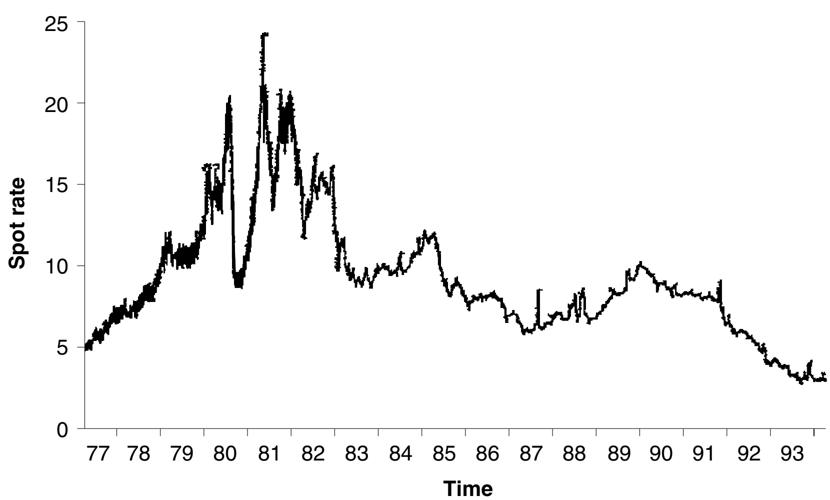


Figure 36.1 US spot rate 1977–1994.

2. By examining the steady-state probability density function for the spot rate we determine the functional form of the drift rate $u(r)$.
3. We examine the slope of the yield curve to determine the market price of risk $\lambda(r)$.

36.4 THE VOLATILITY STRUCTURE

Our first observation is that many popular models take the form

$$dr = u(r)\,dt + \nu r^{\beta}\,dX. \tag{36.4}$$

Examples of such models are the Ho & Lee ($\beta = 0$), Vasicek ($\beta = 0$) and Cox, Ingersoll & Ross[1] ($\beta = 1/2$) models. Because of their popularity with both practitioners and academics there have been a number of empirical studies, trying to estimate the coefficient β from data. Perhaps the most cited of these works is that by Chan, Karolyi, Longstaff & Sanders (1992) on US data. They obtain the estimate $\beta = 1.36$.[2] This agrees with the experiences of many practitioners, who say that in practice the relative spot rate change dr/r is insensitive to the level of r.

We can think of models with

$$dr = \cdots + c\,dX$$

as having a Normal volatility structure and those with

$$dr = \cdots + cr\,dX$$

as having a lognormal volatility structure. In reality it seems that the spot rate is closer to the lognormal than to the Normal models. This puts the BDT model ahead of the others.

Using our US spot rate data we can estimate the best value for β.[3] There are any number of sophisticated methods that one can use. Here I am going to describe a very simple, not at all sophisticated, method. From the time-series data divide the changes in the interest rate, δr, into buckets covering a range of r values. Then calculate the average value of $(\delta r)^2$ for each bucket. If the model (36.4) is correct we would expect

$$E[(\delta r)^2] = \nu^2 r^{2\beta}\delta t$$

to leading order in the time step δt, which for our data is one day. This is the same technique as used in Chapter 53 for estimating the volatility of volatility.

In Figure 36.2 is plotted $\log(E[(\delta r)^2])$ against $\log r$ using the US data. The slope of this 'line' gives an estimate for 2β. We can see that the line is remarkably straight. From this calculation it is estimated that

$$\beta = 1.13 \quad \text{and} \quad \nu = 0.126.$$

This confirms that the spot rate volatility is close to lognormal in nature.

---

[1] In another model, Cox, Ingersoll & Ross (1980) had $\beta = 3/2$ with $u = 0$. They found analytical solutions for some perpetual instruments, but had to make some assumptions again for reasons of tractability. With these same assumptions it is possible to find explicit *similarity* solutions for zero-coupon bonds.

[2] Murphy (1995) finds $\beta = 0.36$ for the UK. The analysis of this chapter could equally well be applied to the UK.

[3] Of course, we don't need to only study power-law volatilities; the technique described here works for any functional form for volatility as long as it only depends on r.

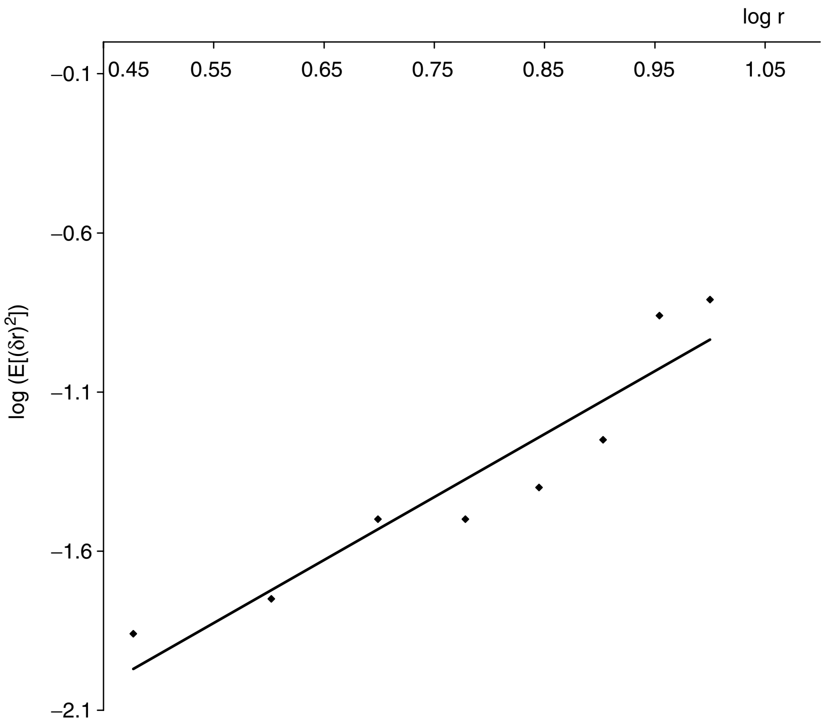


Figure 36.2 Estimation of β.

36.5 THE DRIFT STRUCTURE

It is statistically harder to estimate the drift term from the data; this term is smaller than the volatility term and thus subject to larger relative errors. If we were to use a naive method to determine the drift, we may find ourselves with a model that behaves well for short times but behaves poorly in the long term. We will therefore take an alternative, more stable, approach involving the empirical and analytical determination of the steady-state probability density function for r.

If r satisfies the stochastic differential equation (36.4) then the probability density function $p(r, t)$ for r satisfies the forward Fokker–Planck equation (see Chapter 10)

$$\frac{\partial p}{\partial t} = \tfrac{1}{2}\nu^2 \frac{\partial^2}{\partial r^2}(r^{2\beta} p) - \frac{\partial}{\partial r}(u(r)p). \tag{36.5}$$

In this r and t are the forward variables, usually denoted by r' and t'. It is possible for equation (36.5) to have a steady-state distribution as a solution. This distribution is that to which the probability density function will evolve from any initial condition. We can estimate this steady state from the empirical data and thus find a solution of (36.5). Note that a steady-state distribution would clearly not exist for an equity, since equity prices generally grow (or decay) exponentially. However, a steady state probability density function is a not unreasonable property to assume for interest rates (Ramamurtie, Prezas & Ulman, 1993). This steady state

$p_\infty(r)$ will satisfy

$$\tfrac{1}{2}\nu^2 \frac{d^2}{dr^2}(r^{2\beta} p_\infty) - \frac{d}{dr}(u(r) p_\infty) = 0. \qquad (36.6)$$

If this steady-state probability density function is empirically determined, then by integrating (36.6) we find that[4]

$$u(r) = \nu^2 \beta r^{2\beta-1} + \tfrac{1}{2}\nu^2 r^{2\beta} \frac{d}{dr}(\log p_\infty).$$

Not only is this method for finding the drift more stable in the long run, but also the steady-state probability density function is something simple to focus attention on; we will know that our model cannot behave too outrageously. This probability density function is something that it may be possible to estimate, or at least take an educated guess at. In contrast, it is harder to have an intuitive feel for the drift coefficient $u(r)$. By choosing a model with a sensible steady-state distribution, we can guarantee that the model will not allow the spot rate to do anything unrealistic such as grow unboundedly.

For special choices of β and $p_\infty(r)$ (and later $\lambda(r)$) we recover all the models mentioned in Table 36.1

Again, looking at US data, we can determine a plausible functional form for $p_\infty(r)$ from one-month US LIBOR rates. The steady-state distribution is determined by dividing r into buckets and observing the frequency with which each bucket is reached. This is the same technique as used in Chapter 53 for estimating the drift of volatility.

The results of this analysis are shown in Figure 36.3. This figure represents our estimate for the steady-state probability density function. The shape of this graph is reminiscent of a

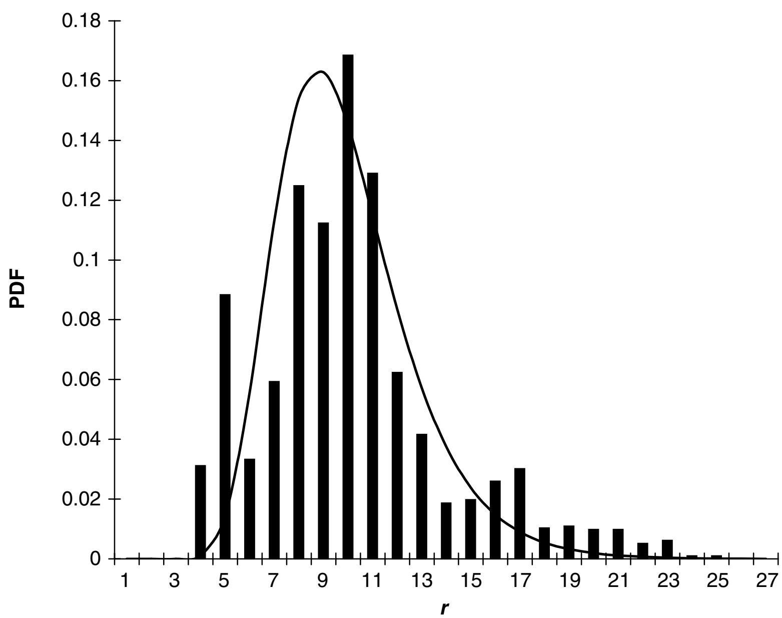


Figure 36.3 Observed and chosen probability density distributions for r.

[4] There are some issues to do with existence, uniqueness, and behavior at $r = 0$ and infinity that I omit for the sake of readability.

lognormal curve. For this reason, and because it has a simple formula with just two parameters,[5] choose $p_\infty(r)$ to be a lognormal curve that best fits the empirical data; this curve is also shown in the figure.

Our choice for $p_\infty(r)$ is

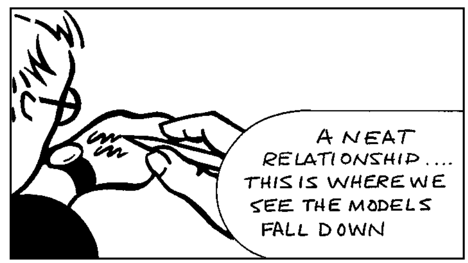
$$\frac{1}{ar\sqrt{2\pi}}\exp\left(-\frac{1}{2a^2}(\log(r/\overline{r}))^2\right)$$

where $a = 0.4$ and $\overline{r} = 0.08$. From this we find that for the US market

$$u(r) = \nu^2 r^{2\beta-1}\left(\beta - \frac{1}{2} - \frac{1}{2a^2}\log(r/\overline{r})\right).$$

The real spot rate is therefore mean-reverting to 8%.[6]

36.6 THE SLOPE OF THE YIELD CURVE AND THE MARKET PRICE OF RISK



Now we have found $w(r)$ and $u(r)$, it only remains for us to find $\lambda(r)$. The model will then be complete. I shall again allow λ to have a spot-rate dependence, but not a time dependence. Note that there is no information about the market price of risk in the spot-rate process; we must look to the yield curve to model this. In particular, we examine the short end of the curve for this information.

Let us expand $Z(r, t; T)$ in a Taylor series about $t = T$; this is the short end of the yield curve. We saw the details of this in Chapter 31. From equation (36.2) we find that

$$Z(r, t; T) \sim 1 - r(T - t) + \tfrac{1}{2}(T - t)^2(r^2 - u + \lambda w) + \dots \quad \text{as } t \to T$$

for any model $w(r)$, $u(r)$ and $\lambda(r)$. From this we have

$$-\frac{\ln Z}{T - t} \sim r + \tfrac{1}{2}(u - \lambda w)(T - t) + \dots \quad \text{as } t \to T. \qquad (36.7)$$

Thus the slope of the yield curve at the short end in this one-factor model is simply $(u - \lambda w)/2$. We can use this result together with time-series data to determine the form for $u - \lambda w$ empirically.

We can calculate a time series for the yield curve slope from one- and three-month US LIBOR data. In Figure 36.4 is plotted the yield curve slope series against the spot rate series, with each data point connected to the following point by a straight line. In Figure 36.5 is the resulting market price of risk, λ, as a function of r. In Figure 36.6 is shown as a function of time. The positive peaks in this figure mean that people are willing to pay to take risk. Note the labeling, fear and greed.

[5] Technically there are four parameters. A mean, a standard deviation, where the distribution starts and whether it goes to the left or right. I have assumed that r is bounded below by zero and that the tail goes to plus infinity.

[6] The mean of the probability density function is *not* the same as the level at which $u = 0$ because of the volatility term in the stochastic differential equation.

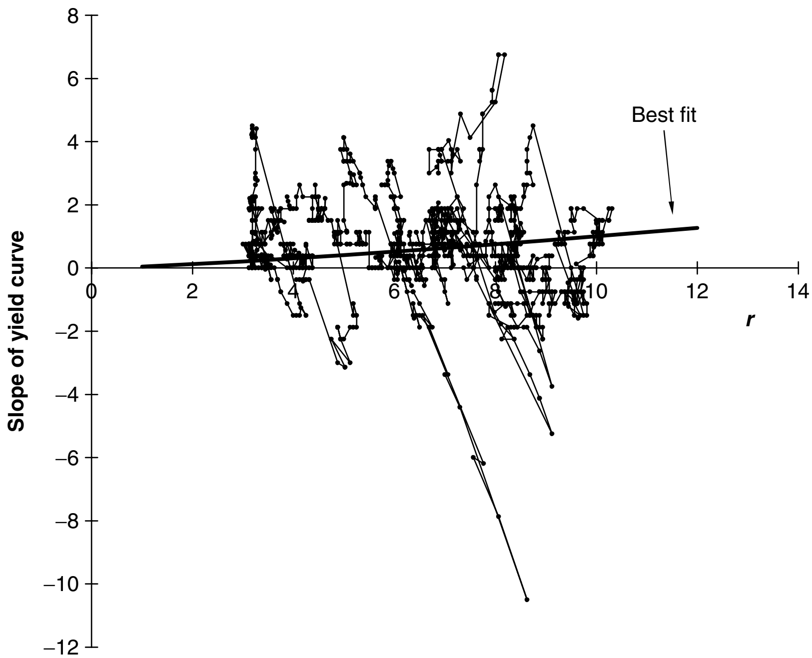


Figure 36.4 Slope of yield curve against spot rate and best fit, US LIBOR data.

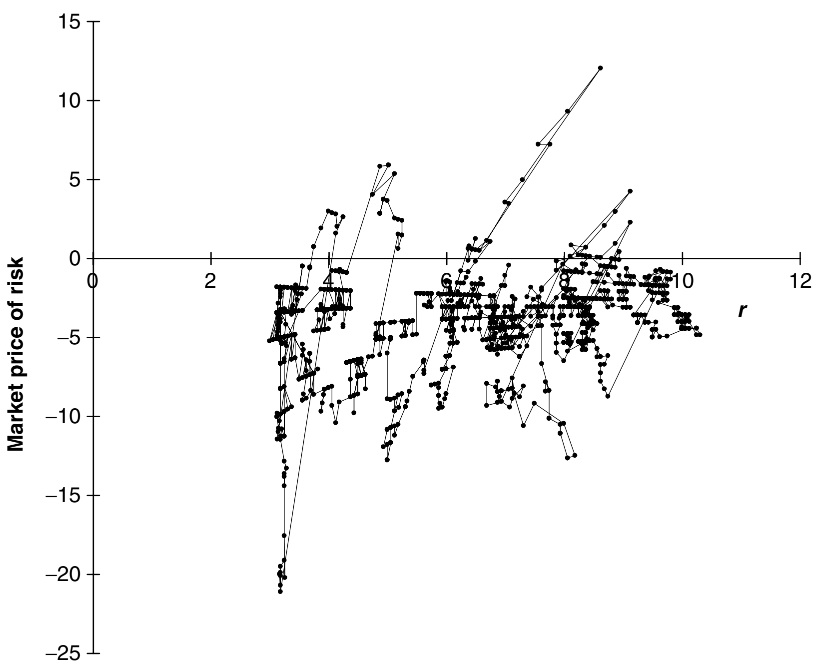


Figure 36.5 Market price of risk against spot rate, US LIBOR data.

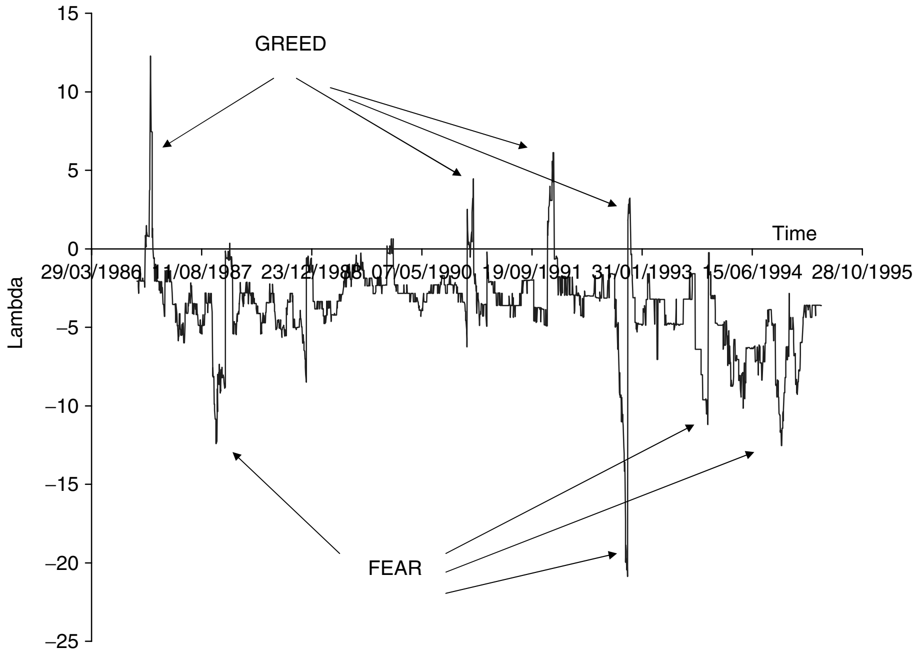


Figure 36.6 Market price of risk versus time, US LIBOR data.

Let us continue with our strategy of finding the best fit with time-independent coefficients. The justification for this is that we are looking at data over a long period, and we choose $\lambda(r)$ to fit Figure 36.4 *on average*; by taking this approach, nothing is brushed under the carpet, we know the limitations of the model.[7]

We find that the simple choice $\lambda(r) = -40r^{\beta-1}$ is a good fit to the average, for US data.

We can now make a comparison between the tractable one-factor models and that presented here. The tractable models require the risk-adjusted drift $u - \lambda w$ to be of the form $a - br$. Our model requires $u - \lambda w$ to be

$$\nu^2 r^{2\beta-1}\left(\beta - \frac{1}{2} - \frac{1}{2a^2}\log(r/\overline{r})\right) - \lambda(r)\nu r^{\beta}.$$

We can immediately see from Figure 36.4 that the form $a - br$ for $u - \lambda w$ is not suitable if a and b are positive. Actually, the BDT risk-adjusted drift rate is much better.

Figure 36.4 is perhaps the most instructive picture. If the plot of $u - \lambda w$ showed it to be single-valued in r then we could easily find $\lambda(r)$. However, it shows no such thing. This calls into question any use of one-factor models. It also strongly suggests that naive fitting of market data gives a false sense of security; as is well known, fitted time-dependent parameters always have to be refitted.[8]

[7] We could instead choose a best/worst $\lambda(r)$ to bound derivative prices. See Chapters 52 and 68 for ideas in this direction.

[8] You can spin all of the facts some of the time, you can spin some of the facts all of the time, but you can't spin all of the facts all of the time. Politicians take note as well.

36.7 WHAT THE SLOPE OF THE YIELD CURVE TELLS US

There are several very important points to note about this result. First, with this choice for λ the drift term u is now much smaller than the λw term. Thus, in this model the risk-adjusted spot rate is mean reverting to a much larger value of the spot rate than 8%. This is a very important observation and highlights the difference between the real and risk-neutral processes. Second, this λ is chosen to match the yield curve slope at the short end. The long rate for such a model is now far too high; the risk-neutral rate reverts to a very high level.

Because of the high yield-curve slope, the conclusion from these observations (and the validity of (36.7) for any model) is that *any* one-factor model is going to hit problems: Either it will be accurate over the short end of the yield curve, with a poor long rate fit, or it will have a good long rate fit with a slope at the short end that is far too shallow. A good comparison to make is between a theoretical forward-rate curve with λ being zero, and a theoretical forward-rate curve with a λ chosen to match the slope. The former gives the *real expected* path of the spot interest rate, with no account being taken of the market price of risk. Generally speaking, empirical evidence suggests that the actual increase or decrease in rates in the short to medium term is much slower than implied by the market yield curve. In Figure 36.7 are shown a typical market forward-rate curve, the expected future path of the short rate using the above model (with zero market price of risk), and the theoretical forward-rate curve when the yield-curve slope is fitted.

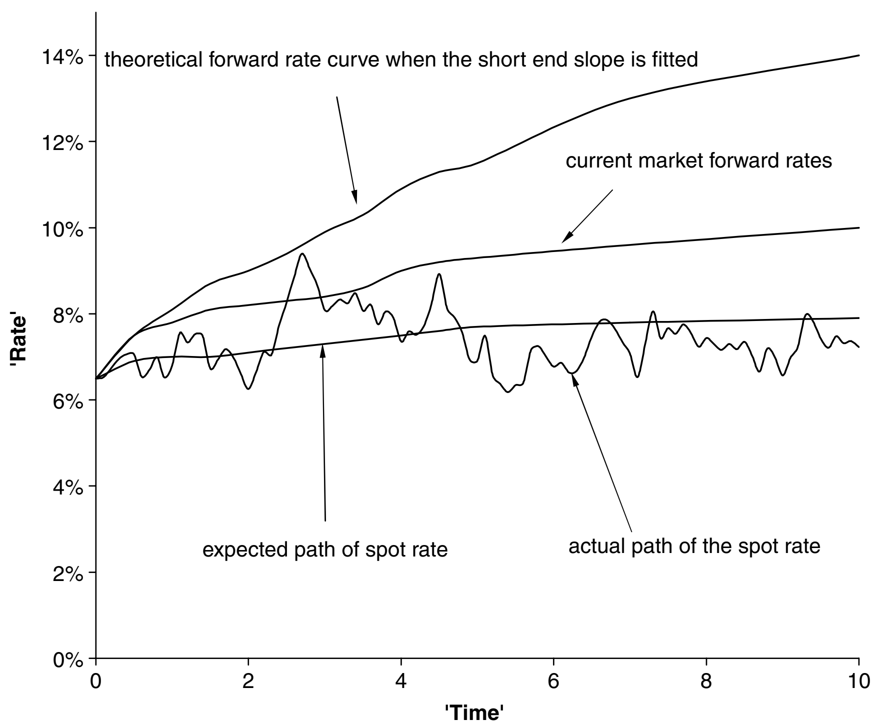


Figure 36.7 Typical market forward rates, the expected future path of the short rate using the empirical model, and the theoretical forward rates when the yield curve slope is fitted at the short end.

When $\lambda = 0$ we are in effect modeling the behavior of the real spot rate. This is a lot easier to do than model the risk-neutral spot rate and the yield curve. The model with $\lambda = 0$ is a good model of the real US spot rate.

Finally, for US data, we find that BDT is the closest popular model to the one we have derived. The BDT volatility structure is realistic and the risk-adjusted drift rate is a fair match to empirical data. The BDT model is close to ours, but the underlying philosophy is fundamentally different—the BDT model was originally chosen for its simple tree structure and to fit data, something which should be treated with care.

36.8 PROPERTIES OF THE FORWARD RATE CURVE 'ON AVERAGE'

Quite reasonably, you might say that the choice of market price of risk should not just depend on the behavior of the yield curve at the short end. What about a more global approach? Continuing with the theme of comparing time-independent models with empirical data, we can look at the behavior of the whole forward rate curve on average. First we need some theory.

From the Fokker–Planck equation we have the following relationship between the real drift, the volatility and the real steady-state probability distribution:

$$u = \frac{1}{2p_\infty}\frac{d}{dr}(w^2 p_\infty).$$

Similarly, if we denote the risk-neutral steady-state distribution by p_∞^* then we have

$$u - \lambda w = \frac{1}{2p_\infty^*}\frac{d}{dr}(w^2 p_\infty^*).$$

Two things follow from these results. First, we can eliminate u between them to get

$$p_\infty^* = p_\infty e^{-2\int^r \frac{\lambda(s)}{w(s)} ds}.$$

The relationship between the real and risk-neutral steady-state distributions obviously depends on the market price of risk. Second, we can eliminate $u - \lambda w$ from the bond pricing equation to get

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2} + \frac{1}{2p_\infty^*}\frac{d}{dr}(w^2 p_\infty^*)\frac{\partial Z}{\partial r} - rZ = 0.$$

This can be written as

$$p_\infty^*(r)\left(\frac{\partial Z}{\partial t} - rZ\right) = -\tfrac{1}{2}\frac{\partial}{\partial r}\left(w^2 p_\infty^*\frac{\partial Z}{\partial r}\right).$$

In terms of the real distribution we have

$$p_\infty(r)e^{-2\int^r \frac{\lambda(s)}{w(s)} ds}\left(\frac{\partial Z}{\partial t} - rZ\right) = -\tfrac{1}{2}\frac{\partial}{\partial r}\left(w^2 p_\infty^*\frac{\partial Z}{\partial r}\right).$$

If we integrate this over the range of permitted r then the right-hand side, being a perfect derivative with nice properties at the end of the range, is zero leaving us with

$$\int_0^\infty p_\infty(r)e^{-2\int^r \frac{\lambda(s)}{w(s)}ds}\left(\frac{\partial Z}{\partial t} - rZ\right)dr = 0.$$

The range need not be zero to infinity but this would be the case for our earlier choices of w and p_∞. Since there is a real probability density function in this integral it is exactly the same as the real expectation, not a risk-neutral expectation. In other words, the average value of

$$e^{-2\int^r \frac{\lambda(s)}{w(s)}ds}\left(\frac{\partial Z}{\partial t} - rZ\right)$$

must be zero. Because the risk-neutral spot rate model is time homogeneous we can write

$$Z(t^*; T) = Z(\tau) \quad \text{where} \quad \tau = T - t^*.$$

We can now write the integral equation as

$$E_\infty\left[g(r)\left(\frac{\partial Z}{\partial \tau} + Z(\tau)\right)\right] = 0 \text{ for all } \tau \tag{36.8}$$

where

$$g(r) = e^{-2\int^r \frac{\lambda(s)}{w(s)}ds}$$

and τ is the maturity. Note that this must be true for all maturities τ. This result follows from the equivalence of time and space averages, the ergodic property. It would not be true if we had any time inhomogeneity.

This result is particularly interesting because it is a property involving both the real and risk-neutral probabilities. Previously we had properties such as the shape of an individual yield curve or its slope that depend on the risk-neutral probabilities, or the distribution of the spot interest rate that depends on the real probabilities. However, the real 'average' of the forward rate curve depends on both kinds of probabilities. Can we use this property to estimate the market price of risk?

From discount factor data we can find the forward rate curve and the spot interest rate. If we have enough data we can estimate reliable averages. Can we find a non-zero function $g(r)$ such that (36.8) is satisfied. The answer is yes. Can we find a *positive* function $g(r)$? This is much, much harder. And we need the function to be positive so that we can take its logarithm to find the market price of risk; we don't want an imaginary market price of risk. If we can't find such a function (and I haven't been able to using US data), then again we must worry about the ability of any one-factor interest rate model to give properties that bear any resemblance to reality.

36.9 IMPLIED MODELING: TWO FACTOR

In the first part of this chapter I showed one way of implying the model for the spot interest rate from empirical data for the spot rate and the slope of the yield curve. In this section I show how to apply this approach to modeling *two* factors instead of one. These are only hints at possible directions for future research. Such an approach is important, however, because it is well known that some popular two-factor models have rather unfortunate properties (such as rates reaching infinity in a finite time).

The general two-factor interest rate model is

$$dr = u\,dt + w\,dX_1$$

and

$$dl = p\,dt + q\,dX_2,$$

for the *real* rates. I am going to assume that l is a long rate, even though this has some internal consistency problems.[9] The correlation coefficient is ρ. I am going to suggest ways in which the functions u, w, p, q, ρ and λ_r can be found from empirical data. Again we assume that these functions are independent of time but, of course, can all be functions of r and l. The following is a list of the steps to be taken to determine the functional forms of all of these six terms.

- The first two terms to be found are the volatility functions, w and q. For simplicity, assume that w and q are of the form $wr^{\alpha}l^{\beta}$ (with different values for w, α and β for each of r and l) and examine spot and long rate volatilities for the best fit. This step is identical to that in the one-factor implied model explained above.
- The next term to be modeled is the correlation ρ. At least in the US, it is a well-known empirical observation that the spot rate and the spread between long and short rates are uncorrelated. This result can be used to determine the ρ.
- Next examine the slope of the yield curve. This gives information about the risk-adjusted drift of the spot interest rate. By performing a Taylor series expansion about short maturities, we find that the risk-adjusted drift for the spot rate is still related to the slope of the yield curve even in a two-factor world. This can be used to model the function λ_r.
- Next examine the real drift rates. If we assume that we have already modeled the volatilities and the correlation, then the empirical steady-state probability density function for r and l gives one piece of information about the two functions u and p, the real drift rates. The relationship between the real drifts and the steady-state probability density function is via a steady-state two-factor forward Fokker–Planck equation.
- The sixth, and last, piece of the puzzle comes from an examination of the model dynamics in the absence of randomness. In other words, what behavior does the model exhibit when we drop the stochastic terms? The answer comes from the solution of the two ordinary differential equations

  $$dr = u(r, l)\,dt$$

  and

  $$dl = p(r, l)\,dt.$$

  From these two equations we can eliminate the time dependence to get the one ordinary differential equation

  $$\frac{dl}{dr} = \frac{p(r, l)}{u(r, l)}.$$

[9] These problems are only associated with risk-neutral rates and not real rates.

The ratio of p to u is chosen to give the solution of this equation similar characteristics to that of the empirical data of l versus r. For example, it is at this stage that we can incorporate economic cycles into the model.

36.10 SUMMARY

Modeling interest rates is very hard. There are few economic rules of thumb to help us. In modeling equities we knew that the return was important, that the return should be independent of the level of the asset. This helped enormously with the modeling of equities, in fact it left us with little modeling to do. We have no such guide with interest rates. All we know is that rates are positive (although there have been a few pathological cases where this was not true) and don't grow unboundedly. That's not much to go on.

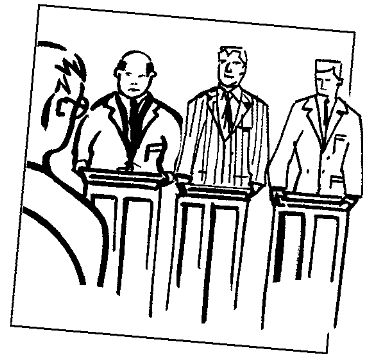
So, in this chapter, I showed how to examine data to see for yourself what a plausible model for rates might be. The reader is encouraged to adopt the approach of modeling for himself, rather than taking some ready-made model from the literature.

The conclusion, as far as one-factor models are concerned, has to be to use them with care. They don't model reality in any quantitative way, and it is not clear whether yield-curve fitting works in the long run or whether it is simply a way of putting in place a delta hedge that will work for just a short time.

FURTHER READING

- Chan, Karolyi, Longstaff & Sanders (1992) examine the US spot rate in great detail and discuss the impact of their results on the validity of popular models.
- See Apabhai, Choe, Khennach & Wilmott (1995) for further details of the approaches to the modeling in this chapter.
- Apabhai (1995) describes the two-factor implied analysis.

CHAPTER 37
the Heath, Jarrow & Morton and Brace, Gatarek & Musiela models



In this Chapter...

- the Heath, Jarrow & Morton (HJM) forward rate model
- the relationship between HJM and spot rate models
- the advantages and disadvantages of the HJM approach
- how to decompose the random movements of the forward rate curve into its principal components
- the Brace, Gatarek & Musiela or LIBOR Market Model

37.1 INTRODUCTION

The **Heath, Jarrow & Morton** approach to the modeling of the whole forward rate curve was a major breakthrough in the pricing of fixed-income products. They built up a framework that encompassed all of the models we have seen so far (and many that we haven't). Instead of modeling a short-term interest rate and deriving the forward rates (or, equivalently, the yield curve) from that model, they boldly start with a model for the whole forward rate curve. Since the forward rates are known today, the matter of yield-curve fitting is contained naturally within their model; it does not appear as an afterthought. Moreover, it is possible to take *real data* for the random movement of the forward rates and incorporate them into the derivative-pricing methodology.

37.2 THE FORWARD RATE EQUATION

The key concept in the HJM model is that we model the evolution of the whole forward rate curve, not just the short end. Write $F(t; T)$ for the forward rate curve at time t. Thus the price of a zero-coupon bond at time t and maturing at time T, when it pays \$1, is

$$Z(t;T) = e^{-\int_t^T F(t;s)\,ds}. \tag{37.1}$$

Let us assume that all zero-coupon bonds evolve according to

$$dZ(t;T) = \mu(t,T)Z(t;T)\,dt + \sigma(t,T)Z(t;T)\,dX. \tag{37.2}$$

This is not much of an assumption, other than to say that it is a one-factor model, and I will generalize that later. In this $d\cdot$ means that time t evolves but the maturity date T is fixed. Note that since $Z(t;t) = 1$ we must have $\sigma(t,t) = 0$. From (37.1) we have

$$F(t;T) = -\frac{\partial}{\partial T}\log Z(t;T).$$

Differentiating this with respect to t and substituting from (37.2) results in an equation for the evolution of the forward curve:

$$dF(t;T) = \frac{\partial}{\partial T}\left(\tfrac{1}{2}\sigma^2(t,T) - \mu(t,T)\right)dt - \frac{\partial}{\partial T}\sigma(t,T)\,dX. \tag{37.3}$$

In Figure 37.1 is shown the forward rate curve today, time t^*, and a few days later. The whole curve has moved according to (37.3).

Where has this got us? We have an expression for the drift of the forward rates in terms of the volatility of the forward rates. There is also a μ term, the drift of the bond. We have seen many times before how such drift terms disappear when we come to pricing derivatives, to be replaced by the risk-free interest rate r. Exactly the same will happen here.

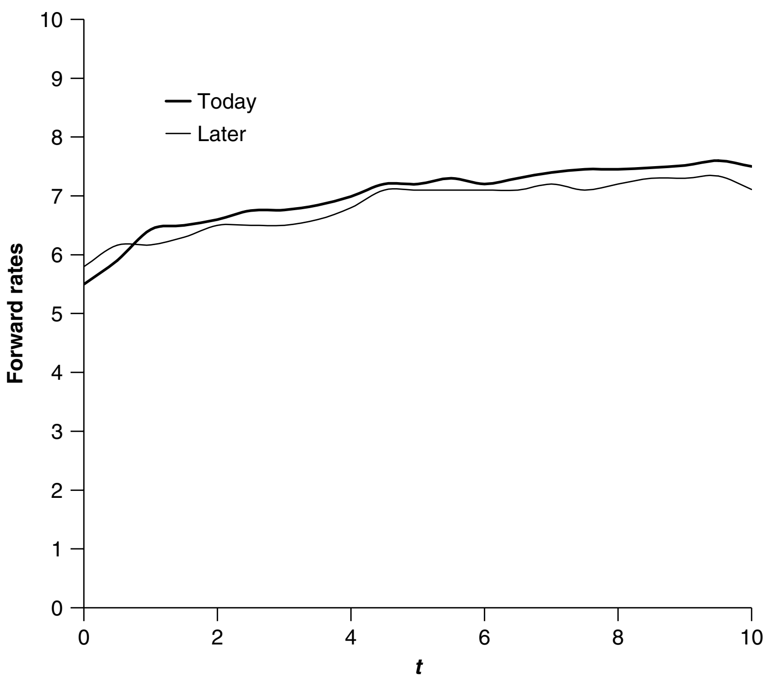


Figure 37.1 The forward rate curve today and a few days later.

37.3 THE SPOT RATE PROCESS

The spot interest rate is simply given by the forward rate for a maturity equal to the current date i.e.

$$r(t) = F(t;t).$$

In this section I am going to manipulate this expression to derive the stochastic differential equation for the spot rate. In so doing we will begin to see why the HJM approach can be slow to price derivatives.

Suppose today is t^* and that we know the whole forward rate curve today, $F(t^*;T)$. We can write the spot rate for *any* time t in the future as

$$r(t) = F(t;t) = F(t^*;t) + \int_{t^*}^{t} dF(s;t).$$

From our earlier expression (37.3) for the forward rate process for F we have

$$r(t) = F(t^*;t) + \int_{t^*}^{t} \left(\sigma(s,t)\frac{\partial \sigma(s,t)}{\partial t} - \frac{\partial \mu(s,t)}{\partial t} \right) ds - \int_{t^*}^{t} \frac{\partial \sigma(s,t)}{\partial t} dX(s).$$

Differentiating this with respect to time t we arrive at the stochastic differential equation for r

$$\begin{aligned} dr = \Bigg(& \frac{\partial F(t^*;t)}{\partial t} - \frac{\partial \mu(t,s)}{\partial s}\Big|_{s=t} \\ & + \int_{t^*}^{t} \left(\sigma(s,t)\frac{\partial^2 \sigma(s,t)}{\partial t^2} + \left(\frac{\partial \sigma(s,t)}{\partial t}\right)^2 - \frac{\partial^2 \mu(s,t)}{\partial t^2} \right) ds \\ & \underline{- \int_{t^*}^{t} \frac{\partial^2 \sigma(s,t)}{\partial t^2} dX(s)} \Bigg) dt - \frac{\partial \sigma(t,s)}{\partial s}\Big|_{s=t} dX. \end{aligned}$$

37.3.1 The Non-Markov Nature of HJM

The details of this expression are not important. I just want you to observe one point. Compare this stochastic differential equation for the spot rate with any of the models in Chapter 30. Clearly, it is more complicated, there are many more terms. All but the last one are deterministic, the last is random. The important point concerns the nature of these terms. In particular, the term underlined depends on the history of σ from the date t^* to the future date t, and *it depends on the history of the stochastic increments dX*. This term is thus highly path-dependent. Moreover, for a general HJM model it makes the motion of the spot rate **non-Markov**. In a **Markov process** it is only the present state of a variable that determines the possible future (albeit random) state. Having a non-Markov model may not matter to us if we can find a small number of extra state variables that contain all the information that we need for predicting the future.[1] Unfortunately, the general HJM model requires an infinite number of such variables to define the present state; if we were to write the HJM model as a partial differential equation we would need an infinite number of independent variables.

[1] Remember Asian option pricing.

At the moment we are in the real world. To price derivatives we need to move over to the risk-neutral world. The first step in this direction is to see what happens when we hold a hedged portfolio.

37.4 THE MARKET PRICE OF RISK

In the one-factor HJM model all stochastic movements of the forward rate curve are perfectly correlated. We can therefore hedge one bond with another of a different maturity. Such a hedged portfolio is

$$\Pi = Z(t; T_1) - \Delta Z(t; T_2).$$

The change in this portfolio is given by

$$\begin{aligned} d\Pi &= dZ(t; T_1) - \Delta dZ(t; T_2) \\ &= Z(t; T_1)\left(\mu(t, T_1)\, dt + \sigma(t, T_1)\, dX\right) - \Delta Z(t; T_2)\left(\mu(t, T_2)\, dt + \sigma(t, T_2)\, dX\right). \end{aligned}$$

If we choose

$$\Delta = \frac{\sigma(t, T_1) Z(t; T_1)}{\sigma(t, T_2) Z(t; T_2)}$$

then our portfolio is hedged, is risk free. Setting its return equal to the risk-free rate $r(t)$ and rearranging we find that

$$\frac{\mu(t, T_1) - r(t)}{\sigma(t, T_1)} = \frac{\mu(t, T_2) - r(t)}{\sigma(t, T_2)}.$$

The left-hand side is a function of T_1 and the right-hand side is a function of T_2. This is only possible if both sides are independent of the maturity date T:

$$\mu(t, T) = r(t) + \lambda(t)\sigma(t, T).$$

As before, $\lambda(t)$ is the market price of risk (associated with the one factor).

37.5 REAL AND RISK NEUTRAL

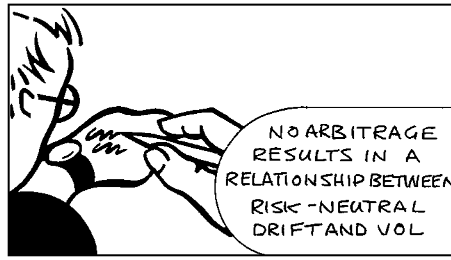
We are almost ready to price derivatives using the HJM model. But first we must discuss the real and risk-neutral worlds, relating them to the ideas in previous chapters.

All of the variables I have introduced above have been *real* variables. But when we come to pricing derivatives we must do so in the risk-neutral world. In the present HJM context, risk-neutral 'means' $\mu(t, T) = r(t)$. This means that in the risk-neutral world the return on any traded investment is simply $r(t)$. We can see this in (37.2). The risk-neutral zero-coupon bond price satisfies

$$dZ(t; T) = r(t) Z(t; T)\, dt + \sigma(t, T) Z(t; T)\, dX.$$

The deterministic part of this equation represents exponential growth of the bond at the risk-free rate. The form of the equation is very similar to that for a risk-neutral equity, except that here the volatility will be much more complicated.

37.5.1 The Relationship Between the Risk-neutral Forward Rate Drift and Volatility



Let me write the stochastic differential equation for the *risk-neutral* forward rate curve as

$$dF(t;T) = m(t,T)\,dt + \nu(t,T)\,dX.$$

From (37.3)

$$\nu(t,T) = -\frac{\partial}{\partial T}\sigma(t,T),$$

is the forward rate volatility and, from (37.3), the drift of the forward rate is given by

$$\frac{\partial}{\partial T}\left(\tfrac{1}{2}\sigma^2(t,T) - \mu(t,T)\right) = \nu(t,T)\int_t^T \nu(t,s)\,ds - \frac{\partial}{\partial T}\mu(t,T),$$

where we have used $\sigma(t,t) = 0$. In the risk-neutral world we have $\mu(t,T) = r(t)$, and so the drift of the risk-neutral forward rate curve is related to its volatility by

$$m(t,T) = \nu(t,T)\int_t^T \nu(t,s)\,ds. \tag{37.4}$$

Therefore

$$dF(t;T) = \nu(t,T)\left(\int_t^T \nu(t,s)\,ds\right)dt + \nu(t,T)\,dX. \tag{37.5}$$

37.6 PRICING DERIVATIVES

Pricing derivatives is all about finding the expected present value of all cashflows in a risk-neutral framework. This was discussed in Chapter 10, in terms of equity, currency and commodity derivatives. If we are lucky then this calculation can be done via a low-dimensional partial differential equation. We have seen the relevant theory in Chapter 10. The one- and two-factor models of Chapters 30 and 35 exploited this. The HJM model, however, is a very general interest rate model and in its full generality one cannot write down a finite-dimensional partial differential equation for the price of a derivative.

Because of the non-Markov nature of HJM in general a partial differential equation approach is infeasible. This leaves us with two alternatives. One is to estimate directly the necessary expectations by simulating the random evolution of, in this case, the risk-neutral forward rates. The other is to build up a tree structure.

37.7 SIMULATIONS

If we want to use a Monte Carlo method, then we must simulate the evolution of the whole forward rate curve, calculate the value of all cashflows under each evolution and then calculate the present value of these cashflows by *discounting at the realized spot rate* $r(t)$.

To price a derivative using a Monte Carlo simulation perform the following steps. I will assume that we have chosen a model for the forward rate volatility, $\nu(t,T)$. Today is t^* when we know the forward rate curve $F(t^*;T)$.

1. Simulate a realized evolution of the whole risk-neutral forward rate curve for the necessary length of time, until T^*, say. This requires a simulation of

$$dF(t;T) = m(t,T)\,dt + \nu(t,T)\,dX,$$

where

$$m(t,T) = \nu(t,T)\int_t^T \nu(t,s)\,ds.$$

After this simulation we will have a realization of $F(t;T)$ for $t^* \le t \le T^*$ and $T \ge t$. We will have a realization of the whole forward rate path.

2. At the end of the simulation we will have the realized prices of all maturity zero-coupon bonds at every time up to T^*.
3. Using this forward rate path calculate the value of all the cashflows that would have occurred.
4. Using the realized path for the spot interest rate $r(t)$ calculate the present value of these cashflows. Note that we discount at the continuously compounded risk-free rate, not at any other rate. In the risk-neutral world all assets have an expected return of $r(t)$.
5. Return to Step 1 to perform another realization, and continue until you have a sufficiently large number of realizations to calculate the expected present value as accurately as required.

The disadvantage of the HJM model is that a Monte Carlo simulation such as this can be slow. On the plus side, because the whole forward rate curve is calculated, the bond prices at all maturities are trivial to find during this simulation.

37.8 TREES

If we are to build up a tree for a non-Markov model then we find ourselves with the unfortunate result that the forward curve after an up move followed by a down is *not* the same as the curve after a down followed by an up. The equivalence of these two paths in the Markov world is what makes the binomial method so powerful and efficient. In the non-Markov world our tree structure becomes 'bushy,' and grows *exponentially* in size with the addition of new time steps.

If the contract we are valuing is European with no early exercise then we don't need to use a tree; a Monte Carlo simulation can be immediately implemented. However, if the contract has some American feature then to price correctly in the early exercise we don't have much choice but to use a tree structure. The exponentially-large tree structure will make the pricing problem very slow.

37.9 THE MUSIELA PARAMETERIZATION

Often in practice the model for the volatility structure of the forward rate curve will be of the form

$$\nu(t,T) = \overline{\nu}(t,T-t),$$

meaning that we will model the volatility of the forward rate at each maturity, one, two, three years, and not at each maturity date, 2006, 2007, ... If we write τ for the maturity period $T - t$ then it is a simple matter to find that $\overline{F}(t; \tau) = F(t, t + \tau)$ satisfies

$$d\overline{F}(t; \tau) = \overline{m}(t, \tau)\, dt + \overline{\nu}(t, \tau)\, dX,$$

where

$$\overline{m}(t, \tau) = \overline{\nu}(t, \tau) \int_0^\tau \overline{\nu}(t, s)\, ds + \frac{\partial}{\partial \tau}\overline{F}(t, \tau).$$

It is much easier in practice to use this representation for the evolution of the risk-neutral forward rate curve.

37.10 MULTI-FACTOR HJM

Often a single-factor model does not capture the subtleties of the yield curve that are important for particular contracts. The obvious example is the spread option, which pays off the difference between rates at two different maturities. We then require a multi-factor model. The multi-factor theory is identical to the one-factor case, so we can simply write down the extension to many factors.

If the risk-neutral forward rate curve satisfies the N-dimensional stochastic differential equation

$$dF(t, T) = m(t, T)\, dt + \sum_{i=1}^{N} \nu_i(t, T)\, dX_i,$$

where the dX_i are uncorrelated, then

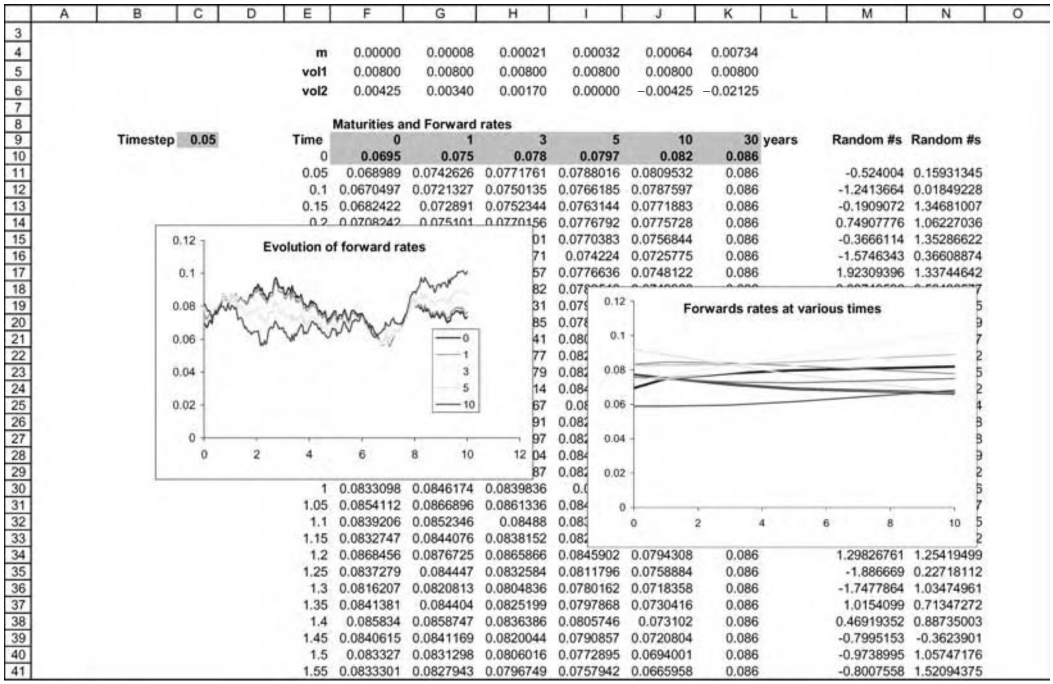
$$m(t, T) = \sum_{i=1}^{N} \nu_i(t, T) \int_t^T \nu_i(t, s)\, ds.$$

Therefore

$$dF(t, T) = \left(\sum_{i=1}^{N} \nu_i(t, T) \int_t^T \nu_i(t, s)\, ds \right) dt + \sum_{i=1}^{N} \nu_i(t, T)\, dX_i. \qquad (37.6)$$

37.11 SPREADSHEET IMPLEMENTATION

The HJM model is very easy to implement on a spreadsheet. Figure 37.2 shows the results of such a simulation for a two-factor model. In this example the Musiela parameterization has been used and $\overline{\nu}$ is a function of $\tau = T - t$ only. This means that the function $\overline{M}$ contains the first volatility term which is just a function of time to maturity (this is row 3 in the figure) and the slope of the forward curve term (this latter is calculated in each of the cells from row 11 down). In this example there are two volatility factors, the first is constant and the second is linear in τ.



	m	vol1	vol2
F	0.00000	0.00800	0.00425
G	0.00008	0.00800	0.00340
H	0.00021	0.00800	0.00170
I	0.00032	0.00800	0.00000
J	0.00064	0.00800	−0.00425
K	0.00734	0.00800	−0.02125

Timestep 0.05

Maturities and Forward rates

Time	0	1	3	5	10	30 years	Random #s	Random #s
0	0.0695	0.075	0.078	0.0797	0.082	0.086		
0.05	0.068989	0.0742626	0.0771761	0.0788016	0.0809532	0.086	-0.524004	0.15931345
0.1	0.0670497	0.0721327	0.0750135	0.0766185	0.0787597	0.086	-1.2413664	0.01849228
0.15	0.0682422	0.072891	0.0752344	0.0763144	0.0771883	0.086	-0.1909072	1.34681007
0.2	0.0708242	0.075101	0.0770156	0.0776792	0.0775728	0.086	0.74907776	1.06227036
				0.0770383	0.0756844	0.086	-0.3666114	1.35286622
				0.074224	0.0725775	0.086	-1.5746343	0.36608874
				0.0776636	0.0748122	0.086	1.92309396	1.33744642
1	0.0833098	0.0846174	0.0839836					
1.05	0.0854112	0.0866896	0.0861336					
1.1	0.0839206	0.0852346	0.08488					
1.15	0.0832747	0.0844076	0.0838152					
1.2	0.0868456	0.0876725	0.0865866	0.0845902	0.0794308	0.086	1.29826761	1.25419499
1.25	0.0837279	0.084447	0.0832584	0.0811796	0.0758884	0.086	-1.886669	0.22718112
1.3	0.0816207	0.0820813	0.0804836	0.0780162	0.0718358	0.086	-1.7477864	1.03474961
1.35	0.0841381	0.084404	0.0825199	0.0797868	0.0730416	0.086	1.0154099	0.71347272
1.4	0.085834	0.0858747	0.0836386	0.0805746	0.073102	0.086	0.46919352	0.88735003
1.45	0.0840615	0.0841169	0.0820044	0.0790857	0.0720804	0.086	-0.7995153	-0.3623901
1.5	0.083327	0.0831298	0.0806016	0.0772895	0.0694001	0.086	-0.9738995	1.05747176
1.55	0.0833301	0.0827943	0.0796749	0.0757942	0.0665958	0.086	-0.8007558	1.52094375

Figure 37.2 Spreadsheet showing results of a two-factor HJM simulation.

37.12 A SIMPLE ONE-FACTOR EXAMPLE: HO & LEE

In this section we make a comparison between the spot rate modeling of Chapter 30 and HJM. One of the key points about the HJM approach is that the yield curve is fitted, by default. The simplest yield-curve fitting spot rate model is Ho & Lee, so we draw a comparison between this and HJM.

In Ho & Lee the risk-neutral spot rate satisfies

$$dr = \eta(t)\,dt + c\,dX,$$

for a constant c. The prices of zero-coupon bonds, $Z(r, t; T)$, in this model satisfy

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2}c^2\frac{\partial^2 Z}{\partial r^2} + \eta(t)\frac{\partial Z}{\partial r} - rZ = 0$$

with

$$Z(r, T; T) = 1.$$

The solution is easily found to be

$$Z(r, t; T) = \exp\left(\tfrac{1}{6}c^2(T-t)^3 - \int_t^T \eta(s)(T-s)\,ds - (T-t)r\right).$$

In the Ho & Lee model $\eta(t)$ is chosen to fit the yield curve at time t^*. In forward rate terms this means that

$$F(t^*; T) = r(t^*) - \tfrac{1}{2}c^2(T - t^*)^2 + \int_{t^*}^{T} \eta(s)\, ds,$$

and so

$$\eta(t) = \frac{\partial F(t^*; t)}{\partial t} + c^2(t - t^*).$$

At any time later than t^*

$$F(t; T) = r(t) - \tfrac{1}{2}c^2(T - t)^2 + \int_{t}^{T} \eta(s)\, ds.$$

From this we find that

$$dF(t; T) = c^2(T - t)\, dt + c\, dX.$$

In our earlier notation, $\sigma(t, T) = -c(T - t)$ and $\nu(t, T) = c$. This is the evolution equation for the risk-neutral forward rates. It is easily confirmed for this model that Equation (37.4) holds. This is the HJM representation of the Ho & Lee model. Most of the popular models have HJM representations.

37.13 PRINCIPAL COMPONENT ANALYSIS

There are two main ways to use HJM. One is to choose the volatility structure $\nu_i(t, T)$ to be sufficiently 'nice' to make a tractable model, one that is Markov. This usually leads us back to the 'classical' popular spot-rate models. The other way is to choose the volatility structure to match data. This is where Principal Component Analysis (PCA) comes in.

In analyzing the volatility of the forward rate curve one usually assumes that the volatility structure depends only on the time to maturity i.e.

$$\nu = \overline{\nu}(T - t).$$

I will assume this but examine a more general multi-factor model:

$$dF(t; T) = m(t, T)\, dt + \sum_{i=1}^{N} \overline{\nu}_i(T - t)\, dX_i.$$

From time series data we can determine the functions $\overline{\nu}_i$ empirically; this is **Principal Components Analysis**. I will give a loose description of how this is done, with more details in the spreadsheets.

If we have forward rate time series data going back a few years we can calculate the covariances between the *changes* in the rates of different maturities. We may have, for example, the one-, three-, six-month, one-, two, three-, five-, seven-, 10- and 30-year rates. The covariance matrix would then be a ten $\times$ ten symmetric matrix with the variances of the rates along the diagonal and the covariances between rates off the diagonal.

	A	B	C	D	E	F	G	H	I	J	K
1	**Forward rates:**				**Changes in rates:**						
2		**1 month**	**3 month**	**6 month**	**1 month**	**3 month**	**6 month**				
3	22-Sep-88	8.25000	8.31250	8.56250							
4	23-Sep-88	8.25000	8.31250	8.56250	0.00000	0.00000	0.00000				
5	26-Sep-88	8.31250	8.37500	8.62500	0.06250	0.06250	0.06250				
6	27-Sep-88	8.31250	8.43750	8.6[illegible]	[illegible]0000	0.06250	0.06250		**1 month**	**3 month**	**6 month**
7	28-Sep-88	8.42188	8.50000	8.81250	0.10938	0.06250	0.12500	**1 month**	0.007501		
8	29-Sep-88	8.37500	8.68750	8.81250	-0.04688	0.18750	0.00000	**3 month**	0.003831	0.004225	
9	30-Sep-88	8.31250	8.62500	8.75000	-0.06250	-0.06250	-0.06250	**6 month**	0.003628	0.004020	0.004997
10	3-Oct-88	8.31250	8.62500	8.68750	0.00000	0.00000	-0.06250				
11	4-Oct-88	8.31250	8.56250	8.68750	0.00000	-0.06250	0.00000	**Scaled covariance matrix:**			
12	5-Oct-88	8.31250	8.56250	8.68750	0.00000	0.00000	0.00000		**1 month**	**3 month**	**6 month**
13	6-Oct-88	8.31250	8.56250	8.68750	0.00000	0.00000	0.00000	**1 month**	0.000189		
14	7-Oct-88	8.31250	8.62500	8.75000	0.00000	0.06250	0.06250	**3 month**	0.000097	0.000106	
15	10-Oct-88	8.25000	8.56250	8.56250	-0.06250	-0.06250	-0.18750	**6 month**	0.000091	0.000101	0.000126
16	11-Oct-88	8.25000	8.56250	8.62500	0.00000	0.00000	0.06250				
17	12-Oct-88	8.31250	8.62500	8.68750	0.06250	0.06250	0.06250				
18	13-Oct-88	8.31250	8.64063	8.68750	0.00000	0.01563	0.00000				
19	14-Oct-88	8.31250	8.62500	8.62500	0.00000	-0.01563	-0.06250				
20	17-Oct-88	8.31250	8.62500	8.62500	0.00000	0.00000	0.00000				
21	18-Oct-88	8.31250	8.62500	8.62500	0.00000	0.00000	0.00000				
22	19-Oct-88	8.31250	8.62500	8.62500	0.00000	0.00000	0.00000				
23	20-Oct-88	8.37500	8.68750	8.68750	0.06250	0.06250	0.06250				
24	21-Oct-88	8.37500	8.68750	8.68750	0.00000	0.00000	0.00000				
25	24-Oct-88	8.37500	8.68750	8.75000	0.00000	0.00000	0.06250				
26	25-Oct-88	8.37500	8.68750	8.75000	0.00000	0.00000	0.00000				
27	26-Oct-88	8.37500	8.68750	8.75000	0.00000	0.00000	0.00000				
28	27-Oct-88	8.37500	8.68750	8.68750	0.00000	0.00000	-0.06250				

=B4-B3

=COVAR(E4:E1721,F4:F1721)

=I8*252/10000

Figure 37.3 One-, three- and sixth-month rates and the changes.

In Figure 37.3 is shown a spreadsheet of daily one-, three- and sixth-month rates, and the day-to-day changes. The covariance matrix for these changes is also shown.

PCA is a technique for finding common movements in the rates, for essentially finding eigenvalues and eigenvectors of the matrix. We expect to find, for example, that a large part of the movement of the forward rate curve is common between rates; that a parallel shift in the rates is the largest component of the movement of the curve in general. The next most important movement would be a twisting of the curve, followed by a bending.

Suppose that we have found the covariance matrix, $\mathbf{M}$, for the changes in the rates mentioned above. This 10 by 10 matrix will have 10 eigenvalues, λ_i, and eigenvectors, $\mathbf{v}_i$ satisfying

$$\mathbf{M}\mathbf{v}_i = \lambda_i \mathbf{v}_i;$$

$\mathbf{v}_i$ is a column vector.

The eigenvector associated with the largest eigenvalue is the first principal component. It gives the dominant part in the movement of the forward rate curve. Its first entry represents the movement of the one-month rate, the second entry is the three-month rate etc. Its eigenvalue is the variance of these movements. In Figure 37.4 we see the entries in this first principal component plotted against the relevant maturity. This curve is relatively flat, when compared with the other components. This indicates that, indeed, a parallel shift of the yield curve is the dominant movement. Note that the eigenvectors are orthogonal; there is no correlation between the principal components.

In this figure are also plotted the next two principal components. Observe that one gives a twisting of the curve and the other a bending.

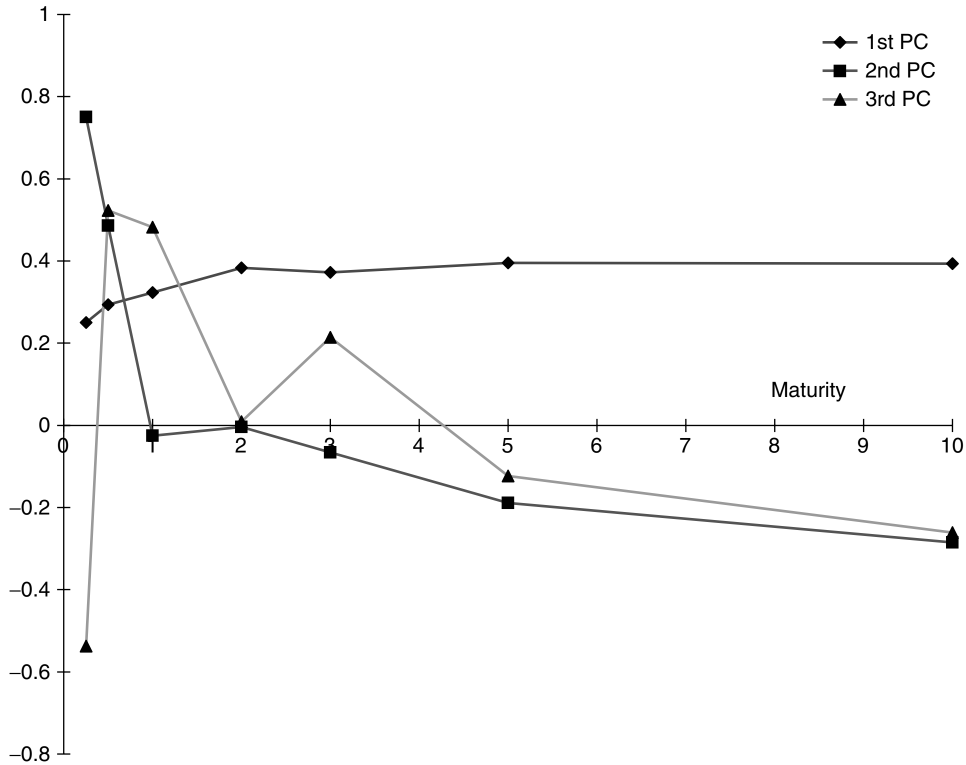


Figure 37.4 The first three principal components for the US forward rate curve. The data run from 1988 until 1996.

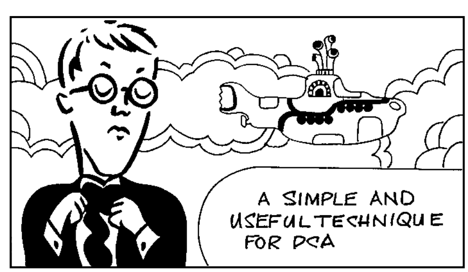
The result of this analysis is that the volatility factors are given by

$$\overline{\nu}_i(\tau_j) = \sqrt{\lambda_i}(\mathbf{v}_i)_j.$$

Here τ_j is the maturity i.e. $1/12$, $1/4$ etc. and $(\mathbf{v}_i)_j$ is the jth entry in the vector $\mathbf{v}_i$. To get the volatility of other maturities will require some interpolation.

The calculation of the covariance matrix is simple, and discussed in Chapter 11. The calculation of the eigenvalues and vectors is also simple if you use the following algorithm.

37.13.1 The Power Method



I will assume that all the eigenvalues are distinct, a reasonable assumption given the empirical nature of the matrix. Since the matrix is symmetric positive definite (it is a covariance matrix) we have all the nice properties we need. The eigenvector associated with the *largest* eigenvalue is easily found by the following iterative procedure. First make an initial guess for the eigenvector, call it $\mathbf{x}^0$. Now iterate using

$$\mathbf{y}^{k+1} = \mathbf{M}\mathbf{x}^k,$$

for $k = 0, \ldots,$ and

$$\beta^{k+1} = \text{element of } \mathbf{y}^{k+1} \text{ having largest modulus}$$

followed by

$$\mathbf{x}^{i+1} = \frac{1}{\beta^{k+1}}\mathbf{y}^{k+1}.$$

As $k \to \infty$, $\mathbf{x}^k$ tends to the eigenvector and β^k to the eigenvalue λ. In practice you would stop iterating once you have reached some set tolerance. Thus we have found the first principal component. It is standard to normalize the vector, and this is our $\mathbf{v}_1$.

To find the next principal component we must define a new matrix by

$$\mathbf{N} = \mathbf{M} - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T.$$

Now use the power method on this new matrix $\mathbf{N}$ to find the second principal component. This process can be repeated until all (10) components have been found.

37.14 OPTIONS ON EQUITIES ETC.

The pricing of options contingent on equities, currencies, commodities, indices etc. is straightforward in the HJM framework. All that we need to know are the volatility of the asset and its correlations with the forward rate factors. The Monte Carlo simulation then uses the risk-neutral random walk for both the forward rates and the asset i.e. zero-coupon bonds and the asset have a drift of $r(t)$. Of course, there are the usual adjustments to be made for dividends, foreign interest rate or cost of carry, amounting to a change to the drift rate for the asset.

The only fly in the ointment is that American-style exercise is difficult to accommodate.

37.15 NON-INFINITESIMAL SHORT RATE

One of the problems with HJM is that there is no guarantee that interest rates will stay positive, nor that the money market account will stay finite. These problems are associated with the use of a continuously compounded interest rate; all rates can be deduced from the evolution of this rate, but the rate itself is not actually observable. Modeling rates with a finite accruals period, such as three-month LIBOR, for example, has two advantages: The rate is directly observable, and positivity and finiteness can be guaranteed. Let's see how this works. I use the Musiela parameterization of the forward rates.

I have said that $\overline{F}(t, \tau)$ satisfies

$$d\overline{F}(t, \tau) = \cdots + \overline{\nu}(t, \tau)\, dX.$$

A reasonable choice for the volatility structure might be

$$\overline{\nu}(t, \tau) = \gamma(t, \tau)\overline{F}(t, \tau)$$

for finite, non-zero $\gamma(t, \tau)$. At first sight, this is a good choice for the volatility, after all, lognormality is a popular choice for random walks in finance. Unfortunately, this model leads to exploding interest rates. Yet we would like to retain some form of lognormality of rates; recall that market practice is to assume lognormality of just about everything.

We can get around the explosive rates by defining an interest rate $j(t,\tau)$ that is accrued m times *per annum*. The relationship between the new $j(t,\tau)$ and the old $\overline{F}(t,\tau)$ is then

$$\left(1+\frac{j(t,\tau)}{m}\right)^m = e^{\overline{F}(t,\tau)}.$$

Now what happens if we choose a lognormal model for the rates? If we choose

$$dj(t,\tau) = \cdots + \gamma(t,\tau)\,j(t,\tau)\,dX,$$

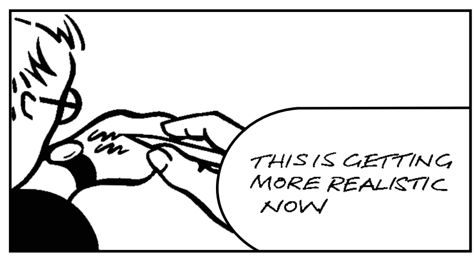
it can be shown that this leads to a stochastic differential equation for $\overline{F}(t,\tau)$ of the form

$$d\overline{F}(t,\tau) = \cdots - m\gamma(t,\tau)\left(1-e^{\overline{F}(t,\tau)/m}\right)\,dX.$$

The volatility structure in this expression is such that all rates stay positive and no explosion occurs.

If we specify the quantity m, we can of course still do PCA to find out the best form for the function $\gamma(t,\tau) = \gamma(\tau)$.

37.16 THE BRACE, GATAREK AND MUSIELA MODEL



The Brace, Gatarek and Musiela[2] model crucially models the actual traded, observable quantities in the fixed-income world. Think of it as a discrete version of the HJM model. This is relatively straightforward, but slightly unnerving if you are used to differential rather than difference equations. The model is popular with practitioners because it can price any contract whose cashflows can be decomposed into functions of the observed forward rates.

We will work in terms of observable, discrete forward rates i.e. rates that really are quoted in the market, rather than the unrealistic continuous forward curve. With $Z(t;T)$ being the price of a zero-coupon bond at time t that matures at time T we have the forward rates at time t, for the period from T_i to the time T_{i+1}, separated by the time step τ, given by

$$1+\tau F(t;T_i,T_{i+1}) = 1+\tau F(t;T_i,T_i+\tau) = 1+\tau F_i(t) = \frac{Z(t;T_i)}{Z(t;T_{i+1})}.$$

This is the definition of forward rates in terms of zero-coupon bond prices, or vice versa. For simplicity of notation in what follows we shall use F_i to mean $F(t;T_i,T_{i+1})$, and Z_i to mean $Z(t;T_i)$. These are, of course, stochastic quantities, varying with time.

The discount factor in terms of the forward rate is just

$$\frac{1}{1+\tau F_i(t)}.$$

This is the discount factor for present valuing from T_{i+1} back to T_i.

[2] The BGM or BGM/J model, the J standing for Jamshidian

Note that here we are using the discrete compounding definition of an interest rate, consistent with that introduced in Chapter 13, rather than our usual continuous definition, which is more often used in the equity world.

Let's suppose that we can write the dynamics of each forward rate, F_i, as

$$dF_i = \mu_i F_i\, dt + \sigma_i F_i\, dX_i.$$

This looks like a lognormal model but, of course, the μs and σs could be hiding more Fs. So really it's a lot more general than that. Similarly suppose that the zero-coupon bond dynamics are given by

$$dZ_i = r Z_i\, dt + Z_i \sum_{j=1}^{i-1} a_{ij}\, dX_j \tag{37.7}$$

where

$$Z_i(t) = Z(t; T_i).$$

There are several points to note about this expression. First of all, we are clearly in a risk-neutral world with the drift of the *traded* asset Z being the risk-free rate. If the bond wasn't traded then we couldn't say that its drift was r.[3] That's why we couldn't say that the drift of F_i was r, since the forward rate is *not* a traded instrument.

Actually, we'll see shortly that we don't need a model for r, it will drop out of the analysis. Also, it looks like a lognormal model but again the as could be hiding more Zs.

Finally, the zero-coupon bond volatility is only given in terms of the volatilities of forward rates of shorter maturities. This is because volatility after its maturity will not affect the value of a bond, and that's why there is no a_{ii} term in Equation (37.7).

We can write

$$Z_i = (1 + \tau F_i) Z_{i+1}.$$

Applying Itô's lemma to this we get

$$dZ_i = (1 + \tau F_i) dZ_{i+1} + \tau Z_{i+1}\, dF_i + \tau \sigma_i F_i Z_{i+1} \left(\sum_{j=1}^{i} a_{i+1,j}\ \rho_{ij} \right) dt, \tag{37.8}$$

where ρ_{ij} is the correlation between dX_i and dX_j.

Think of Equation (37.8) as containing three types of terms: dX_i; dX_j for $j = 1, \ldots, i-1$; dt.

Equating coefficients of dX_i in Equation (37.8) we get

$$0 = (1 + \tau F_i) a_{i+1,i} Z_{i+1} + \tau Z_{i+1} \sigma_i F_i,$$

($a_{ii} = 0$ remember), that is

$$a_{i+1,i} = -\frac{\sigma_i F_i \tau}{1 + \tau F_i}.$$

[3] The drift isn't r, of course, in the real world. But we are in the risk-neutral world for pricing, and in that world all traded instruments have growth r.

Equating the other random terms, dX_j for $j = 1, \dots, i-1$, gives

$$a_{ij} Z_i = (1 + \tau F_i) Z_{i+1} a_{i+1,j},$$

that is

$$a_{i+1,j} = a_{ij} \quad \text{for } j < i.$$

It follows that

$$a_{i+1,j} = -\frac{\sigma_j F_j \tau}{1 + \tau F_j} \quad \text{for } j < i.$$

Finally, equating the dt terms we get

$$r Z_i = (1 + \tau F_i) r Z_{i+1} + \tau Z_{i+1} \mu_i F_i + \tau \sigma_i F_i Z_{i+1} \sum_{j=1}^{i} a_{i+1,j}\ \rho_{ij}.$$

From the definition of F_i the terms including r cancel, leaving

$$\mu_i = -\sigma_i \sum_{j=1}^{i} a_{i+1,j} \rho_{ij} = \sigma_i \sum_{j=1}^{i} \frac{\sigma_j F_j \tau \rho_{ij}}{1 + \tau F_j}.$$

And we are done.

The stochastic differential equation for F_i can be written as

$$dF_i = \left(\sum_{j=1}^{i} \frac{\sigma_j F_j \tau \rho_{ij}}{1 + \tau F_j} \right) \sigma_i F_i \, dt + \sigma_i F_i \, dX_i. \tag{37.9}$$

Equation (37.9) is the discrete BGM version of Equation (37.6) for HJM.[4]

Assuming that we can measure or model the volatilities of the forward rates (σ_i) and the correlations between them (ρ_{ij}) then we have found the correct risk-neutral drift. We are on familiar territory; Monte Carlo simulations using these volatilities, correlations and drifts can be used to price interest rate derivatives. In practice, one estimates the volatility functions from the market prices of caplets. Invariably, this is how one calibrates the time-dependent functions, rather than estimating them from historical data, say.

37.17 SIMULATIONS

We can write (37.9) as

$$d(\log(F_i)) = \left(\sigma_i \sum_{j=1}^{i} \frac{\sigma_j F_j \tau \rho_{ij}}{1 + \tau F_j} - \tfrac{1}{2}\sigma_i^2 \right) dt + \sigma_i \, dX_i.$$

---

[4] Although the notation is slightly different. In the HJM analysis we wrote the random component in terms of uncorrelated dX_i, here we have a different dX_i for each forward rate, but they are all potentially correlated.

In simulating this random walk we would typically divide time up into equal intervals; we would assume, in order to integrate this expression from one interval to the next, that the Fs (and the σs and ρs) were all piecewise constant during each time interval. Simulation then becomes relatively straightforward.

What is not so obvious, however, is how to present value the cashflows. We know from the risk-neutral concepts that there are two aspects to calculating the expectation that is the contracts value, and they are

- simulating the risk-neutral random walk, and
- present valuing the cashflows.

Well, I've explained the first of these, what about the second?

37.18 PVING THE CASHFLOWS

Before present valuing the cashflows (in anticipation of later averaging, and hence pricing) we must be able to write them in terms of the quantities we have simulated, that is the forward rates. That may be simple or hard. For the simpler instruments they are already defined in terms of these quantities. The more complicated contracts might have cashflows that are, in the sense of our exotic option classification, higher order. In the HJM framework we present valued these cashflows using the average spot rate r up until each cashflow. In the BGM model we don't have such an r, of course.

In the BGM model we must present value using the discount factors applicable (for each realization) from one accrual period to the next. That is, we present value each cashflow back to the present through all of the dates T_i using the one-period discount factor at each period:[5]

$$\frac{1}{1+\tau F_i(T_i)}.$$

This is the discrete version of the present valuing we do with the average spot rate in the HJM model. Indeed, if you take the limit as $\tau \to 0$ in all of the above equations you will get back to the HJM model; all the sums become integrals for example.

37.19 SUMMARY

The HJM and BGM approaches to modeling the whole forward rate curve in one go are very powerful. For certain types of contract it is easy to program the Monte Carlo simulations. For example, bond options can be priced in a straightforward manner. On the other hand, the market has its own way of pricing most basic contracts, such as the bond option, as we discussed in Chapter 32. It is the more complex derivatives for which a model is needed. Some of these are suitable for HJM/BGM.

[5] Notice that I didn't use the word 'step' instead of 'period' here. 'Step' would refer to the small time step in the Monte Carlo simulation; period refers to the time between T_{i-1} and T_i.

FURTHER READING

- See the original paper by Heath, Jarrow & Morton (1992) for all the technical details for making their model rigorous.
- For further details of the finite maturity interest rate process model see Sandmann & Sondermann (1994), Brace, Gatarek & Musiela (1997) and Jamshidian (1997).

CHAPTER 38
fixed-income term sheets

In this Chapter...

- the Chooser Range Note
- the Index Amortizing Rate Swap

38.1 INTRODUCTION

We now take a close, detailed, look at a couple of particularly interesting fixed income contracts. You will fully appreciate the Visual Basic code after you have read about numerical methods at the end of the book.

38.2 CHOOSER RANGE NOTE

Figure 38.1 shows the term sheet for a chooser range note. The vanilla range note has cashflows linked to the number of days that the reference rate (typically a LIBOR rate) lies within a specified band. In the Chooser Range Note (CRN), the band is not pre-specified in the contract but is chosen by the contract holder at the start of each period. In the example of the term sheet shown here there are four decisions to be made, one at the start of each period. And that decision is not of the simple binary type ('Do I exercise or not,' 'Do I pay the instalment or not') but is far more complex. At the start of each period the holder must choose a range, represented by, say, its mid point. Thus there is a continuous and infinite amount of possibilities.

38.2.1 Optimal Choice of Ranges?

Deciding on the optimal ranges is not as complicated as it seems, if approached correctly. The contract is priced from the hedger's perspective and the ranges are chosen so as to give the contract the highest possible value. The hedging writer of the contract is exposed to risk-neutral interest rates, and the forward curve; the contract holder will choose ranges depending on his view on the direction of real rates. Since forward rates contain a component of 'market price of risk' and since actual rates rarely show the same dramatic slope in rates and curvature as shown in the forward curve, then it is unlikely that the holder of the contract will choose the range that coincides with that giving the contract its highest value.

GBP 2YR Chooser Range Accrual Note Linked to 6 month GBP LIBOR

The note pays a coupon based on the number of days that 6-month LIBOR sets within an 80bps range. The range is chosen by the buyer at the beginning of each coupon period.

Issue Date	24th March 2000
Maturity Date	24th March 2002
Issue Price	100%
Coupon	[6 month LIBOR + 1.00%] x N/D
N	Number of business days that 6-month LIBOR is within the RANGE
D	Number of business days in the OBSERVATION PERIOD
RANGE	Determined by the buyer two days prior to the beginning of each OBSERVATION PERIOD
OBSERVATION PERIOD	Period 1: 24th March 2000–24th September 2000 Period 2: 24th September 2000–24th March 2001 Period 3: 24th March 2001–24th September 2001 Period 4: 24th September 2001–24th March 2002

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 38.1 Term sheet for a chooser range note.

38.2.2 Pricing

Introduce M as the mid point of the chosen range. To price this contract we just ask how does its value vary with M.

Since the payoff depends primarily on the level of a short-term interest rate we can probably just use a one-factor model. This contract is, unusually, one that has embedded decisions yet can be priced simply in either a partial differential equation/finite-difference manner or by Monte Carlo simulations. The former technique is often preferred when a contract contains decisions, as discussed in Chapter 24. Because of this we can work in either a classical risk-neutral one-factor, Vasicek etc., world, or HJM or BGM. Whichever model we choose let's just write the value of a *non-chooser* range note starting at time t when the spot rate is r and having mid point of the range M as $V^f(r, M, t)$.[1]

We will start by valuing the first leg of the contract, the part expiring, in the example, after just six months.

Choose M and value a non-chooser rate note with the same characteristics. Now vary M. Typically you will end up with results that resemble those shown in Figure 38.2. This function is $V^f(r_0, M, t_0)$ with r_0 being today's spot rate and t_0 being today's date.

Clearly, there is an value of M, M^* in the figure, which gives the contract its highest value, V^*. This is how much we must sell the first leg of the contract for. If we sell it for less than this we run the risk of the holder of the contract choosing this very M^*, so that we would lose money. However, by selling for this amount we can only benefit if the holder chooses a

[1] The r represents either a short rate or the random factor in HJM or BGM.

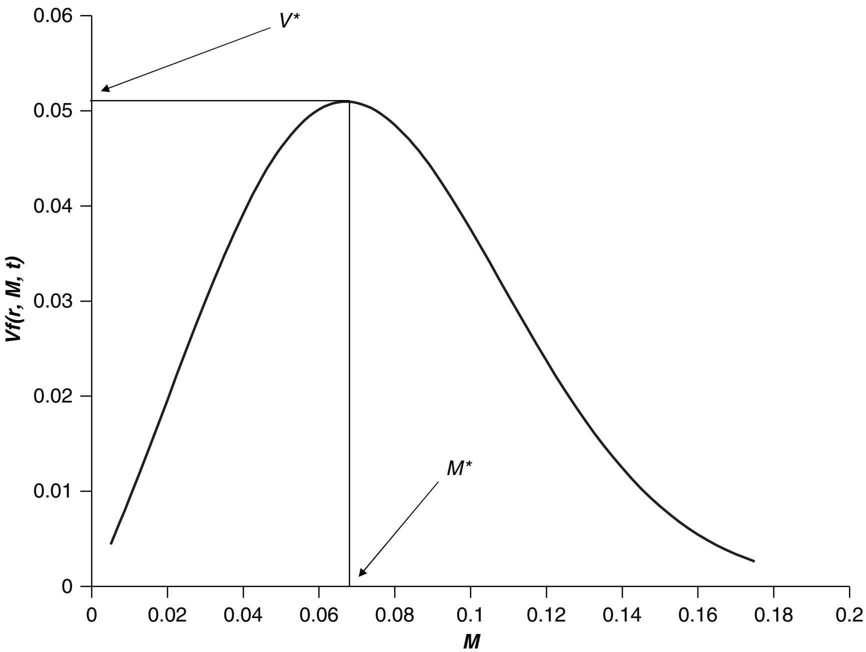


Figure 38.2 How the first leg of the contract varies with *M*.

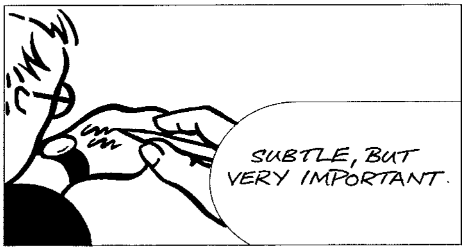
value different from M^*. Remember we are valuing in the risk-neutral world because we will be hedging, whereas the holder is more concerned with the real behavior of rates.

The second, third, and fourth legs are slightly more complicated. We were able to value the first leg because we know r today. To value the second etc. legs we have to imagine ourselves at different levels of r at the start of the range period. For example, to find the value of the second leg now, six months before it starts, we must value a fixed range note with different starting values for r, as explained above (by introducing the mid point M). This function of r then becomes the final condition for a differential equation or simulation over the six months from now to the start of the second leg.

If we denote the start of the later legs by t_1, t_2 and t_3, we must find the functions $V^f(r, M, t_i)$. This is the value of a non-chooser range note. We then take its maximum

$$V(r, t_i) = \max_M \left(V^f(r, M, t_i)\right).$$

We use this as the final condition for valuing back from time t_i to the present, t_0.



38.2.3 Differences Between Optimal for the Writer and the Buyer

This contract requires the holder to make four decisions during its life. Each of these four decisions involves choosing the mid point of an interest rate range, a continuous spectrum of possibilities.

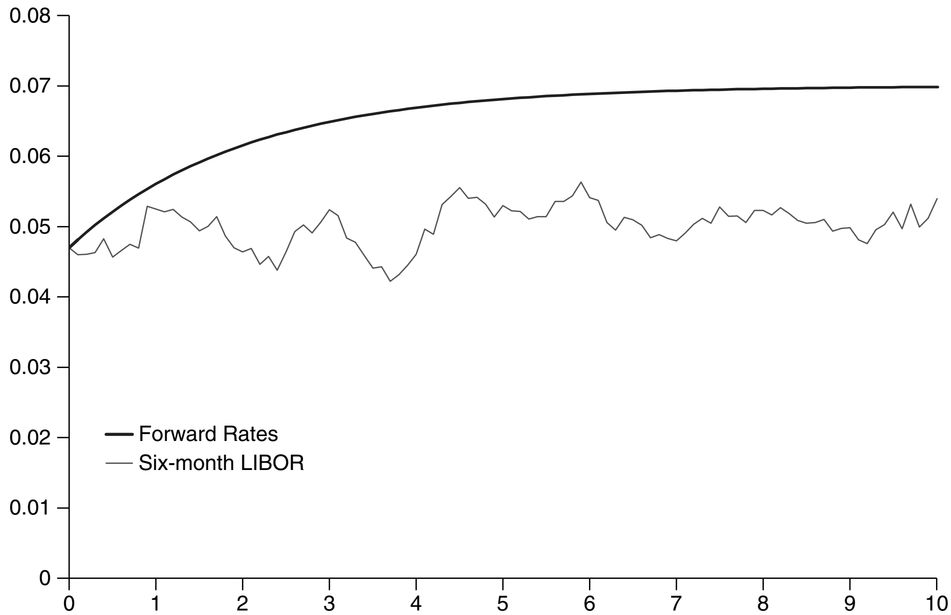


Figure 38.3 Typical forward rate curve at the start of the contract's life, and typical evolution of actual short-term interest rates over its life.

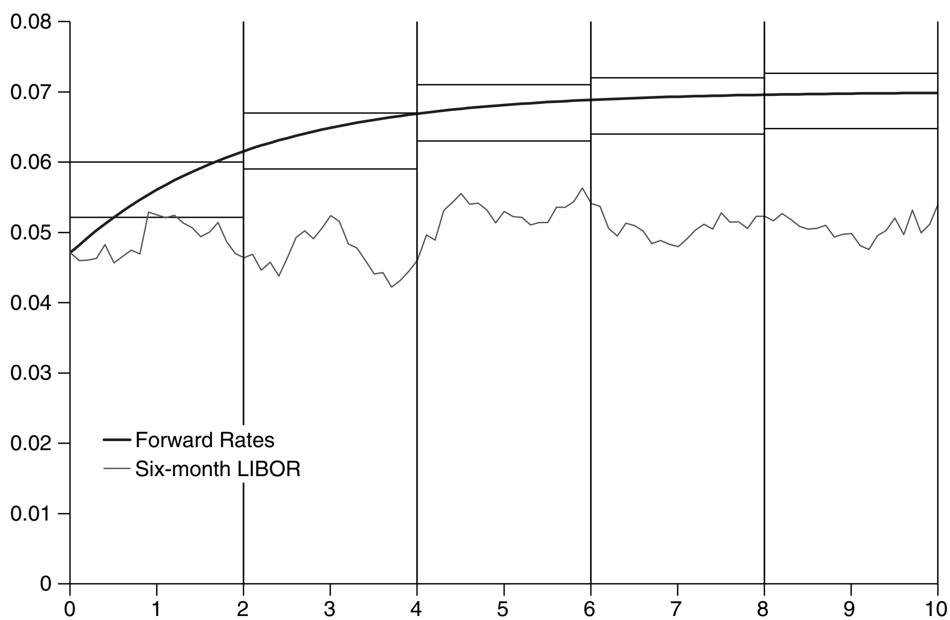


Figure 38.4 The price-maximizing ranges will depend on the risk-neutral, forward rate curve. (Schematic only, the choice will also depend on the volatility of the curve.)

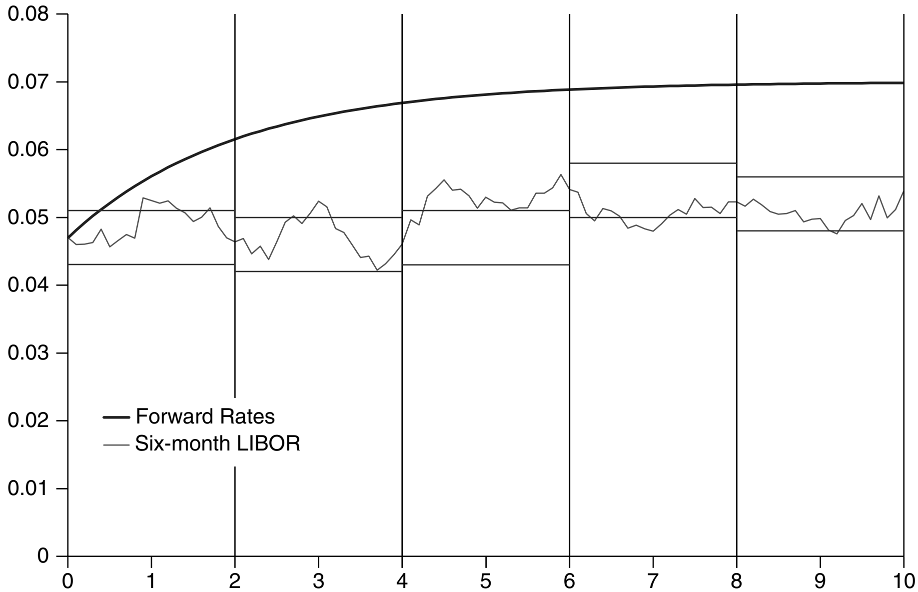


Figure 38.5 The ranges chosen by the holder are more likely to represent the best guess at the evolution of actual rates.

In the first figure, Figure 38.3, we see the forward rate curve as it might be at the start of the contract's life. The shape of this curve is more often than not upward sloping, representing adjustment for the price of risk. One expects a higher return for holding something for a longer term.

The figure also shows a possible evolution of short-term interest rates. It is this path which determines, in part, the final payoff. Notice how the path of rates does not follow the forward curve. Obviously it is stochastic, but it does not usually exhibit the rapid growth at the short end.

The second figure. Figure 38.4, shows a plausible choice of price-maximizing ranges. These will naturally be dependent upon the forward curve. (The figure is schematic only. The actual ranges 'chosen' by the writer when maximizing the price will depend on the volatility of interest rates as well.)

The final figure, Figure 38.5, shows the ranges as chosen by the holder of the contract. He makes a decision about each range at the start of each new period. Of course, his choice will be closely related to where the short-term rate is at that time, with some allowance for his view.

Clearly there is great scope for a significant difference between the price-maximizing choice and the final choices made by the holder. Our concept applies equally well to this case as to the exercise of American options, discussed in Chapter 63. The writer of the option can expect a windfall profit which depends on the difference between the holder's strategy and the price-maximizing strategy.

38.3 INDEX AMORTIZING RATE SWAP

In Figure 38.6 we see a term sheet for an index amortizing rate swap as described in Chapter 32. Here we see the mathematics and coding for such a contract.

USD Index Amortizing Swap

Counterparties	XXXX The Customer
Notional Amount	USD 50 millions, subject Amortization Schedule
Settlement Date	Two days after Trade Date
Maximum Maturity Date	Five years after Trade Date
Early Maturity Date	On any Fixing Date leading to a Notional Amount equal to 0
Payments made by Customer	**USD 6m LIBOR** paid semiannually, A/360
Payments made by XXXX	**In USD X% p.a.** paid semiannually, 30/360
Index Rate	USD 6m LIBOR
Base Rate	[ ]%
Amortization Schedule (after 1st coupon period)	see schedule below

USD 6m LIBOR – Base Rate	Amortization
−3%	−[ ]%
−2%	−[ ]%
−1%	−[ ]%
0	−[ ]%
1%	0%
2%	0%

NB If the observed difference falls between two entries of this schedule, the amortization amount is interpolated

Fixing Dates	2 business days before each coupon period
USD 6m LIBOR	The USD 6m LIBOR rate as seen on Telerate page 3750 at noon, London time, on each Fixing Date
Documentation	ISDA
Governing Law	English

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 38.6 Term sheet for a USD index amortizing swap.

Valuing such an index amortizing rate swap is simple in the framework that we have set up, once we realize that we need to introduce a new state variable. This new state variable is the current level of the principal and we denote it by P. Thus the value of the swap is $V(r, P, t)$.

The variable P is *not* stochastic: It is deterministic and jumps to its new level at each resetting (every quarter in the above example). Since P is piecewise constant, the governing differential equation for the value of the swap is, in a one-factor interest rate world, simply

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV = 0,$$

where

$$dr = u\,dt + w\,dX$$

and λ is the market price of interest rate risk. Of course, in this, V is a function of the three variables, r, P and t.

Two things happen simultaneously at each reset date: There is an exchange of interest payments and the principal amortizes. If we use t_i to denote the reset dates and r_f for the fixed interest rate, then the swap jumps in value by an amount $(r - r_f)P$. Subsequently, the principal P becomes $g(r)P$ where the function $g(r)$ is the representation of the amortizing schedule. Thus we have the jump condition

$$V(r, P, t_i^-) = V(r, g(r)P, t_i^+) + (r - r_f)P.$$

At the maturity of the contract there is one final exchange of interest payments, thus

$$V(r, P, T) = (r - r_f)P.$$

38.3.1 Similarity Solution

Since the amortizing schedule is linear in P, the simple index amortizing rate swap described above has a similarity reduction, taking the problem from three down to two dimensions and so reducing the computational time. The solution takes the form

$$V(r, P, t) = PH(r, t).$$

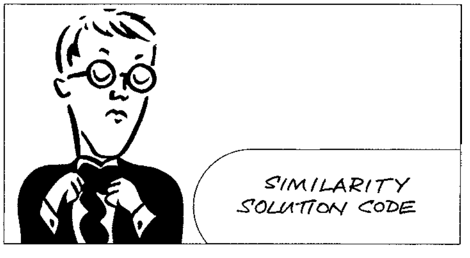
The function $H(r, t)$ satisfies

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 H}{\partial r^2} + (u - \lambda w)\frac{\partial H}{\partial r} - rH = 0,$$
$$H(r, t_i^-) = g(r)H(r, t_i^+) + r - r_f$$

with

$$H(r, T) = r - r_f.$$

38.3.2 The Code



The following code solves for the similarity solution $H(r, t)$. Results are shown in Figure 38.7. The risk-neutral interest rate model is mean-reverting with volatility depending linearly on the rate. Note that the amortizing schedule is a separate function at the end.

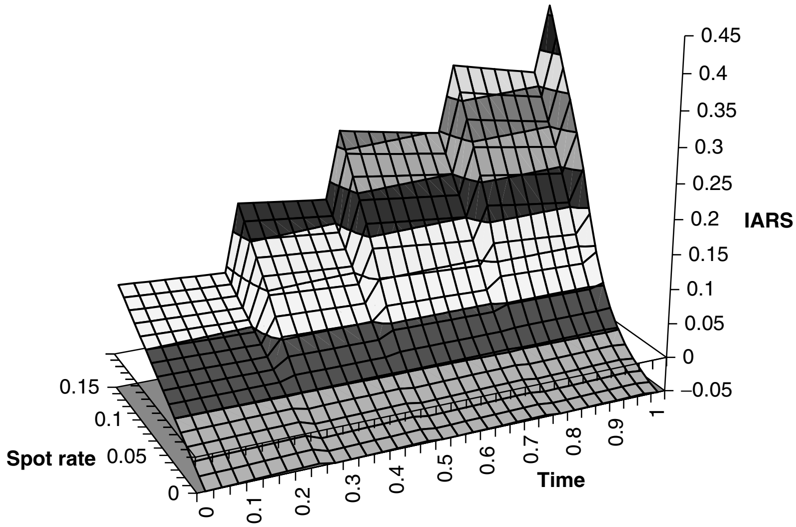


Figure 38.7 The function $H(r, t)$ for an index amortizing rate swap.

```
Function IARS3D(alpha, beta, vol, expry, fixedrate, period, NRS, NTS)

ReDim r(0 To NRS) As Double
ReDim H(0 To NRS, 0 To NTS + 1) As Double
dt = expry / NTS
dr = 3 * fixedrate / NRS

For i = 0 To NRS
r(i) = i * dr
H(i, 0) = r(i) - fixedrate
Next i

For k = 0 To NTS - 1
tim = k * dt
'Engine
    For i = 1 To NRS - 1
    Delta = (H(i + 1, k) - H(i - 1, k)) / 2 / dr
    gamma = (H(i + 1, k) - 2 * H(i, k) + H(i - 1, k)) / dr / dr
    Theta = -0.5 * vol * vol * r(i) * r(i) * gamma - _
                (alpha - beta * r(i)) * Delta + r(i) * H(i, k) _
                'Bond pricing equation
    H(i, k + 1) = H(i, k) - dt * Theta
    Next i
'Boundary conditions
Slopey = (H(1, k) - H(0, k)) / dr
timederiv = -alpha * Slopey
H(0, k + 1) = H(0, k) - dt * timederiv
H(NRS, k + 1) = 2 * H(NRS - 1, k + 1) - H(NRS - 2, k + 1)

'Jump condition
If Int(tim / period) < Int((tim + dt) / period) - 0.001 Then
    For i = 0 To NRS
    H(i, k + 1) = Amortizing_Schedule(r(i)) * H(i, k + 1) + r(i) - _
                fixedrate
```

```
    Next i
End If

Next k

IARS3D = H

End Function


Function Amortizing_Schedule(r)
Amortizing_Schedule = 1
If r < 0.1 Then Amortizing_Schedule = 10 * r
End Function
```

PART FOUR
credit risk

This part of the book addresses the wider issues of risk, its measurement and management. Derivatives theory is typically based in a Black–Scholes world where all options are delta hedged, no risk is taken, and all portfolios earn the risk-free rate of interest. This is a two-dimensional cross-section of a more interesting three-dimensional world. I will now flesh out that third dimension.

Chapter 39: Value of the Firm and Risk of Default Most businesses borrow for expansion. But lending money is risky; will the business be able pay off the debt? As an introduction to the theory of defaultable contracts we look at how to assign a value to a firm. Along the way we see how likely the firm is to go bankrupt.

Chapter 40: Credit Risk Hedging plays a minor, if any, role when there is a real risk of default. How should you price a bond when you suspect that the issuer may default? I describe popular models for risk of default, including endogenous and exogenous default and change of credit rating. The subject of credit risk is immensely important, but from a modeling viewpoint, still in its infancy.

Chapter 41: Credit Derivatives If you hold a bond which may default you can buy insurance that pays you in that event. This is the simplest of credit derivatives. More complex structures exist that pay off upon a change of credit rating for example. These and other contracts are discussed in this chapter.

Chapter 42: RiskMetrics and CreditMetrics In order to make the measurement of risk easier, and consistent from one bank to another, there have grown up several methodologies. I describe a couple of these here. One is for basic value at risk measurement in normal market situations and one is for credit risk.

Chapter 43: CrashMetrics A methodology for measuring risk in the event of market crashes is described in this chapter. The ideas and math are the ultimate in simplicity.

Chapter 44: Derivatives ** Ups** Derivatives, good or evil? In the wrong hands. . .

CHAPTER 39
value of the firm and the risk of default

In this Chapter...

- models for default based on the 'value of the firm'
- modeling the value of a firm based on measurable parameters and variables

39.1 INTRODUCTION

So far, the products on which we have concentrated have all had cashflows that are *guaranteed*. We have assumed that these cashflows, coupons, payoffs and redemption values, are from a completely creditworthy source or underwritten in such a way that the income is certain. Options bought through an exchange are usually considered free of the risk of default because of the way that exchanges are underwritten, and partly because of the requirement for a margin to be deposited.

In practice, many bonds have no such guarantee. Perhaps they are issued by a company as a form of borrowing for expansion. In this case, the issuing company may declare bankruptcy before all of the cashflows have been paid. Alternatively, they may be issued by a government with a record for irregular payment of debt. Over the counter (OTC) options can have significant counterparty risks. For this reason there has grown up over the past decade a considerable body of rules and regulations governing **capital adequacy**, to ensure that banks are covered in the event of extreme market movements that might otherwise lead to collapse (see Chapter 19 on Value at Risk).

In this chapter we discuss the subject of modeling when there is **risk of default** or **credit risk**. There are two main approaches to the modeling of default. One revolves around modeling the value of the issuing firm (or country). The other models an exogenous risk of default. The former is appealing because it is clearly closer to reality. The downside is that these models are usually more complicated to solve, with parameters that are difficult to measure. I describe the most popular such model and then describe a similar model, one with more easily-measured parameters.

39.2 THE MERTON MODEL: EQUITY AS AN OPTION ON A COMPANY'S ASSETS

The Merton model shows very elegantly how the equity of a company can be thought of as a simple call option on the assets of the company. He starts off by assuming that the assets of the company A follow a random walk

$$dA = \mu A\,dt + \sigma A\,dX.$$

Clearly, the value of the equity equals assets less liabilities:

$$S = A - V.$$

Here S is the value of the equity (just the share price, I've assumed that there is only one share for simplicity) and V the value of the debt.

At maturity of this debt

$$S(A, T) = \max(A - D, 0) \quad \text{and} \quad V(A, T) = \min(D, A),$$

where D is the amount of the debt, to be paid back at time T.

We can now apply ideas from the Black–Scholes model for options, namely the ideas of delta hedging and no arbitrage. Set up a portfolio consisting of the debt and short a quantity Δ of the equity.

$$\Pi = V - \Delta S.$$

We are going to hedge the debt with the stock. (Don't worry about buying and selling a fraction of the one share, in reality there would be millions of shares issued.)

This portfolio changes by

$$d\Pi = dV - \Delta\,dS.$$

So applying Itô's lemma

$$\begin{aligned} d\Pi = {} & \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 A^2 \frac{\partial^2 V}{\partial A^2} - \Delta \frac{\partial S}{\partial t} - \tfrac{1}{2}\Delta\sigma^2 A^2 \frac{\partial^2 S}{\partial A^2}\right) dt \\ & + \left(\frac{\partial V}{\partial A} - \Delta \frac{\partial S}{\partial A}\right) dA. \end{aligned}$$

Now choose

$$\Delta = \frac{\partial V/\partial A}{\partial S/\partial A}$$

to eliminate risk, and set the return $d\Pi$ equal to the risk-free return $r\Pi\,dt\ldots$

The end result is, for the current value of the debt V

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 A^2 \frac{\partial^2 V}{\partial A^2} + rA \frac{\partial V}{\partial A} - rA = 0$$

subject to

$$V(A, T) = \min(D, A)$$

and exactly the same partial differential equation for the equity of the firm S but with

$$S(A,T) = \max(A - D, 0).$$

Recognize this? The problem for S is exactly that for a call option, but now we have S instead of the option value, the underlying variable are the assets A and the strike is D, the debt. The solution for the equity value is precisely that for a call option.

39.2.1 Default Before Maturity

Creditors may be able to force liquidation of the company's assets before maturity should the value of its assets fall below a critical level. The critical level may be time-dependent:

$$A = K(t).$$

This would introduce a boundary condition on $A = K(t)$ making this problem very similar to that for a barrier option. On the (possibly moving) barrier the value of the debt is simply whatever the assets are at that time:

$$V(K(t), t) = \min(K(t), D).$$

39.2.2 Probability of Default

The probability of default before maturity $P(A,t)$ is equivalent to the probability of the asset value reaching the critical level $K(t)$ before maturity:

$$\frac{\partial P}{\partial t} + \tfrac{1}{2}\sigma^2 A^2 \frac{\partial^2 P}{\partial A^2} + \mu A \frac{\partial P}{\partial A} = 0$$

with

$$P(K(t), t) = 1 \quad \text{and} \quad P(A,T) = 0.$$

39.2.3 Stochastic Interest Rates

We can make the model more realistic by introducing an interest rate model into the problem—after all, if there is no credit risk we would like to return to the simpler world of pricing non-risky debt. I will be vague about the choice of interest rate model and just write

$$dr = u(r,t)\,dt + w(r,t)\,dX_1.$$

Still we will assume that

$$dA = \mu A\,dt + \sigma A\,dX_2.$$

There will be a correlation of ρ between the two random walks.

Now the value of the debt V, say, is a function of three variables; we have $V(A, r, t)$. We have seen in earlier chapters how to derive the equation for a bond value in the absence of default risk. The result is a diffusion equation in r and t. The equation for V will be similar but now there will be some A dependence and derivatives with respect to A.

To find the equation satisfied by V we construct a portfolio of our risky bond, and hedge it with the stock and a riskless zero-coupon bond with price $Z(r, t)$:

$$\Pi = V(A, r, t) - \Delta S - \Delta' Z(r, t).$$

From this, we calculate $d\Pi$, choose Δ and Δ' to eliminate risk and set the return equal to the risk-free rate. You know the drill, so cutting to the chase

$$\frac{\partial V}{\partial t} + \frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + \rho\sigma w A\frac{\partial^2 V}{\partial r\,\partial A} + \frac{1}{2}\sigma^2 A^2\frac{\partial^2 V}{\partial A^2} + (u - \lambda w)\frac{\partial V}{\partial r} + rA\frac{\partial V}{\partial A} - rV = 0$$

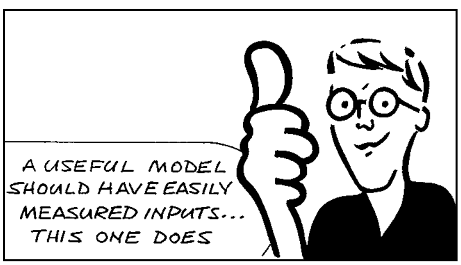
where λ is again the price of interest rate risk.

The final condition for this equation is

$$V(A, T) = \min(D, A)$$

representing the payment of the debt at maturity.

The main criticism of this model is that it is very difficult to measure variables and parameters. Nevertheless, it can be useful as a phenomenological model, perhaps for estimating the *relative* value of different types of debt issued by the same business, or businesses with the same credit rating.



39.3 MODELING WITH MEASURABLE PARAMETERS AND VARIABLES

In this section I describe a model that takes easily measurable inputs. We will concentrate on valuing the debt when interest rates are deterministic; the model could easily be extended to stochastic interest rates at the cost of additional complexity and computing time.

We will value the debt of a company having a very simple operating procedure: They sell their product, they pay their costs and they put any profit into the bank. The key quantity we will model is the earnings of the company. Think of these earnings as being the gross income from the sale of the product. The net earnings or profit will be the gross earnings less the costs. Assume that the gross annualized earnings E of the company are random:

$$dE = \mu E\,dt + \sigma E\,dX.$$

(Of course, we need not choose a lognormal model, but it is the traditional starting point.)

We assume that the company has fixed costs of E^* per annum and floating costs of kE. The profit of $E - E^* - kE = (1 - k)E - E^*$ is put into a bank earning a fixed rate of interest r. If we denote by C the cash in this bank account then this is given by

$$C = \int_0^t \left((1 - k)E(\tau) - E^*\right) e^{r(t-\tau)} d\tau.$$

This expression represents the accumulation of income together with any bank interest. Differentiating this gives the stochastic differential equation satisfied by C:

$$dC = \left((1 - k)E - E^* + rC\right) dt.$$

I have chosen to model the earnings of the business rather than the firm's value since the former are far easier to measure, requiring only an examination of the firm's accounts perhaps. We shall see how the value of the firm is then an *output* of the model.

The well-being of the company is determined by its earnings and bank account balance at any time, i.e. by E and C. The owners of the company hope that $(1-k)E > E^*$, but even if it is not (at the start of trading, say), then perhaps the growth in earnings will eventually take the company into profit.

Suppose that the company owes D which must be repaid at time T. We make the simple assumption that if the company has D in the bank at time T then it will repay the debt, if it has less than D in the bank it will repay everything that it has, and if there is a negative amount in the bank then they repay nothing. This gives a repayment of

$$\max(\min(C, D), 0). \tag{39.1}$$

Again, this would be more complicated if we were to incorporate partial repayment or refinancing.

The value of the debt will be a function of E, C and t. Introduce the quantity $V(E, C, t)$ as the present value of the *expected* amount in (39.1). This function satisfies the differential equation

$$\boxed{\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 E^2 \frac{\partial^2 V}{\partial E^2} + \mu E \frac{\partial V}{\partial E} + \big((1-k)E - E^* + rC\big)\frac{\partial V}{\partial C} - rV = 0,}$$

with final condition

$$V(E, C, T) = \max(\min(C, D), 0).$$

In Figure 39.1 we see a plot of the value of the debt according to this model when there is an amount \$100,000 to be repaid in two years, with a risk-free interest rate of 5%. The firm has fixed costs of \$30,000 p.a. and variable costs of 7%. The drift of the earnings is 10% and the volatility 25%. Observe that for both large C and E the value approaches the value of a risk-free zero-coupon bond. In Figure 39.2 is the equivalent yield, defined as

$$-\tfrac{1}{2}\log\left(\frac{V}{100{,}000}\right).$$

The $\frac{1}{2}$ being in front since it is a two-year loan. Subtract the risk-free 5% from this and you get the credit spread.

As it stands this problem can be usefully used to value risky debt: The value of the debt is simply $V(E, C, t)$ with today's values for E, C and t. However, it can be modified quite simply to accommodate more sophisticated operating procedures for the company. One possibility is to say that the company closes down immediately that it goes into the red. This could be modeled by the boundary condition

$$V(E, 0, t) = 0.$$

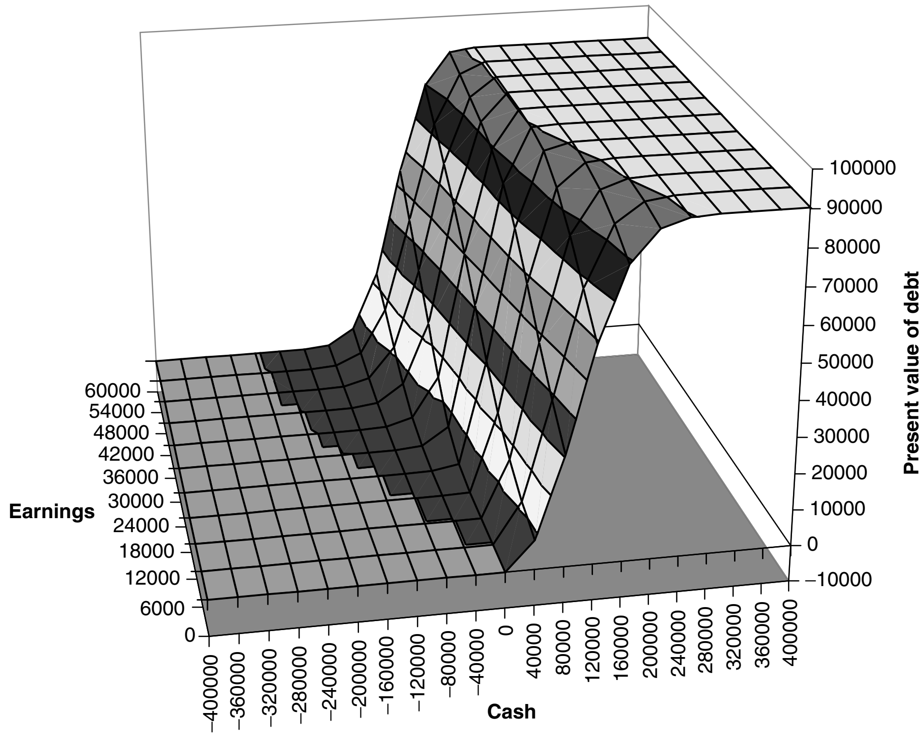


Figure 39.1 Value of debt issued by a limited liability company as a function of annual earnings and retained cash.

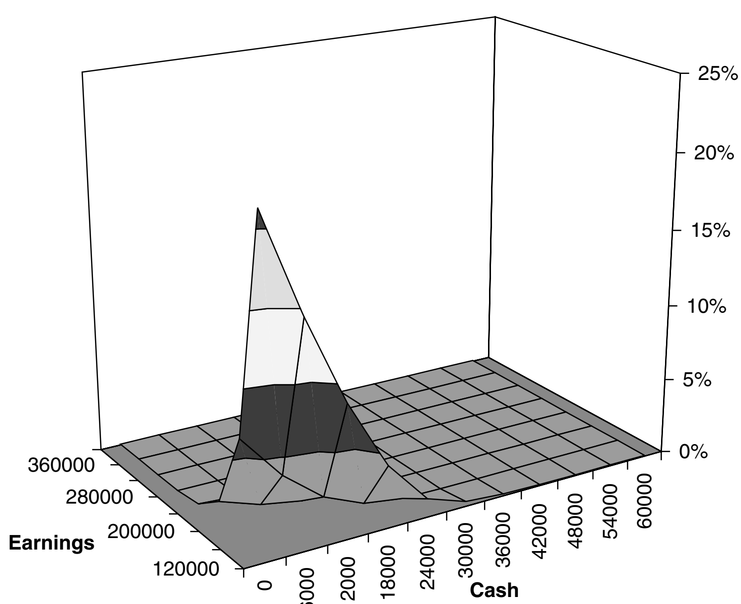


Figure 39.2 Effective yield.

39.4 CALCULATING THE VALUE OF THE FIRM

By changing final and boundary conditions it is a very simple matter to use the above model to value the firm, and to examine the effects of various business strategies on that value. For example, suppose that we take the value of the business to be the present value of the expected cash in the bank at some time T_0 in the future. Such a finite time horizon is a common assumption when we are estimating the present value of a potentially infinite sum of cashflows which are (one hopes) growing faster than the interest rate. In this case, the firm value $V(E, C, t)$ satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 E^2 \frac{\partial^2 V}{\partial E^2} + \mu E \frac{\partial V}{\partial E} + \big((1-k)E - E^* + rC\big) \frac{\partial V}{\partial C} - rV = 0.$$

As an example of the flexibility of this approach, consider the different final conditions applying to the two different problems valuing a limited liability company and valuing an unlimited liability partnership.

Limited liability

If the business has no liability when it has a negative amount in the bank at time T_0, then

$$V(E, C, T_0) = \max(C, 0).$$

Partnership

If the owners of the business are liable for the debts of the business then

$$V(E, C, T_0) = C.$$

In the former case, if the business expires in the red then the company directors declare bankruptcy and walk away from the debt (assuming that they have not acted negligently).

Not only can we use this model to examine the legal standing of the company, but also to study the effects of various operating procedures on its value. Here is an example.

Optimal close-down

If the model gives a value of \$3,000,000 to your company but you currently have \$5,000,000 in the bank then the model is trying to tell you something: Bad times are just around the corner. In such a situation it is not wise to keep trading; you will be better off closing down the business. The decision to close down the business can be optimized by a constraint of the form

$$V(E, C, t) \geq C$$

with continuity of the first derivatives. This is just like an American option problem and the justification is similar. An example of the company valuation problem with this constraint is shown in Figure 39.3.

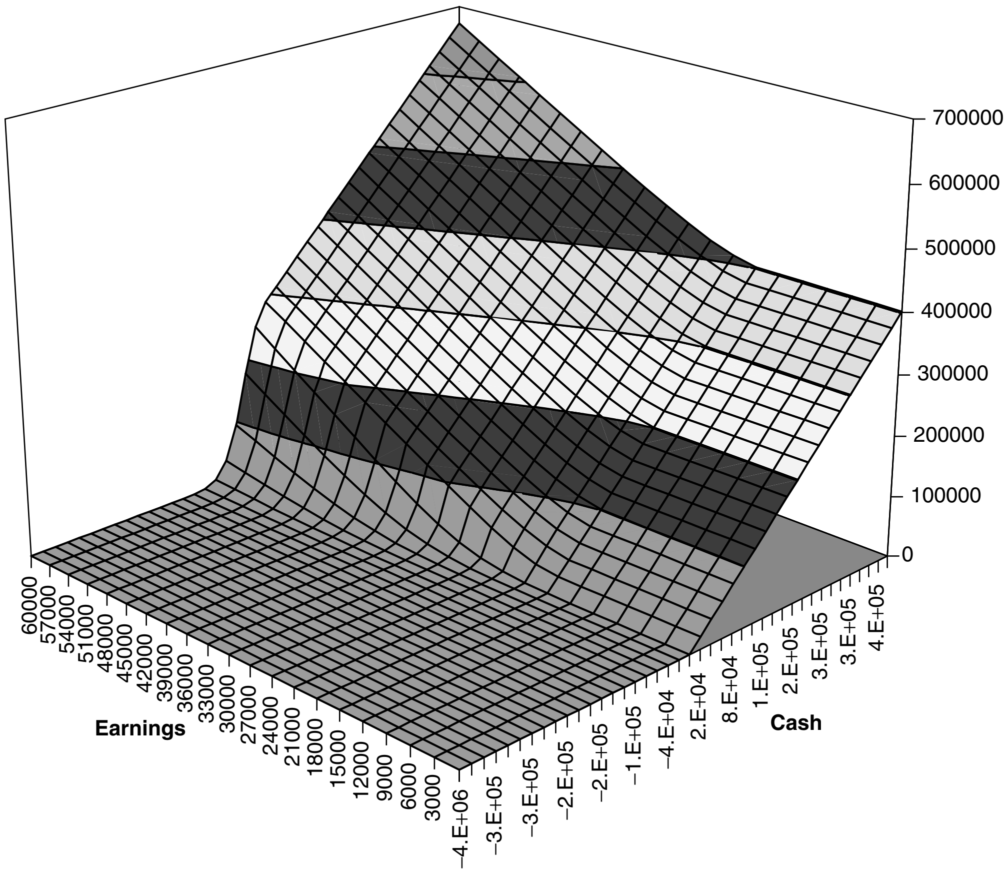


Figure 39.3 Company valuation with optimal close-down. Parameter values are the same as in the previous figure.

39.5 SUMMARY

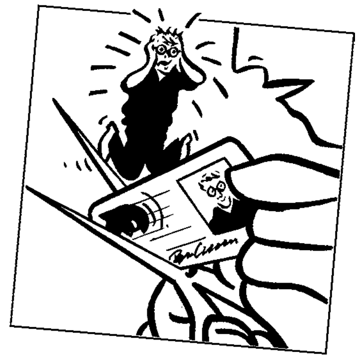
I've introduced firm valuation via a credit risk problem. There are many other reasons why one might want to know the value of a firm and these are discussed in depth in Chapter 73. But next, in Chapter 40 we examine another approach to modeling credit risk, one that supposedly requires no detailed knowledge of the firm issuing the debt.

FURTHER READING

- See Black & Scholes (1973), Merton (1974), Black & Cox (1976), Geske (1977) and Chance (1990) for a treatment of the debt of a firm as an option on the assets of that firm. See Longstaff & Schwartz (1994) for more recent work in this area.
- See Apabhai *et al*. (1998) for more details of the company and debt valuation model, especially for final and boundary conditions for various business strategies. Epstein *et al*.

(1997a, b) describe the firm valuation models in detail, including the effects of advertising and market research.

- The classic reference, and a very good read, for firm-valuation modeling is Dixit & Pindyck (1994). For a non-technical POV see Copeland, Koller & Murrin (1990).
- See Kim (1995) for the application of the company valuation model to the question of company mergers and some suggestions for how it can be applied to problems in company relocation and tax status.



CHAPTER 40
credit risk

In this Chapter...

- models for instantaneous and exogenous risk of default
- stochastic risk of default and implied risk of default
- credit ratings
- how to model change of rating
- how to model risk of default in convertible bonds

40.1 INTRODUCTION

In the previous chapter I described some ways of looking at default via models for the creditworthiness of the firm issuing the debt. This is a nice approach *if you have access to all the data*. A more recent approach is to model default as a completely exogenous event i.e. a bit like the tossing of a coin or the appearance of zero on a roulette wheel, and having nothing to do with how well the company or country is doing. Typically, one then infers from risky bond prices the probability of default as perceived by the market. I'm not wild about this idea but it is very popular.

Later in this chapter I describe the rating service provided by Standard & Poor's and Moody's, for example. These ratings provide a published estimate of the relative creditworthiness of firms.

40.2 RISKY BONDS

If you are a company wanting to expand, but without the necessary capital, you could borrow the capital, intending to pay it back with interest at some time in the future. There is a chance, however, that before you pay off the debt the company may have got into financial difficulties or even gone bankrupt. If this happens, the debt may never be paid off. Because of this risk of default, the lender will require a significantly higher interest rate than if he were lending to a very reliable borrower such as the US government.

The real situation is, of course, more complicated than this. In practice it is not just a case of all or nothing. This brings us to the ideas of the seniority of debt and the partial payment of debt.

Firms typically issue bonds with a legal ranking, determining which of the bonds take priority in the event of bankruptcy or inability to repay. The higher the priority, the more likely the debt

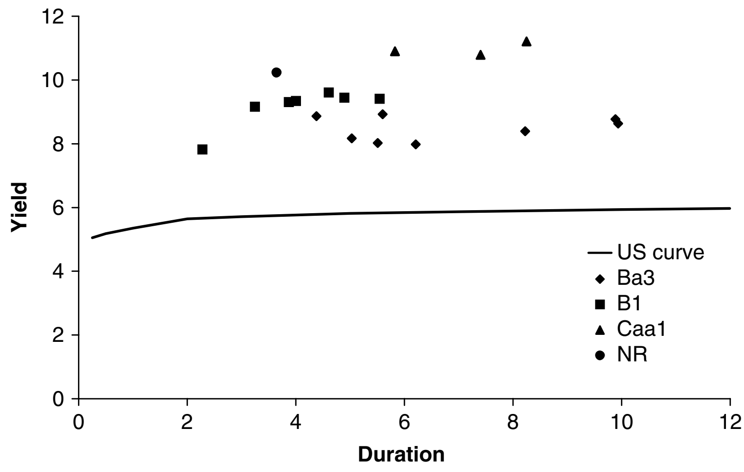


Figure 40.1 Yield versus duration for some risky bonds.

is to be repaid, the higher the bond value and the lower its yield. In the event of bankruptcy there is typically a long, drawn out battle over which creditors get paid. It is usual, even after many years, for creditors to get *some* of their money back. Then the question is how much and how soon? It is also possible for the repayment to be another risky bond; this would correspond to a refinancing of the debt. For example, the debt could not be paid off at the planned time so instead a new bond gets issued entitling the holder to an amount further in the future.

In Figure 40.1 is shown the yield versus duration, calculated by the methods of Chapter 13, for some risky bonds. In this figure the bonds have been ranked according to their estimated riskiness. We will discuss this later, for the moment you just need to know that Ba3 is considered to be less risky than Caa1 and this is reflected in its smaller yield spread over the risk-free curve.

The problem that we examine in this chapter is the modeling of the risk of default and thus the fair value of risky bonds. Conversely, if we know the value of a bond, does this tell us anything about the likelihood of default?

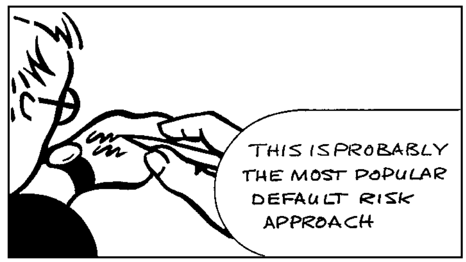
40.3 MODELING THE RISK OF DEFAULT

The models that I have described or will describe here fall into two categories, those for which the likelihood of default depends on the behavior of the issuing firm and those for which the likelihood of default is exogenous. The former is appealing because it is clearly closer to reality. The downside is that these models are usually more complicated to solve, with parameters that are difficult to measure.

The instantaneous risk of default models are simpler to use and are therefore the most popular type of credit risk models. In its simplest form the time at which default occurs is completely exogenous. For example, we could roll a die once a month, and when a 1 is thrown the company defaults. This illustrates the exogeneity of the default and also its randomness; a Poisson process is a typical choice for the time of default. We will see that when the intensity of the Poisson process is constant (as in the die example), the pricing of risky bonds amounts to the addition of a time-independent spread to the bond yield. We will also see models for which the intensity is itself a random variable.

A refinement of the modeling that we also consider is the regrading of bonds. There are agents, such as Standard & Poor's and Moody's, who classify bonds according to their estimate of their risk of default. A bond may start life with a high rating, for example, but may find itself regraded due to the performance of the issuing firm. Such a regrading will have a major effect on the perceived risk of default and on the price of the bond. I will describe a simple model for the rerating of risky bonds.

40.4 THE POISSON PROCESS AND THE INSTANTANEOUS RISK OF DEFAULT



A popular approach to the modeling of credit risk is via the **instantaneous risk of default**, p. If at time t the company has not defaulted and the instantaneous risk of default is p then the probability of default between times t and $t + dt$ is $p\,dt$. This is an example of a Poisson process, as described in detail in Chapter 57; nothing happens for a while, then there is a *sudden* change of state. This is a continuous-time version of our earlier model of throwing a die.

The simplest example to start with is to take p constant. In this case we can easily determine the risk of default before time T. We do this as follows.

Let $P(t; T)$ be the probability that the company does not default before time T given that it has not defaulted at time t. The probability of default between later times t' and $t' + dt'$ is the product of $p\,dt$ and the probability that the company has not defaulted up until time t'. Thus,

$$P(t' + dt', T) - P(t', T) = p\,dt \;\; P(t', T).$$

Expanding this for a small time step results in the ordinary differential equation representing the rate of change of the required probability:

$$\frac{\partial P}{\partial t'} = pP(t'; T).$$

If the company starts out not in default then $P(T; T) = 1$. The solution of this problem is

$$e^{-p(T-t)}.$$

The value of a zero-coupon bond paying \$1 at time T could therefore be modeled by taking the present value of the *expected* cashflow. This results in a value of

$$e^{-p(T-t)}Z, \tag{40.1}$$

where Z is the value of a riskless zero-coupon bond of the same maturity as the risky bond. Note that this does not put any value on the risk taken. The yield to maturity on this bond is now given by

$$-\frac{\log(e^{-p(T-t)}Z)}{T-t} = -\frac{\log Z}{T-t} + p.$$

Thus the effect of the risk of default on the yield is to add a spread of p. In this simple model, the spread will be constant across all maturities.

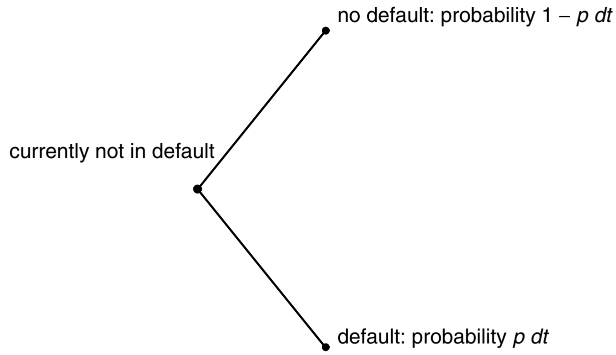


Figure 40.2 A schematic diagram showing the two possible situations: default and no default.

Now we apply this to derivatives, including risky bonds. We will assume that the spot interest rate is stochastic. For simplicity we will assume that there is no correlation between the diffusive change in the spot interest rate and the Poisson process.

Construct a 'hedged' portfolio:

$$\Pi = V(r, p, t) - \Delta Z(r, t).$$

Consider how this changes in a time step. See Figure 40.2 for a diagram illustrating the analysis below.

There is a probability of $(1 - p\,dt)$ that the bond does not default. In this case the change in the value of the portfolio during a time step is

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}\right)dt + \frac{\partial V}{\partial r}dr + -\Delta\left(\left(\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2}\right)dt + \frac{\partial Z}{\partial r}dr\right). \tag{40.2}$$

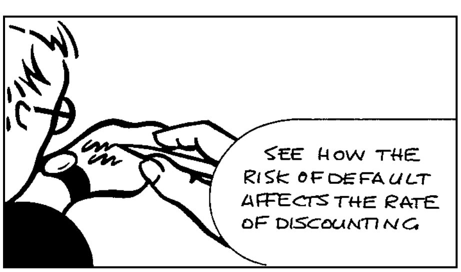
Choose Δ to eliminate the risky dr term.

On the other hand, if the bond defaults, with a probability of $p\,dt$, then the change in the value of the portfolio is

$$d\Pi = -V + O(dt^{1/2}). \tag{40.3}$$

This is due to the sudden loss of the risky bond; the other terms are small in comparison.

Taking expectations and using the bond-pricing equation for the riskless bond, we find that the value of the risky bond satisfies



$$\boxed{\begin{aligned}&\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r}\\ &\quad - (r + p)V = 0.\end{aligned}} \tag{40.4}$$

Observe that the spread has been added to the discounting term.

The portfolio is only hedged against moves in the interest rate (assuming that the market price of risk is known, which is a big assumption) and not the event of default. I'll come back to this point below.

This model is the most basic for the instantaneous risk of default. It gives a very simple relationship between a risk-free and a risky bond. There is only one new parameter to estimate, p.

To see whether this is a realistic model for the expectations of the market we take a quick look at the valuation of Brady bonds. In particular we examine the market price of Latin American Par bonds, described in full later in this chapter. For the moment, we just need to know that these bonds have interest payments and final return of principal denominated in US dollars. If the above is a good model of market expectations with constant p then we would find a very simple relationship between interest rates in the US and the value of Brady bonds. To find the Brady bond value perform the following:

1. Find the risk-free yield for the maturity of each cashflow in the risky bond.
2. Add a constant spread, p, to each of these yields.
3. Use this new yield to calculate the present value of each cashflow.
4. Sum all the present values.

Conversely, the same procedure can be used to determine the value of p from the market price of the Brady bond: this would be the **implied risk of default**. In Figure 40.3 are shown the

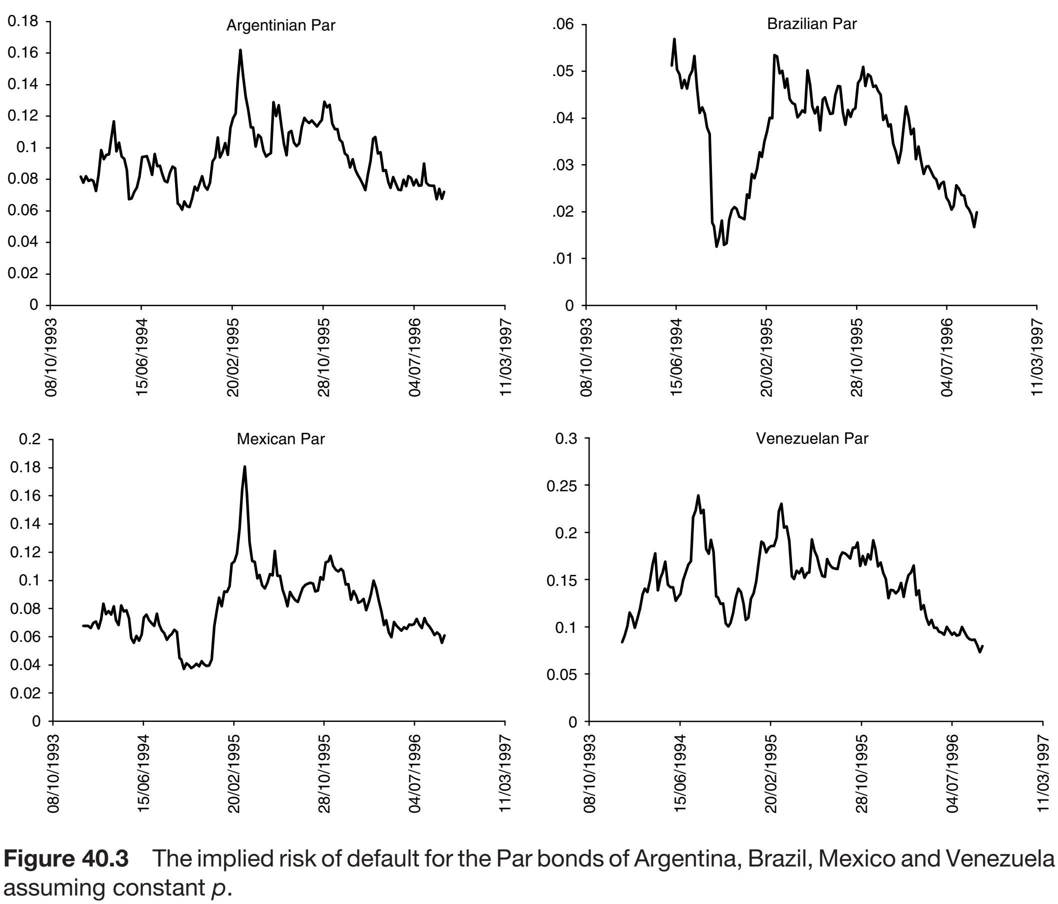


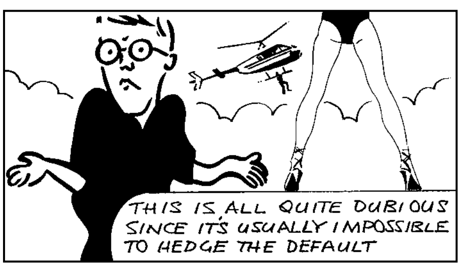
Figure 40.3 The implied risk of default for the Par bonds of Argentina, Brazil, Mexico and Venezuela assuming constant p.

implied risks of default for the Par bonds of Argentina, Brazil, Mexico and Venezuela using the above procedure and assuming a constant p.

In this simple model we have assumed that the instantaneous risk of default is constant (different for each country) through time. However, from Figure 40.3 we can see that, if we believe the market prices of the Brady bonds, this assumption is incorrect: The market prices are inconsistent with a constant p. This will be our motivation for the stochastic risk of default model which we will see in a moment. Nevertheless, supposing that the figure represents, in some sense, the views of the market (and this constant p model *is* used in practice) we draw a few conclusions from this figure before moving on.

The first point to notice in the graph is the perceived risk of Venezuela, which is consistently greater than the three other countries. Venezuela's risk peaked in July 1994, nine months before the rest of South America, but this had absolutely no effect on the other countries.

The next, and most important, thing to notice is the 'Tequila effect' in all the Latin markets. The Tequila crisis began with a 50% devaluation of the Mexican peso in December 1994. Markets followed suit by plunging. Before December 1994 we can see a constant spread between Mexico and Argentina and a contracting spread between Brazil and Argentina. The consequences of Tequila were felt through all the first quarter of 1995 and had a knock-on effect throughout South America. In April 1995 the default risks peaked in all the countries apart from Venezuela, but by late 1996 the default risk had almost returned to pre-Tequila levels in all four countries. By this time, Venezuela's risk had fallen to the same order as the other countries.



40.4.1 A Note on Hedging

In the above we have not hedged the event of default. This can sometimes be done (kinda), provided we can hedge with other bonds that will default at exactly the same time, and later we'll see how this introduces a market price of default risk term, as might be expected. Usually, though, there are so few risky bonds that hedging default is not possible. In that case it may be better to examine the expected return and some measure of the residual risk, as we did in Chapter 59. For the moment I'm going to work in terms of real expectations on the understanding that in the world of default risk, perfect hedging is rarely possible.

40.5 TIME-DEPENDENT INTENSITY AND THE TERM STRUCTURE OF DEFAULT

Suppose that a company issues risky bonds of different maturities. We can deduce from the market's prices of these bonds how the risk of default is perceived to depend on time. To make things as simple as possible let's assume that the company issues zero-coupon bonds and that in the event of default in one bond, all other outstanding bonds also default with no recovery rate.

If the risk of default is time-dependent, $p(t)$, and uncorrelated with the spot interest rate, then the real expected value of a risky bond paying \$1 at time T is just

$$Ze^{-\int_t^T p(\tau)\,d\tau}.$$

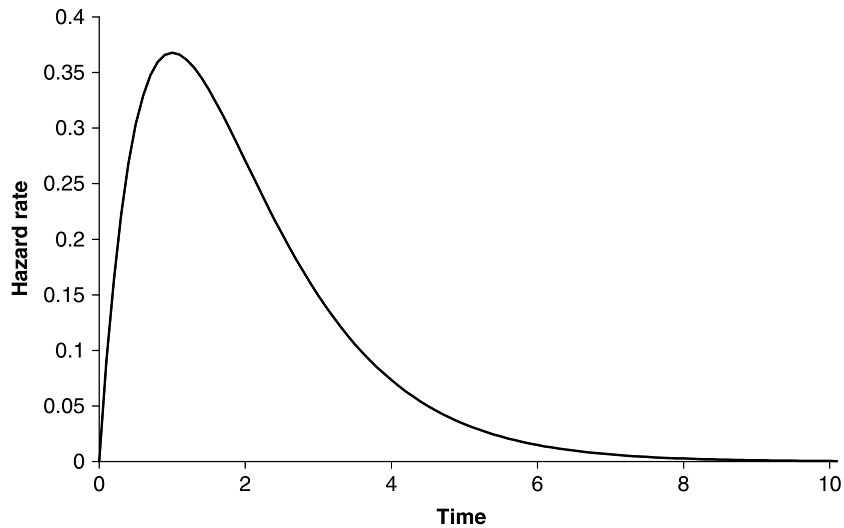


Figure 40.4 A plausible structure for a time-dependent hazard rate.

If the market value of the risky bond is Z^* then we can write

$$\int_t^T p(\tau)\,d\tau = \log\left(\frac{Z}{Z^*}\right).$$

Differentiating this with respect to T gives the market's view at the current time t of the **hazard rate** or risk of default at time T. A plausible structure for such a hazard rate is given in Figure 40.4. This figure shows a very small chance of default initially, rising to a maximum before falling off. The company is clearly expected to be around for at least a little while longer, and in the long term it will either have already expired or become very successful. If the area under the curve is finite then there is a finite probability of the company never going bankrupt.

40.6 STOCHASTIC RISK OF DEFAULT

To 'improve' the model, and make it consistent with market prices, we now consider a model in which the instantaneous probability of default is itself random. We assume that it follows a random walk given by

$$dp = \gamma(r, p, t)\,dt + \delta(r, p, t)\,dX_1,$$

with interest rates still given by

$$dr = u(r, t)\,dt + w(r, t)\,dX_2.$$

It is reasonable to expect some interest rate dependence in the risk of default, but not the other way around.

To value our risky zero-coupon bond we construct a portfolio with one of the risky bonds with value $V(r, p, t)$ (to be determined), and short Δ of a riskless bond, with value $Z(r, t)$ (satisfying our earlier bond pricing equation):

$$\Pi = V(r, p, t) - \Delta Z(r, t).$$

In the next time step the bond either defaults or it doesn't. There is a probability of default of $p\,dt$. We must consider the two cases: Default and no default in the next time step. As in the two models above, we take expectations to arrive at an equation for the value of the risky bond.

First, suppose that the bond does not default; this has a probability of $(1 - p\,dt)$. In this case the change in the value of the portfolio during a time step is

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + \rho w\delta\frac{\partial^2 V}{\partial r\,\partial p} + \tfrac{1}{2}\delta^2\frac{\partial^2 V}{\partial p^2}\right)dt + \frac{\partial V}{\partial r}dr + \frac{\partial V}{\partial p}dp$$
$$- \Delta\left(\left(\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2}\right)dt + \frac{\partial Z}{\partial r}dr\right),$$

where ρ is the correlation between dX_1 and dX_2. Choose Δ to eliminate the risky dr term.

On the other hand, if the bond defaults, with a probability of $p\,dt$, then the change in the value of the portfolio is

$$d\Pi = -V + O(dt^{1/2}).$$

This is due to the sudden loss of the risky bond; the other terms are small in comparison. It is at this point that we could put in a recovery rate, discussed in the next section: As it stands here default means no return whatsoever.

Taking expectations and using the bond-pricing equation for the riskless bond, we find that the value of the risky bond satisfies

$$\boxed{\begin{aligned}&\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + \rho w\delta\frac{\partial^2 V}{\partial r\,\partial p} + \tfrac{1}{2}\delta^2\frac{\partial^2 V}{\partial p^2} + (u - \lambda w)\frac{\partial V}{\partial r}\\ &\quad + \gamma\frac{\partial V}{\partial p} - (r + p)V = 0.\end{aligned}} \tag{40.5}$$

This equation has final condition

$$V(r, p, T) = 1,$$

if the bond is zero coupon with a principal repayment of \$1.

Equation (40.5) again shows the similarity between the spot interest rate, r, and the hazard rate, p. The equation is remarkably symmetrical in these two variables, the only difference is in the choice of the model for each. In particular, the final term includes a discounting at rate r and also at rate p. These two variables play similar roles in credit risk equations.

As a check on this result, return to the simple case of constant p. In the new framework this case is equivalent to $\gamma = \delta = 0$. The solution of (40.5) is easily seen to be

$$e^{-p(T-t)}Z(r,t),$$

as derived earlier.

If γ and δ are independent of r and the correlation coefficient ρ is zero then we can write

$$V(r,p,t) = Z(r,t)H(p,t),$$

where H satisfies

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}\delta^2 \frac{\partial^2 H}{\partial p^2} + \gamma \frac{\partial H}{\partial p} - pH = 0,$$

with

$$H(p,T) = 1.$$

In this special, but important case, the default risk decouples from the bond pricing.

40.7 POSITIVE RECOVERY

In default there is usually *some* payment, not all of the money is lost. In Table 40.1, produced by Moody's from historical data, is shown the mean and standard deviations for recovery according to the seniority of the debt. These numbers emphasize the fact that the rate of recovery is itself very uncertain. How can we model a positive recovery?

Suppose that on default we know that we will get an amount Q. This will change the partial differential equation. To see this we return to the derivation of Equation (40.4). If there is no default we still have Equation (40.2). However, on default Equation (40.3) becomes instead

$$d\Pi = -V + Q + O(dt^{1/2});$$

we lose the bond but get Q. Taking expectations results in

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2 \frac{\partial^2 V}{\partial r^2} + \rho w \delta \frac{\partial^2 V}{\partial r \partial p} + \tfrac{1}{2}\delta^2 \frac{\partial^2 V}{\partial p^2} + (u - \lambda w)\frac{\partial V}{\partial r} + \gamma \frac{\partial V}{\partial p} - (r+p)V + pQ = 0.$$

Now you are faced with the difficult task of estimating Q, or modeling it as another random variable. It could even be a fraction of V. The last is probably the most sensible approach.

Table 40.1 Rate of recovery. Source: Moody's.

Class	Mean (%)	Std Dev. (%)
Senior secured	53.80	26.86
Senior unsecured	51.13	25.45
Senior subordinated	38.52	23.81
Subordinated	32.74	20.18
Junior subordinated	17.09	10.90

40.8 SPECIAL CASES AND YIELD CURVE FITTING

We saw in Chapter 30 that some spot interest rate models lead to explicit solutions for bond prices, for example the Vasicek model, the CIR model and in general the affine model with four time-dependent parameters. We can find simpler equations than the two-factor diffusion equation for the value of a risky bond in the above framework if we choose the functions $u - \lambda w$, w, γ, δ and ρ carefully. Obviously, from the analysis of Chapter 30, we must choose $u - \lambda w$ and w^2 to be linear in r. For simplifications of (40.5) to exist we also require γ and δ^2 to be linear in r and p. The form of the correlation coefficient is more complicated so we shall choose it to be zero.

With these choices for the functions in the two stochastic differential equations we find that the solution of (40.5) with $V(r, p, T) = 1$ is

$$V = \exp\left(A(t;T) - B(t;T)r - C(t;T)p\right)$$

where A, B and C satisfy non-linear first-order ordinary differential equations. I leave it as an exercise for the reader to derive them. In some cases these equations can be solved explicitly, usually in terms of special functions, but in others they must be solved numerically. Such a solution will of course be much quicker than the numerical solution of the two-factor diffusion equation.

If we allow the spot interest rate model to have some simple time dependence then we have the freedom to fit the yield curve. This concept is discussed in detail in Chapter 31 and the principle is the same here. In practice this may mean that we would restrict our attention to the extended Vasicek model with constant spot rate volatility. Similarly, if there is time dependence in the model for the hazard rate, and the model is sufficiently tractable, then you can also fit a hazard rate term structure.

40.9 A CASE STUDY: THE ARGENTINE PAR BOND

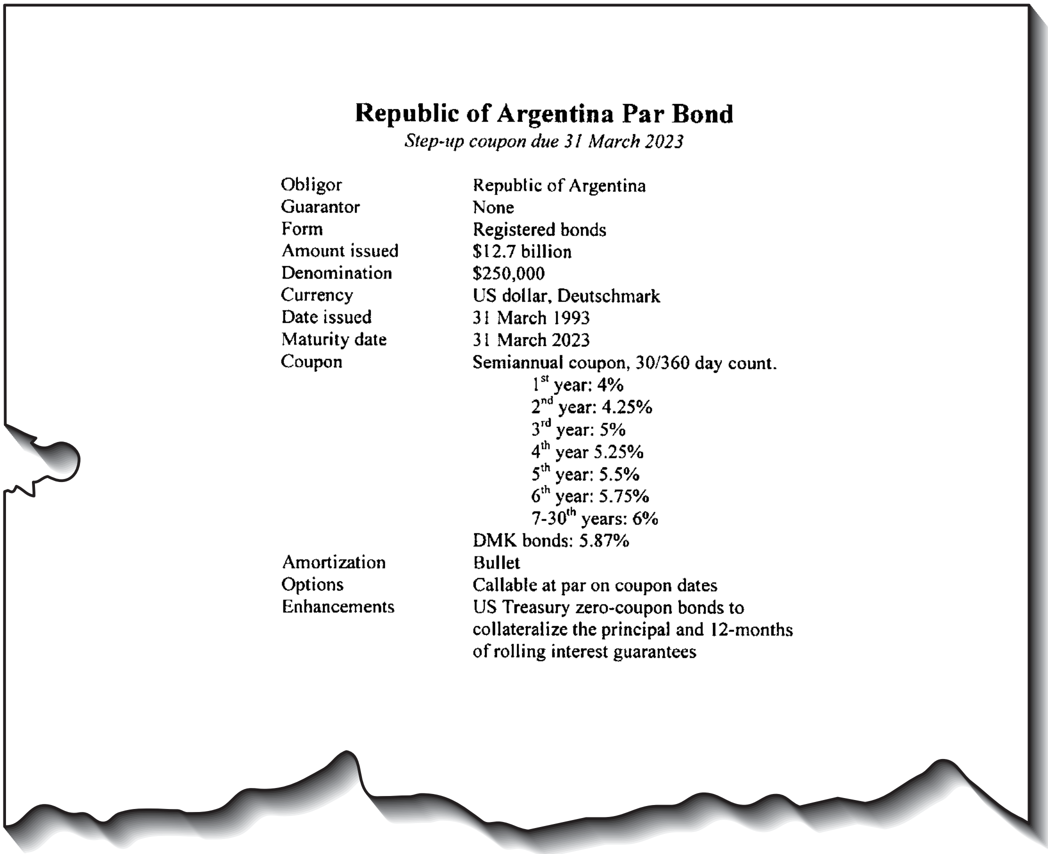
The Brady Plan was conceived in 1989 by former US Treasury Secretary Nicholas Brady. The plan consists of the repackaging of commercial bank debt into tradable fixed-income securities. Creditor banks either lower their interest on the debt or reduce the principal. Debtor countries, in exchange, are committed to make macroeconomic adjustments. Most Brady bonds are dollar denominated with maturities of longer than ten years and either fixed or floating coupon payments. At the time of writing many countries were beginning to buy back their Brady bonds. For up-to-date information see `www.bradynet.com`.

The US dollar-denominated Argentine Par bond has the specifications shown in Figure 40.5.

From time-series data for real risky bond prices and a suitable model, such as described above, we can calculate the value of the instantaneous risk of default for each data point that is needed for the model to give a theoretical value equal to the market value of the bond. This number is the implied instantaneous risk of default and plays a role in default risk analysis that is similar to that played by implied volatility for options: It is used as a trading indicator or as a measure of relative value.

In the analysis of the Argentine Par bond the risk of default was assumed to satisfy

$$dp = (f - hp)\,dt + jp^{1/2}\,dX_2, \tag{40.6}$$



Republic of Argentina Par Bond

Step-up coupon due 31 March 2023

Obligor	Republic of Argentina
Guarantor	None
Form	Registered bonds
Amount issued	\$12.7 billion
Denomination	\$250,000
Currency	US dollar, Deutschmark
Date issued	31 March 1993
Maturity date	31 March 2023
Coupon	Semiannual coupon, 30/360 day count. 1st year: 4% 2nd year: 4.25% 3rd year: 5% 4th year 5.25% 5th year: 5.5% 6th year: 5.75% 7-30th years: 6% DMK bonds: 5.87%
Amortization	Bullet
Options	Callable at par on coupon dates
Enhancements	US Treasury zero-coupon bonds to collateralize the principal and 12-months of rolling interest guarantees

Figure 40.5 Specifications of the Argentinian Par bond.

and to be uncorrelated with the spot interest rate. In this model the risk of default is mean reverting. I have not included any interest rate dependence in this because, provided $f > j^2/2$, this precludes the possibility of negative risk of default. It also makes the yield-curve fitting very simple because of the decomposition of the present value of each cashflow into the product of risk-free bond value and a function of just p and t. In other words, an interest rate model is not needed if a risk-free yield curve is given. The speed of reversion is determined by h. The parameters were chosen to be $h = 0.5$, $f = 0.045$ and $j = 0.03$. This choice was made partly so that the resulting time series for the implied p had the right theoretical properties (deduced from (40.6)) and partly using common sense.

Finally, the value for p was chosen daily so that the market price of the bonds and their theoretical price coincided.

The period chosen (end December 1993 to end September 1996) is a particularly exciting one because of the 'Tequila Effect' and it could easily be argued that there was a dramatic change of market conditions (and hence model parameters) at that time. However, I have kept the same parameters for the whole of this period since it was risk of default causing the Tequila effect and this should therefore be accounted for in these parameters.

The Tequila effect took place in December 1994 but its consequences lasted much longer, in some countries up to three and four months. In the case of Argentina, we can observe in Figure 40.6 the minimum price of the Par bond at the end of March 1995, dipping below \$35. Since then a steady recovery can be seen.

In Figure 40.7 we can observe that the Tequila effect was accompanied by a sharp increase in long rates in the US, which knocked Brady bond prices even further. The highest long rate over this period was 8% and it occurred in March 1995.

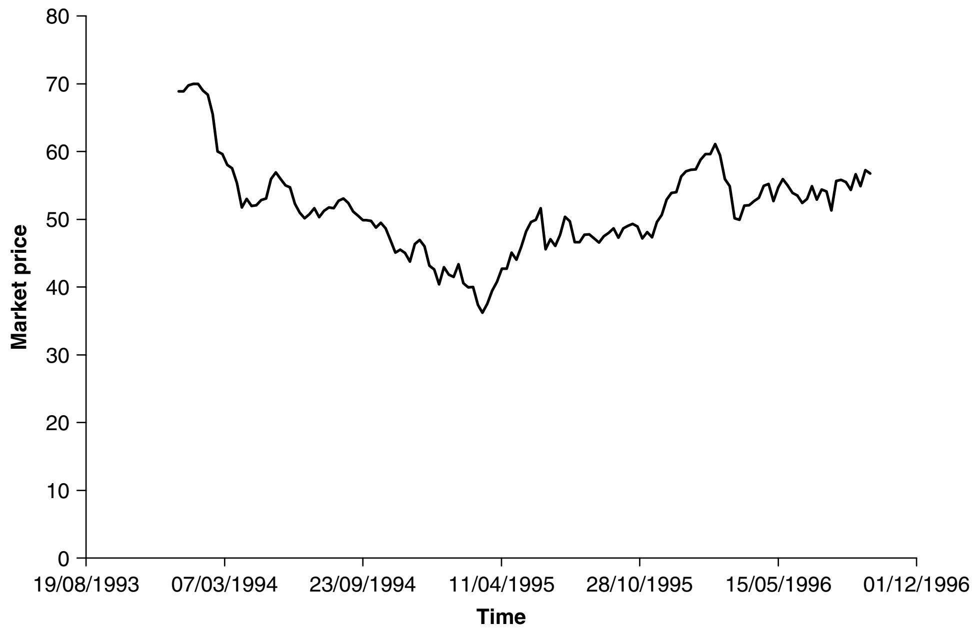


Figure 40.6 Market price of the Argentinian Par bond from end December 1993 to end September 1996.

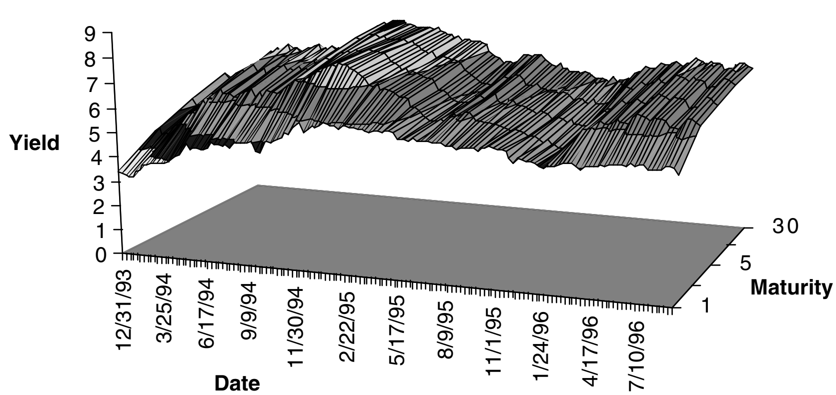


Figure 40.7 US yield curve from end December 1993 to end September 1996.

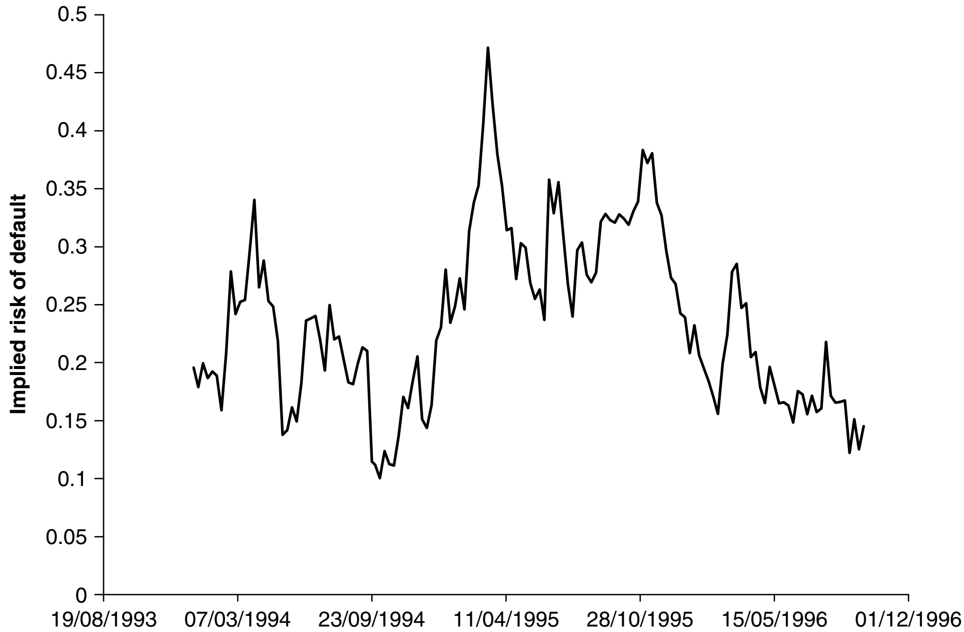


Figure 40.8 The implied risk of default for the Argentinian Par, see text for description of the stochastic model.

In Figure 40.8 is shown the implied instantaneous risk of default for Argentinian Par bonds over the period end December 1993 to end September 1996. As expected, the highest probability of default took place at the end of March 1995 when the Tequila Effect was at its worst. Since then there has been a steady, but obviously not monotonic, decrease in the risk of default implied by this model.

40.10 HEDGING THE DEFAULT

In the above we used riskless bonds to hedge the random movements in the spot interest rate. Can we introduce another risky bond or bonds into the portfolio to help with hedging the default risk? To do this we must assume that default in one bond automatically means default in the other.

Assuming that the risk of default p is constant for simplicity, consider the portfolio

$$\Pi = V - \Delta Z - \Delta_1 V_1,$$

where both V and V_1 are risky.

The choices

$$\Delta_1 = \frac{V}{V_1} \quad \text{and} \quad \Delta = \frac{V_1 \dfrac{\partial V}{\partial r} - V \dfrac{\partial V_1}{\partial r}}{V_1 \dfrac{\partial Z}{\partial r}}$$

eliminate both default risk and spot rate risk. The analysis results in

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+(u-\lambda w)\frac{\partial V}{\partial r}-(r+\lambda_1(r,t)p)V=0.$$

Observe that the 'market price of default risk' λ_1 is now where the probability of default appeared before, thus we have a risk-neutral probability of default. This equation is the risk-neutral version of Equation (40.4).

Can you imagine what happens if the risk of default is stochastic? There are actually three sources of randomness:

- the spot rate (the random movement in r);
- the probability of default (the random movement in p);
- the event of default (the Poisson process kicking in).

This means that we must hedge with *three* bonds, two other risky bonds and a risk-free bond, say. Where will you find market prices of risk? Can you derive a risk-neutral version of Equation (40.5)?

40.11 IS THERE ANY INFORMATION CONTENT IN THE MARKET PRICE?

A bond has a nominal value of 100, an amount to be received in one day's time. Yet the value in the market is 96. What does this mean? The implied p from this is an astronomical 1000%. Conclusion: Some people in the market know something. Default is certain.

Assume that a certain fraction of the market 'knows' what a mess the company is in. Call that small fraction ϵ. Assume that this fraction of the market assigns a value of zero to the bond, while the remaining $1-\epsilon$ give it the default-free value. Equate the average value with the value in the market to get

$$\epsilon\times 0+(1-\epsilon)\times e^{-r(T-t)}=e^{-(r+p)(T-t)}.$$

From this we get

$$p=-\frac{\log(1-\epsilon)}{T-t}\approx\frac{\epsilon}{T-t}\quad\text{if }\epsilon\text{ is small.}$$

Don't worry about the details of this concept, just observe how p grows inversely with time to maturity.

40.11.1 Implied Hazard Rate and Duration

There is an interesting relationship between the implied hazard rate and the duration of a coupon-bearing bond.

Again, suppose that a fraction ϵ of the market are 'in the know.' Equate the average value with the market value assuming a constant hazard rate model

$$\epsilon\times 0+(1-\epsilon)\times\sum_{i=1}^{N}c_i e^{-r(t_i-t)}=\sum_{i=1}^{N}c_i e^{-(r+p)(t_i-t)}.$$

(The notation is obvious, and I've lumped the principal in as just another coupon.)

If p is small we can write the right-hand side as

$$\sum_{i=1}^{N} c_i e^{-r(t_i-t)} - p\left(\sum_{i=1}^{N} c_i(t_i-t)e^{-r(t_i-t)}\right) + \cdots.$$

Equating the two sides we find that p is just ϵ divided by the duration of the bond when assumed risk free.

40.12 CREDIT RATING

There are many **Credit Rating Agencies** who compile data on individual companies or countries and estimate the likelihood of default. The most famous of these are **Standard & Poor's** and **Moody's**. These agencies assign a **credit rating** or **grade** to firms as an estimate of their credit-worthiness. Standard & Poor's rate businesses as one of AAA, AA, A, BBB, BB, B, CCC or Default. Moody's use Aaa, Aa, A, Baa, Ba, B, Caa, Ca, C. Both of these companies also have finer grades within each of these main categories. The Moody grades are described in Table 40.2.

In Figure 40.9 is shown the percentage of defaults over the past eighty years, sorted according to their Moody's credit rating.

The credit rating agencies continually gather data on individual firms and will, depending on the information, grade/regrade a company according to well-specified criteria. A change of rating is called a **migration** and has an important effect on the price of bonds issued by the company. Migration to a higher rating will increase the value of a bond and decrease its yield, since it is seen as being less likely to default.

Clearly there are two stages to modeling risky bonds under the credit-rating scenario. First we must model the migration of the company from one grade to another and second we must price bonds taking this migration into account.

Figure 40.10 shows the credit rating for Eastern European countries by rating agency.

Table 40.2 The meaning of Moody's ratings.

Aaa	Bonds of best quality. Smallest degree of risk. Interest payments protected by a large or stable margin.
Aa	High quality. Margin of protection lower than Aaa.
A	Many favorable investment attributes. Adequate security of principal and interest. May be susceptible to impairment in the future.
Baa	Neither highly protected nor poorly secured. Adequate security for the present. Lacking outstanding investment characteristics. Speculative features.
Ba	Speculative elements. Future not well assured.
B	Lack characteristics of a desirable investment.
Caa	Poor standing. May be in default or danger with respect to principal or interest.
Ca	High degree of speculation. Often in default.
C	Lowest-rated class. Extremely poor chance of ever attaining any real investment standing.

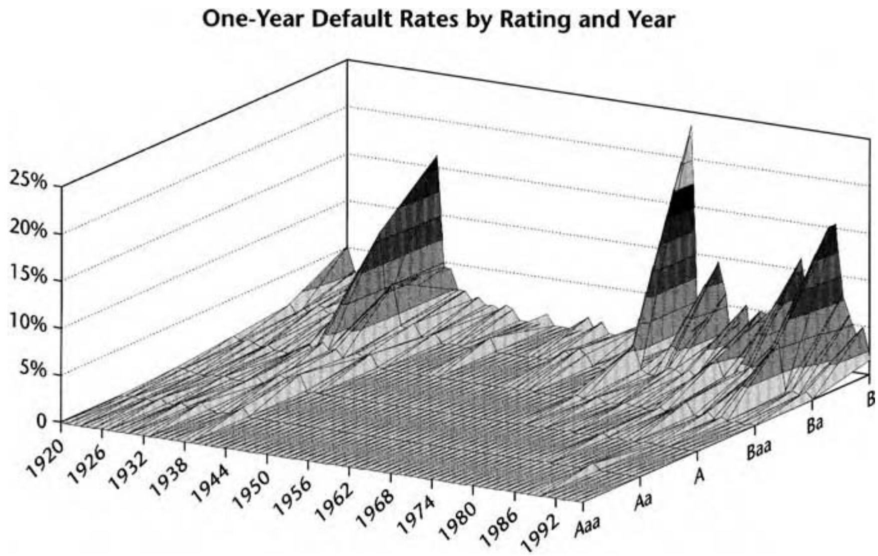


Figure 40.9 Percentage of defaults according to rating. Source: Moody's. Reproduced by permission of Moody's Investors Services.

<HELP> for explanation. DL18 Corp **CSDR**
Enter # <GO> for historical ratings.

Page 1/2

Foreign Currency LT Debt
Region – Eastern Europe

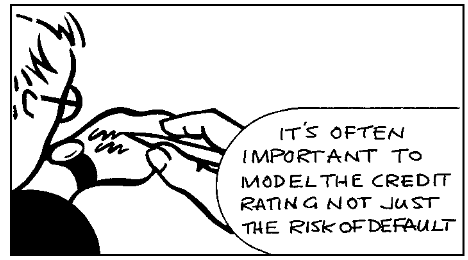
	MOODY'S	S&P	DCR	FI
Bulgaria	1) B2	15) B	29) NR	43) B+
Croatia	2) Baa3	16) BBB-	30) NR	44) BB+
Cyprus	3) A2	17) A+	31) NR	45) NR
Czech Republic	4) Baa1	18) A-	32) A-	46) BBB+
Estonia	5) Baa1	19) BBB+	33) NR	47) BBB
Hungary	6) Baa1	20) BBB	34) BBB	48) BBB
Latvia	7) Baa2	21) BBB	35) NR	49) BBB
Lithuania	8) Ba1	22) BBB-	36) NR	50) BB+
Moldova	9) B2 *-	23) NR	37) NR	51) B
Poland	10) Baa1	24) BBB	38) BBB-	52) BBB+
Romania	11) B3	25) B-	39) NR	53) B-
Russia	12) B3	26) SD	40) B- *-	54) CCC
Slovakia	13) Ba1	27) BB+	41) NR	55) BB+
Slovenia	14) A3	28) A	42) NR	56) A-

COLOR DENOTES A RATING CHANGE WITHIN THE LAST 30 DAYS (Pos/Neg/Neutral)

Bloomberg

Figure 40.10 Ratings for Eastern European countries by rating agency. Source: Bloomberg L.P.

40.13 A MODEL FOR CHANGE OF CREDIT RATING



Company XYZ is currently rated A by Standard & Poor's. What is the probability that in one year's time it will still be rated A? Suppose that it is 91.305%. Now what is the probability that it will be rated AA or even AAA, or in default? We can represent these probabilities over the one year time horizon by a **transition matrix**. An example is shown in Table 40.3.

This table is read as follows. Today the company is rated A. The probability that in one year's time it will be at another rating can be seen by reading across the A row in the table. Thus the probability of being rated AAA is 0.092%, AA 2.42%, A 91.305% etc. The highest probability is of no migration. By reading down the rows, this table can be interpreted as either a representation of the probabilities of migration of *all* companies from one grade to another, or of company XYZ had it started out at other than A. Whatever the grade today, the company must have some rating at the end of the year even if that rating is default. Therefore the probabilities reading across each row must sum to one. And once a company is in default, it cannot leave that state, therefore the bottom row must be all zeros except for the last number which represents the probability of going from default to default, i.e. 1.

This table or matrix represents probabilities over a finite horizon. But during that time a bond may have gone from A to BBB to BB; how can we model this sequence of migrations? This is done by introducing a transition matrix over an infinitesimal time period. We can model continuous-time transitions between states via **Markov chains**.

We will model migrations over the short time period from t to $t + dt$. Since this time period is very short, the chance of any migration at all is small. The most likely event is that there is no migration. I am going to scale the probability of a change of state with the size of the time step dt; any other scaling will lead to a meaningless or trivial model. If the transition matrix over the time step is $\mathbf{P}_{dt}$ then I can write

$$\mathbf{P}_{dt} = \mathbf{I} + dt\,\mathbf{Q},$$

for some matrix $\mathbf{Q}$ and where $\mathbf{I}$ is the identity matrix. The sum of the entries in each row of $\mathbf{Q}$ must sum to zero, and the bottom row must only contain zeros since default is an absorbing state. I will use $\mathbf{P}(t, t')$ to denote the transition matrix over a *finite* time interval from t until t'.

Table 40.3 An example of a transition matrix

	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	0.90829	0.08272	0.00736	0.00065	0.00066	0.00014	0.00006	0.00012
AA	0.00665	0.90890	0.07692	0.00583	0.00064	0.00066	0.00029	0.00011
A	0.00092	0.02420	0.91305	0.05228	0.00678	0.00227	0.00009	0.00041
BBB	0.00042	0.00320	0.05878	0.87459	0.04964	0.01078	0.00110	0.00149
BB	0.00039	0.00126	0.00644	0.07710	0.81159	0.08397	0.00970	0.00955
B	0.00044	0.00211	0.00361	0.00718	0.07961	0.80767	0.04992	0.04946
CCC	0.00127	0.00122	0.00423	0.01195	0.02690	0.11711	0.64479	0.19253
Default	0	0	0	0	0	0	0	1

40.13.1 The Forward Equation

By considering how one can change from state to state during the time step dt and the relevant probabilities we find that the relationship between $\mathbf{P}(t, t')$ and $\mathbf{P}_{dt}$ is simply

$$\mathbf{P}(t, t' + dt) = \mathbf{P}(t, t')\mathbf{P}_{dt}.$$

In terms of $\mathbf{Q}$ this is

$$\mathbf{P}(t, t' + dt) = \mathbf{P}(t, t')(\mathbf{I} + dt\,\mathbf{Q}).$$

Subtracting $\mathbf{P}(t, t')$ from both sides and dividing by dt we get

$$\frac{\partial \mathbf{P}(t, t')}{\partial t'} = \mathbf{P}(t, t')\mathbf{Q}.$$

This ordinary differential equation is the **forward equation** and must be solved with

$$\mathbf{P}(t, t) = \mathbf{I}.$$

The solution of this *matrix* equation for constant $\mathbf{Q}$ is

$$\mathbf{P}(t, t') = e^{(t'-t)\mathbf{Q}}. \tag{40.7}$$

The exponential of a matrix is defined via an infinite sum so that

$$e^{(t'-t)\mathbf{Q}} = \sum_{i=0}^{\infty} \frac{1}{i!}(t' - t)^i \mathbf{Q}^i.$$

We can use Equation (40.7) in several ways. First, suppose that at time $t = 0$ company XYZ is rated A. Supposing that we know $\mathbf{Q}$, how can we find the probability of being in any particular state at the future time T? This is simple. We just need to find the third row down in the matrix $\mathbf{P}(0, T)$, with $\mathbf{e}_i$ to denote the row vector with zeros everywhere except in the ith column, corresponding to the initial state. In our case $i = 3$. The answer to the question is

$$\mathbf{e}_i\mathbf{P}(0, T) = \mathbf{e}_i e^{T\mathbf{Q}}.$$

Another way to use the solution of the forward equation is to deduce the matrix $\mathbf{Q}$ from the transition matrix over a finite time horizon. In other words, we can solve

$$e^{T\mathbf{Q}} = \mathbf{P}(0, T)$$

for $\mathbf{Q}$. Why might we want to do this? One reason is that some rating agencies, and other firms, publish the transition matrix for a time horizon of one year, for example Table 40.3. If you want to know what might happen for shorter timescales than that (and you believe the one-year matrix) then you should find $\mathbf{Q}$.

Suppose that we can **diagonalize** the matrix $\mathbf{Q}$ in the form

$$\mathbf{Q} = \mathbf{M}\mathbf{D}\mathbf{M}^{-1},$$

where $\mathbf{D}$ is a diagonal matrix. If we can do this then the entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{Q}$. We can then write

$$\begin{aligned}
\mathbf{P}(0,T) = e^{T\mathbf{Q}} &= \sum_{i=0}^{\infty} \frac{1}{i!} T^i \mathbf{Q}^i \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} T^i (\mathbf{M}\mathbf{D}\mathbf{M}^{-1})^i \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} T^i \underbrace{(\mathbf{M}\mathbf{D}\mathbf{M}^{-1})\dots(\mathbf{M}\mathbf{D}\mathbf{M}^{-1})}_{i} \\
&= \mathbf{M} \sum_{i=0}^{\infty} \frac{1}{i!} T^i \mathbf{D}^i \mathbf{M}^{-1}.
\end{aligned}$$

But since $\mathbf{D}$ is diagonal, when it is raised to the ith power the result is another diagonal matrix with each diagonal element raised to the ith power:

$$\mathbf{D}^i = \begin{pmatrix} d_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 \\ 0 & 0 & d_3 & 0 & 0 \\ 0 & 0 & 0 & d_4 & 0 \\ 0 & 0 & 0 & 0 & d_5 \end{pmatrix}^i = \begin{pmatrix} d_1^i & 0 & 0 & 0 & 0 \\ 0 & d_2^i & 0 & 0 & 0 \\ 0 & 0 & d_3^i & 0 & 0 \\ 0 & 0 & 0 & d_4^i & 0 \\ 0 & 0 & 0 & 0 & d_5^i \end{pmatrix}.$$

From this it follows that

$$\mathbf{P}(0,T) = \mathbf{M} e^{T\mathbf{D}} \mathbf{M}^{-1},$$

where $e^{T\mathbf{D}}$ is the matrix with diagonal elements e^{Td_i}. The eigenvalues of the two matrices $\mathbf{P}(0,T)$ and $\mathbf{Q}$ are closely related. The strategy for finding $\mathbf{Q}$ is to first diagonalize $\mathbf{P}(0,T)$ to find $\mathbf{M}\ e^{T\mathbf{D}}$, from which it is a simple matter to determine the matrix $\mathbf{Q}$.

40.13.2 The Backward Equation

The **backward equation**, which has a similar meaning to the backward equation for diffusion problems, can be derived in a similar manner. The equation is

$$\frac{\partial \mathbf{P}(t,t')}{\partial t} = -\mathbf{Q}\mathbf{P}(t,t'). \tag{40.8}$$

40.14 THE PRICING EQUATION

Having built up a model for rating migration, let's look at how to price risky bonds. We will concentrate on zero-coupon bonds. In the previous section we derived forward and backward equations for the transition matrix. The link between the backward equation and contract prices in the Brownian motion world is retained in the Markov chain world, so I will skip most of the details.

40.14.1 Constant Interest Rates

The price of the risky bond depends on the credit rating of the company. We will therefore need one value per rating. The column vector $\mathbf{V}$ will have as its entries the bond value for each of the credit states. Assuming for the moment that interest rates are constant, this vector will be a function of t only. In the same way that the value of an option is related to the backward equation for the transition density function, we now have the following equation for the bond value:

$$\boxed{\frac{d\mathbf{V}}{dt} + (\mathbf{Q} - r\mathbf{I})\mathbf{V} = 0,}$$

This is just the backward equation (40.8) with an extra discounting term. The final condition for this equation is

$$\mathbf{V}(T) = \mathbf{1},$$

where $\mathbf{1}$ is the column vector consisting of 1s in all the rows, except the last where there is a zero.

How does the equation change if there is a recovery on default?

Where would the market price of risk appear if you were able to hedge against change of rating?

40.14.2 Stochastic Interest Rates

The extension to stochastic interest rate is quite straightforward. The governing equation is

$$\boxed{\frac{\partial \mathbf{V}}{\partial t} + \tfrac{1}{2}\beta^2 \frac{\partial^2 \mathbf{V}}{\partial r^2} + (\alpha - \lambda\beta)\frac{\partial \mathbf{V}}{\partial r} + (\mathbf{Q} - r\mathbf{I})\mathbf{V} = 0,}$$

40.15 CREDIT RISK IN CBS

The risk of default can be very important for the convertible bond (CB), discussed in some depth in Chapter 33 but with little reference to credit issues. The CB is like a bond in that it pays its owner coupons during its life and a principal at maturity. However, the holder may convert the bond at specified times into a number of the underlying stock. This feature makes the CB very like an option. The reader is referred to Chapter 33 for all the details and the notation.

We can combine the ideas in the present chapter with those from Chapter 33 to derive a model for CBs with risk of default priced in. I will present two possible approaches.

40.15.1 Bankruptcy when Stock Reaches a Critical Level

We can model the default of the issuing company by saying that should its stock fall to a level S_b then it will default. Such a model is similar in spirit to that described in Section 39.2. We only need to add the condition

$$V(S_b, t) = 0,$$

to our favorite (no credit risk) CB model.

40.15.2 Incorporating the Instantaneous Risk of Default

Another possibility, in line with the instantaneous risk of default model described above, is to have an exogenous default triggered by a Poisson process, as in Section 40.4. In a two-factor CB setting we arrive at

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma S w \frac{\partial^2 V}{\partial S \partial r} + \tfrac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2}$$
$$+ (r + p) S \frac{\partial V}{\partial S} + (u - \lambda w) \frac{\partial V}{\partial r} - (r + p) V = 0.$$

In this model we have stochastic r and stochastic S. We can have p as constant, a known function of time, or even a known function of r and S.

This approach has the advantage that it reduces to common market pricing practice in the absence of any stock dependence. This allows us to price instruments in a consistent fashion.

When the stock price is very small the above model will yield a CB price that is close to the price of a non-convertible bond. It is market experience, however, that in such a situation the price of the CB falls dramatically. The market sees the low stock price as an indicator of a very sick company. This can be modeled by having the instantaneous risk of default being dependent on the stock price, $p(S)$. If p goes to infinity sufficiently rapidly as $S \to 0$ we find that the CB value goes to zero. Note that there is no more effort involved, computationally, in solving such a problem since we must anyway solve for the CB value as a function of S. In Figure 40.11 is shown the value of a CB with and without credit risk taken into account in a one-factor model, with deterministic interest rates.

In Figure 40.12 is shown the output of a two-factor model for the CB with risk of default. The interest rate model was Vasicek, fitted to a flat 7% yield curve.

40.16 MODELING LIQUIDITY RISK

Liquidity risk affects the price of a bond through the yield spread. The yield spread is the difference between the bid and ask price expressed as a percentage of the mid-price. For example, if the bid price is \$0.97, and the ask price is \$1.03 so that the mid-price is \$1.00 then the yield spread is 6%. Measuring the yield spread in percentage terms rather than dollar terms ensures consistency between bonds with different prices.

A bond is liquid if it is traded in large volume. There are many buyers and sellers and so the dealer can easily match both sides of the transaction. The uncertainty faced by dealers when matching buyers to sellers is small, leading to a relatively small bid/ask spread.

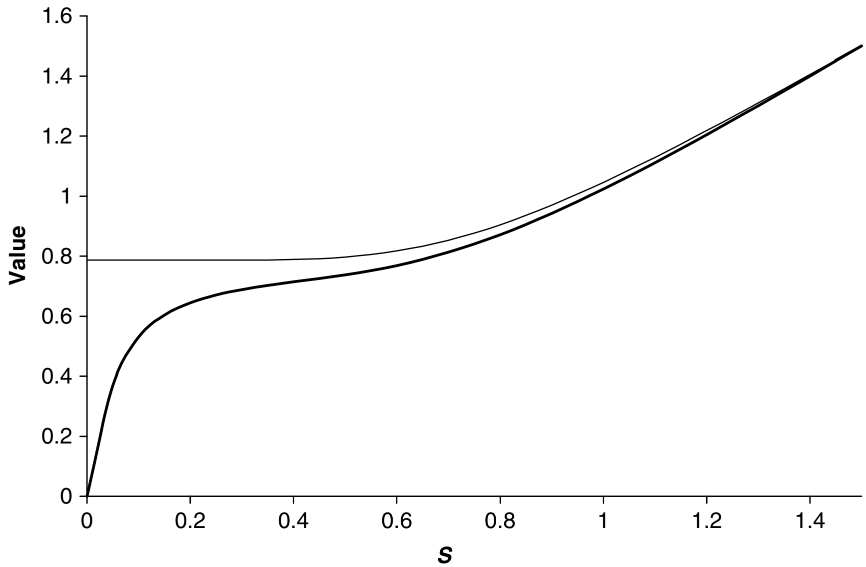


Figure 40.11 Value of a CB against underlying with (lower curve) and without (upper curve) risk of default.

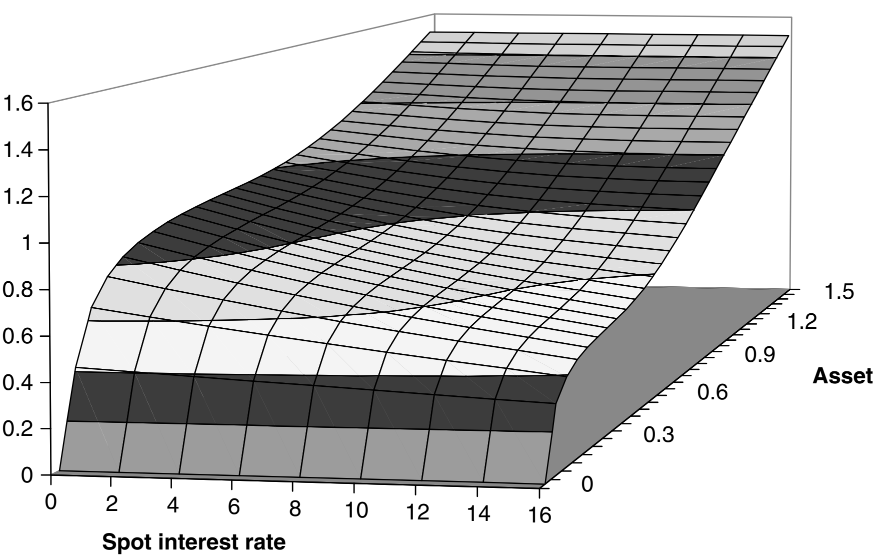


Figure 40.12 Value of a CB against underlying and interest rate with risk of default taken into account.

If a bond is illiquid it is only traded in small volume and hence there are few buyers and/or sellers in the market. In this case the dealer faces much uncertainty when matching buyers to sellers. In order to compensate for this risk the spread for illiquid bonds tends to be large.

The liquidity of a bond can change over time. A usually liquid issue may go through temporary periods of poor liquidity.

In times of poor liquidity the yield spread widens making it more expensive to buy or sell bonds. Possible reasons for periods of illiquidity are the following.

- A change in the issuing company's credit rating making the bond less attractive.
- Other company specific factors, such as a company reporting low profits.
- Liquidity tends to dry up for bonds that are close to maturity.
- Generally we would expect a small issue to have lower liquidity than a large issue because there are fewer bonds to buy and sell.
- Placement of unusually large orders.
- Macroeconomic factors could lead to a positive or negative view on specific companies, industries or the entire bond market. For example, a booming equity market may divert funds away from the bond market.

A period of poor liquidity may only last for a brief time or it may last until maturity. If a bond holder decides to sell a large number of bonds then illiquidity may only last for a couple of days until they have finished selling. In the case of a credit-rating downgrade, liquidity may be reduced until the credit rating returns to its original level. This may not happen prior to maturity. Hence, a period of illiquidity due to a credit rating change may last indefinitely.

Investors only incur costs due to illiquidity if they choose to trade bonds during a period of poor liquidity. An investor who wishes to hold a bond until maturity faces no costs due to poor liquidity and thus need not be concerned about liquidity risk. Investors who wish to trade bonds on the secondary market face potential liquidity risk. In order to assess the cost of illiquidity, such investors need to consider the following:

- likelihood of a period of poor liquidity
- size of change in bid/ask spread during a period of poor liquidity
- probability that they will need to trade during a period of poor liquidity

For the sake of simplicity we shall focus on changes in liquidity due to changes in credit rating. If a company's credit rating is downgraded then the bid/ask spread on its corporate bonds tends to increase. This relationship is highlighted in Table 40.4 which shows movements in the bid/ask spread after changes in credit rating for Korea Development Bank corporate bonds.

Let's model the spread s as

$$ds = a_i(s)\,dt + b_i(s)\,dX + \sum_{j=1}^{8} c_{i,j}(s)\,dq_{i,j}. \tag{40.13}$$

Here i is the credit rating, 1 is the highest rating and 8 is default.

Table 40.4 Changes in rating and spread for KDB.

Date of rerating	Upgrade or downgrade	Credit rating before rerating	Credit rating after rerating	% change in bid/ask spread
24/10/97	Downgrade	AA−	A+	10.02
25/11/97	Downgrade	A+	A−	NA
11/12/97	Downgrade	A−	BBB−	45.46
22/12/97	Downgrade	BBB−	B+	12.20
18/2/98	Upgrade	B+	BB+	−3.88
25/1/99	Upgrade	BB+	BBB−	−5.41

The term $a_i(s)$ is the drift of the spread while the bond is i rated. A suitable form for this function would be the mean reverting

$$a_i(s) = \alpha_i(e_i - s)$$

where e_i is some equilibrium, or average, level for the spread while i rated.

The volatility of the spread is given by b. A suitable form would be

$$b_i(s) = \beta_i s^{1/2}.$$

The Poisson terms $dq_{i,j}$ represent jumps from grade i to grade j. The intensity of these jump processes are given by the earlier $\mathbf{Q}$ matrix. The simplest form for the jump magnitudes $c_{i,j}(s)$ would be

$$c_{i,j}(s) = e_j - e_i,$$

i.e. the jump in spread is the difference between the equilibrium spreads for the two ratings.

A model such as (40.13) can be used to model, for example, expected level of the spread at any time in the future.

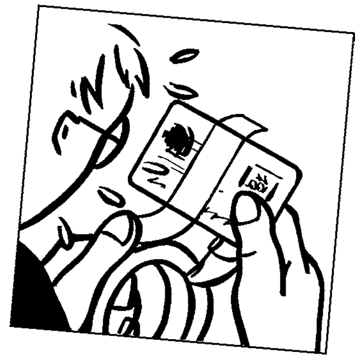
40.17 SUMMARY

As can be seen from this chapter, credit risk modeling is a very big subject. I have shown some of the popular approaches, but they are by no means the only possibilities. To aid with the assessment of credit risk, the bank JP Morgan has created CreditMetrics, a methodology for assessing the impact of default events on a portfolio. This is described in Chapter 42.

As a final thought for this chapter, suppose that a company issues just the one risky bond so that there is no way of hedging the default. If you believe that the market is underpricing the bond because it overestimates the risk of default then you might decide to buy it. If you intend holding it until expiry, then the market price in the future is only relevant in so far as you may change your mind. But you really do care about the likelihood of default and will pay very close attention to news about the company. On the other hand, if you buy the bond with the intention of only holding it for a short time your main concern should be for how the market is behaving and the real risk of default is irrelevant. You may still watch out for news about the company, but now your concern will be for how the market reacts to the news, not the news itself.

FURTHER READING

- Important work on the instantaneous risk of default model is by Jarrow & Turnbull (1990, 1995), Litterman & Iben (1991), Madan & Unal (1994), Lando (1994a), Duffie & Singleton (1994a, b) and Schönbucher (1996).
- See Blauer & Wilmott (1998) for the instantaneous risk of default model and an application to Latin American Brady bonds.
- See Duffee (1995) for other work on the estimation of the instantaneous risk of default in practice.
- The original work on change of credit rating was due to Lando (1994b), Jarrow, Lando & Turnbull (1997) and Das & Tufano (1994). Cox & Miller (1965) describe Markov chains in a very accessible manner.
- See Ahn, Khadem & Wilmott (1998) for the rather sensible use of utility theory in credit risk modeling.
- Current market conditions and prices for Brady bonds can be found at `www.bradynet.com`.
- See `www.emgmkts.com` for financial news from emerging markets.
- Clark (2002) uses the implied volatility as a measure of country risk.



CHAPTER 41
credit derivatives

In this Chapter...

- definitions and uses of credit derivatives
- credit derivatives triggered by default
- derivatives of the yield spread
- payment on change of rating
- examples of credit derivatives, including term sheets
- pricing credit derivatives using different models

41.1 INTRODUCTION

In this chapter I continue with the theme of pricing risky contracts. Here 'risky' means that there is some exposure to default risk. With credit derivatives that risk of default is explicitly acknowledged, and in many cases the owner of the credit derivative will benefit in the case of default. Thus a credit derivative may be thought of as insurance for another risky contract such as a simple bond; the holder of a bond always loses out on default.

There are many different types of credit derivative and the business is currently growing very rapidly. I will only discuss the main issues in the pricing of these contracts, giving examples as much as space will allow.

41.2 WHAT ARE CREDIT DERIVATIVES?

Credit derivatives are contracts that involve cashflows/payoffs between two parties that are linked to the credit characteristics of some third party.

The possible defining characteristics are as follows:

- Payoffs triggered by default event;
- Payoffs linked to mark-to-market return on specific security;
- Payoffs linked to independent credit rating assessment.

A credit derivative will almost certainly still have **market risk**, that is, the risk of market underlying moving against you. But they will also have **credit risk**, the risk that a counterparty won't pay up, or that someone's non-payment triggers a payoff, or that someone else's estimate of risk of default changes.

```
                                                          DL18 Corp   ODE5
Enter 11 <GO> for Involved Parties
                     SECURITY DISPLAY                     PAGE    1/ 4
ANGUS LIMITED    ANGUSFloat 08/00      N O T   P R I C E D

                                         ISSUER INFORMATION
SECURITY INFORMATION           SERIES: EMTN  NAME   ANGUS LIMITED
CPN FREQ        SEMI-AN        NOT SEASONED  TYPE   SPECIAL PURPOSE
CPN TYPE        FLOATING       IDENTIFICATION #'s   REDEMPTION INFO
MTY/REFUND TYP NORMAL          ISIN    XS0089628639 MATURITY DT   8/17/00
CALC TYP (198)NO CALC-FLOATERS MLNUM 99999          REFUNDING DT
DAY COUNT( 2) ACT/360                               NEXT CALL DT
MARKET ISS      EURO MTN       COMMON   008962863   WORKOUT DT    8/17/00
COUNTRY/CURR    KY  /USD                            RISK FACTOR
COLLATERAL TYP UNSUBORDINATED  ISSUANCE INFO
AMT ISSUED            30,000(M) ANNOUNCE DT   8/17/98    RATINGS
AMT OUTSTAND          30,000(M) 1ST SETTLE DT 8/17/98  MOODY   NA
MIN PC/INC  1,000,000/ 1,000,000 1ST CPN DT   2/17/99  S & P   NA
PAR AMT           1,000,000.00  INT ACCRUE DT 8/17/98  COMP    NR
LEADMGR/UWRTR                   PRICE @ ISSUE 100      DCR     NA
LISTED        LUXEMBOURG                               FI      NA
Hit 1 <GO> to See Other Avail Info
NOTES  NO PROSPECTUS, TAX CHANGES CALL
CPN RATE=6MO US$LIBOR +265BP. MOD FOLLOW BUS DAY CNVTN:IT, EN &US. SEC'D BY SWAP
AGRMNT.CALL IF CRED EVENT ON UNDERLYING OCCURS.UNSUB.UNSEC'D.SERIES 48,TRANCHE 1
Copyright 1999 BLOOMBERG L.P.  Frankfurt:69-920410  Hong Kong:2-2977-6000  London:171-330-7500  New York:212-318-2000
Princeton:609-279-3000   Singapore:226-3000   Sydney:2-9777-8686   Tokyo:3-3201-8900   Sao Paulo:11-3048-4500
                                                             I741-53-0 08-Sep-99 19:20:35
Bloomberg
```

Figure 41.1 Bond becomes callable if a credit event occurs. Source: Bloomberg L.P.

Figure 41.1 shows details of a bond that becomes callable if a ‘credit event’ occurs. In this example **credit event** was defined as one or more of the following:

- A failure to make payments when due;
- Distressed rescheduling of payment;
- An event of bankruptcy, debt restructuring.

Credit derivatives can offer *protection against* credit risks:

- Default risk;
- Credit spread risk;
- Downgrade risk.

Credit derivatives can be used to *gain an exposure* to credit risks:

- Enhance yields;
- Take on particular exposures.

Credit derivatives offer a way to ‘unbundle’ credit risk from other risks, to strip out credit risk from market risk, interest rate risk, etc. Previously you had to take a view on all of these together whether you wanted them or not, the classic example being convertible bonds. Credit derivatives are Over-the-Counter (OTC) products.

Aside: When considering the pricing of credit risk instruments don't forget that there is also credit exposure to whoever sold you the 'protection.'

41.2.1 Uses of Credit Derivatives: Banks

Banks are major players in the credit derivatives market. The reasons for their interest in these instruments are as follows:

- Hedging specific default risk
- Diversifying loan portfolio credit risk
- Hedging geographic, industry, etc. risk
- Freeing credit lines
- Tailoring exposures
- Lay-off risks from certain participations
- Managing capital

41.2.2 Uses of Credit Derivatives: Investors

Investors, and this includes hedge funds, have the following purposes in mind when they purchase credit derivatives:

- Access new asset classes
- Bank loans, foreign and emerging markets
- Create synthetic high-yield debt
- Arbitrage mispricing
- Decouple credit risk and market risk
- Provide leverage possibilities
- Exploit view of investor

41.2.3 Uses of Credit Derivatives: Corporates

Then, of course, there are the real people (an increasingly threatened minority!) who have an interest in these products:

- Hedging creditor, customer/supplier risk
- Hedging project risk (e.g. sovereign risk)

41.3 POPULAR CREDIT DERIVATIVES

The world of credit derivatives is expanding at an enormous rate. Here are the names of a few popular contracts, and we'll take a brief look at each of these.

- Asset swap
- Credit Default Swap
- Total Return Swaps
- Derivatives triggered by default

- Limited recourse note
- Derivatives of the yield spread
- Default calls and puts
- Exchange options
- Credit spread options
- Payment on change of rating

41.3.1 Asset Swap

Asset swaps are a basic building block for structured credit. Counterparty A owns a risky bond, but wants to swap the credit risk. He gives the bond to counterparty B in return for interest payments of LIBOR plus some suitable fixed spread. B will pay this interest until maturity life of the bond, even after default. The risk of default has passed from A to B. This is an **asset swap**. This contract can be made more complicated by including call or put provisions.

How the asset swap works

Purchase the fixed-rate asset (Figure 41.2) and simultaneously enter into a swap (Figure 41.3), fixed to floating.

Details

- Term of swap matches maturity of bond.
- Swap can include options if there are options in the bond (embedded calls, for example).
- Can include a written call, so that contract can be terminated if interest rates fall below a specified strike.

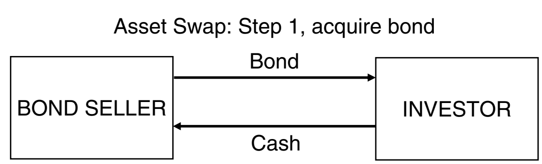


Figure 41.2 Purchase the fixed-rate asset.

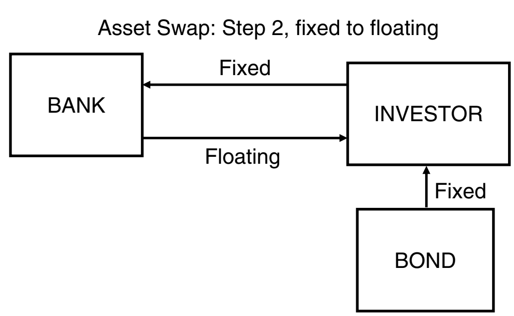


Figure 41.3 Enter into the swap.

Advantages

- Simple, transparent, flexible.
- Easy to isolate credit risk.

Disadvantages

- Many investors cannot enter into derivatives transactions (regulatory, accounting or investment policy restrictions).
- If the underlying bond defaults the investor continues to hold the swap. (The investor can continue payments or close the position at the then prevailing rate.)
- Investor is exposed to credit worthiness of swap counterparty.

Asset swap swap/switch

The asset swap swap is for the exchange of one asset swap for another. Typically the two assets are uncorrelated and the switch is triggered by the credit spread on one or both of the asset swaps.

41.3.2 Total Return Swaps or Total Rate of Return Swaps

Total return swaps were among the earliest credit derivatives. They existed before default swaps, but now default swaps are the more commonly traded instruments. The difference between a total return swap and a default swap is that a default swap simply transfers credit risk, by reference to some designated asset whereas a total return swap transfers all the risks of owning the designated asset.

How the TRS works

This is an off-balance-sheet transaction in which

- **Total Return Payer:** Transfers the cashflows plus any positive value change of a reference asset to the other party. . .interest payments, appreciation, coupons, etc.
- **Total Return Receiver:** Transfers a LIBOR + spread or fixed rate return plus any negative value change. This would be settled at specified intervals, every three months, say.

The maturity is typically less than the maturity of the underlying instrument. A TRS therefore provides a means of packaging and transferring *all* of the risks associated with a reference obligation, including credit risk (see Figure 41.4).

The Payer has stripped out all exposure to the market and credit but hasn't had to sell the asset. In return he gets LIBOR + spread. The Receiver gets the exposure but without having to purchase the asset. (Think of renting a part of someone else's balance sheet.)

Contrast this with a vanilla interest rate swap of fixed for floating. In such a swap there is a large netting of cashflows, even though the nominal may be large the cashflows may not be. In the TRS it is possible for the Receiver to have double the payment to make, the LIBOR + spread + capital appreciation if this is negative. And, of course, there is the risk of the receiver defaulting.

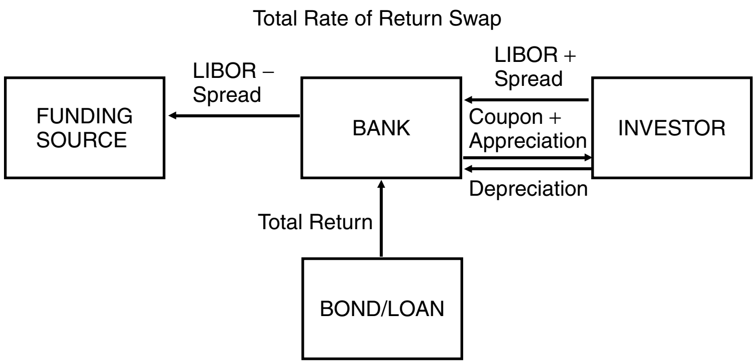


Figure 41.4 A Total Return Swap.

Why enter into a total return swap?

- TRSs are more flexible than transactions in the underlyings. For example, varying the terms of the swap contract allows the creation of synthetic assets that may not be otherwise available.
- Short-selling is often difficult, but can be achieved with a TRS.
- The swap receiver never has to make the outlay to buy the security. Even after posting collateral and paying a high margin, the resulting **leverage** and **enhanced return on regulatory capital** can be large. *This is the most important reason behind the popularity of TRSs.*

41.4 DERIVATIVES TRIGGERED BY DEFAULT

The most basic form of credit derivatives are those that pay off in the event of default by the issuing company or country. Technically, the definition of default is any non-compliance with the exact specifications of a contract, so that a coupon paid just one day late would count as a default event. Examples are

- credit default swap
- limited recourse note
- first to default
- nth to default

41.4.1 Basic Definitions

- **Credit Event:** Failure to meet payment obligations, plus...
- **Reference entity/credit:** A specified entity (sovereign, financial institution, corporation, or one among a basket of such specified entities)
- **Reference security/asset:** A security issued by the reference entity
- **Protection Buyer:** Also the 'Buyer,' the party paying a (periodic) fee in return for a contingent payment by the other party following a credit event of the reference entity

- **Protection Seller:** Also the ‘Seller,’ the party receiving a (periodic) fee in return for making a contingent payment

41.4.2 What Defines a ‘Credit Event’?

Credit default is any non compliance with the exact specifications of a contract in terms of compulsory cashflows, so that a coupon paid just one day late may count as a default event. Typically declaration of bankruptcy by the issuer/reference entity would come under this category. There may be ‘grace periods’ or limits put on the amount of outstanding money owed.

Also included are receivership, bankruptcy, insolvency, winding up etc. of the reference entity. Other credit events interfering with the terms of the contract such as war or revolution may also be included as criteria in the contract specifying certain action. General banking suspension may be included in the terms.

The definition of default also includes ‘materiality’ clause. There must be an accompanying significant move in the reference bond. That is, the market believes that the default event is ‘true,’ and not just an error in payment. There is thus a legal test and a market test for the default event.

41.4.3 Credit Default Swap

- In the **credit default swap** counterparty A pays interest to counterparty B for a prescribed time until default of the underlying instrument. In the event of default B pays A the principal.

This is the simplest example of a credit derivative and can be thought of as insurance on the underlying instrument, bought by A from B (see Figure 41.5).

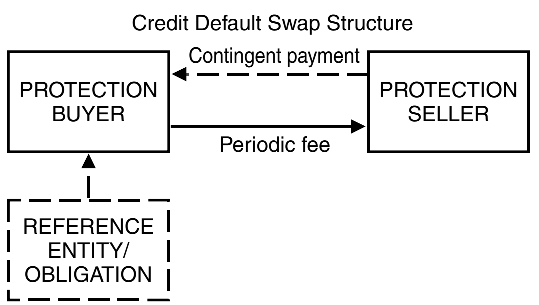


Figure 41.5 The Credit Default Swap.

Example 1

Vanilla credit default swap on a coupon-bearing bond. If the issuer of the bond misses a coupon payment then the buyer of the protection receives from the seller the specified amount.

Example 2

Very similar to the above, counterparty A pays LIBOR plus a fixed premium until default of the underlying instrument, and B pays A LIBOR for the whole life of the underlying instrument. In this case there is no exchange of principal. But again A is simply buying insurance against default.

Details of the CDS

The CDS is the dominant credit derivative in the structured credit market. The premium is usually paid periodically (quoted in basis points per notional). Premium can be an up-front payment, for short-term protection.

On the credit event, settlement may be the delivery of the reference asset in exchange for the contingent payment or settlement may be in cash (that is, value of the instrument before default less value after, recovery value.)

The mark-to-market value of the CDS depends on changes in credit spreads. So they can be used to get exposure to or hedge against changes in credit spreads.

If a large company/country defaults (or looks like defaulting) then credit default swaps become popular for all companies (same sector)/countries. Therefore there is a large amount of correlation between instruments.

Uses

The CDS is used to insure against overexposure to some credit risk. Can be used to get exposure to or hedge against maturities that are otherwise not available in the market. The source of the credit risk does not even need to know of the existence of derivatives on it.

As credit spreads fall it becomes harder to make money. The risk of default may not have changed but perceived risk or supply and demand affects the spread.

Factors affecting prices

- Probability of default of reference entity/credit rating
- Maturity of trade
- Expected recovery value
- Correlation between reference entity and swap counterparty[1]

41.4.4 Limited Recourse Note

The **limited recourse note** has a lesser exposure to default than the above contracts. Typically, there is an exposure to two underlying instruments. Interest is paid at a rate of r_1 while neither of the instruments has defaulted. When/if the first instrument is defaulted the interest is reduced to r_2, and reduced again to r_3 on the default of the second instrument.

41.4.5 First to Default

These are the very simplest of basket credit derivatives. The basket usually consists of 10 or so names:

- Structure is similar to a CDS;
- But with several reference entities;
- First entity to default triggers the payoff;
- Settlement is the same as ordinary CDS.

[1] For example, don't buy protection against default of one Japanese bank from another Japanese bank. (How do we know if there is a correlation? We won't know about exposures of the swap counterparty since these are off balance sheet items.) For pricing you need to know the joint probability of default.

41.4.6 *n*th to Default

Similar to the above, but here the contract gives the purchaser protection against default of the nth out of these names.

The holder pays an ongoing premium, and perhaps an upfront premium as well. This premium continues until maturity of the contract or the default of the nth company in the basket. Upon this default, of the nth name, the holder receives the loss suffered by that defaulting entity.

41.5 DERIVATIVES OF THE YIELD SPREAD

Other derivatives do not require default for there to be a payoff. These can be thought of as derivatives of the spread in the risky yield above the yield for an equivalent risk-free contract with the same, but guaranteed, cashflows.

- Default calls and puts
- Exchange options
- Credit spread options

41.5.1 Default Calls and Puts

Options on the price of a credit product, such as a bond or a loan.

- **Default calls** and **puts** are respectively the rights to buy and sell the underlying instrument.

Clearly much of the Black–Scholes theory is relevant here, but crucially the risk of default must be included in some way.

41.5.2 Exchange Options

Similar to the basic default options are the options giving you the right to exchange one bond for another. For example, the option may give you the right to give back the risky bond in exchange for a smaller quantity of the riskless bond. The ratio of riskless to risky bonds would be set at the start of the contract.

In the following * means the risky bond.

Payoff for this exchange option takes the form

$$\max(qZ - Z^*, 0),$$

where q is the prearranged ratio.

41.5.3 Credit Spread Options

Calculate the yield to maturity of a risky bond, call this Y^*, and calculate the yield to maturity of the equivalent riskless bond, call this Y. The difference between these two yields, $Y^* - Y$, is the spread.

- A contract having a payoff that depends on this spread would be a credit spread option.

The payoff could be the spread between two risky bonds, perhaps from the same issuer and perhaps not, in which case the spread could change sign. Typically the payoff would then take the form

$$\max(Y_1^* - Y_2^*, 0).$$

41.6 PAYMENT ON CHANGE OF RATING

The final kind of default derivative that I want to discuss are the derivatives that have a dependence on the rating of the issuer. The issuer begins with a certain rating and during the life of the contract the rating changes, possibly triggering some payment to the holder of the derivative.

Example: Korea Development Bank Note

In Figure 41.7 are the contract specifications for a puttable floating rate note issued by the Korea Development Bank in June 1997. This is a US dollar-denominated note that pays US three-month LIBOR plus 18.75 basis points every quarter. The contract was to mature in 2002. The 18.75 basis points excess over LIBOR is compensation for the risk of default. It can be sold back to the Korea Development Bank at par, i.e. for the amount of the principal, should the rating of the bank fall below A− (S&P) or A3 (Moody's).

What happened next?

Moody's rating was cut from A3 to Baa2 on 10th December 1997, as can be seen in the output from Bloomberg's on 11th December 1997 (Figure 41.8). The S&P rating was also dropped from A− to BBB−. This triggered the put feature, and indeed most of the bonds were sold back at par, at a cost to the Korea Development Bank of $300m. Although this put even more pressure on South Korea at a time when it was facing a dramatic liquidity crisis, the bank did not default. Later the KDB ratings fell even further, Korea was demoted to **junk bond** status, BB/Ba or lower.

There are many other types of derivatives that depend on the rating of the issuer. The payoff, or new feature, as in the above example, might be triggered as soon as the rerating occurs. Alternatively, the payment may depend on the rating at a specific time, so that a downgrade followed by an upgrade to the initial level will result in no action. One can imagine subtle path dependency in the contract or a contract that pays off if there is a downgrade of two notches, without the intermediate notch being realized. One example of this is the contract that pays off in the event of default, without there having been any prior rerating.

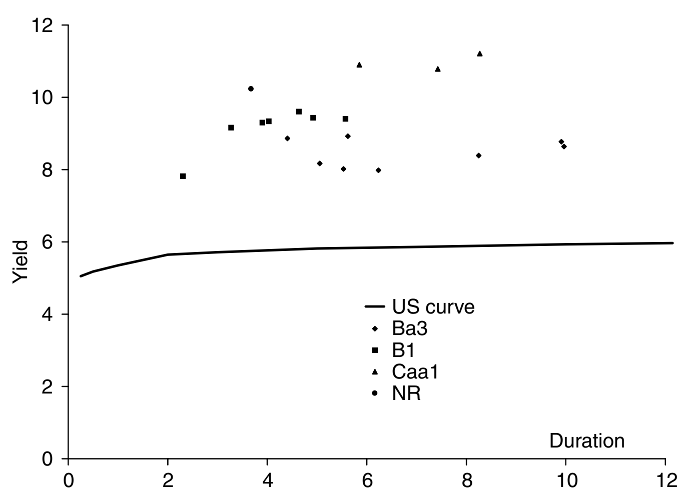


Figure 41.6 Yield versus duration for some risky bonds.

KOREA DEVELOPMENT BANK NOTE

CPN FREQ	QUARTERLY
CPN TYPE	FLOATING
MTY	CALL/PUT
DAY COUNT	ACT/360
COUNTRY / CURR	SK / USD

CPN RATE=3MO US$LIBOR + 18.75BP. PUT W/10 BUS DAYS NOTICE IF S&P RATING FALLS BELOW A– OR MOODY'S RATING FALLS BELOW A3.

RATINGS:
MOODY Baa2
S & P A–

Figure 41.7 Korea Development Bank Note.

```
DES                                                                DL17 CORP  DES
Enter 10 <Go> for News, 11 <Go> for Involved Parties
                S E C U R I T Y   D I S P L A Y                     PAGE 1/ 6
KOREA DEV BANK   KDB Float 06/02  N O T  P R I C E D

                                                   ISSUER INFORMATION
SECURITY INFORMATION                SERIES: EMTN   NAME  KOREA DEVELOPMENT BANK
                                    NOT SEASONED   TYPE  BANK
CPN FREQ         QUARTLY             IDENTIFICATION #'s      REDEMPTION INFO
CPN TYPE         FLOATING            ISIN    XS0077584307   MATURITY DT      6/30/ 2
MTY/REFUND TYP CALL/PUT              MLNUM FDLD9            REFUNDING DT
CALC TYP ( 21)FLOAT RATE NOTE                               NEXT CALL DT     6/30/ 0
DAY COUNT ( 2)ACT/360                COMMON   007758430     WORKOUT DT       6/30/ 2
MARKET ISS       EURO MTN                                   RISK FACTOR
COUNTRY/CURR     S.K./USD              ISSUANCE INFO
COLLATERAL TYP   SR UNSUB            ANNOUNCE DT     6/16/97          RATINGS
AMT ISSUED           300,000(M)      1ST SETTLE DT 6/30/97     MOODY     Baa2/*-
AMT OUTSTAND         300,000(M)      1ST CPN DT      9/30/97   S & P     A-  /*-
MIN PC/INC     10,000/  10,000       INT ACCRUE DT 6/30/97     COMP      BBB1
PAR AMT              10,000.00       PRICE @ ISSUE 99.95       DCR       NA
LEADMGR/UWRTR  GS INTL                                         FITCH     NA
LISTED         LONDON                                          IBCA      NA

NOTES NO PROSPECTUS
CPN RATE=3MO US$LIBOR + 18.75BP. UNSEC'D. PUT W/10 BUS DAYS NOTICE IF S&P RATING
FALLS BELOW A- OR MOODY'S RATING FALLS BELOW A3.
      Bloomberg-all rights reserved  Frankfurt:69-920410  Hong Kong:2-521-3000  London:171-330-750
      New York:212-318-2000     Princeton:609-279-3000     Singapore:226-3000     Sydney:2-9777-8600
                                              Tokyo:3-3201-8900  Sao   Paulo:11-3048-4500
                                                           G194-79-13   12-Dec-97  8:53:41
```

Figure 41.8 Contract specifications for a puttable floating rate note issued by the Korea Development Bank. Source: Bloomberg L.P.

41.7 USING DEFAULT SWAPS IN CB ARBITRAGE

Convertible bonds are exposed to

- stock price
- interest rates
- credit risk

The first and third are the most important. Convertible bond arbitrage involves exploiting a volatility forecast and/or credit view. Exploiting the volatility view can be done via delta hedging, or hedging with exchange-traded options (see Figure 41.9).

41.7.1 Exploiting your Credit Risk View

Step 1: Stock risk

Eliminate exposure to stock-specific risk by delta hedging.

Step 2: Decisions

Are you intending to exploit absolute or relative risk of default, that is default versus change in market's view of credit risk?

- Absolute versus relative: Is your view on the market's estimate of the credit risk of the company or on its relative risk with respect to another (similar) company?
- Are you expecting the company to default or for the market's estimate of risk to change (credit spreads to change)?

Step 3: Buy/sell protection

The answer to Step 2 will determine which side to be on and what kind(s) of contract to buy/sell.

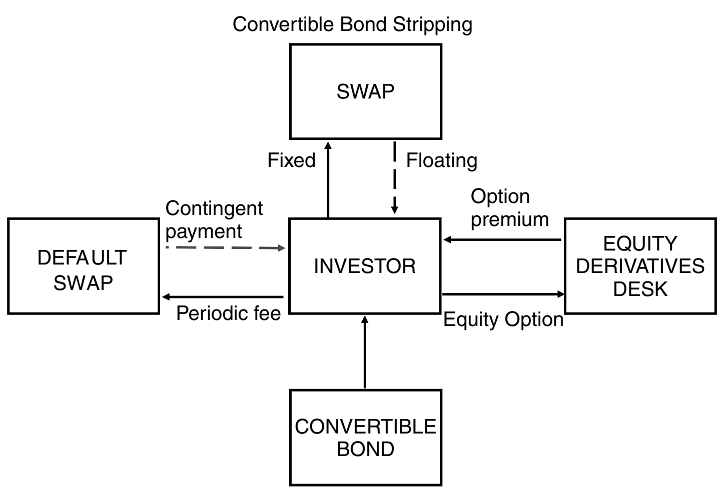


Figure 41.9 Use of credit instruments in CB arbitrage.

41.8 TERM SHEETS

Several of the following term sheets and many more can be found in the excellent book by Nelken (Nelken, 1999).

41.8.1 Put on Credit Spread on XYZ Bond

Spread Put Buyer	DB
Spread Put Seller	Investor
Notional Principal	\$20MM
Exercise Date	One year
Underlying Index	XYZ bonds maturing dd/mm/yyyy
Reference US	Offer yield on US Treas. 5.25% due d′d′/m′m′/yyyy
Index Credit Spread	YTM on underlying (Bid) minus Reference US 12:00pm EST two days before Exercise Date
Current Spread	1.85%
Spread Put Strike	2.1%
Put Payment	Notional Principal * max (DUR * (Index Credit Spread −Spread Put Strike), 0)
DURation	8
Premium	1.2% of Notional Principal

Notes

1. Bank buys an option on the credit spread between XYZ bond and US Treasury
2. Maturities are closely matched
3. Appears to be a call . . . but the XYZ bond and spread are inversely related
4. DUR relates change in a rate to change in a value: It is an approximation
5. Sheet specifies use of bid/offer prices

41.8.2 Binary Payoff Bond

Issuer	AAA-rated SPV
Underlying Asset	XYZ 7.25% dd/mm/yyyy
Notional Principal	\$10MM
Maturity Date	d′d′/m′m′/yyyy
Price	Par
Coupon	9%
Coupon Payment Dates	. . .
Principal Redemption	Par, unless provision is activated
Default Provision	If at any time during the life of the note XYZ fails to make a payment due on the Underlying Asset interest payments will cease and the note will be redeemed at zero.

Notes

1. Suitable for investors who hold the bond to maturity and believe it will not default
2. Differs from CDS in that there is no recovery value

41.8.3 Digital Spread Option, One-year Note Linked to Venezuelan Par Bond

Issuer	AA or better
Principal Amount	\$75MM
Settlement	Two weeks from today
Maturity Date	One year from settlement date
Price	Par
Underlying Index	Venezuelan Brady Bonds due June 2023
Reference US Treasury	US Treas. 7.125% due Feb. 2023
Index Credit Spread	The yield to maturity of the underlying index using the flat bid price (net of accrued interest) minus the offer yield on the reference US Treasury at 12:00pm NY EST two days before exercise date
Coupon	0%
Principal Redemption Value	If index credit spread $\geq$ 4% and $\leq$ 6% then 111% otherwise 102% of par
Current Index Spread	Par

Notes

1. The word ‘digital’ usually refers to a discontinuous payoff

41.8.4 Basket Credit-Linked Note

Issuer	AAA-rated SPV
Principal	\$25MM
Maturity Date	dd/mm/yyyy
Price	Par
Coupon	Six-month LIBOR + 1.5%
Coupon Payment Dates	...(Every six months)
Principal Redemption	Par, unless provision is activated
Default Provision	If, at any time during the life of the note one of the bonds in the basket defaults on principal or interest, the investor will receive 80% of the face value of the bond or any other bond in the basket. The issuer will have no further obligations.
Basket	ABC DEF GHI ...

Notes

1. Investor receives something in the event of default yet still receives the higher coupon
2. On default the investor gets 80% of the defaulted bond, of course, which has a market value that is much less than the face value
3. What is the role of correlation and diversification? The investor wants correlated assets. This is an easy calculation if bonds are uncorrelated: What is probability of none defaulting? This is an easy calculation if bonds are perfectly correlated: If one defaults they all default

41.8.5 Asset Swap Put Option on One Year ABC 6.65% Bonds 10/2006 Asset Swap

Swap details:	
Underlying Issue	ABC 6.65% Bonds 10/2006
Notional Principal Amount	$20MM
Asset Swap Settlement	One year from today
Asset Swap Maturity Date	Bond Maturity Date
Asset Swap Price	$100
Asset Swap Coupon	3 month LIBOR + 0.6% quarterly, Act/360
Current Asset Swap Level	3 month LIBOR + 0.55%
Credit Spread Option Details:	
Put Option Buyer	Bank
Put Option Seller	Investor
Put Credit Spread Option	The bank has the right but not the obligation to sell the investor the assets swap on the determination date
Option Payment by the Bank	0.4% of the notional principal amount
Determination Date	Five days prior to the contingent asset swap settlement date
Cash Settlement Provision	The investor may choose to cash settle on the determination date at the present value of the difference between the asset swap coupon and the then current trading level of the underlying asset.

Notes

1. Bank pays $0.4\% \times \$20\text{MM} = \$80{,}000$
2. In one year the bank can exercise the option to enter into an asset swap. If the bank exercises its option then

- the bank sells the asset to the investor for par
- the investor passes the 6.65% coupon to the bank
- the bank pays to the investor LIBOR + 0.6%

41.9 PRICING CREDIT DERIVATIVES

We are going to examine two credit derivatives in detail to see possible approaches to the modeling. The first example is of an option to exchange a risky bond for a quantity of the equivalent riskless bond. To model this contract we will use a stochastic hazard rate model. The second example is a contract that pays off in the event of a rerating; for this we must obviously use a model that explicitly captures the possible change of rating.

41.10 AN EXCHANGE OPTION

An option to exchange a risky zero-coupon bond for a riskless zero-coupon bond at time T has a payoff

$$\max(qZ - Z^*, 0),$$

for some fixed q.

There are various levels of sophistication on which we can address the pricing of this bond.

The first level is to assume that risk-free rates and hazard rate are deterministic. We then arrive at a completely deterministic price for both the risk-free and the risky bonds. This is unsatisfactory because randomness is very important in the pricing of any non-linear payoff such as the one we have here.

The second level of sophistication is to assume that one, but not both, of the interest rate and the risk of default is stochastic. For example, assume that interest rates are random but that the risk of default is constant. This is a common approach, resulting often in the simple addition of a fixed spread to yields as seen in Chapter 40. Let's see how this approach works with the exchange option. (Warning: This approach is not going to give us any useful results, but the analysis is none the less instructive.)

The first stage is to price the risk-free bond. If this matures at time T_B when it receives the principal of \$1, then it will satisfy

$$\frac{\partial Z}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z}{\partial r^2} + (u - \lambda w)\frac{\partial Z}{\partial r} - rZ = 0$$

with

$$Z(r, T_B) = 1$$

and where $u - \lambda w$ and w are the risk-adjusted drift and the volatility of the spot rate respectively.

The risky bond will similarly satisfy

$$\frac{\partial Z^*}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 Z^*}{\partial r^2} + (u - \lambda w)\frac{\partial Z^*}{\partial r} - (r + p)Z^* = 0$$

with

$$Z^*(r, T_B) = 1.$$

The solution of this equation is simply

$$Z^*(r, t) = e^{-p(T_B - t)}Z(r, t).$$

This completely deterministic relationship between the two bonds, which is a result of the assumption of constant hazard rate, clearly scuppers the pricing of the exchange option. The subtlety in the pricing of this contract is due to the *randomness in the risk of default*. Usually, as in this example, the assumption of constant hazard rate is not appropriate for credit derivatives.

Still on the second level of complexity, a better assumption would be that interest rates are given by the forward rates and that the hazard rate, p, satisfies some stochastic differential equation. This approach should make more sense than the above for our contract.

Assuming that

$$dp = \gamma\, dt + \delta\, dX$$

and, for simplicity that interest rates are constant, we have

$$Z = e^{-r(T_B - t)}$$

and

$$\frac{\partial Z^*}{\partial t} + \tfrac{1}{2}\delta^2 \frac{\partial^2 Z^*}{\partial p^2} + \gamma \frac{\partial Z^*}{\partial p} - (r+p)Z^* = 0 \qquad (41.1)$$

with

$$Z^*(p, T_B) = 1.$$

This problem for $Z^*(p, t)$ is mathematically identical to the earlier problem for $Z^*(r, t)$ when rates were stochastic but probability constant. Again we see the similarity between risk of default and the rate of interest. But now the payoff for our exchange option, having value $V(p, t)$, is

$$\max\left(qe^{-r(T_B - T)} - Z^*(p, T), 0\right).$$

The first term inside the parentheses is a constant and so this problem looks exactly like a put option on a zero-coupon bond, as discussed in Chapter 32. We could even go so far as to take the factor $e^{-r(T_B - t)}$ out from Z^* so losing it completely from Equation (41.1). The only difference then would be that you see a p instead of the usual r in the partial differential equation. Of course, there still remains the choice of functions γ and δ, but these are often chosen in practice so that we can find explicit solutions; again we are in territory familiar from Chapter 30.

The next level of sophistication is to have both stochastic interest rates and stochastic hazard rate. Both Z and Z^* satisfy the equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2 \frac{\partial^2 V}{\partial r^2} + \rho w \delta \frac{\partial^2 V}{\partial r\, \partial p} + \tfrac{1}{2}\delta^2 \frac{\partial^2 V}{\partial p^2} + (u - \lambda w)\frac{\partial V}{\partial r} + \gamma \frac{\partial V}{\partial p} - (r+p)V = 0. \qquad (41.2)$$

However, the risk-free bond is independent of the risk of default so that we have $Z(r, t)$, with no p dependence. The risky bond does depend on the risk of default and is therefore a function of three variables, $Z^*(r, p, t)$.

We must first solve for the underlying bonds using

$$Z(r, T_B) = Z^*(r, p, T_B) = 1.$$

And then solve for the exchange option $V(r, p, t)$ which again satisfies (41.2) with

$$V(r, p, T) = \max(qZ(r, T) - Z^*(r, p, T), 0).$$

Because this exchange option is a second-order contract the price may be quite sensitive to the model.

41.11 DEFAULT ONLY WHEN PAYMENT IS DUE

We are now going to look briefly at what happens if the company only defaults when it needs to. The model follows from the key observation that the concept of default prior to a coupon payment is completely meaningless. What company is going to announce default before it becomes strictly necessary? After all, perhaps the company will recover in the meantime. So, any model that talks about hazard rates might not capture what is really going on.[2]

41.11.1 The Market's Estimate of Default Risk

Let's start with a defaultable zero-coupon bond, market price Z^*, in a world of constant interest rate r. We can relate the market price of the risky bond to the market's estimate of the risk of default (in a sense) by

$$p = \frac{\log\left(\dfrac{Z}{Z^*}\right)}{T-t}.$$

The notation is the usual. I say 'in a sense' because this comes from a hazard rate model, which I have just said may not capture reality. Never mind, here I use p as a proxy for the market price of the bond; it just happens to be easier to have this as a variable than Z^*.

We won't worry for the moment whether or not this estimate has any relationship to reality. Anyway, we can at least measure p and estimate its stochastic process. By whatever means, we find

$$dp = \mu\, dt + \sigma\, dX.$$

The coefficients can be p and t dependent if we wish.

41.11.2 Hedging

Now let's hedge a credit derivative with the underlying bond:

$$\Pi = V - \Delta Z^*.$$

Can we do this? We certainly can because Z^* is traded and no default is allowed to happen prior to the bond's maturity. In this model perfect hedging is always possible.

The change in the portfolio value is ... eliminating risk gives us

$$\Delta = -\frac{\dfrac{\partial V}{\partial p}}{(T-t)e^{-(r+p)(T-t)}}$$

... setting return equal to the risk free rate ... and the answer is[3]

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2 V}{\partial p^2} + \left(\frac{p}{T-t} + \tfrac{1}{2}\sigma^2(T-t)\right)\frac{\partial V}{\partial p} - rV = 0. \tag{41.3}$$

We just need boundary conditions, final conditions and we're there.

[2] Ok, maybe that's going a bit far!
[3] The '...' means skipping the usual stuff.

41.12 PAYOFF ON CHANGE OF RATING

More subtle than a simple payoff in default is payoff on change of rating, for example, the Korea Development Bank note described above. I will describe the pricing of two distinct styles of such a contract. In the first example there will be a payment if the rating takes certain values at expiry and in the second there will be a payoff if certain ratings are realized at any time before expiry.

Suppose that an issuer is currently rated AAA and a contract specifies that a fixed amount will be paid to the owner of the contract if the issuer is rated only AA on a certain date. Clearly, to price this contract we need a model that explicitly allows for rating migration. Let us take the Markov chain model of Chapter 40 and assume that interest rates are constant. The equation to be solved is

$$\frac{d\mathbf{V}}{dt} + (\mathbf{Q} - r\mathbf{I})\mathbf{V} = 0.$$

The contract specification of payment if the grade is AA must be incorporated into the final condition. Since there is no payment unless the issuer is rated AA the final condition is simply

$$\mathbf{V}(T) = \mathbf{e}_{\mathrm{AA}},$$

where $\mathbf{e}_{\mathrm{AA}}$ is the column vector with a zero in all the rows, except for the row corresponding to the rating AA where there is a 1.

The contract that is triggered by a downgrade to AA at *any time* is more interesting, and not much harder to price. It might help if you think of this contract as being like an ‘in’ barrier option. In this option the payment is triggered by the underlying hitting a given level. We have a similar situation with the present credit derivative, with the level of the credit rating playing the role of the underlying.

Again we must solve

$$\frac{d\mathbf{V}}{dt} + (\mathbf{Q} - r\mathbf{I})\mathbf{V} = 0.$$

with the final condition

$$\mathbf{V}(T) = \mathbf{e}_{\mathrm{AA}}.$$

But now we have an extra condition, corresponding to the barrier boundary condition in the knock-in:

$$\mathbf{V}_{\mathrm{AA}} = 1 \quad \text{for all} \quad t < T,$$

where $\mathbf{V}_{\mathrm{AA}}$ is the entry in the vector $\mathbf{V}$ corresponding to the AA rating. In other words, the minute that the level AA is reached we receive a payment of 1. In such a contract it is common to limit the times when the trigger is active. In such a contract the condition on $\mathbf{V}_{\mathrm{AA}}$ is only switched on when the trigger is active.

Let’s take a look at the Korea Development Bank note again. This is very similar to the last problem above. The difference is that instead of getting a fixed amount on a downgrade we are allowed to put the note back to the issuer. This results in the constraint

$$\mathbf{V}_i \geq 1 \quad \text{for all} \quad i \leq i^*,$$

where i^* is the index for the grade triggering the put option. (When the interest rate is non-diffusive the continuity of derivatives of the option price, as seen in American options, is no longer necessary.)

Finally, the **rating protected note** pays regular coupons whose amount depends on the then credit rating. Typically, the lower the rating the higher the coupon. This is priced via a jump condition of the form

$$\mathbf{V}(t_i^-) = \mathbf{V}(t_i^+) + \mathbf{c}.$$

The entries in $\mathbf{c}$ are the coupons to be paid for each rating level.

41.13 MULTI-FACTOR DERIVATIVES

Figure 41.10 shows details of a bond that becomes callable on a credit event. Payoff at maturity is linked to the growth in an index, the JP Morgan EMBI+. To price this we must model the likelihood of a credit event occurring to the issuing company and also the behavior of the index.

DL18 Corp **ODE5**

Enter 11 <GO> for Involved Parties
Enter 1 <GO> To See Other Available Info

STRUCTURED NOTE DESCRIPTION Page 1 of 3

BETA FINANCE CRP BETA 0 ½ 05/01 N O T P R I C E D

SERIES EMTN

SECURITY INFORMATION		IDENTIFICATION #'s		ISSUER INFORMATION	
COUPON FREQ	ANNUAL	ISIN	XS0086414470	BETA FINANCE CORP LTD	
COUPON TYPE	FIXED	MLNUM	99999	FINANCIAL	
MTY/REFUND TYP	CALLABLE			**REDEMPTION INFO**	
CALC TYP(751)	ACCRUED ONLY	COMMON	008641447	MATURITY	05/08/01
DAY COUNT(20)	ISMA-30/360			NEXT CALL	
COUNTRY/CURR	KY /USD			WORKOUT	05/08/01
AMT ISSUED(M)	35,000			**RATINGS INFO**	
AMT OUTSTAND(M)	35,000	**ISSUANCE INFO**		MOODY	Aaa
MIN PC/INC	1,000,000/ 1,000,000	ANNOUNCE	05/06/98	S & P	AAA
PAR AMOUNT	1,000,000.00	1st SETTLE	05/08/98	COMP	AAA
LEADMGR/UWRTR	JPM-sole	1st COUPON	05/08/99	DCR	NA
EXCHANGE	LONDON	INT ACCRUAL	05/08/98	FI	NA
MARKET ISSUE	EURO MTN	PRICE @ ISS	100.0000		
COLLATERAL TYP	SECURED				

NOTES: HAVE PROSPECTUS, TAX CHANGES CALL
RDMPTN LINKED TO THE JP MORGAN EMBI+ COMPOSITE NON-LATIN INDEX. SEC'D BY ASSETS.
CALL IF CREDIT EVENT ON UNDERLYING OCCURS. SERIES 615.

Bloomberg

Figure 41.10 Payoff depends on an index. Source: Bloomberg L.P.

41.14 COPULAS: PRICING CREDIT DERIVATIVES WITH MANY UNDERLYINGS

Although mentioned briefly above, credit derivatives with many underlyings have become very popular of late. As a natural skeptic when it comes to modeling in finance generally, I have to say that some of these instruments and models being used for these instruments fill me with some nervousness and concern for the future of the global financial markets. Not to mention mankind, as well. Never mind, it's probably just me. We'll see the most important of these ideas and instruments now.

41.14.1 The Copula Function

The technique now most often used for pricing credit derivatives when there are many underlyings is that of the **copula**. The copula[4] function is a way of simplifying the default dependence structure between many underlyings in a relatively transparent manner. The clever trick is to separate the distribution for default for each individual name from the dependence structure between those names. So you can rather easily analyze names one at a time, for calibration purposes, for example, and then bring them all together in a multivariate distribution. Mathematically the copula way of representing the dependence (one marginal distribution per underlying, and a dependence structure) is no different from specifying a multivariate density function. But it can simplify the analysis.

The copula approach in effect allows us to readily go from a single-default world to a multiple-default world almost seamlessly. And by choosing the nature of the dependence, the copula function, we can explore models with richer 'correlations' than we have seen so far in the multivariate Gaussian world. For example, having a higher degree of dependence during big market moves is quite straightforward.

41.14.2 The Mathematical Definition

Take N uniformly distributed random variables $U_1, U_2, \dots, U_N$, each defined on [0, 1]. The copula function (see Li, 2000) is defined as

$$C(u_1, u_2, \dots, u_N) = \text{Prob}(U_1 \le u_1, U_2 \le u_2, \dots, U_N \le u_N).$$

Clearly we have

$$C(u_1, u_2, \dots, 0, \dots, u_N) = 0,$$

and

$$C(1, 1, \dots, u_i, \dots, 1) = u_i.$$

So that's the copula function. The way it links many univariate distributions with a single multivariate distribution is as follows.

Let $x_1, x_2, \dots, x_N$ be random variables with cumulative distribution functions (so-called **marginal** distributions) of $F_1(x_1), F_2(x_2), \dots, F_N(x_N)$. Combine the Fs with the copula function,

$$C(F_1(x_1), F_2(x_2), \dots, F_N(x_N)) = F(x_1, x_2, \dots, x_N)$$

[4] From the Latin for 'join.'

and it's easy to show that this function $F(x_1, x_2, \dots, x_N)$ is the same as

$$\text{Prob}(X_1 \le x_1, X_2 \le x_2, \dots, X_N \le x_N).$$

A key reference in this subject is Sklar (1959) who showed that any multivariate distribution can be written in terms of a copula function, i.e. the converse of the above.

In pricing basket credit derivatives we would use the copula approach by simulating default times of each of the constituent names in the basket. Then we would perform many such simulations in order to be able to analyze the statistics, the mean, standard deviation, distribution, etc., of the present value of resulting cashflows.

41.14.3 Examples of Copulas

Here are some examples of bivariate copula functions. They are readily extended to the multivariate case.

Bivariate normal

$$C(u, v) = N_2\left(N_1^{-1}(u), N_1^{-1}(v), \rho\right), \quad -1 \le \rho \le 1,$$

where N_2 is the bivariate Normal cumulative distribution function, and N_1^{-1} is the inverse of the univariate Normal cumulative distribution function.

Frank

$$C(u, v) = \frac{1}{\alpha} \log\left(1 + \frac{(e^{\alpha u} - 1)(e^{\alpha v} - 1)}{e^{\alpha} - 1}\right), \quad -\infty < \alpha < \infty.$$

Fréchet–Hoeffding upper bound

$$C(u, v) = \min(u, v).$$

Gumbel–Hougaard

$$C(u, v) = \exp\left(-\left((-\log u)^{\theta} + (-\log v)^{\theta}\right)^{1/\theta}\right), \quad 1 \le \theta < \infty.$$

This copula is good for representing extreme value distributions.

Product

$$C(u, v) = uv$$

Other copulas include Archimedean, Clayton and Student. See Cherubini, Luciano and Vecchiato (2004) for a comprehensive list.

One of the simple properties to examine with each of these copulas, and which may help you decide which is best for your purposes, is the **tail index**. Examine

$$\lambda(u) = \frac{C(u, u)}{u}.$$

This is the probability that an event with probability less than u occurs in the first variable given that at the same time an event with probability less than u occurs in the second variable. Now look at the limit of this as $u \to 0$,

$$\lambda_L = \lim_{u \to 0} \frac{C(u, u)}{u}.$$

This tail index tells us about the probability of both extreme events happening together.

41.15 COLLATERALIZED DEBT OBLIGATIONS

Whereas most of the credit instruments described above have a single underlying name, the **collateralized debt obligation** or **CDO** is an instrument designed to give protection against losses in a portfolio, and usually a portfolio containing 100s of individual companies. As with other credit derivatives there is the protection buyer and the seller.

The CDO gives protection, not against an entire portfolio, but against bits of it, so called **tranches**. First define the aggregate loss in the portfolio as the sum of all losses due to default. As more and more companies default so the aggregate loss will increase. Now specify hurdles, as percentages of notional, and these define the tranches. For example, there may be the 0–3% tranche, and the 3–7% tranche etc. As the aggregate loss increases past each of the 3%, 7%, etc. hurdles so the owner of that tranche will begin to receive compensation, at the same rate as the losses are piling up. You will only be compensated once your **attachment point** has been reached, and until the **detachment point**.

So there are two types of cashflows in this instrument:

- Premiums: The protection buyer will pay periodic premiums to the seller, quoted in basis points. There may also be an upfront premium.
- 'Compensation:' The holder of the instrument, the protection buyer, receives the loss suffered by his tranche.

To clarify, losses are assigned initially to the first tranche until the threshold is attained, and then to the second tranche until its detachment point is reached, and so on. An example is given in Table 41.1.

Table 41.1 A typical CDO and its tranches. Source: St Pierre (2004)

Tranche	Upfront premium	Ongoing premium (bps)
0–3%	42%	500
3–7%	0%	331
7–10%	0%	126
10–15%	0%	54
15–30%	0%	16

To price these instruments you obviously have to consider the probability of each name defaulting and the correlation structure between them. A model often used in practice is the Gaussian copula approach. So far so good. To make the problem tractable, given the large number of underlyings and parameters, it is also often assumed that the structure of the correlation is described by a single random factor. Each has its own stock-specific source of randomness but there is a factor common to all names. Then, to simplify matters further, it is also often assumed that there is a single number to represent all of these correlations; that is, there is just one single correlation parameter. Probably (at least) one simplification too far. One of the reasons for such assumptions is the usual; that it leads to simple pricing formulae.

Working within the framework of a single parameter, how is this parameter used? Suppose we want to price a new CDO, and we have the price of a traded CDO that is not dissimilar. We can back out the **implied** or **compound correlation** from the CDO with the known price and plug it into the model for the other CDO. This is, of course, what is often done in the derivatives world, equity, FX or fixed income, with volatility. It's just another form of calibration.

There is a major problem with this. As happens in the derivatives world, non-credit, we find that we need a different implied correlation for each tranche. Typically, you might find that the implied correlation falls from first tranche to second, and then rises for third and subsequent tranches (Finger, 2004). This is akin to the implied volatility smile and skews that we have seen before. This suggests that the market is somehow allowing for things not within the one-size-fits-all correlation assumption. Slightly concerning, but not too much. No, the real worry is that you may find that there are *two* possible implied correlations, or no implied correlation at all. This means that a market price is either consistent with two possible constant correlations, or there is no constant correlation which matches theoretical and market prices. Generally speaking, if we were to plot value of a tranche versus implied correlation then we might find a non-monotonic curve. Why is this a worry? Think back to examples we've seen with implied volatility, and in particular, the up-and-out call option that we analyzed in some detail (see Chapters 7 and 23). We had the same problem there. Before I expand on this in one paragraph's time, let's see one final way of pricing CDOs.

Using the example in the table, you can think of the 3–7% tranche as long a 0–7% tranche and short the 0–3% tranche. So you can value a non-existent 0–7% tranche very easily. Back out from market data for the 0–3% tranche the implied correlation for that first tranche. Now back out an implied correlation for the 0–7% tranche that you have created. This is called the **base correlation**. Now use these base correlations to price other CDOs with tranches 0–$x\%$, and then construct the $y\%$–$x\%$ tranches accordingly.

Now back to the problem with all of this. Do you remember what the non uniqueness in the implied volatility for the up-and-out call meant? It was because the price versus (constant) volatility graph was non monotonic. It rose to a peak and then fell. If we have that behavior with a CDO then alarms bells should start ringing. In the knockout example the non monotonicity was because there were two competing effects going on. If we were out of the money we wanted high volatility so we moved into the money and hence got a nice payoff. But once in the money we wanted low volatility so that we didn't knock out and lose the option. So sometimes volatility was good for us and sometimes bad. In Chapter 52 we see how to relate this to the sign of gamma, and how to examine more scenarios than just a simple change in constant volatility. We find that really bad (and really good) scenarios are related to a richer volatility structure than simple parallel shifts, for example a skew can have a very, very big effect on a barrier option's value. Now think bigger than barrier options: Two competing effects, one good, one bad. There must be something similar going on with the CDO. Sometimes correlation is good, sometimes bad. What if the good thing didn't kick in and we were only left with the bad? That's what we look at when we examine the barrier option; we say what if the good volatility, out of the money, didn't appear, but volatility only rose in the money where it hurt us. Net result, much, much lower barrier option prices than if volatility were to rise uniformly everywhere. Exactly the same thing is happening with CDOs. What if the good correlation didn't appear, what if only the bad correlation rose? Ouch. Prices could be much, much lower than you would expect.

With so many potential variables and parameters because of the 100s of underlyings it is going to be much harder to examine more interesting scenarios. But a step in the right direction would be to at least have two sources of randomness leading to correlations between the underlying names. Or to vary correlation with burn through (aggregate loss).

41.16 SUMMARY

The current state of default-risk modeling is, in my opinion, far from satisfactory. The problems associated with modeling and parameter estimation are enormous. In later chapters we will see how to get round uncertainty in parameters in the options world. These elegant methods, such as worst-case scenario analysis, are unlikely to be of much use here because of the extreme nature of the worst case. A possible direction for further work would be to distinguish between one's own estimation and pricing of default risk and the market's.

It is very difficult to assess the risk of default, by either analysis of time series data or fundamental analysis. And the results of your effort may not even be relevant to the matter of pricing. Suppose, for example, that you buy a risky bond, intending to hold it for only one year during which time you think there is no risk of default. When you come to sell it, hopefully at a profit after including any cashflows, how much is it then worth? The answer to this depends on the market's subjective view of the likelihood and timing of default and not on any objective risk of default. On the other hand, if you buy a bond intending to hold it to maturity then the risk of default is all that matters; who cares what the market thinks? This example illustrates the importance of distinguishing between the market's 'model' and your own model of reality. Some of the consequences of this will be discussed in Chapter 59.

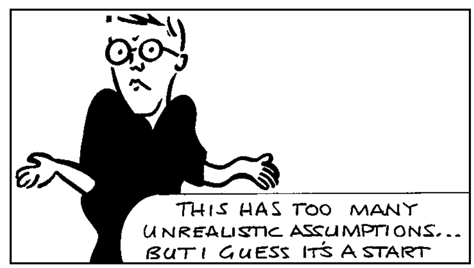
FURTHER READING

- See the papers by Das (1995), Schönbucher (1996, 1997a, b, 1998) and Schönbucher & Schlögl (1996) for the pricing and, importantly, the hedging of credit derivatives.
- The best book on the technical aspects of pricing and hedging credit derivatives is by Schönbucher (2003).
- See Penaud & Selfe (2003) for an analysis of first-to-default swaps.
- A very well written and comprehensive study of problems in pricing CDOs is by Finger (2004).
- Li (2000) is a very clear exposition of the copula-function approach.
- For a detailed, technical explanation of copulas see the books Nelsen (1999) and Cherubini, Luciano & Vecchiato (2004).

CHAPTER 42
RiskMetrics and CreditMetrics

In this Chapter...

- the methodology of RiskMetrics for measuring value at risk
- the methodology of CreditMetrics for measuring a portfolio's exposure to default events



42.1 INTRODUCTION

In Chapter 19 I described the concept of the 'Value at Risk' (VaR) of a portfolio. I repeat the definition of VaR here: VaR is 'an estimate, with a given degree of confidence, of how much one can lose from one's portfolio over a given time horizon.' In that chapter I showed ways of calculating VaR (and some of the pitfalls of such calculations). Typically the data required for the calculations are parameters for the 'underlyings' and measures of a portfolio's current exposure to these underlyings. The parameters include volatilities and correlations of the assets, and, for longer time horizons, drift rates. The exposure of the portfolio is measured by the deltas and, if necessary, the gammas (including cross derivatives) and the theta of the portfolio. The sensitivities of the portfolio are obviously best calculated by the owner of the portfolio, the bank. However, the asset parameters can be estimated by anyone with the right data. In October 1994 the American bank JP Morgan introduced the system **RiskMetrics** as a service for the estimation of VaR parameters.

JP Morgan has also proposed a similar approach, together with a data service, for the estimation of risks associated with risk of default: **CreditMetrics**. CreditMetrics has several aims; two of these are the creation of a benchmark for measuring credit risk and the increase in market liquidity. If the former aim is successful then it will become possible to measure risks systematically across instruments and, at the very least, to make relative value judgments. From this follows the second aim. Once instruments, and in particular the risks associated with them, are better understood they will become less frightening to investors, promoting liquidity.

42.2 THE RISKMETRICS DATASETS

The RiskMetrics datasets are extremely broad and comprehensive. They consist of three types of data: One used for estimating risk over the time horizon of one day, the second having a one-month time horizon and the third has been designed to satisfy the requirements in the proposals from the Bank for International Settlements on the use of internal models to estimate market risk. The datasets contain estimates of volatilities and correlations for almost 400 instruments, covering foreign exchange, bonds, swaps, commodities and equity indices. Term-structure information is available for many currencies.

42.3 CALCULATING THE PARAMETERS THE RISKMETRICS WAY

A detailed technical description of the method for estimating financial parameters can be found at the JP Morgan website. Here I only give a brief outline of major points.

42.3.1 Estimating Volatility

The volatility of an asset is measured as the annualized standard deviation of returns. There are many ways of taking this measurement. The simplest is to take data going back a set period, three months, say, calculate the return for each day (or over the typical timescale at which you will be rehedging) and calculate the sample standard deviation of these data. This will result in a time series of three-month volatility.[1] This approach gives equal weight to all of the observations over the previous three months. This estimate of volatility on day i is calculated as

$$\sigma_i = \sqrt{\frac{1}{\delta t(M-1)} \sum_{j=i-M+1}^{i} (R_j - \overline{R})^2},$$

where δt is the time step (typically one day), M is the number of days in the estimate (approximately 63 in three months), R_j is the return on day j and $\overline{R}$ is the average return over the previous M days. If δt is small then we can in practice neglect $\overline{R}$.

This measurement of volatility has two major drawbacks. First, it is not clear how many days' data we should use; what happened three months ago may not be relevant to today. But the more data we have the smaller will be the sampling error if the volatility really has not changed in that period. Second, a large positive or negative return on one day will be felt in this historical volatility for the next three months. At the end of this period the volatility will apparently drop suddenly, yet there will have been no underlying change in market conditions; the drop will be completely spurious. Thus, the volatility measured in this way will show 'plateauing.' A typical thirty-day volatility plot, with plateauing, is shown in Figure 42.1.

In RiskMetrics the volatility is measured as the square root of a variance that is an exponential moving average of the square of price returns. This ensures that any individual return has a

[1] Which may or may not bear any resemblance to three-month implied volatility.

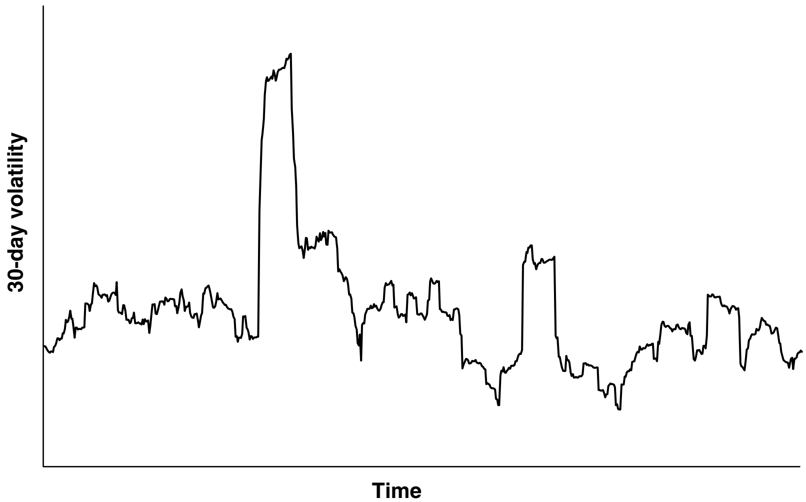


Figure 42.1 Thirty-day volatility.

gradually decreasing effect on the estimated volatility, and plateauing does not occur. This volatility is estimated according to

$$\sigma_i = \sqrt{\frac{1-\lambda}{\delta t} \sum_{j=-\infty}^{i} \lambda^{i-j} R_j^2},$$

where λ represents the weighting attached to the past volatility versus the present return (and we have neglected the mean of R, assuming that the time horizon is sufficiently small). This difference in weighting is more easily seen if we write the above as

$$\sigma_i^2 = \lambda \sigma_{i-1}^2 + (1-\lambda)\frac{R_i^2}{\delta t}.$$

The parameter λ has been chosen by JP Morgan as either 0.94 for an horizon of one day and 0.97 for an horizon of one month. Another possibility is to choose λ to minimize the difference between the squares of the historical volatility and an implied volatility.

The spreadsheet in Figure 42.2 shows how to calculate such an exponentially-weighted volatility. As can be seen from the plot, the exponentially weighted volatility is much 'better behaved' than the uniformly weighted version.

42.3.2 Correlation

The estimation of correlation is similar to that of volatility. To calculate the covariance σ_{12} between assets 1 and 2 we can take an equal weighting of returns from the two assets over the last M days:

$$\sigma_{12_i} = \sqrt{\frac{1}{\delta t(M-1)} \sum_{j=i-M+1}^{i} (R_{1_j} - \overline{R}_1)(R_{2_j} - \overline{R}_2)}.$$

	A	B	C	D	E	F	G	H
1	**Start volatility**	0.3		**Date**	**Stock**	**Returns**	**Vol^2**	**Volatility**
2	**Lambda**	0.97		1-Jan-85	218.32		0.09	0.3
3				2-Jan-85	217.16	−0.005313	0.087513	0.295827
4		=(E3−E2)/E2		3-Jan-85	215.24	−0.008841	0.085479	0.292368
5				4-Jan-85	215.24	0.000000	0.082915	0.287949
6		=B1*B1		7-Jan-85	217.16	0.008920	0.081029	0.284655
7				8-Jan-85	220.25	0.014229	0.080129	0.28307
8	= $B$2*G7 + (1−$B$2)*F8*F8*252			Jan-85	224.87	0.020976	0.081051	0.284695
9				10-Jan-85	224.87	0.000000	0.07862	0.280392
10				11-Jan-85	224.1	−0.003424	0.07635	0.276314
11		= SQRT(G11)		14-Jan-85	219.09	−0.022356	0.077838	0.278994
12				15-Jan-86	219.09	0.000000	0.075502	0.274777
13				16-Jan-85	222.17	0.014058	0.074731	0.273371
14				17-Jan-85	226.02	0.017329	0.07476	0.273422
15				18-Jan-85	226.02	0.000000	0.072517	0.26929
16				21-Jan-85	226.79	0.003407	0.070429	0.265385
17				22-Jan-85	234.49	0.033952	0.077031	0.277545
18					57	−0.008188	0.075227	0.274275
19					65	0.013243	0.074296	0.272573
20					27	0.019605	0.074973	0.273812
21					.5	−0.003205	0.072802	0.269818
22					72	−0.024134	0.075021	0.273899
23					35	0.041203	0.085605	0.292583
24					12	0.003164	0.083112	0.288292
25					04	0.007865	0.081086	0.284757
26					04	0.000000	0.078654	0.280453
27					39	0.015648	0.078145	0.279545
28					97	−0.007683	0.076247	0.276129
29					.97	0.000000	0.07396	0.271956
30				8-Feb-85	249.12	0.004638	0.071904	0.268149
31				11-Feb-85	243.35	−0.023162	0.073802	0.271666
32				12-Feb-85	238.34	−0.020588	0.074792	0.273482
33				13-Feb-85	239.5	0.004867	0.072728	0.269681
34				14-Feb-85	238.34	−0.004843	0.070723	0.265938
35				15-Feb-85	235.65	−0.011286	0.069565	0.263751
36				18-Feb-85	232.57	−0.013070	0.068769	0.262239
37				19-Feb-85	234.49	0.008256	0.067221	0.259271
38				20-Feb-85	235.65	0.004947	0.06539	0.255714
39				21-Feb-85	236.42	0.003268	0.063509	0.252009
40				22-Feb-85	235.65	−0.003257	0.061684	0.248362
41				25-Feb-85	234.49	−0.004923	0.060016	0.244982
42				26-Feb-85	235.65	0.004947	0.058401	0.241663
43				27/02/85	232.57	−0.013070	0.05794	0.240708

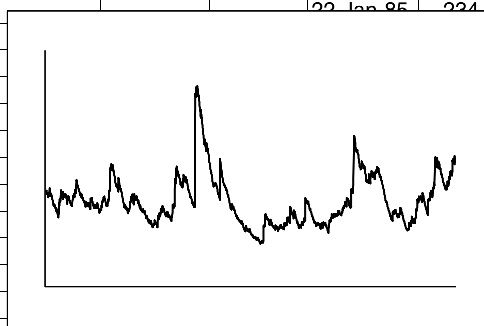
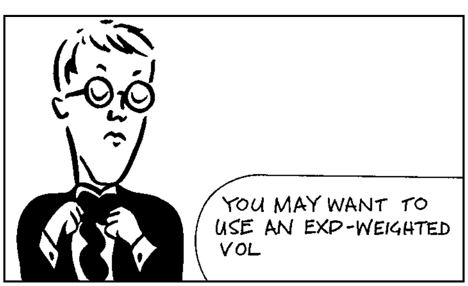


Figure 42.2 Spreadsheet to calculate an exponentially-weighted volatility.



Again, this measure shows spurious sudden rises and falls because of the equal weighting of all the returns.

Alternatively, we can use an exponentially weighted estimate

$$\sigma^2_{12_i} = \lambda\sigma^2_{12_{i-1}} + (1-\lambda)\frac{R_{1_i}R_{2_i}}{\delta t}.$$

There are problems with the estimation of covariance due to the synchronicity of asset movements and measurement. Two assets may be perfectly correlated but because of their measurement at different times they may appear to be completely uncorrelated. This is a problem when using data from markets in different time zones. Moreover, there is no guarantee that the exponentially-weighted covariances give a positive-definite matrix.

42.4 THE CREDITMETRICS DATASET

The CreditMetrics dataset is available from `www.jpmorgan.com`. The CreditMetrics methodology is also described in great detail at that site. The dataset consists of four data types: Yield curves, spreads, transition matrices and correlations. Before reading the following sections the reader should be comfortable with the concept of credit rating (see Chapter 40).

42.4.1 Yield Curves

The CreditMetrics yield curve dataset consists of the *risk-free* yield to maturity for major currencies. In Figure 42.3 is shown an example of these risk-free yields. The dataset contains yields for maturities of one, two, three, five, seven, 10 and 30 years. For example, from the yield curve dataset we have information such as the yield to maturity for a three-year US dollar bond is 6.12%.

42.4.2 Spreads

For each credit rating, the dataset gives the spread above the riskless yield for each maturity. In Figure 42.3 is shown a typical riskless US yield curve, and the yield on AA and BBB bonds. For example, we may be given that the spread for an AA bond is 0.54% above the riskless yield for a three-year bond. Observe that the riskier the bond the higher the yield; the yield on the BBB bond is everywhere higher than that on the AA bond which is in turn higher than the risk-free yield. This higher yield for risky bonds is compensation for the possibility of not receiving future coupons or the principal.

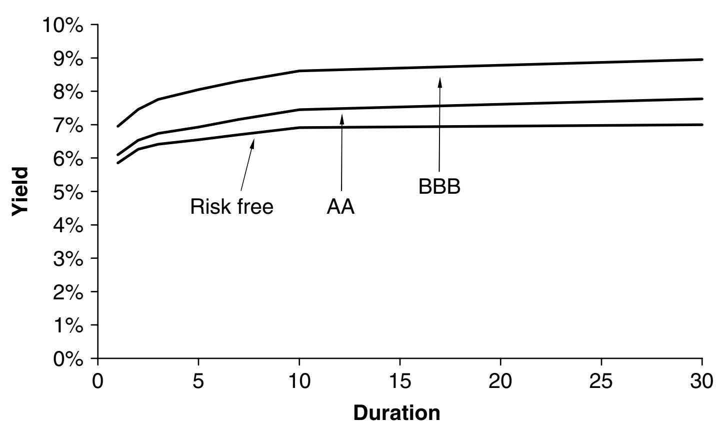


Figure 42.3 Risk-free and two risky yield curves.

42.4.3 Transition Matrices

The concept of the transition matrix has been discussed in Chapter 40. In the CreditMetrics frame-work, the transition matrix has as its entries the probability of a change of credit rating at the end of a given time horizon, for example, the probability of a upgrade from AA to AAA might be 5.5%. The time horizon for the CreditMetrics dataset is one year. Unless the time horizon is very long, the largest probability is typically for the bond to remain at its initial rating; let's say that the probability of staying at AA is 87% in this example.

42.4.4 Correlations

In the risk-free yield, the spreads and the transition matrix, there is sufficient information for the CreditMetrics method to derive distributions for the possible future values of a single bond. I show how this is done in the next section. However, when we come to examine the behavior of a portfolio of risky bonds, we must consider whether there is any relationship between the rerating or default of one bond and the rerating or default of another. In other words, are bonds issued by different companies or governments in some sense correlated? This is where the CreditMetrics correlation dataset comes in. This dataset gives the correlations between major indices in many countries.

Each company issuing bonds has the return on its stock decomposed into parts correlated with these indices and a part which is specific to the company. By relating all bond issuers to these indices we can determine correlations between the companies in our portfolio. We will see how this is used in practice later in this chapter.

42.5 THE CREDITMETRICS METHODOLOGY

The CreditMetrics methodology is about calculating the possible values of a risky portfolio at some time in the future (the time horizon) and estimating the probability of such values occurring. Let us consider just a single risky bond currently rated AA. Suppose that the bond is zero-coupon, with a maturity of three years and we want to know the possible state of this investment in one year's time. The yield to maturity on this instrument might be 6.12%, for a three-year riskless bond, plus 0.54% for the spread for a three-year AA-rated bond. The total yield is therefore 6.66%, giving a price of 0.819.

The value of the bond will fluctuate between now and one year's time for each of three reasons: The passage of time, the evolution of interest rates and the possible regrading of the bond. Let us take these three points in turn.

First, because of the passage of time our three-year bond will be a two-year bond in one year. But what will be the yield on a two-year bond in one year's time? This is the second point. The assumption that is made in CreditMetrics is that the forward rates do not change between today and the time horizon ('rolling down the curve'). From the yields that we have today we can calculate the forward rates that apply between now and one year, between one and two years, between two and three years etc. This calculation is described in Chapter 13. We can calculate the value of the bond after one year; suppose it is 0.882. But why should the bond still be rated AA at that time? This is the third point. From our transition matrix we see that the probability of the bond's rating staying AA is 87%. So, there is an 87% chance that the bond's value will be 0.882. We can similarly work out the value of the bond in one year if it is rated AAA, A,

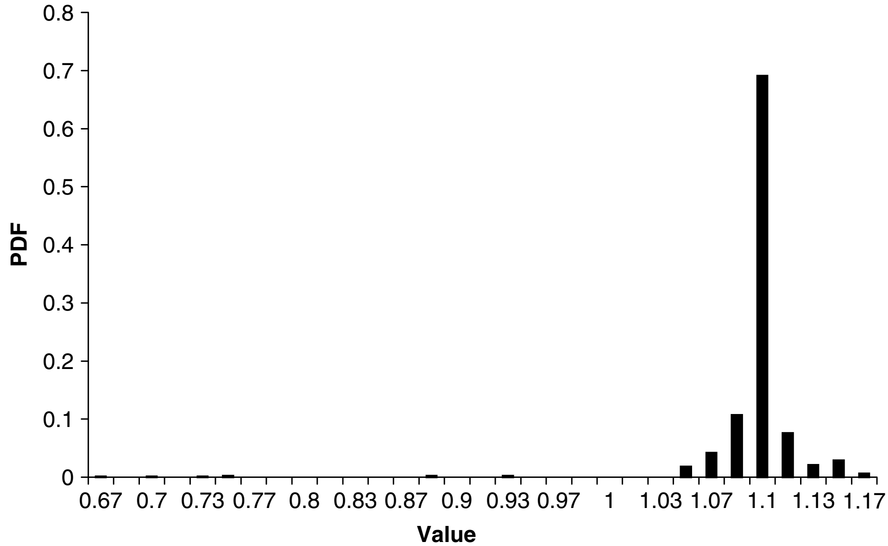


Figure 42.4 The probability distribution for the bond's value after one year.

BBB etc. using the relevant forward rates and spreads that we assume will apply in one year's time. And each of these has a probability of occurring that is given in the transition matrix. A probability distribution of the possible bond values is shown in Figure 42.4.

This, highly skewed, distribution tells us all we need to know to determine the risk in this particular bond. We can, for example, calculate the expected value of the bond.

42.6 A PORTFOLIO OF RISKY BONDS

We have seen how to apply the CreditMetrics methodology to a single risky bond, to apply the ideas to a portfolio of risky bonds is significantly harder since it requires the knowledge of any relationship between the different bonds. This is most easily measured by some sort of correlation.

Suppose that we have a portfolio of two bonds. One, issued by ABC, is currently rated AA and the other, issued by XYZ, is BBB. We can calculate, using the method above, the value of each of these bonds at our time horizon for each of the possible states of the two bonds. If we assume that each bond can be in one of eight states (AAA, AA, ... , CCC, Default) there are $8^2 = 64$ possible joint states at the time horizon. To calculate the expected value of our portfolio and standard deviation we need to know the probability of each of these joint states occurring. This is where the correlation comes in.

There are two stages to determining the probability of any particular future joint state:

1. Calculate the correlations between bonds;
2. Calculate the probability of any joint state.

Stage 1 is accomplished by decomposing the return on the stock of each issuing company into parts correlated with the major indices.

42.7 CREDITMETRICS MODEL OUTPUTS

CreditMetrics is, above all, a way of measuring risk associated with default issues. From the CreditMetrics methodology one can calculate the risk, measured by standard deviation, of the risky portfolio over the required time horizon. Because of the risk of default the distribution of returns from a portfolio exposed to credit risk is highly skewed, as in Figure 42.4. The distribution is far from being Normal. Thus ideas from simple portfolio theory must be used with care. Although, it may not be a good absolute measure of risk in the classical sense, the standard deviation is a good indicator of relative risk between instruments or portfolios.

42.8 SUMMARY

This chapter has outlined some of the methodologies for competing and complementary Value at Risk measures. With something as important as Value at Risk there is an obvious case to be made for exploring all of the possible VaR measures to build up as accurate a profile as possible of the dangers lurking in your portfolio. In the next chapter we look at measuring and reducing the risk in stock market crashes.

FURTHER READING

- Alexander (1996a) is a critique of RiskMetrics as a risk measurement tool.
- Shore (1997) describes and implements the CreditMetrics methodology.

CHAPTER 43
CrashMetrics

In this Chapter...

- the methodology of CrashMetrics for measuring a portfolio's exposure to sudden, unhedgeable market movements
- Platinum Hedging
- crash coefficients
- margin hedging
- counterparty risk
- the CrashMetrics Index for measuring the magnitude of crashes

43.1 INTRODUCTION

The final piece of the jigsaw for estimating risk in a portfolio is **CrashMetrics**. If Value at Risk is about normal market conditions then CrashMetrics is the opposite side of the coin; it is about 'fire sale' conditions and the far-from-orderly liquidation of assets in far-from-normal conditions. CrashMetrics is a dataset and methodology for estimating the exposure of a portfolio to extreme market movements or crashes. It assumes that the crash is unhedgeable and then finds the worst outcome for the value of the portfolio. The method then shows how to mitigate the effects of the crash by the purchase or sales of derivatives in an optimal fashion, so-called Platinum Hedging. Derivatives have sometimes been thought of as being a dangerous component in a portfolio, but in the CrashMetrics methodology they are put to a benign use.

43.2 WHY DO BANKS GO BROKE?

There are two main reasons why banks get into serious trouble. The first reason is the lack of suitable or sufficient control over the traders. Through misfortune, negligence or dishonesty, large and unmanageable positions can be entered into. The consequences are either that the trader concerned becomes a hero and the bank makes a fortune, or the bank loses a fortune, the trader makes a run for it and the bank goes under. The odds are fifty-fifty. The second causes of disaster are the extreme, unexpected and unhedgeable moves in the stock market; the crashes.

43.3 MARKET CRASHES

In typical market conditions one's portfolio will fluctuate rapidly, but not dramatically. That is, it will rise and fall, minute by minute, day by day, but will not collapse. There are times, say once a year on average, when that fluctuation is dramatic . . . and usually in the downward direction. These are extreme market movements or market crashes. VaR can tell us nothing about these and they must be analyzed separately.

What's special about a crash? Two things spring to mind. Obviously a crash is a sudden fall in market prices, too rapid for the liquidation of a portfolio. But a crash isn't just a rise in volatility. It is also characterized by a special relationship between individual assets. During a crash, all assets fall together. There is no such thing as a crash where half the stocks fall and the rest stay put. Technically this means that all assets become perfectly correlated. In normal market conditions there may be some relationship between stocks, especially those in the same sector, but this connection may not be that strong. Indeed, it is the weakness of these relationships that allows diversification. A small insurance company will happily insure your car, because they can diversify across individuals. Insuring against an earthquake is a different matter. A high degree of correlation makes diversification impossible. This is where traditional VaR falls down, at exactly the time when it is needed most. Figure 43.1 shows the behavior of the correlation of several constituents of the S&P500 around the time of the 1987 stock market crash.

All is not lost. VaR is a very recent concept, created during the 1990s and fast becoming a market standard, with known drawbacks. Many researchers in universities and in banks are turning their thoughts to analyzing and protecting against crashes. Some of these researchers

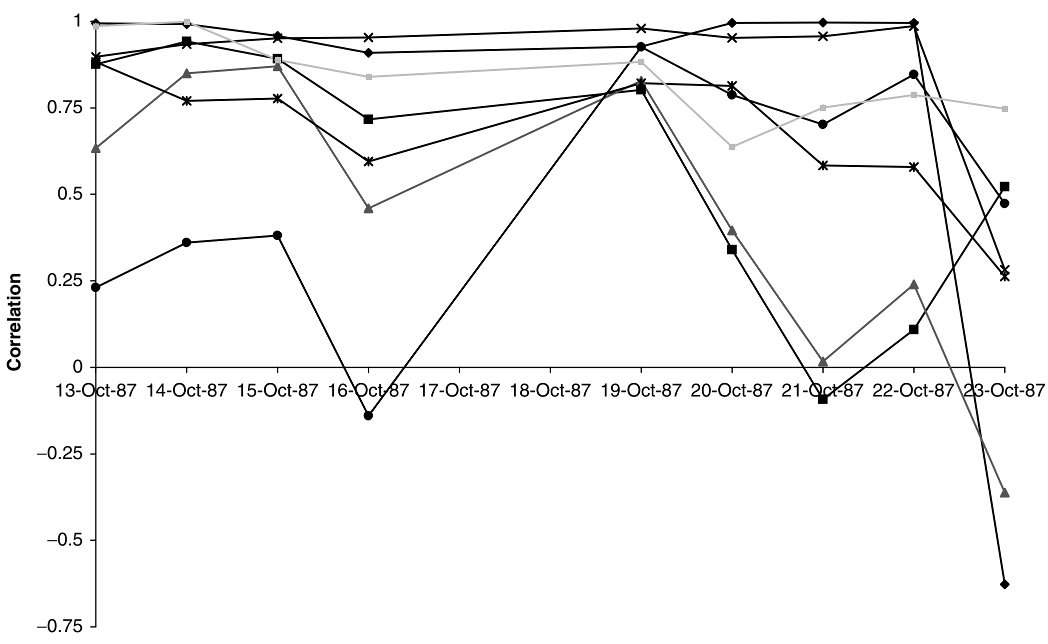


Figure 43.1 The correlation between several assets for a few days before and after the 1987 crash. When the correlation is close to one all assets move in the same direction.

are physicists who concern themselves with examining the tails of returns distributions; are crashes more likely than traditional theory predicts? The answer is a definite yes.

My personal preference though is for models that don't make any assumptions about the likelihood of a crash. One line of work is that of 'worst-case scenarios.' Given that a crash could wipe out your portfolio, wouldn't you like to know what is the worst that could realistically happen, or would you be happy knowing what you would lose on average? CrashMetrics is used to analyze worst cases, and provide advice about how to hedge or insure against a crash.

43.4 CRASHMETRICS

CrashMetrics is a methodology for evaluating portfolio performance in the event of extreme movements in financial markets. It is not part of the JP Morgan family of performance measures. In its most complex and sophisticated form, the concept and mathematics are explained in full in Chapter 58. There we see how the portfolio of financial instruments is valued under a worst-case scenario with few assumptions about the size of the market move or its timing. The only assumptions made are that the market move, the 'crash', is limited in size and that the number of such crashes is limited in some way. There are no assumptions about the probability distribution of the size of the crash or its timing.

The simpler method, used for day-to-day portfolio protection, is concerned with the extreme market movements that may occur when we are not watching, or that cannot be hedged away. These are the fire sale conditions. This is the method I will explain for the rest of this chapter. There are many nice things about the method such as its simplicity and ease of generalization, and no explicit dependence on the annoying parameters volatility and correlation.

43.5 CRASHMETRICS FOR ONE STOCK

To introduce the ideas, let's consider a portfolio of options on a single underlying asset. For the moment think in terms of a stock, although we could equally well talk about currencies, commodities or interest rate products.

If the stock changes dramatically by an amount δS how much does the portfolio of options on that stock behave? There will be a relationship between the change in the portfolio value $\delta\Pi$ and δS:

$$\delta\Pi = F(\delta S).$$

The function $F(\cdot)$ will simply be the sum of all the formulae, expressions, numerical solutions...for each of the individual contracts in the portfolio. Think of it as the sum of lots of Black–Scholes formulae with lots of different strikes, expiries, payoffs. If there is no change in the asset price there will be no change in the portfolio, so $F(0) = 0$. (There will be a small time decay, which we'll come back to later.) Figure 43.2 shows a possible portfolio change against underlying change.

If we are lucky, and we are not too near to the expiries and strikes of the options then we could approximate the portfolio by the Taylor series in the change in the underlying asset:

$$\delta\Pi = \Delta\ \delta S + \tfrac{1}{2}\Gamma\ \delta S^2 + \cdots. \tag{43.1}$$

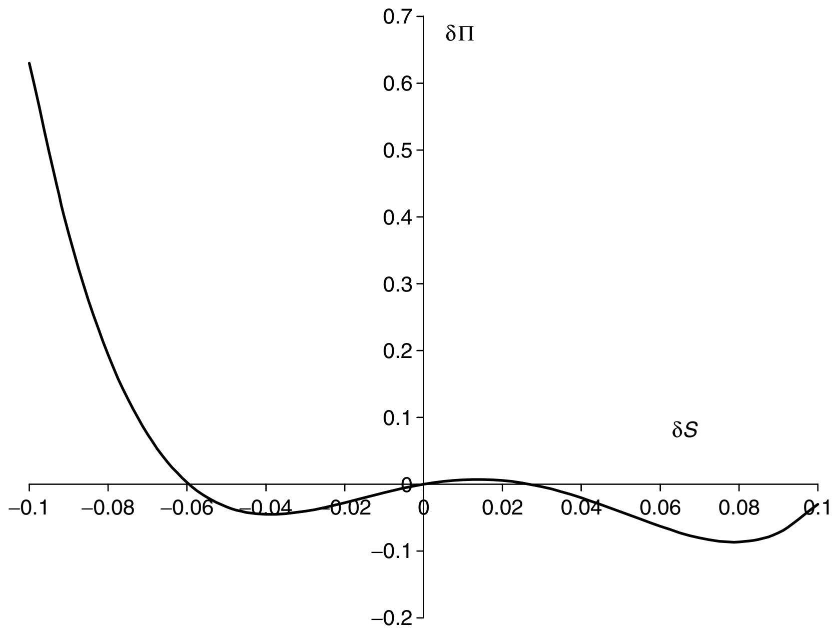


Figure 43.2 Size of portfolio change against change in the underlying.

In practice this won't be a good enough approximation. Imagine having some knock-out options in the portfolio, we really will have to use the relevant formula or numerical method to capture the sudden drop in value of this contract at the barrier. A simple delta-gamma approximation is not going to work.

However, as far as the math is concerned I'm going to show you both the general Crash-Metrics methodology and the simple Taylor series version.

Now let's ask what is the worst that could happen to the portfolio overnight say? We want to find the minimum of $F(\delta S)$.

In practice we would only want to look at crashes/rallies of up to a certain magnitude. For this reason we may want to constrain the move in the underlying by

$$-\delta S^- < \delta S < \delta S^+.$$

If we can't use the greek approximation (Taylor series) then we're looking for

$$\min_{-\delta S^- < \delta S < \delta S^+} F(\delta S).$$

Figure 43.2 shows an example where there is one local minimum as well as a global one; it's the global one we want.

In Figure 43.3 we see a plot of the change in the portfolio against δS assuming that a Taylor approximation is valid. Note that it is zero at $\delta S = 0$. If the gamma is positive the portfolio change (43.1) has a minimum at

$$\delta S = -\frac{\Delta}{\Gamma}.$$

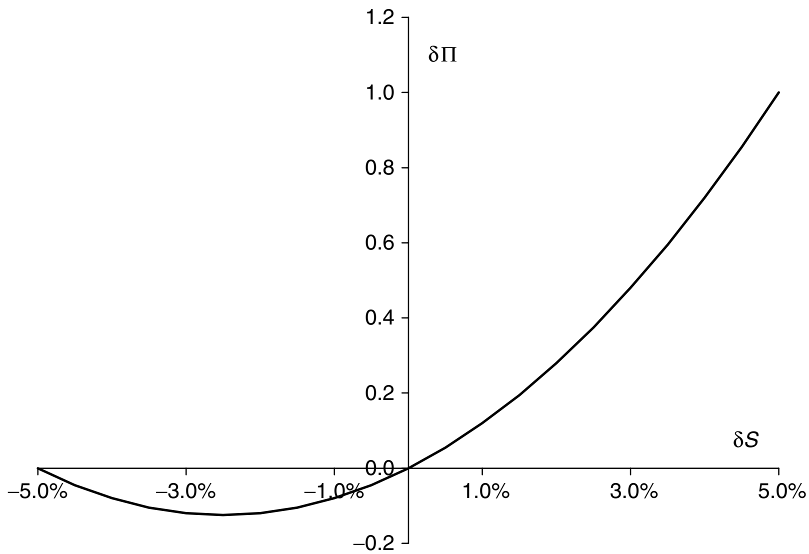


Figure 43.3 Size of portfolio change against change in the underlying, Taylor approximation.

The portfolio change in this worst-case scenario is

$$\delta\Pi_{\text{worst}} = -\frac{\Delta^2}{2\Gamma}.$$

This is the worst case given an arbitrary move in the underlying. If the gamma is small or negative the worst case will be a fall to zero or a rise to infinity, both far too unrealistic.

43.6 PORTFOLIO OPTIMIZATION AND THE PLATINUM HEDGE

Having found a technique for finding out what could happen in the worst case, it is natural to ask how to make that worst case not so bad. This can be done by optimal static hedging.

To start with, I'll assume the Taylor expansion and then generalize.

Suppose that there is a contract available with which to hedge our portfolio. This contract has a bid-offer spread, a delta and a gamma. I will call the delta of the hedging contract Δ^*, meaning the sensitivity of the hedging contract to the underlying asset. The gamma is similarly Γ^*. Denote the bid-offer spread by $C > 0$, meaning that if we buy (sell) the contract and immediately sell (buy) it back we lose this amount.

Imagine that we add a number λ of the hedging contract to our original position. Our portfolio now has a first-order exposure to the crash of

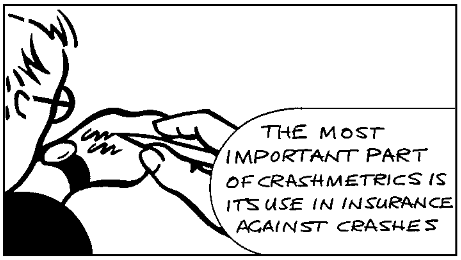
$$\delta S\left(\Delta + \lambda\Delta^*\right)$$

and a second-order exposure of

$$\tfrac{1}{2}\delta S^2\left(\Gamma + \lambda\Gamma^*\right).$$

Not only does the portfolio change by these amounts for a crash of size δS but also it loses a *guaranteed* amount

$$|\lambda|\, C$$



just because we cannot close our new position without losing out on the bid-offer spread.

The total change in the portfolio with the static hedge in place is now

$$\delta\Pi = \delta S \left(\Delta + \lambda \Delta^*\right) + \tfrac{1}{2}\delta S^2 \left(\Gamma + \lambda \Gamma^*\right) - |\lambda|\, C.$$

In general, the optimal choice of λ is such that the worst value of this expression for $-\delta S^- \le \delta S \le \delta S^+$ is as high as possible. Thus we are exchanging a guaranteed loss (due to bid-offer spread) for a reduced worst-case loss. This is simply insurance and the optimal choice gives the **Platinum Hedge**, named for the plastic card that comes after green and gold. For the optimal choice of the λ Figure 43.4 shows the change in the portfolio value as a function of δS. Note that it no longer goes through $(0, 0)$.

In Figure 43.5 is shown a simple spreadsheet for finding the worst-case scenario and the Platinum Hedge when there is a single asset.

If we can't use the Taylor approximation, as will generally be the case, we must look for the worst case of

$$F(\delta S) + \lambda F^*(\delta S) - |\lambda| C.$$

Here $F^*(\cdot)$ is the 'formula' for the change in value of the hedging contract.

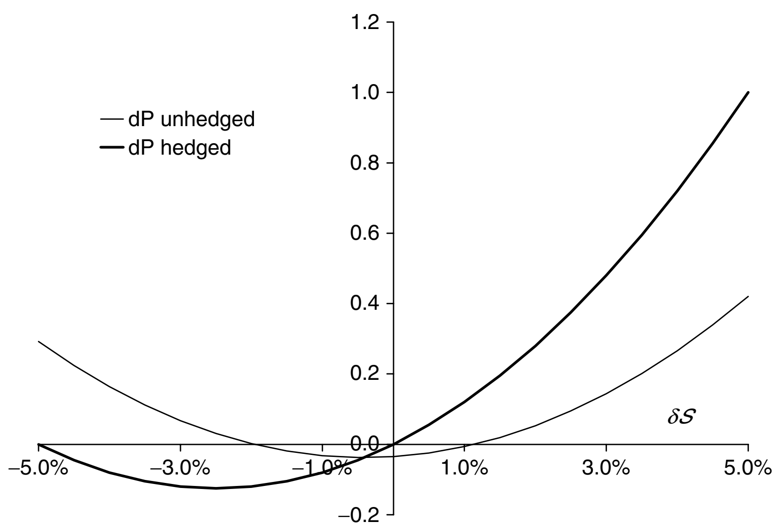


Figure 43.4 Size of portfolio change against δS after optimal hedging, Taylor approximation.

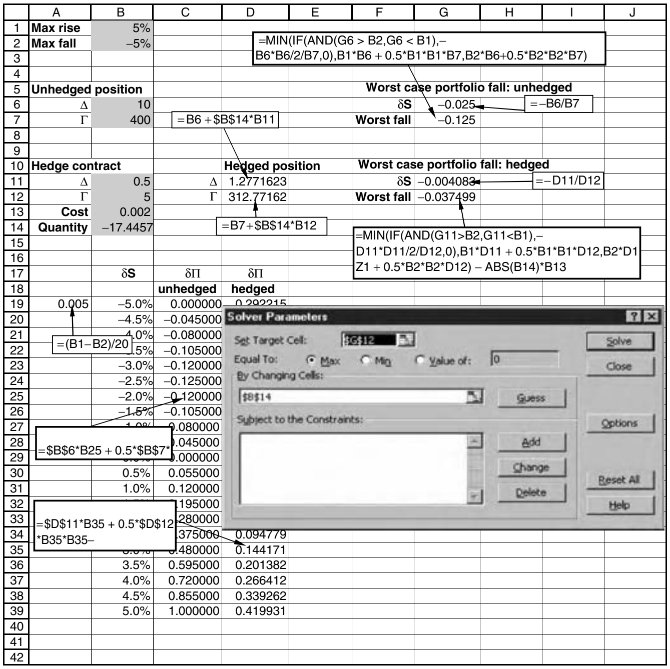
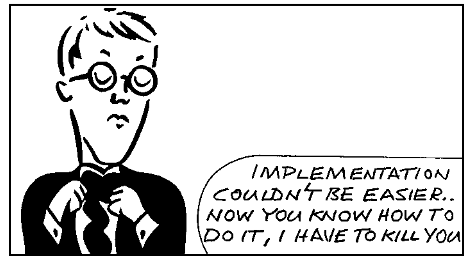


Figure 43.5 Spreadsheet for implementing basic CrashMetrics in one asset.



Having found the worst case, we just make this as painless as possible by optimizing over the hedge ratio λ.

Of course, there won't just be the one option with which to hedge statically, there will be many. How does this change the optimization? We'll find out soon.

43.6.1 Other 'Cost' Functions

In the above we have said that the 'cost' associated with the above is the bid-offer spread on the static hedge part of the portfolio. This is but one of several choices. The reason for this choice is that we may decide tomorrow to get rid of the static hedge, and hence that would genuinely be our cost.

We could also say that the cost function was the value of the bought options, in the sense that we are writing off the 'insurance.' Another choice would be to subtract of a cost representing the difference between what we value the static hedge at and the value in the market. This would be suitable for a portfolio that has been constructed for its volatility arbitrage possibilities. A option bought for static hedging that is correctly priced would therefore have no associated cost.

43.7 THE MULTI-ASSET/SINGLE-INDEX MODEL

A bank's portfolio has many underlyings, not just the one. How does CrashMetrics handle them? This is done via an index or benchmark.

We can measure the performance of a portfolio of assets and options on these assets by relating the magnitude of extreme movements in any one asset to one or more **benchmarks** such as the S&P 500. The relative magnitude of these movements is measured by the **crash coefficient** for each asset relative to the benchmark. If the benchmark moves by $x\%$ then the ith asset moves by $\kappa_i x\%$. Estimates of the κ_i for the constituents of the S&P 500, with that index as the benchmark, may be downloaded free of charge from `www.crashmetrics.com`. Note that the benchmark need not be an index containing the assets, but can be any representative quantity. Unlike the RiskMetrics and CreditMetrics datasets, the CrashMetrics dataset does not have to be updated frequently because of the rarity of extreme market movements.

Tables 43.1–43.5 give the crash coefficients for a few constituents of major indices in several countries. The crash coefficients have been estimated using the tails of the daily return distributions from the beginning of 1985 until the end of 1997, and so include the Black Monday crash of October 1987 and the rice/dragon/sake/Asian 'flu' effect starting in October 1997. For example, in Table 43.1 we see the 10 largest positive and negative daily returns in the S&P 500 during that period. In this table are also shown the returns on the same days for several constituents of the index.

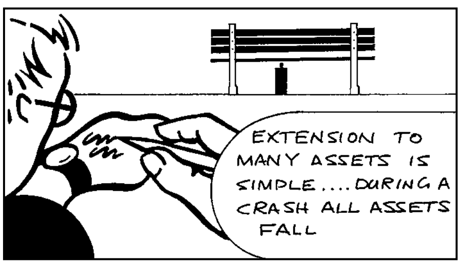


Figure 43.6 uses the same data as in Chapter 18 for the calculation of the beta for Disney. The fine line in this figure has slope beta. On the figure is shown the line with zero intercept that fits the largest 20 rises and falls in the S&P500, this is the bold line.

In Figure 43.7 are the returns on the Hong Kong and Shanghai Hotel group versus returns on the Hang Seng and in

Table 43.1 The 10 largest positive and negative moves in several constituents of the S&P500 against the moves in the S&P500 on the same days.

Date	S&P500	% change	ABBOTT LABS.	ADOBE SYS.	ADVD.MICR. DEVC.	AEROQUIP-VICKERS	AETNA	AHMANSON (H.F.)
19-Oct-87	225	−20.4	−10.5	−22.2	−36.1	−36.6	−15.3	−20.8
26-Oct-87	228	−8.3	−7.3	−20.0	−14.3	−15.2	−4.5	−4.3
27-Oct-97	877	−6.9	−5.3	−6.1	−19.8	−6.7	−8.5	−6.5
08-Jan-88	243	−6.8	−3.8	−14.3	−6.8	−13.5	−7.1	−6.2
13-Oct-89	334	−6.1	−8.2	−12.5	−5.8	−9.3	−5.5	−3.7
16-Oct-87	283	−5.2	−4.6	0.0	−5.3	−6.4	−1.5	−1.3
11-Sep-86	235	−4.8	−5.2	−50.0	−4.7	−5.3	−3.3	−2.9
14-Apr-88	260	−4.4	−4.0	−12.5	−5.6	−2.6	−4.4	−4.8
30-Nov-87	230	−4.2	−6.7	−16.7	−1.4	−9.8	−2.7	−6.1
22-Oct-87	248	−3.9	−4.6	0.0	−5.9	−6.5	−1.9	−1.4
21-Oct-87	258	9.1	4.3	0.0	4.1	11.5	8.0	9.5
20-Oct-87	236	5.3	−0.6	−14.3	6.5	14.3	−3.7	3.3
28-Oct-97	921	5.1	5.2	4.3	19.0	−0.6	4.3	7.4
29-Oct-87	244	4.9	2.4	33.3	10.3	15.5	−1.9	3.2
17-Jan-91	327	3.7	4.3	0.0	4.6	8.0	2.8	3.0
04-Jan-88	255	3.6	0.8	0.0	7.6	−0.4	1.9	−0.7
31-May-88	262	3.4	2.7	0.0	5.3	0.9	3.9	2.5
27-Aug-90	321	3.2	5.4	8.3	4.5	2.2	1.4	3.1
02-Sep-97	927	3.1	3.8	2.6	3.3	0.1	2.5	2.3
21-Aug-91	391	2.9	3.1	4.2	3.5	0.0	−3.3	2.6

Table 43.2 The 10 largest positive and negative moves in several constituents of the FTSE100 against the moves in the FTSE100 on the same days.

Date	FTSE100	% change	ALLIED DOMECQ	ASDA FOODS	ASSD.BRIT.	BAA	BANK OF SCOTLAND	BARCLAYS
20-Oct-87	1802	−12.2	−7.1	−9.0	−10.1	−7.1	−12.7	−13.2
19-Oct-87	2052	−10.8	−11.5	−9.6	−3.1	−3.7	−3.8	−11.4
26-Oct-87	1684	−6.2	−7.6	−2.4	−3.0	−3.1	−9.0	−7.5
22-Oct-87	1833	−5.7	−4.0	−7.7	−4.7	−2.6	−1.8	−1.0
30-Nov-87	1580	−4.3	−3.0	−4.3	−2.3	−3.8	−2.1	−6.3
05-Oct-92	2446	−4.1	−2.0	−2.9	2.3	−3.4	−6.8	−4.6
03-Nov-87	1653	−4.0	−3.0	−1.7	−2.0	−0.8	−2.0	−7.8
09-Nov-87	1565	−3.4	−3.9	−5.5	−1.0	−4.5	−1.6	−2.2
29-Dec-87	1730	−3.4	−2.3	−1.8	−2.3	−2.1	−0.6	−2.6
16-Oct-89	2163	−3.2	−3.2	−5.0	−3.4	−3.3	0.0	−2.2
21-Oct-87	1943	7.9	7.7	4.9	2.6	2.7	1.5	8.7
10-Apr-92	2572	5.6	8.2	3.3	3.5	7.1	13.1	7.6
17-Sep-92	2483	4.4	3.5	3.6	2.4	2.5	9.6	15.5
11-Nov-87	1639	4.2	2.3	1.8	4.4	2.1	1.0	3.4
30-Oct-87	1749	4.0	1.5	5.0	6.9	3.2	4.6	2.1
12-Nov-87	1702	3.9	1.8	−0.6	−1.0	4.1	0.8	2.2
05-Oct-90	2143	3.6	5.9	0.0	−1.0	3.2	9.3	9.8
18-Sep-92	2567	3.3	6.7	3.4	4.9	3.0	4.2	3.0
26-Sep-97	5226	3.2	1.8	−0.3	1.9	3.2	9.5	8.9
31-Dec-91	2493	3.0	4.0	7.8	1.7	1.3	4.5	2.7

Table 43.3 The 10 largest positive and negative moves in several constituents of the Hang Seng against the moves in the Hang Seng on the same days.

Date	Hang Seng	% change	AMOY PROPS	BANK OF E. ASIA	CHEUNG KONG.	CHINA LT.& POW.	FIRST PACIFIC	GREAT EAGLE
26-Oct-87	2241	−33.3	−37.4	−37.0	−29.2	−32.2	−28.3	−57.3
05-Jun-89	2093	−21.7	−36.4	−22.9	−27.0	−16.6	−28.6	−41.3
28-Oct-97	9059	−13.7	−4.8	−12.7	−9.8	−10.8	−14.7	−19.3
19-Oct-87	3362	−11.1	−18.9	0.0	−9.6	−10.5	−10.7	−20.2
22-May-89	2806	−10.8	−18.6	−8.2	−10.7	−6.2	−14.9	−15.4
23-Oct-97	10426	−10.4	−14.0	−8.9	−13.4	−7.4	−26.9	−6.3
25-May-89	2752	−8.5	−16.6	−7.6	−10.8	−5.0	−9.5	−14.9
19-Aug-91	3722	−8.4	−11.6	−4.7	−6.8	−6.8	−10.2	−9.4
03-Dec-92	4978	−8.0	−2.4	−13.4	−4.8	−9.9	−7.5	−11.3
06-Aug-90	3107	−7.4	−10.3	−6.1	−6.1	−8.2	−6.9	−6.0
29-Oct-97	10765	18.8	9.2	4.9	17.8	20.2	17.3	18.0
23-May-89	3067	9.3	11.4	7.5	9.9	5.2	11.4	14.3
06-Nov-87	2113	7.8	8.5	1.2	12.2	5.4	2.2	9.7
12-Jun-89	2440	7.6	11.7	8.7	10.5	7.0	7.0	16.9
03-Sep-97	14713	7.1	6.0	6.2	5.3	11.9	11.2	5.0
24-Oct-97	11144	6.9	4.1	2.0	9.5	5.9	12.6	7.1
27-Oct-87	2395	6.9	20.1	−2.0	3.8	9.0	−11.4	1.8
03-Nov-97	11255	5.9	5.3	7.8	7.4	−2.9	13.7	6.8
14-Jan-94	10774	5.9	6.1	3.7	6.6	3.6	3.9	4.9
19-Jun-85	1510	5.8	0.0	6.1	6.1	6.2	0.0	13.2

Table 43.4 The 10 largest positive and negative moves in several constituents of the Nikkei against the moves in the Nikkei on the same days.

Date	Nikkei	% change	AJINOMOTO	ALL NIPPON AIRWAYS	AOKI	ASAHI BREW.	ASAHI CHEM.	ASAHI DENKA KOGYO
20-Oct-87	21910	−14.9	−14.4	−17.9	−18.9	−18.1	−15.7	−10.8
02-Apr-90	28002	−6.6	−3.2	−6.3	−13.7	−5.0	−1.2	−5.8
19-Aug-91	21456	−6.0	−12.1	−7.9	−8.4	−1.6	−5.5	−4.7
23-Aug-90	23737	−5.8	−7.6	−12.0	−9.6	−5.4	−8.4	−7.3
23-Jan-95	17785	−5.6	−8.0	−7.2	−4.9	−1.9	−5.3	−5.2
19-Nov-97	15842	−5.3	−9.0	−3.5	−16.3	−3.4	−3.2	−5.6
24-Jan-94	18353	−4.9	−4.3	−6.0	−6.8	−1.6	−4.1	−7.6
23-Oct-87	23201	−4.9	−4.0	−6.8	−0.9	−4.1	−6.8	−3.4
26-Sep-90	22250	−4.7	−5.2	−2.4	−5.4	−1.6	−2.7	−0.6
03-Apr-95	15381	−4.7	−2.2	−4.1	−2.3	−2.0	−8.4	−6.6
02-Oct-90	22898	13.2	16.0	14.7	9.8	11.4	10.0	15.8
21-Oct-87	23947	9.3	13.5	16.3	11.6	14.7	9.3	5.2
17-Nov-97	16283	8.0	7.3	6.1	0.0	7.4	11.8	9.9
31-Jan-94	20229	7.8	8.5	3.6	13.6	4.2	4.6	9.6
10-Apr-92	17850	7.5	10.5	3.9	13.3	0.0	6.1	3.5
07-Jul-95	16213	6.3	9.8	6.3	9.2	3.0	5.0	1.2
21-Aug-92	16216	6.2	7.9	3.3	21.2	5.6	2.9	2.5
27-Aug-92	17555	6.1	5.6	−1.0	9.4	10.1	5.8	1.8
06-Jan-88	22790	5.6	5.0	6.3	6.4	2.7	6.6	2.4
15-Aug-90	28112	5.4	2.2	6.9	5.4	3.2	3.8	3.8

Table 43.5 The 10 largest positive and negative moves in several constituents of the Dax against the moves in the Dax on the same days.

Date	Dax	% change	ALLIANZ HLDG.	BASF	BAYER	BAYER HYPBK.	BAYERISCHE VBK.	BMW
16-Oct-89	1385	−12.8	−11.3	−10.0	−7.0	−16.1	−13.8	−13.1
19-Aug-91	1497	−9.4	−9.9	−4.8	−5.8	−11.2	−11.2	−10.0
19-Oct-87	1321	−9.4	−10.4	−9.4	−8.5	−7.0	−3.6	−8.2
28-Oct-97	3567	−8.0	−4.7	−8.3	−9.6	−8.6	−7.8	−14.8
26-Oct-87	1193	−7.7	−10.2	−4.0	−5.3	−9.0	−5.5	−6.4
28-Oct-87	1142	−6.8	−7.1	−2.6	−3.2	−8.9	−5.7	−7.5
22-Oct-87	1287	−6.7	−4.2	−4.6	−6.9	−4.6	−5.8	−7.5
10-Nov-87	945	−6.5	−8.9	−4.8	−4.1	−7.4	−6.6	−7.8
04-Jan-88	943	−5.6	−9.9	−6.9	−5.7	−2.1	−5.6	−3.3
06-Aug-90	1740	−5.4	−3.4	−4.3	−6.1	−5.5	−5.2	−6.4
17-Jan-91	1422	7.6	7.8	5.3	6.9	5.8	9.6	9.3
12-Nov-87	1061	7.4	16.6	4.7	7.6	7.6	10.8	6.4
30-Oct-87	1177	6.6	10.2	3.6	8.1	7.8	2.7	7.4
17-Oct-89	1475	6.5	5.8	3.7	2.5	5.9	9.2	5.9
01-Oct-90	1420	6.4	7.8	7.4	9.0	6.0	9.5	7.0
05-Jan-88	1004	6.4	8.6	5.1	4.7	1.7	4.6	4.4
29-Oct-97	3791	6.3	3.5	8.4	8.2	5.9	6.8	12.0
27-Aug-90	1654	6.1	3.9	8.7	5.7	5.0	2.5	6.1
21-Oct-87	1379	5.9	10.9	1.5	8.8	8.7	4.1	4.2
08-Oct-90	1465	5.3	9.1	5.0	5.2	5.1	1.8	3.7

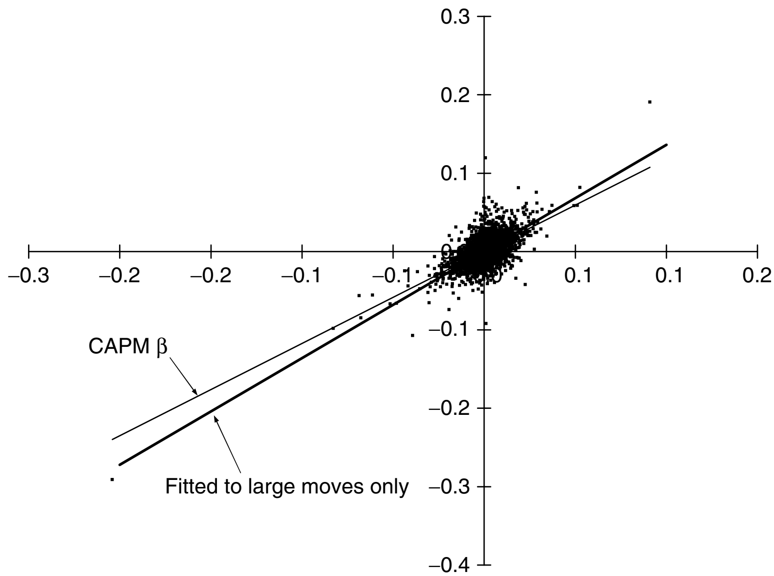


Figure 43.6 The returns on Disney versus returns on the S&P500. Also shown are the line with slope beta, fitted to all points, and the line with slope κ fitted to the 40 extreme moves and having zero intercept.

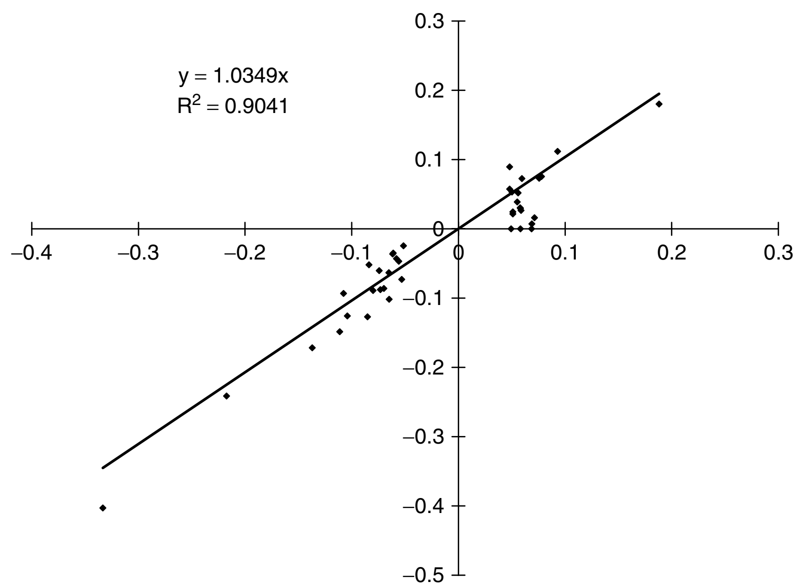


Figure 43.7 The returns on the Hong Kong and Shanghai Hotel group versus returns on the Hang Seng. Also shown is the line with slope κ fitted to the 40 extreme moves and having zero intercept.

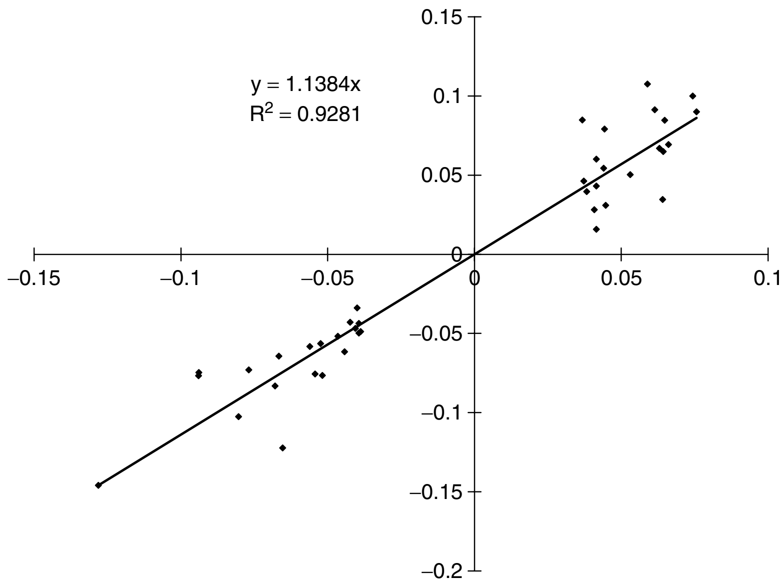


Figure 43.8 The returns on Daimler–Benz versus returns on the Dax. Also shown is the line with slope κ fitted to the 40 extreme moves and having zero intercept.

Figure 43.8 are the 40 extreme moves in Daimler–Benz versus returns on the Dax. It is important to note at this stage that the crash coefficient is not the same as the asset's beta with respect to the index. Not only is the number different, but preliminary results suggest that the crash coefficient is more stable than the beta. Moreover, for large moves in the index, the stock and the index are far more closely correlated than under normal market conditions. In other words, when there is a crash all stocks move together.

Figure 43.9 shows the different regimes for the movement of the market as a whole and individual names. There are three scenarios:

1. Typical day-to-day non-event; small movements in stock and market; perhaps some weak correlation
2. Stock-specific event; big rise or fall in individual name; market as a whole is oblivious
3. Market crash or dramatic rally; market and stock both move dramatically

Because of the shape of this figure, I call it the **rings of Saturn** effect.

43.7.1 Assuming Taylor Series for the Moment

Let's use these ideas, first assuming a Taylor expansion for the portfolio change.

In the single-index, multi-asset model we can write the change in the value of the portfolio as

$$\delta\Pi = \sum_{i=1}^{N} \Delta_i \, \delta S_i + \tfrac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \Gamma_{ij} \, \delta S_i \, \delta S_j \tag{43.2}$$

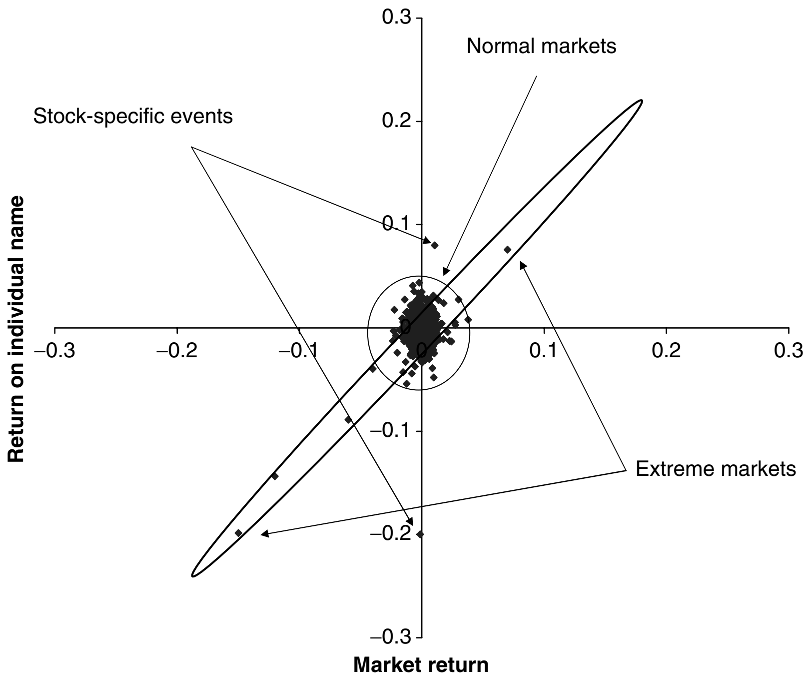


Figure 43.9 Schematic diagram showing the various returns regimes.

with the obvious notation. (In particular, observe the cross gammas.) We assume that the percentage change in each asset can be related to the percentage change in the benchmark, x, when there is an extreme move:

$$\delta S_i = \kappa_i x S_i.$$

This simplifies (43.2) to

$$\begin{aligned}\delta\Pi &= x\sum_{i=1}^{N}\Delta_i\kappa_i S_i + \tfrac{1}{2}x^2\sum_{i=1}^{N}\sum_{j=1}^{N}\Gamma_{ij}\kappa_i S_i\kappa_j S_j\\ &= xD + \tfrac{1}{2}x^2 G.\end{aligned}$$

Observe how this contains a first- and a second-order exposure to the crash. The first-order coefficient D is the **crash delta** and the second-order coefficient G is the **crash gamma**.

Now we constrain the change in the benchmark by

$$-x^- \le x \le x^+.$$

The worst-case portfolio change occurs at one of the end points of this range or at the internal point

$$x = -\frac{D}{G}.$$

In this last case the extreme portfolio change is

$$\delta\Pi_{\text{worst}} = -\frac{D^2}{2G}.$$

We can also calculate the crash delta and gamma at this worst point.

All of the ideas contained in the single-asset model described above carry over to the multi-asset model, we just use x instead of δS to determine the worst that can happen to our portfolio.

If we can't use the delta-gamma Taylor series expansion then we must look for the worst case of an expression such as

$$\delta\Pi = F(\delta S_1, \dots, \delta S_N) = F(\kappa_1 x S_1, \dots, \kappa_N x S_N).$$

This is not hard, or even time consuming, as long as we have formulae for the options in our portfolio.

Suppose you want to look at the exposure of your portfolio of N underlyings in various scenarios; then generally you'd want to plot an $N+1$-dimensional graph, value versus the N underlyings. However, if you want to plot the value of the portfolio in the event of a crash, using the above analysis, this $N+1$-dimensional graph collapses to a more manageable two dimensions, value versus the index.

43.8 PORTFOLIO OPTIMIZATION AND THE PLATINUM HEDGE IN THE MULTI-ASSET MODEL

Suppose that there are M contracts available with which to hedge our portfolio. Let us call the deltas of the kth hedging contract Δ_i^k, meaning the sensitivity of the contract to the ith asset, $k = 1, \dots, M$. The gammas are similarly Γ_{ij}^k. Denote the bid-offer spread by $C_k > 0$, meaning that if we buy (sell) the contract and immediately sell (buy) it back we lose this amount.[1]

Imagine that we add a number λ_k of each of the available hedging contracts to our original position. Our portfolio now has a first-order exposure to the crash of

$$x\left(D + \sum_{k=1}^{M} \lambda_k \sum_{i=1}^{N} \Delta_i^k \kappa_i S_i\right)$$

and a second-order exposure of

$$\tfrac{1}{2}x^2\left(G + \sum_{k=1}^{M} \lambda_k \sum_{i=1}^{N} \sum_{j=1}^{N} \Gamma_{ij}^k \kappa_i S_i \kappa_j S_j\right).$$

Not only does the portfolio change by these amounts for a crash of size x but also it loses a guaranteed amount

$$\sum_{k=1}^{M} |\lambda_k| \, C_k$$

[1] Of course, as explained above, other cost functions could be used.

just because we cannot close our new positions without losing out on the bid-offer spread.[2]

The total change in the portfolio with the static hedge in place is now

$$\delta\Pi = x\left(D + \sum_{k=1}^{M}\lambda_k \sum_{i=1}^{N}\Delta_i^k \kappa_i S_i\right) + \frac{1}{2}x^2\left(G + \sum_{k=1}^{M}\lambda_k \sum_{i=1}^{N}\sum_{j=1}^{N}\Gamma_{ij}^k \kappa_i S_i \kappa_j S_j\right) - \sum_{k=1}^{M}|\lambda_k|\,C_k.$$

And if we can't use the Taylor series expansion? We must examine

$$\delta\Pi = F(\kappa_1 x S_1, \ldots, \kappa_N x S_N) + \sum_{k=1}^{M}\lambda_k F_k(\kappa_1 x S_1, \ldots, \kappa_N x S_N) - \sum_{k=1}^{M}|\lambda_k|\,C_k.$$

Here F is the original portfolio and the $F_k s$ are the available hedging contracts.

From now on I'll stick to the delta-gamma approximation and leave it to you to do the more robust and realistic whole-formulae approach.

43.8.1 The Marginal Effect of an Asset

We can separate the contribution to the portfolio movement in the worst case into components due to each of the underlying:

$$\delta\Pi_i = x^*\Delta_i\,\delta S_i + \frac{1}{2}x^{*2}\sum_{j=1}^{N}\Gamma_{ij}\,\delta S_i\,\delta S_j,$$

where x^* is the value of x in the worst case. This has, rather arbitrarily, divided up the parts with exposure to two assets (when the cross gamma is non-zero) equally between those assets. The ratio

$$\frac{\delta\Pi_i}{\delta\Pi_{\text{worst}}}$$

measures the contribution to the crash from the ith asset.

43.9 THE MULTI-INDEX MODEL

In the same way that the CAPM model can accommodate multiple indices, so we can have a multiple index CrashMetrics model. I will skip most of the details; the implementation is simple.

We fit the extreme returns in each asset to the extreme returns in the indices according to

$$\delta S_i = \sum_{j=1}^{n}\kappa_i^j x_j,$$

where the n indices are denoted by the j sub/superscript.

[2] We can get an idea of which options are good value for Platinum hedging purposes by looking at the ratio $\kappa/\sigma_{\text{imp}}$. The cost of an at-the-money option is roughly proportional to its implied volatility while its effectiveness in a crash depends on its crash coefficient.

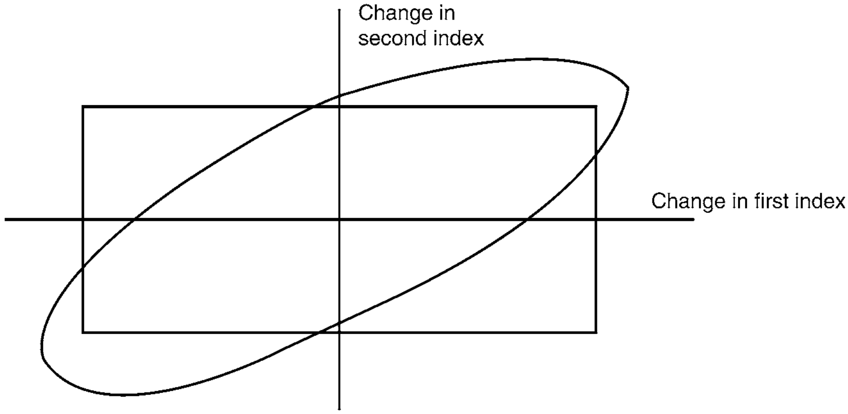


Figure 43.10 Regions of interest in the two-index model.

The change in value of our portfolio of stocks and options is now quadratic in all of the x_js. At this point we must decide over what range of index returns do we look for the worst case. Consider just the two index case, because it is easy to draw the pictures. One possibility is to allow x_1 and x_2 to be independent, to take any values in a given range. This would correspond to looking for the minimum of the quadratic function over the rectangle in Figure 43.10. Note that there is no correlation in this between the two indices; fortunately this difficult-to-measure parameter is irrelevant. Alternatively if you believe that there is some relationship between the size of the crash in one index and the size of the crash in the other you may want to narrow down the area that you explore for the worst case. An example is given in the figure.

43.10 INCORPORATING TIME VALUE

Generally we are interested in the behavior over a longer period than overnight. Can we examine the worst case over a finite time horizon? We can expand the portfolio change in a Taylor series in both δS and δt, the time variable, to get:

$$\delta\Pi - r\Pi\ \delta t = (\Theta - r\Pi)\,\delta t + \Delta\,\delta S + \tfrac{1}{2}\Gamma\,\delta S^2. \tag{43.3}$$

Observe that we examine the portfolio change in excess of the return at the risk-free rate. We must now determine the lowest value taken by this for

$$0 < \delta t < \tau \text{ and } -\delta S^- < \delta S < \delta S^+,$$

where τ is the horizon of interest. Since the time and asset changes decouple, the problem for the worst-case asset move is exactly the same as the above, overnight, problem. The worst-case time decay up to the horizon will be

$$\min\left((\Theta - r\Pi)\,\tau, 0\right).$$

The idea of Platinum Hedging carries over after a simple modification. The modification we need is to the theta. We must incorporate the Θ^* for each of the hedging contracts, suitably multiplied by the number of contracts.

43.11 MARGIN CALLS AND MARGIN HEDGING

Stock market crashes are more common than one imagines, if one defines a crash as any unhedgeable move in prices. Although we have focused on the change in value of our portfolio during a crash this is not what usually causes trouble. One of the reasons for this is that in the long run stock markets rise significantly faster than the rate of interest, and banks are usually net long the market. What causes banks, and other institutions, to suffer is not the paper value of their assets but the requirement suddenly to come up with a large amount of cash to cover an unexpected margin call. Banks can weather extreme markets provided they do not have to come up with large amounts of cash for margin calls. For this reason it makes sense to be 'margin hedged.' Margin hedging is the reduction of future margin calls by buying/selling contracts so that the net margin requirement is insensitive to movements in underlyings. In the worst-case crash scenario discussed here, this means optimally choosing hedging contracts so that the worst-case margin requirement is optimized. Typically, over-the-counter (OTC) options will not play a role in the optimal margin hedge since they do not usually have margin call requirements.

Recent examples where margin has caused significant damage are Metallgesellschaft and Long Term Capital Management.[3]

I now show how the basic CrashMetrics methodology can be easily modified to estimate and ameliorate worst-case margin calls.

43.11.1 What is Margin?

Writing options is very risky. The downside of buying an option is just the initial premium, the upside may be unlimited. The upside of writing an option is limited, but the downside could be huge. For this reason, to cover the risk of default in the event of an unfavorable outcome, the clearing houses that register and settle options insist on the deposit of a margin by the writers of options. Clearing houses act as counterparty to each transaction.

Margin comes in two forms, the initial margin and the maintenance margin. The initial margin is the amount deposited at the initiation of the contract. The total amount held as margin must stay above a prescribed maintenance margin. If it ever falls below this level then more money (or equivalent in bonds, stocks etc.) must be deposited. The levels of these margins vary from market to market.

43.11.2 Modeling Margin

The amount of margin that must be deposited depends on the particular contract. Obviously, we are not too concerned with the initial margin since this is known in advance of the purchase/sale of the contract. It is the variation margin that will concern us since a dramatic market move could result in a sudden large margin call that may be difficult to meet.

[3] The latter suffered after a 'once in a millennium...10 sigma event.' Unfortunately it happened in only their fourth year of trading.

We will model the margin call as a percentage of the change in value of the contract. We denote that percentage by the Greek letter χ. Note that for an over-the-counter (OTC) contract there is usually no margin requirement so $\chi = 0$.

With the same notation as above, the change in the value of a single contract with a single underlying is

$$\delta\Pi = \Delta\ \delta S + \tfrac{1}{2}\Gamma\ \delta S^2.$$

Therefore the margin call would be

$$\delta M = \chi\left(\Delta\ \delta S + \tfrac{1}{2}\Gamma\ \delta S^2\right).$$

The reader can imagine the details of extending this formula to many underlyings and many contracts, and to include time decay.

The final result is simply

$$\delta M - rM\ \delta t = \left(\sum_{i=1}^{N}\overline{\Theta}_i - rM\right)\delta t + \sum_{i=1}^{N}\overline{\Delta}_i\ \delta S_i + \tfrac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\overline{\Gamma}_{ij}\ \delta S_i\ \delta S_j.$$

This assumes that interest is received by the margin. Here

$$\begin{aligned}
\overline{\Theta}_i &= \textbf{Margin Theta} \text{ of all options with } S_i \text{ as the underlying}\\
&= \sum \chi\Theta, \text{ where the sum is taken over all options}\\
\overline{\Delta}_i &= \textbf{Margin Delta} \text{ of all options with } S_i \text{ as the underlying}\\
&= \sum \chi\Delta, \text{ where the sum is taken over all options}\\
\overline{\Gamma}_{ij} &= \textbf{Margin Gamma} \text{ of all options with } S_i \text{ and } S_j \text{ as the underlyings}\\
&= \sum \chi\Gamma, \text{ where the sum is taken over all options}\\
r &= \text{Risk-free interest rate}\\
\delta t &= \text{Time horizon}\\
\delta S_i &= \text{Change in value of } i\text{th asset}
\end{aligned}$$

The notation is self explanatory.

The conclusion is that the CrashMetrics methodology will carry over directly to the analysis of margin provided that the greeks are suitably redefined. We have therefore introduced the new greeks, $\overline{\Theta}$, $\overline{\Delta}$ and $\overline{\Gamma}$, margin theta, margin delta and margin gamma, respectively.

The reader who is aware of the Metallgesellschaft fiasco will recall that they were delta hedged but not margin hedged.

I've assumed a Taylor series/delta-gamma approximation that almost certainly won't be realistic during a crash. We are lucky when modeling margin that virtually every contract on which there is margin has a nice formula for its price. The complex products which require numerical solution are typically OTC contracts with no margin requirements at all. I leave it to the reader to go through the details when using formulae rather than greek approximations.

43.11.3 The Single-index Model

As in the original CrashMetrics methodology we relate the change in asset value to the change in a representative index via

$$\delta S_i = \kappa_i x S_i.$$

Thus we have

$$\begin{aligned}\delta M - rM\ \delta t &= \left(\sum_{i=1}^{N}\overline{\Theta}_i - rM\right)\delta t + x\sum_{i=1}^{N}\overline{\Delta}_i\kappa_i S_i + \tfrac{1}{2}x^2\sum_{i=1}^{N}\sum_{j=1}^{N}\overline{\Gamma}_{ij}\kappa_i S_i\kappa_j S_j\\ &= \overline{Th}\ \delta t + x\overline{D} + \tfrac{1}{2}x^2\overline{G}.\end{aligned}$$

Here $\overline{Th}$ is the portfolio margin theta *in excess of the risk-free growth*. Observe how this expression contains a first- and a second-order exposure to the crash. The first-order coefficient $\overline{D}$ is the crash margin delta and the second-order coefficient $\overline{G}$ is the crash margin gamma.

We've seen worst-case scenarios, Platinum Hedging and the multi-index model applied to portfolio analysis and hedging. All of these carry over unchanged to the case of margin analysis, provided that the greeks are suitably redefined. In particular Platinum Margin Hedging is used to reduce the size of margin calls optimally in the event of a market crash.

43.12 COUNTERPARTY RISK

If OTC contracts do not have associated margin calls, they do have another serious kind of risk: Counterparty risk. During extreme markets counterparties may go broke, having a knock-on effect on other banks. For this reason, one should divide up one's portfolio by counterparty initially and examine the worst-case scenario counterparty by counterparty. Everything that we have said above about worst-case scenarios and Platinum Hedging carries over to the smaller portfolio associated with each counterparty.

43.13 SIMPLE EXTENSIONS TO CRASHMETRICS

In this section I want briefly to outline ways in which CrashMetrics has been extended to other situations and to capture other market effects. Because of the simplicity of the basic form of CrashMetrics, many additional features can be incorporated quite straightforwardly.

First of all, I haven't described how the CrashMetrics methodology can be applied to interest rate products. This is not difficult; simply use a yield (or several) as the benchmark and relate changes in the values of products to changes in the yield via durations and convexities. The reader can imagine the rest.

A particularly interesting topic is what happens to parameter values after a crash. In this case there is usually a rise in implied volatility and an increase in bid-offer spread. The rise in volatility can be incorporated into the methodology by including vega terms, dependent also on the size of the crash. This is conceptually straightforward, but requires analysis of option price data around the times of crashes. If you are long vanilla options during a crash, you will benefit from this rise in volatility. Similarly, crash-dependent bid-offer spread can be incorporated but

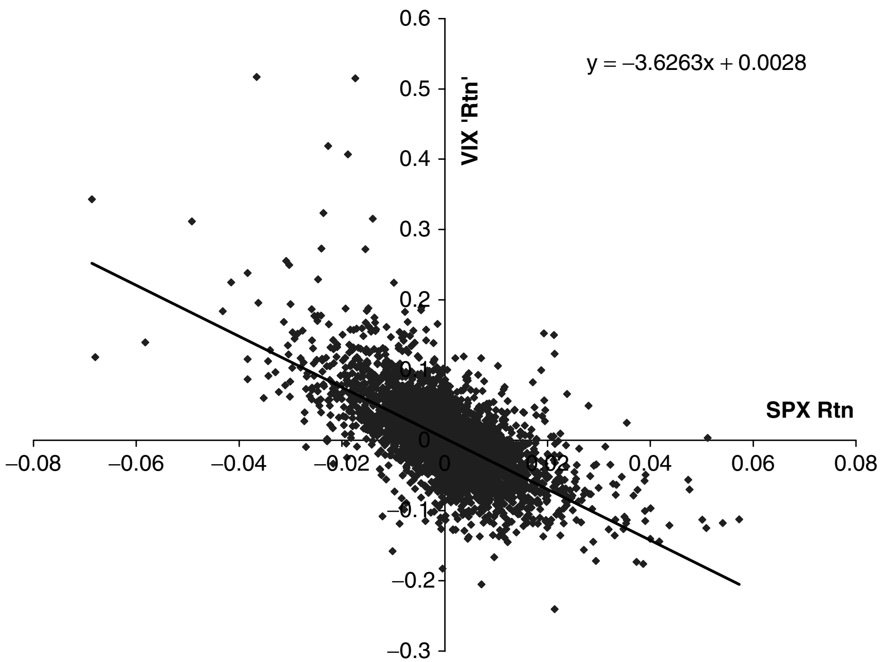


Figure 43.11 VIX return versus SPX return.

again requires historical data analysis to model the relationship between the size of a crash and the increase in the spread (see Figure 43.11).

Finally, it is common experience that shortly after a crash stocks bounce back, so that the real fall in price is not as bad as it seems. Typically 20% of the sudden loss is recovered shortly afterwards, but this is by no means a hard and fast rule. You can see this in the earlier data tables; a date on which there is a very large fall is followed by a date on which there is a large rise. To incorporate such a dynamic effect into the relatively static CrashMetrics is an interesting task.

43.14 THE CRASHMETRICS INDEX (CMI)

The results and principles of CrashMetrics have been applied to a **CrashMetrics Index** (**CMI**). This is an index that measures the magnitude of market moves and whether or not we are in a crash scenario. It's like a Richter scale for the financial world. Unlike most measures of market movements this one is *not* a volatility index in disguise, it is far more subtle than that. However, being proprietary, I can't tell you how it's defined. Sorry. I can give you some clues: It's based on a logarithmic scale; it has only one timescale (unlike volatility which needs a long timescale such as thirty days, and a short one, a day, say); it exploits the effect shown in Figure 43.1. Figure 43.12 shows a time series of the CMI applied to the S&P500.

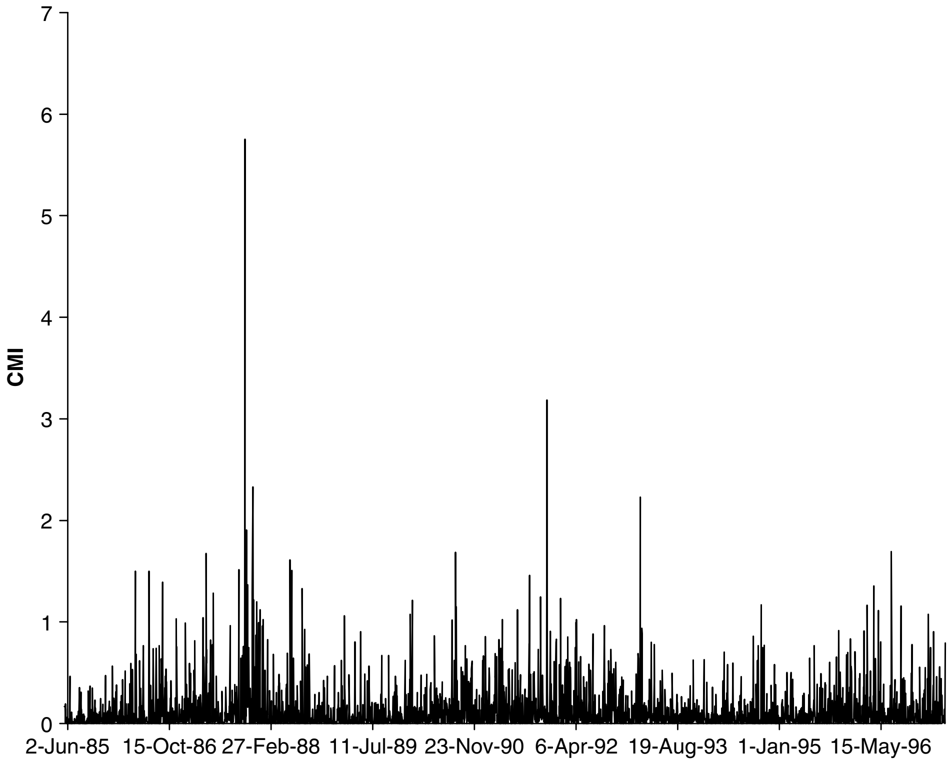


Figure 43.12 The S&P500 CMI.

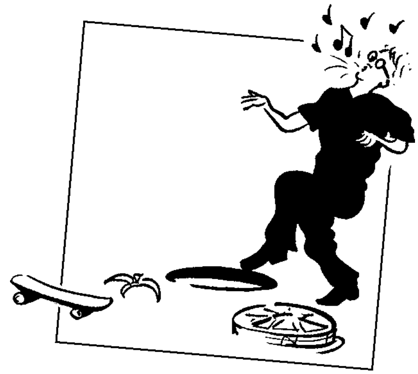
43.15 SUMMARY

This chapter described a VaR methodology that is specifically designed for the analysis of and protection against market crashes. Such analysis is fundamental to the well-being of financial institutions and for that reason I have taken a non-probabilistic approach to the modeling.

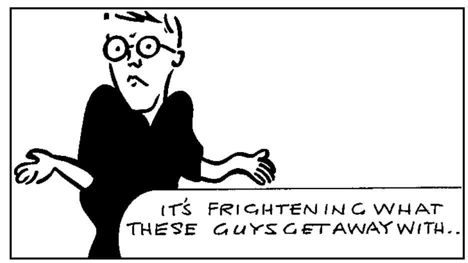
FURTHER READING

- Download the CrashMetrics technical documents, datasets and demonstration software for CrashMetrics from `www.crashmetrics.com`.

CHAPTER 44
derivatives **** ups



In this Chapter...



- Orange County
- Proctor and Gamble
- Metallgesellschaft
- Gibson Greetings
- Barings
- LTCM

44.1 INTRODUCTION

Derivatives, in the wrong hands, can be dangerous weapons. They can destroy careers and institutions. In this chapter we take a look of some of the more well publicized cases of derivatives **** ups. In many cases the details are not in the public domain but where possible I include what *is* known, and offer some analysis.

44.2 ORANGE COUNTY

Orange County is in California. For the first half of the 1990s the County Treasurer was the aptly named Robert Citron. He was in charge of the County investment fund, a pool of money into which went various taxes of the townsfolk. From 1991 until the beginning of 1994 Citron steadily made money, totaling some $750 million, by exploiting low interest rates. This was a very good return on investment, representing approximately 400 basis points above US government rates. How was this possible? He had invested the good people's money in leveraged inverse floating rate notes.

A floating rate note is a bond that pays coupons linked to a floating interest rate such as three- or six-month LIBOR. Normally, the coupon rises when rates rise, not so with inverse floaters. Typically inverse floaters have a coupon of the form

$$\max(\alpha r_f - r_L, 0)$$

where r_f is a fixed rate, r_L some LIBOR rate and $\alpha > 1$ a multiplicative factor. As rates rise, the coupons fall but with a floor at zero; the bond holder never has to return money. Citron

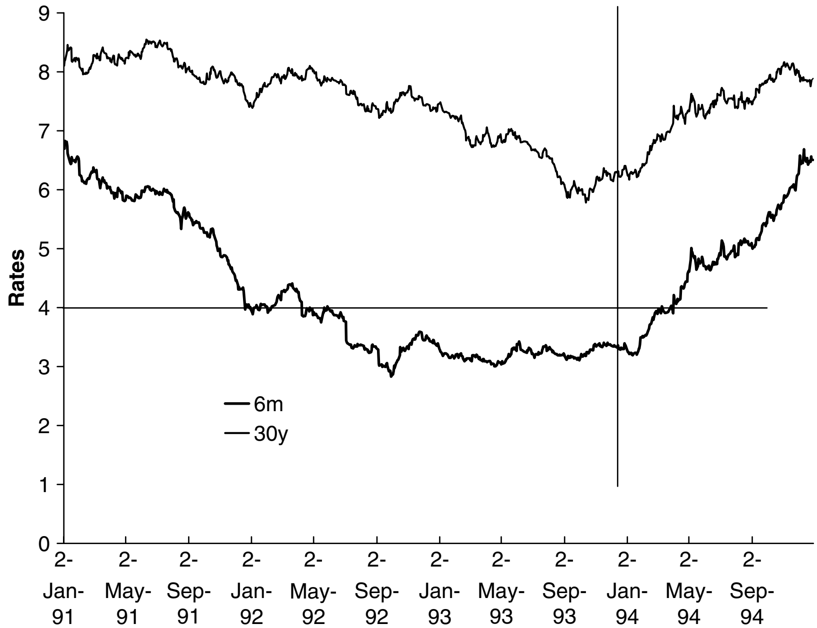


Figure 44.1 US rates in the early 1990s.

bought leveraged inverse floaters[1] having coupons of the form

$$\max(\alpha r_f - \beta r_L, 0)$$

with $\beta > 1$. (If $\beta < 1$ it is a deleveraged note.) While rates are low, coupons are high. This was the situation in the early 90s, and Citron and Orange County benefitted. But what if rates rise? These notes have a high degree of gearing.

Citron gambled that rates would stay low.

In mid-1994 US rates rose dramatically, by a total of 3%. The leveraging in the notes kicked in big time and Orange County lost some \$1.6 billion. Figure 44.1 shows US interest rates during the early 1990s. For about 18 months short-term yields had been below 4%, the horizontal line in the figure. But in early 1994 they started a rise that was dramatic in comparison with the previous period of 'stability.' Citron (and others, see below) were 'caught out' by this rise (to the right of the vertical line in the figure).

Orange County didn't have to declare bankruptcy, they were still in the black and money was still pouring into the fund. But as part of the tactics in their damage suit against the brokers of the deal, they declared bankruptcy on 6th December 1994.

It seems that no one had much of a clue how the losses were piling up. There was no frequent marking to market necessary. Had there been, presumably the losses would have been anticipated and someone would have taken action. At the time of the bankruptcy filing S&P's rated Orange County as AA and Moody's as Aa1; very high ratings and thought very unlikely to default.

[1] Actually, he bought a lot of other things as well. He was not a lucky boy, our Bob.

'Leveraged inverse floating rate note' is a very long name for a very simple instrument. The risks should be obvious. After all, how difficult is it to understand $\max(\alpha r_f - \beta r_L, 0)$?

If Orange County lost who gained? The counterparties selling the floaters were various US government housing agencies. Amusingly, the money flowed from the Orange County taxpayers to the US taxpayers everywhere. Hee, hee.

Citron was found guilty of violating state investment laws and was sentenced to one year of community service. During the sentencing phase psychologists found that he had the math skills of a seventh grader and that he was in the lowest 5% of the population in terms of ability to think and reason.

44.3 PROCTOR AND GAMBLE

Proctor and Gamble (P&G) is a major multinational company that manufactures beauty and health care products, food and beverages, and laundry and cleaning products. They have a large exposure to interest rates and to exchange rates. To reduce this exposure they use interest rate and currency swaps.

In late 1993 P&G wanted to enter into a swap from fixed to floating, having the view that rates then low would remain low. A vanilla swap would be fine as long as rates didn't rise, but what if they did? Bankers Trust (BT), the counterparty to the deal, suggested some modifications to the swap that satisfied P&G's concerns.

The deal, struck on 2nd November 1993 was a five-year swap on a notional $200 million. It contained something a little out of the ordinary, but not outrageously so. P&G had sold BT something like a put on long-term bond prices.

The deal went like this. BT pays P&G a fixed rate of interest on the $200 million for five years. In return P&G pays BT a fixed rate for the first six months, thereafter a rate defined by

$$r_C - 0.0075 + 0.01 \times \max\left(\frac{98.5}{5.78}Y_5 - P_{30}, 0\right) \tag{44.1}$$

where r_C was the rate on P&G's own corporate bonds, Y_5 the five-year Treasury yield and P_{30} the price of the 30-year Treasury bond. The Treasury yield and price would be known at the time of the first payment, 2nd May 1994, *at which time it would be fixed in the formula*. In other words, the yield and price pertaining on that date would be locked in for the remaining four and a half years.

The best that P&G could achieve would be for rates to stay near the level of November 1993 for just a few more months in which case they would benefit by

$$0.0075 \times \$200 \text{ million} \times 5 = \$7.5 \text{ million}.$$

Not a vast amount in the scheme of things. Five- and 30-year rates had been falling fairly steadily for the whole of the 1990s so far (see Figure 44.1); perhaps they would continue to do so, matching the stability of the short-term yields.

However, if rates were to rise between November and May ...

Expression (44.1) increases as the five-year yield increases and decreases if the 30-year bond rises in value. But, of course, if the 30-year yield rises the bond price falls and (44.1) increases. Although there is some small exposure to the slope of the yield curve, the dominant effect is due to the level of the yield curve.

Table 44.1 Effect of parallel shift in yield curve on P&G's losses.

Parallel shift (bps)	0	50	100	150	200
Five-year yield	4.95%	5.45%	5.95%	6.45%	6.95%
Price of 30-year bond	103.02	97.77	93.04	88.74	84.82
30-year yield	5.97%	6.47%	6.97%	7.47%	7.97%
Total loss over 4.5 years $m	0	0	75	190	302

In November 1993 the 6.25% coupon bond maturing in August 2023 had a price of about 103.02, corresponding to a yield of approximately 5.97%. The five-year rate was around 4.95%. With those values expression (44.1) was safely the required $r_C - 0.0075$. However, rates rose at the beginning of 1994 and the potential $7.5 million was not realized, instead P&G lost close to $200 million.

Subsequently, P&G sued BT on the grounds that they failed to disclose pertinent information. The case was settled out of court.

The following was taken from P&G Corporate News Releases

P&G Settles Derivatives Lawsuit With Bankers Trust
May 9, 1996

> CINCINNATI, May 9, 1996–The Procter & Gamble Company today reached an agreement to settle its lawsuit against Bankers Trust. The suit involves two derivative contracts on which Bankers Trust claimed P&G owed approximately $200 million. Under the terms of the agreement, P&G will absorb $35 million of the amount in dispute, and Bankers Trust will absorb the rest, or about 83% of the total.
>
> 'We are pleased with the settlement and are glad to have this issue resolved,' said John E. Pepper, P&G chairman and chief executive.

It's not difficult to work out the potential losses *a priori* from a shift in the yield curve, and I've done just that in Table 44.1 assuming a parallel shift. P&G start to lose out after about a 70 bps rise in the yield curve. Thereafter they lose about $2.3 million per basis point. On 2nd May 1994 the five-year and 30-year rates were 6.687% and 7.328% respectively, an average rate rise of over 150 bps.

In Figure 44.2 is shown the distribution of changes in US five-year rates over a six-month period during the ten years prior to November 1993, data readily available at the time that the contract was signed.[2] These historical data suggest that there is a 14% chance of rates rising more than the 70 bps at which P&G start to lose out (the black bars in the figure). There is a 3% chance of a 150 bps or worse rise. Using these data to calculate the expected profit over the five-year period one finds that it is −$8.7 million, rather than the hoped for +$7.5 million.

When you get to the end of this chapter I want you to do a little exercise. Perform the above parallel shift calculation on a spreadsheet. Time yourself to find out how long it would have taken you to save $200 million. (If you use Excel's built-in spreadsheet functions to calculate yields and prices then it should take less than 10 minutes of typing, and that includes switching on your PC and a comfort break.)

[2] I've cheated a bit in using overlapping data. Using non-overlapping data gives the same, or slightly worse for P&G, results.

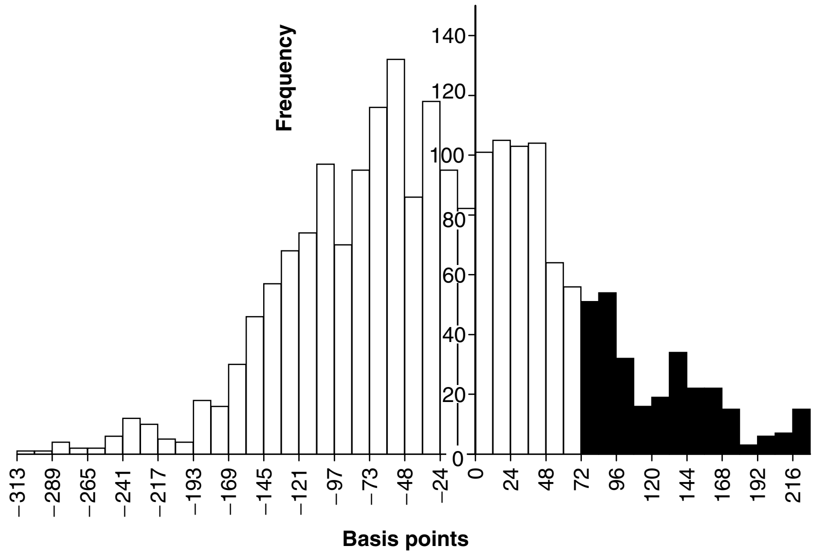


Figure 44.2 Distribution of changes in US five-year rates over a six-month period covering the 10 years prior to November 1993.

The following, taken from the P&G website (`www.pg.com`), seems a decent enough principle (I don't know when it was written):

> Integrity: We always try to do the right thing. We are honest and straightforward with each other. We operate within the letter and spirit of the law. We uphold the values and principles of P&G in every action and decision. We are data-based and intellectually honest in advocating proposals, including *recognizing risks*.

The italics are mine.

44.4 METALLGESELLSCHAFT

Metallgesellschaft (pron. Met Al gazelle shaft, emphasis on 'Al') is a large German conglomerate with a US subsidiary called MG Refining and Marketing (MGRM). In 1992 MGRM issued forward contracts to its clients, locking in the price of heating oil and gasoline for 10 years. The forward price was fixed at about $3 above the then spot prices. Each month the client received a delivery of oil, paying fixed price. The contract also allowed for both parties to close the position. For example, the client could cancel the contract at any time that the shortest-dated oil futures price exceeded the fixed price. On exercise they would receive 50% of the difference in price between the short future and the fixed price, multiplied by the total volume of oil remaining on the contract. MGRM could also close some of the contracts if the short futures price exceeded some prescribed exit price.

These contracts proved popular with clients because they were the only long-dated contracts available with which to lock in a fixed price. No such contracts existed on an exchange. The total volume of oil in the contracts amounted to some 180 million barrels, the equivalent of 85 days of the output of Kuwait.

Since these contracts were OTC forwards no money changed hands until each delivery. At which time the net cash flow to MGRM was fixed minus spot. The lower the spot price the more that the contracts were of value to MGRM.

MGRM naturally wanted to hedge the oil price risk. But the only exchange traded contracts available were one- to three-month futures. MGRM implemented a strategy of hedging the long-term short OTC forward position with long positions in short-term futures traded on the New York Mercantile Exchange (NYMEX). Because the short-term contracts expired every few months the position had to be rolled over; as one position expired so another was entered into for the next shortest maturity. Theoretically this strategy would have been successful provided that MGRM had a good model for interest rates and the cost of carry. We will return to this point in a moment. Even if MGRM did have a decent model they hit problems because of the important distinction between futures and OTC forwards.

Oil prices fell during the latter half of 1993. Because they held long positions in the futures, a fall in price had to be met on a daily basis by an increase in margin. Futures are marked to market. The extra margin requirements amounted to \$900 million during 1993, a large sum of money. MGRM turned to its parent company to help with the funding. Metallgesellschaft responded by taking control of MGRM, installing a new management. In December 1993 the new management closed out half of the short-term contracts. Oil prices then started to rise in early 1994 so that MGRM started to lose out on the long-term contracts. Their response was to cancel the OTC forward contracts and close all positions. By this time losses had amounted to \$1.3 billion.

But were they really losses? MGRM were losing out on the short-term futures positions as oil prices fell but remember that these contracts were for hedging purposes. As they lost money on the hedging position they also made money on the OTC forward contracts. The problem was that because these were forward contracts the profit on them was not realized until the positions expired. Marking to model should have resulted in a net flat position regardless of what the oil price did. Think back to margin hedging, discussed in Chapter 43.

Some say that the MGRM management panicked, they say otherwise.

44.4.1 Basis Risk

There was a slight complication in this story due to the behavior of futures prices. The relationship between forward prices and spot prices is not as simple in the commodity markets as it is in the FX markets, for example. Arbitrage considerations lead to the theoretical result

$$F = Se^{(r+q)(T-t)}$$

where T is the maturity, r is the relevant risk-free yield and q is the cost of carry. This relationship leads to forward prices that are higher than spot prices, the graph of forward prices as a function of maturity would be upward sloping. In this case the market is said to be in **contango**.

In practice the strategy required to take advantage of the arbitrage opportunity is so impractical, involving the buying, transporting, storing, transporting, ... of the commodity, that the theoretical arbitrage is irrelevant. It is therefore possible, and even common, for the forward curve to be downward sloping. The market is said to be in **backwardation**.

In the Metallgesellschaft story the oil markets were in a state of backwardation initially. By rolling over the futures contract it was possible to make a profit, benefitting from the slope of the forward curve at the short end. During 1993 the spot oil price fell sharply and the market moved into contango. Now the rolling over led to losses, amounting to $20 million per month. The question remains, did MGRM price into the OTC forward contracts this possible behavior of effectively the cost of carry? If they did (which seems unlikely, or even impossible *a priori*) then their hedging strategy would have been successful had it not been cut off in its prime.

It is difficult to build an accurate interest rate and/or cost of carry model. MGRM was therefore exposed to the risk of hedging one instrument with an imperfectly correlated one. Such a risk is generally termed **basis risk**.

44.5 GIBSON GREETINGS

> 'I think we should use this as an opportunity. We should just call [the Gibson contact], and maybe chip away at the differential a little more. I mean we told him $8.1 million when the real number was 14. So now if the real number is 16, we'll tell him that it is 11. You know, just slowly chip away at the differential between what it really is and what we're telling him.
>
> ' ... when there's a big move, you know, if the market backs up like this, and he is down another 1.3, we can tell him he is down another 2 ... If the market really rallies like crazy, and he's made back a couple of million dollars, you can say you have only made back a half a million.'
>
> February 23, 1994, BT Securities tape of a BT Securities manager discussing the BT internal valuation of the Gibson's positions and the valuation given to Gibson.

Gibson Greetings is a US manufacturer of greetings cards. In May 1991 they issued $50 million worth of bonds with a coupon of 9.33% and with maturities of between four and 10 years. In the early 1990s interest rates fell and Gibson were left paying out a now relatively high rate of interest. To reduce the cost of their debt they entered into vanilla interest rate swaps with Bankers Trust in November 1991.

Swap 1

Notional $30 million, two-year maturity, BT pay Gibson six-month LIBOR and Gibson pay BT 5.91%.

Swap 2

Notional $30 million, five-year maturity, BT pay 7.12% and Gibson pay six-month LIBOR.

For the first two years the LIBOR cashflows cancel and Gibson receive $7.12 - 5.91 = 1.21\%$.

In July 1992 both swaps were canceled. Shortly afterwards Gibson entered into a more leveraged swap contract, and so began a sequence of buying and canceling increasingly complex products. Some of these products are described below.

Ratio swap

For five years Gibson were to pay

$$\frac{50}{3}\ r_L^2$$

to BT every six months and receive 5.5% on a notional of $30 million. Here r_L is six-month LIBOR. Since the first fixing of LIBOR was at approximately 3.08% the first payment by Gibson was just 1.581%. The second payment was 1.893% corresponding to LIBOR of 3.4%.

Interest rates were starting to rise. And judging by the implied forward curve at the time the market was 'expecting' rates to go higher. Should six-month LIBOR reach 5.7% ($=\sqrt{0.06 \times 0.055}$) the net cashflow would be towards BT from Gibson. Fearing that rates would go beyond this level, the swap contract was amended three times and finally canceled in April 1993, six months after its initiation.

LIBOR did not rise as rapidly as the forward curve had suggested (it usually doesn't). With hindsight it probably would have been in Gibson's best interests to have retained the swap in its original form.

Periodic floor

For five years BT were to pay 28 bps above six-month LIBOR on a notional of \$30 million. Gibson were to pay

$$r_L.$$

The BT payments were only to be paid while $r_L > r'_L - 0.15\%$ where r'_L is six-month LIBOR measured at the previous swap date. This is a path-dependent contract. At the end of 1992 BT informed Gibson that the value of the periodic floor was negative and it was canceled nine months after its initiation. The loss was incorporated into a later contract.

Treasury-linked swap

This swap was entered into in exchange for BT decreasing the maturity of the above ratio swap. Gibson were to pay BT LIBOR and BT were to pay Gibson LIBOR plus 2%. The catch was in the amount of the principal repayment at maturity. Gibson were to pay \$30 million while BT were to pay

$$\min\left(\$30.6\text{ million}, \$30\text{ million} \times \left(1 + \frac{103 \times Y_2}{4.88} - \frac{P_{30}}{100}\right)\right),$$

where Y_2 is the two-year Treasury yield and P_{30} is the 30-year Treasury price. Does this formula look familiar?

Knock-out call option

To reduce its exposure under the terms of the Treasury-linked swap due to an anticipated small principal repayment by BT, Gibson entered into a knock-out call option in June 1993. In return for an up-front premium Gibson were to receive at expiry

$$12.5 \times \$25\text{ million} \times \max(0, 6.876\% - Y_{30})$$

where Y_{30} is the yield on 30-year Treasuries. The downside for Gibson was that the contract expired if the 30-year Treasury yield fell below 6.48%. Long-term rates would have had to stay remarkably stable since the payoff would only be received if $6.48\% < Y_{30} < 6.876\%$ at expiry with Y_{30} *never* having fallen below 6.48%.

Table 44.2 Ranges for LIBOR in the Corridor Swap.

6 Aug 1993–6 Feb 1994	3.1875–4.3125%
6 Feb 1994–6 Aug 1994	3.2500–4.5000%
6 Aug 1994–6 Feb 1995	3.3750–5.1250%
6 Feb 1995–6 Aug 1995	3.5000–5.2500%

The time swap/corridor swap

In August 1993 Gibson entered into a corridor swap with BT. The maturity of the contract was three years and the notional $30 million. BT were to pay Gibson six-month LIBOR plus 1%. Gibson's payments were as follows:

$$r_L + 0.05 \times N\%$$

where r_L is six-month LIBOR and N is the number of days during each six-month calculation period that r_L was outside the range specified in Table 44.2.

After much modification of the multiplier (0.05 initially), the ranges and the termination date, the contract was canceled in January 1994.

In all there were approximately 29 contracts. Gibson's losses amounted to over $20 million, equivalent to almost a year's profits. They sued BT and in an out-of-court settlement agreed in November 1994 they were released from paying BT $14 million outstanding from some swap contracts, only paying BT $6.2 million. Gibson's argument was that they had been misled by BT as to the true value of their contracts. As the tape recording transcribed at the start of this section shows, BT had their own internal models for the value of the contracts but consistently understated the losses to Gibson.

44.6 BARINGS

The Barings story is actually very simple, and the role of derivatives is relatively small. The reasons for all the fuss are the magnitude of the losses ($1.3 billion), that it involved a very staid, 200-year old bank, and that its main protagonist, 28-year old Nick Leeson, did a runner.

Nick Leeson was a trader for Barings Futures, working in Singapore. In 1994 he was reported to have made $20 million for Barings and was expecting to be rewarded with a bonus of $1 million on top of his $150,000 basic. It is likely that he had in fact lost a considerable amount already, with the real position only appearing in the infamous 'Error Account 88888.' Leeson had been given too much freedom. At the time, late 1994 and early 1995, Leeson had control over both a trading desk and back office operations; in effect he was allowed to police his own activities.

The trades that caused the downfall of Barings involved the Japanese stock market, specifically futures and options on the Nikkei 225 index. In November and December 1994 Leeson had been selling straddles on the Nikkei with strikes in the region of 18500–20000. The Nikkei index was then trading in a similar range. As long as the index was stable and volatility remained low he would profit from this almost delta-neutral strategy. On 17th January 1995 an earthquake hit Kobe and the Nikkei started to fall (see Figure 44.3). The figure also shows Leeson's trading range. Over the next few days Leeson began buying index futures with March expiry. One of the reasons behind this strategy was the belief that the magnitude of his trades would

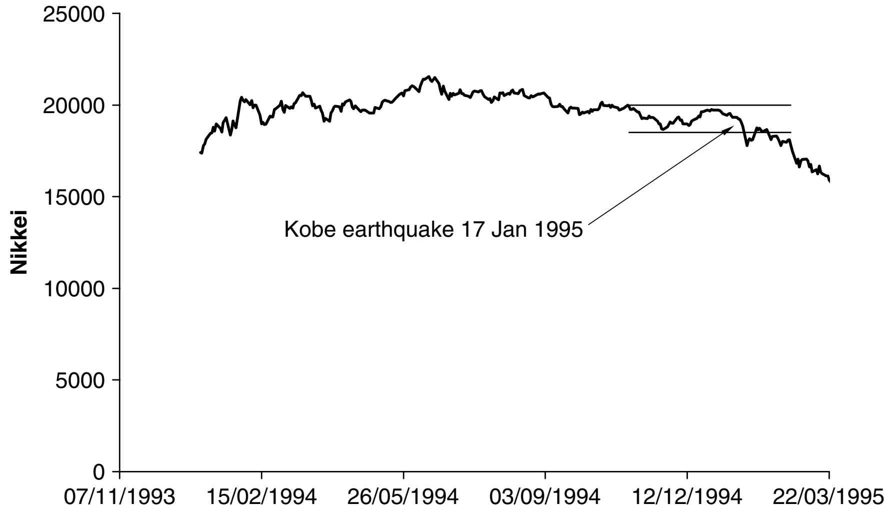


Figure 44.3 The Nikkei 225 index from the beginning of 1994 until March 1995.

act to reverse the declining market. If he could bring the Nikkei up to the pre-earthquake levels his option positions would be safe.

Although his trades had a significant impact on the index they could not hold back the fall. Over the next month the index was to fall to 17400. As the index continued its fall, Leeson increased his trades to shore up the index. Margin was required for the futures mark-to-market and vast sums of money were transferred from Barings in London. Finally, the margin calls became too much to cover. On February 23 1995 Nick Leeson went on the lam, fleeing Singapore. Finally, he took a flight to Frankfurt, where he was arrested. On December 1 1995 he pleaded guilty to two offenses of deceiving Barings' auditors in a way likely to cause harm to their reputation as well as cheating SIMEX. The next day he was sentenced to six and a half years in prison.

Leeson was released from the Tanah Merah wing of Changi jail on 3rd July 1999. While in prison he developed cancer of the colon, his wife divorced him and remarried, and they made a movie of the story starring Ewan McGregor and Anna Friel.

Meanwhile Barings went broke. The Sultan of Brunei was approached to bail them out but declined. The Dutch bank ING 'was finally persuaded to take on the corpse of Barings,' as Richard Thomson put it, for the grand sum of £1.

44.7 LONG-TERM CAPITAL MANAGEMENT

Long-Term Capital Management (LTCM) is a hedge fund. Hedge funds are supposed to hedge, you'd think. Yet there are few regulations governing their activities. The term 'hedge fund' came about because these funds take short as well as long positions, but this is *not* the same as hedging. LTCM was founded by John Meriwether, ex-Salomon bond arbitrage team, with Nobel Laureates Myron Scholes and Robert Merton as partners in the firm. (The excellent book by Lewis (1989) is an account of dodgy dealings at Salomon's during Meriwether's time there.)

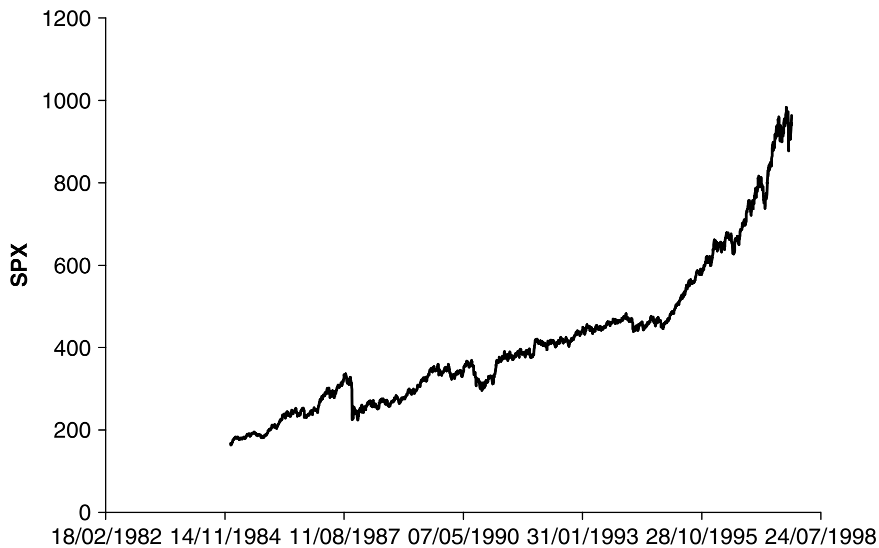


Figure 44.4 The S&P 500.

Edward Thorp (of Blackjack fame, see Chapter 18) started one of the first hedge funds in 1969. It has been a very successful hedge fund and Thorp was invited to invest in LTCM in 1994 when it was founded. Thorp declined: 'I didn't want to have anything to do with it because I knew these guys were just dice rollers . . . It was just a mutual admiration society at Long-Term and nobody was focusing clearly enough on the model.'

If Thorp didn't want to invest, he was in the minority. It was a glamorous line-up and many who should have known better got sucked in. For the first couple of years they made good returns, of the order of 40%. But look at Figure 44.4: *everyone* was making 40%. And in 1997 they made 27%.

The hype surrounding LTCM, its legendary team and the friends they had in high places meant that they had three benefits that other firms would kill to have:

- They were able to leverage \$4.8 billion into \$100 billion. Its notional position in swaps was at one time \$1.25 trillion, 5% of the entire market;[3]
- They were excused collateral on many deals;
- When they finally went under, the Federal Reserve organized a bail out.

During 1998 LTCM took huge leveraged bets on the relative value of certain instruments. They were expecting a period of financial calm and convergence of first and emerging world interest rates and credit risk. Here are some of the strategies LTCM had in place, what they gambled would happen and what actually did happen. Observe how many of these trades come in pairs; LTCM were making relative value trades.

[3] Bill Ziemba has estimated that they were overbetting to about twice the Kelly level. Recall from Chapter 17 that this is the level at which expected growth rate starts to become negative.

European government bonds

LTCM sold German bonds and bought other European bonds in the expectation of interest rate convergence in the run up to EMU. They did ultimately converge but not before first diverging further.

Emerging market bonds and US treasuries

LTCM had long positions in Brazilian and Argentinian bonds and short positions in US Treasuries. They expected credit spreads to decrease; instead they widened to as much as 2000 bps.

Russian GKOs and Japanese bonds

They expected Russian yields to fall and Japanese to rise, and so bought Russian GKOs and sold Japanese bonds. They were wrong about the direction of Japanese yields, and Russia defaulted.

Long- and short-term German bonds

LTCM bought 30-year German bonds, sold 10-year bonds, expecting the German yield curve to flatten. Instead, demand for the short-term bonds caused the yield curve to steepen.

Long- and short-dated swaption straddles

They bought long-dated swaption straddles and sold short-dated swaption straddles. This is an almost delta-neutral strategy and therefore a volatility trade. Short-term volatility rose, resulting in losses.

The main point to note about these trades is that all seem to have the same view of the world market. They were expecting a period of relative stability, with emerging markets in particular benefitting. In fact, the default by Russia on August 17th 1998 sent markets into a panic, investors 'fled to quality' and all the above trades went wrong. LTCM had most definitely made a big bet on one aspect of the market, and were far from being diversified. Along the way the partners lost 90% of their investment and a couple of Nobel Prize winners had very red faces. They say on Wall Street that if you lose \$5 million you'll never work on the Street again, lose \$50 million and you'll walk into a new job the next day. But lose billions? It's good to have friends.

Figure 44.5 shows the price of the $9\frac{1}{4}\%$ 2001 USD Argentine bond. Prior to the August Russian default the price is quite stable. But then the price plummeted in the overall panic. As things calmed down the price returned. LTCM may have been correct in their views *long term*, but if you can't weather the storm in the *short term* your correct market view is irrelevant. As hedge fund is a misnomer, so, perhaps, is LTCM … a more accurate name might be STCMM.

On 21st August 1998 they managed to lose \$550 million, and there was more to come.

It was deemed too dangerous for LTCM to be allowed to fail completely, as the impact on the US economy could have been disastrous. So the New York Federal Reserve organized, in September 1998, a bailout in which 14 banks invested a total of \$3.6 billion in return for a 90% stake in LTCM.

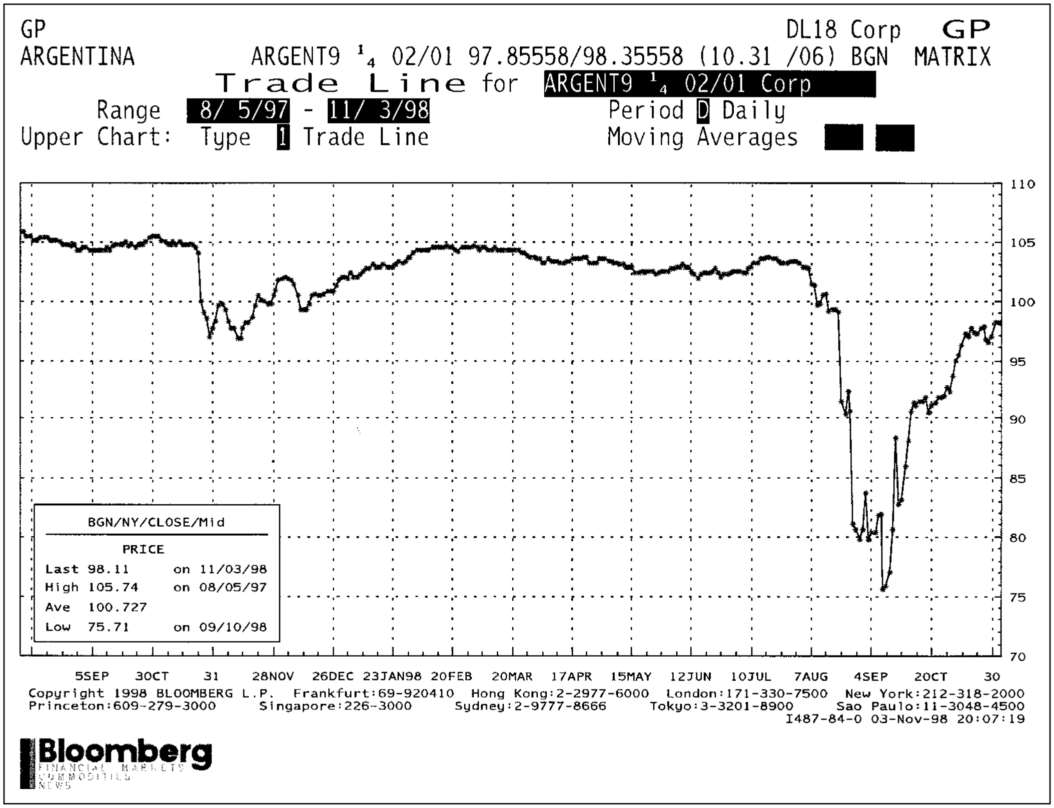


Figure 44.5 Time series of the $9\frac{1}{4}$% 2001 USD Argentine bond. Source: Bloomberg L.P.

Much of the LTCM story is still unknown; they were so glamorous that many of their investors didn't even know what they were up to. However, it seems that LTCM were using only rudimentary VaR estimates, with little or no emphasis on stress testing. They estimated that daily swings in the portfolio should be of the order of $45 million. But this didn't allow for extreme market moves and problems with liquidity. In trying to reduce risks LTCM even sold off liquid assets, leaving the (theoretically) more profitable illiquid trades to stay on their books. Not a sensible move at times of crisis. Simple VaR as described in Chapter 19 is fine as far as it goes, but cannot deal with global financial meltdowns.

'When we examine banks, we expect them to have systems in place that take account of outsized market moves.' Alan Greenspan, chairman of the Federal Reserve.

There's another aspect to the LTCM tale that should make anyone using simple quantitative analysis a little bit wary, and that is the matter of liquidity. In August 1998 there was a worldwide drying up of liquidity. Without liquidity it is impossible to offload your positions and the idea of 'value' for any product becomes meaningless. You just have to wait until the market decides to loosen up. Part of the problem was that many of the banks they dealt with knew of LTCM's trades. Many of these were simply copying LTCM's strategies. This wouldn't have mattered if LTCM and all these banks were dealing in small quantities. However, they weren't and in some cases they completely cornered the market. It's one thing to corner the market in one of the necessities of life, such as beer, but to corner the market in something obscure makes you a sitting duck when you want to sell.

44.8 SUMMARY

Most of these stories have similar themes: Over confidence, lack of understanding of the risks, pure speculation at inappropriate times or for inappropriate reasons, or over-gearing.

We're only human but what is the point of the math modeling when any profits get thrown away by a few individuals who don't know what they are doing? If they *do* know what they are doing, I would assume they are crooks.

I'll end with a note of caution sounded by Robert Merton in 1993. 'Any virtue can become a vice if taken to extreme, and just so with the application of mathematical models in finance practice... At times the mathematics of the models become too interesting and we lose sight of the models' ultimate purpose. The mathematics of the models are precise, but the models are not, being only approximations to the complex, real world... The practitioner should therefore apply the models only tentatively, assessing their limitations carefully in each application.' Doh!

FURTHER READING

- Leeson (1997) explains the Barings disaster from his own perspective.
- Chew (1996) is an excellent book covering the risks of derivative transactions. She discusses the technical side of contracts, the legal side and the morality.
- Partnoy (1998) is an cracking good tale of goings on at Morgan Stanley (coincidentally one of my least favorite banks).
- For background on Meriwether, the LTCM partner, see Lewis (1989).
- Dunbar (1998), Kolman (1999) and Jorion (1999) are nice accounts of LTCM.
- Miller (1997) discusses Metallgesellschaft in detail.
- Thomson (1998) has plenty of inside gen on many of the derivatives stories.
- Read Merton (1995) for his thoughts on the influence of mathematical models on the finance world.

PART FIVE
advanced topics

This part of the book concerns extensions to the Black–Scholes world. All of the assumptions in the Black–Scholes model are incorrect to a greater or lesser degree. I will show how to incorporate ideas aimed at improving the modeling of the underlying asset. Some of these ideas are relatively standard, fitting easily into a classical Black–Scholes-type framework. Others take us far from the well-trodden path into new and uncharted territories. One of the most important points to watch out for is whether or not a model is linear or non linear; is the value of a portfolio of options the same as the sum of the values of the individual components? I believe that non-linear models capture some important features that are neglected by the more common linear models. I will expand on this point at various times throughout the rest of the book.

Chapter 45: Financial Modeling A tongue-in-cheek look at the process of modeling in quantitative finance; what to watch out for when choosing your model.

Chapter 46: Defects in the Black–Scholes Model Many deficiencies in the simple Black–Scholes model are brought together in this chapter for a general discussion. The issues raised here will be expanded upon in later chapters.

Chapter 47: Discrete Hedging Even if all of the other assumptions that go into the Black–Scholes model were correct, the impossibility of delta hedging continuously in time would have an enormous effect on the actual profit and loss experienced during the business of hedging.

Chapter 48: Transaction Costs This was one of the first areas of research for me when I discovered mathematical finance. Unfortunately, most of the advanced work in this subject cannot easily be condensed into one chapter of a book. For that reason I will summarize many of the ideas and results, giving details only for the Leland model for vanilla options and its Hoggard–Whalley–Wilmott extension to arbitrary contracts.

Chapter 49: Overview of Volatility Modeling The next few chapters focus on the most important financial unknown, volatility. So this chapter puts the following into context.

Chapter 50: Deterministic Volatility Surfaces Volatility is the most important parameter in the valuation of options. It is difficult to measure statistically from time series data, however, we can tell from the prices of traded instruments what the market thinks is the correct value or structure for volatility.

Chapter 51: Stochastic Volatility It is natural to model volatility as a stochastic variable because it is so clearly not constant and not predictable.

Chapter 52: Uncertain Parameters The difficulties associated with observing, estimating and predicting important parameters means that we should try to reduce the dependence of a price on such quantities. One way of doing this is to allow parameters to lie in a prescribed range; we don't even need to know their value at any point in time if we price in a worst-case scenario. This is the Avellaneda, Levy, Parás and Lyons model for uncertain volatility.

Chapter 53: Empirical Analysis of Volatility It is not too difficult to find a decent stochastic model for volatility if one knows how to examine the data. I make a few suggestions in that direction.

Chapter 54: Stochastic Volatility and Mean-variance Analysis Can you really hedge away volatility risk?

Chapter 55: Asymptotic Analysis of Volatility Models are often chosen for their tractability, meaning that they are simple to solve. Here I show a very useful technique for solving, approximately, very general models. Now you don't have to be restricted in your choice of model.

Chapter 56: Volatility Case Study: The Cliquet Option The cliquet option is a particularly interesting exotic option. It looks quite harmless, but as we see, it has a very subtle dependence on the volatility model.

Chapter 57: Jump Diffusion One of the real-life features not captured by the Black–Scholes world is that of discontinuous asset price paths. I describe the Merton model for jump-diffusive random walks.

Chapter 58: Crash Modeling The Merton model for jumps requires quite a few parameters to be input, as well as an estimate of the distribution of jump sizes. This, together with the impossibility of hedging in the Merton world, makes the model somewhat unsatisfactory. If we price assuming the worst-case scenario then most of these difficulties vanish.

Chapter 59: Speculating with Options Option theory is built on the idea of hedging. But what about those people who use options for speculation? Very little is ever said about how they can make investment decisions. I will try to remedy that in this chapter.

Chapter 60: Static Hedging The nonlinearity in many of the preceding models means that the value of a portfolio of contracts is not the same as the sum of the values of the individual components. Thus the value of a contract depends on what you hedge it with. The beautiful consequences of this are discussed in this chapter.

Chapter 61: The Feedback Effect of Hedging in Illiquid Markets This chapter describes a highly non-classical detour away from the Black–Scholes world. What if there is so much trade in the underlying asset for the hedging of derivatives that the normal causal link between the underlying and the option gets confused?

Chapter 62: Utility Theory What if you can't hedge or don't want to? You'll be left with some uncertainty/randomness/risk. Utility theory is a way of assigning a value to a random outcome. This topic is popular with economists but not with practitioners. I used not to like the subject as well but now I feel that sometimes you don't have a choice but to *carpe urticum*.

Chapter 63: More About American Options and Related Matters There is some confusion concerning when it is best to exercise an American option. I try to clear up this confusion in this chapter. You've seen the theory from the point of view of the option writer in Chapter 9, now read about the option holder's strategy.

Chapter 64: Advanced Dividend Modeling Stock dividends can have a big impact on the value of an option. In this chapter we look at sophisticated models for dividends to try and better capture the effect.

Chapter 65: Serial Autocorrelation in Returns Serial autocorrelation means that an asset price return is dependent on the past, for example, today's return depends on whether the stock went up or down yesterday.

Chapter 66: Asset Allocation in Continuous Time The models and ideas presented in a previous chapter concern one-period investments. Faced with the possibility of buying and selling assets at times of your choosing, how is the portfolio optimally chosen?

Chapter 67: Asset Allocation Under Threat of a Crash We build on the ideas of portfolio allocation and crash modeling to examine how to allocate money to an investment when you are worried that the investment might suddenly plummet.

Chapter 68: Interest-Rate Modeling Without Probabilities There are many theoretical and practical problems and inconsistencies in classical stochastic interest rate models. In this chapter I describe the Epstein–Wilmott short-rate model that sidesteps many of these problems by being non Brownian. Instead of modeling the short rate as a stochastic process, we model it in a deterministic fashion but with uncertain parameters. This is again a worst-case scenario analysis.

Chapter 69: Pricing and Optimal Hedging of Derivatives, the Non-probabilistic Model Cont'd The Epstein–Wilmott model can be used to price interest rate derivatives. We look at the pricing of some representative contracts and make comparisons with more traditional models. Because the model is non linear we can optimize static hedges to get the best price.

Chapter 70: Extensions to the Non-probabilistic Interest-rate Model Because the EW model is so simple, it can be extended in many, many ways. In this chapter we look at some of those extensions.

Chapter 71: Modeling Inflation We examine data for inflation so as to develop a model for instruments such as inflation-linked bonds.

Chapter 72: Energy Derivatives Energy is a commodity and is traded, in the form of oil, gas or electricity. There are some issues that are specific to these markets, that we have not seen so far.

Chapter 73: Real Options Option theory can be applied to many business decisions, whenever there is some choice such as whether to invest in R&D for a new product, whether to open a new branch of a shop, etc. In this chapter I explain some of the math behind this subject of Real options.

Chapter 74: Life Settlement and Viaticals A chapter on life insurance policies, and portfolios thereof. We use the same ideas as with traditional financial instruments but instead of default we have death.

Chapter 75: Bonus Time Many traders are paid outrageous amounts in 'compensation.' Is this sensible, does it correctly motivate the trader? How can you tell whether a trader is skillful or plain lucky?

CHAPTER 45
financial modeling

In this Chapter...

- some advice on modeling from an old cynic
- the 'Find-and-Replace' school of mathematical modeling

45.1 INTRODUCTION

> I really must say that you are an ignorant person, friend Greybeard, if you know nothing of this enigmatic business which is at once the fairest and most deceitful in Europe, the noblest and the most infamous in the world, the finest and most vulgar on earth. It is a quintessence of academic learning and a paragon of fraudulence; it is a touchstone for the intelligent and a tombstone for the audacious, a treasury of usefulness and a source of disaster, and finally a counterpart of Sisyphus who never rests as also of Ixion who is chained to a wheel that turns perpetually.
>
> Joseph de la Vega in *Confusión de Confusiones*, 1688.

This book is about modeling of financial quantities and instruments. It starts out covering much of the classical foundations, and leads on to pretty stratospheric, cutting-edge stuff. Within the classical foundations the book covers lognormal random walks, dynamic hedging, the binomial model and the Black–Scholes model, as well as basic portfolio management. But these topics are covered in a near-infinite number of finance textbooks already, with varying degrees of success. One level up from the classical basics, we also look at the more advanced topics such as volatility surfaces, stochastic volatility, jump diffusion, etc. These may be more advanced but they are still subjects that are covered in some of the better-quality books. And they are topics dear to the heart of the more 'sophisticated' practicing quant. What you won't find in many other books are the more interesting cutting-edge material, the non-linear models, the data analysis, the 'research' topics. These chapters are my personal favorites. It is these chapters that make this book what it is, a journey from pretty straightforward material to places that most researchers never even dream of going. And what I have tried to achieve is to make this journey as seamless, and unfrightening, as possible. You should never be aware of the times when we leave the well-trodden path into the more uncertain terrain.

In designing the map and the guide book for this journey I have always assumed a certain level of intelligence of the reader. There is no spoonfeeding in this book. Simultaneously, I have tried to give the best and most straightforward explanations of ideas and methods. There is a lot of material in this book and to make it as readable as possible, I have eliminated all padding. That means that I have tried to miss out bits appropriately. The word ‘appropriately’ is key here. Once you’ve seen the application of Itô’s lemma and delta hedging a few times then we can safely start cutting to the chase; but then I may linger on more subtle points in an argument, points which may seem to be of no great importance. Well, there is method in my madness, and any such lingering is usually because of some point of theory that the real world fails to live up to, something which can be of tremendous importance to the practitioner or anyone wanting to do a decent job of capturing reality with their models.

45.2 WARNING: MODELING AS IT IS CURRENTLY PRACTICED

In this book we see the classical models and the not so classical. We see financial theories as currently used in practice and models that no one touches. At first you may find it difficult to decide which models are the best, and which ones are used most often. So let me give you a few tips that I have found useful. They may not help you with modeling, but they may make your career path less rocky.

Rule 1: The simpler the model the more popular

So far so good. I like this rule. The best math is the simplest math. In practice I find that the robustness and transparency of a simple model (like the Black–Scholes model) can often outweigh the supposed accuracy of the more sophisticated models.[1]

Rule 2: Closed-form solutions are more popular than models requiring numerical solution

Not so keen on this. Given everyone’s access to phenomenal computing power I see this as just people being frightened of number crunching. (And this view is backed by the popularity of the binomial method for pricing, the lowest common denominator of numerical methods.) We aren’t going to be frightened in this book, are we?

Rule 3: Never use real data to test your model

This shocks me, as someone who analyzes data on a daily basis. But think about it, what’s the upside? The fact of the matter is that all financial models are inaccurate, so there’s no incentive to demonstrate the bleedin’ obvious by testing your models. Erm? But surely you should at least find out *how* bad your model is? And which ones are better than the others. Apparently not. Anyway, it turns out that very few people bother with the messy bit of actual testing, at least on the sell side of investment banks. I imagine that there’s somewhat more analysis done on the sensible buy side, the hedge funds for example.

Rule 4: People don’t like models that are ‘different’

‘It is better to fail conventionally than to succeed unconventionally’ as Keynes put it. Most of the models you will see in this book are of the Brownian motion/diffusion/parabolic partial differential equation type. (Those words all mean just about the same thing mathematically.) No one is threatened by a model of this type, everyone understands them. There are other models out there, and we’ll see a few, that inhabit other mathematical worlds, but they aren’t

[1] Einstein’s law of modeling says make things as simple as possible, but no simpler. Wilmott’s law says if you must complicate things, do so one step at a time.

popular. They aren't popular because they are different. If ever a subject had scope for creativity in modeling it is quantitative finance. But, human nature being such as it is ... This leads naturally on to the process of modeling.

45.2.1 Models, a Personal View

In 1973 Black, Scholes and Merton published their work on pricing options. This seminal work rightly changed the face of finance. Starting with a stochastic differential equation model for a stock price S, the now well-established lognormal random walk, they derived an equation and formulae for basic options. Let's give them a warm hip-hip-hooray for their work, and seven marks out of ten for their model, it's pretty accurate but, as I've said, cannot be perfect.

In 1977 Vasicek applied the Black, Scholes, Merton ideas to pricing interest rate derivatives. His starting point was ... a stochastic differential equation model again, but this time for r, a short-term interest rate. It was a very elegant idea, building a framework upon the earlier work and leading to similar mathematics. Not as accurate as the equity price model (there is still no consensus as to what is the best interest rate model) so we will give this five marks out of ten.

45.2.2 The Find-and-Replace School of Mathematical Modeling

In the decade that followed various people used ... stochastic differential equation models again, but this time for credit risk, with the random variable being the 'hazard rate' or 'instantaneous risk of default' p. The resulting models were not just similar to the interest rate models, they were exactly the same. The only conceivable difference was that everywhere there was an r in an interest rate model it was replaced by a p in the credit model. So was born the **Find-and-Replace School of Mathematical Modeling** (Figure 45.1). With the right software, suddenly the pressure for academics to publish disappeared, just dust off your old fixed-income research papers, go to Edit | Replace and you've got yourself a whole new publication. As practiced by ... you know who you are.

Now that I have given away one of the big secrets of quantitative finance modeling, let's look at the stages in producing a model.

Figure 45.1 Quantitative financial modeling 101: How to double the length of your CV.

Stage 1: Determinism

Start off by modeling some basic quantity as 'deterministic,' meaning totally predictable with no uncertainty or randomness at all. For example, you might say that the stock price which is currently \$100 will be at \$137 in two months' time. You can ask some questions now, which your model may be able to answer. Is that a good investment?; What will the value of an option be?

The simplest models of interest rates assume that they are predictable, a subject to which we devoted a chapter.

Problem is, when you look at the stock price two months later it's probably not going to be \$137. Not only was your model 'wrong' (which is bad enough) but it is also demonstrably wrong (a worse sin, remember the 11th commandment, 'Don't Get Caught').

A more sophisticated model would acknowledge the fact that you can't predict prices accurately. In particular it might include an element of randomness. Hence Stage 2 of modeling.

Stage 2: Stochasticity

Introduce something random. Still working with equities, why don't we make the stock price itself random and see where that takes us? This is where 'stochasticity' or randomness comes in. We model the stock price as random via a stochastic differential equation and base trading decisions and option price models on that. This can work very well, producing excellent models, all you need to do is choose a model that captures equity price behavior as well as possible. You might choose a lognormal random walk model. If you did, then the option model based on that is the Black–Scholes model.

You might find your models successful and profitable for many years. However, invariably the day comes when you are dissatisfied with your model. Perhaps it starts giving prices that don't agree with traders' intuition, or maybe there are some new instruments which it doesn't price well. (And here 'well' often means 'the same as other banks,' it can be a relative measure rather than an absolute one. In absolute terms yours may be the best model on the Street but if it gives different numbers from everyone else who are you going to believe? The answer to that boils down to personality types.)

How can we get 'better' prices? Easy ...

Stage 2′: Iterate

Make something else random.

Your model doesn't give the same prices as another bank? Just introduce a new random variable, let's say volatility. Now you've got a whole extra variable, with quite a few extra parameters and hence lots more degrees of freedom to get any price you want! By choosing to make volatility stochastic you have also killed two birds with one stone. Not only do you have more scope for matching someone else's prices but because volatility is unobservable no one can say that your model is wrong without them going into a great deal of statistical analysis.

The second stochastic factor may give you a year or two's grace in which your model is 'competitive.' But again the day comes when it is seen as too simple, it's not up to scratch when compared with the models of the top-tier banks. Do not despair, simply go back to Stage 2′ again, and iterate ... introduce a third stochastic variable, perhaps credit risk this time. More variables, more parameters, you can price everything the same as the next person ... I don't know how good those prices will be but at least you can't be fired.

Currently three is the optimum number of stochastic variables. This is not for any sound scientific or financial reason but simply that if your model has three random factors then you are seen as being 'sophisticated,' you are *the man*. Four factors is perceived as too many, you've lost the plot.

Should your models still leave something to be desired even with all these variables and parameters at your command, maybe you don't quite match implied volatility skews, then you have to go somewhere else other than Brownian motion for your next fix.

Stage 3: Jumps

The place to go is the Poisson process, the simple model for jumps. Allow one of your variables to gap, to jump, to have discontinuous paths. What you now have is a jump-diffusion process. And you are back on top of the game.

A popular model at the moment is to have stochastic asset and volatility and to allow the asset to jump discontinuously. Such a model has enough parameters to satisfy almost all modeling junkies.

Stage 3′: Iterate

And if one jumping variable is not enough, just iterate ... introduce another discontinuous variable, such as volatility. Now you have stochastic asset and volatility with jumps in both of these. So many variables, so many parameters, surely that will satisfy you?

I'm not sure, did that convey the right amount of sarcasm?

Now you know what to look out for in a mathematical model you are prepared to enter into the new and exciting non-Black–Scholes world. Read on!

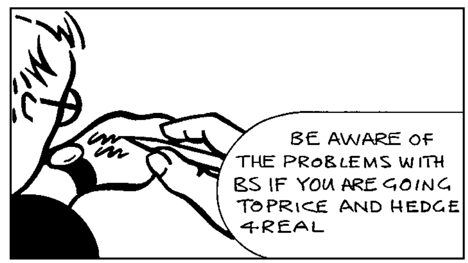
45.3 SUMMARY

There goes the promising career.

CHAPTER 46
defects in the Black–Scholes model

In this Chapter...

- why the Black–Scholes assumptions are wrong
- how to improve the Black–Scholes model

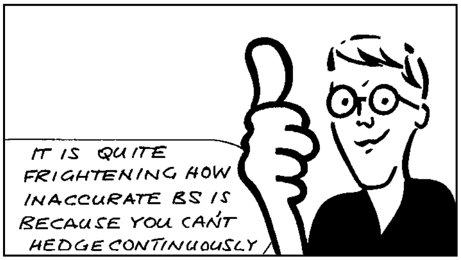


46.1 INTRODUCTION

Before pointing out some of the flaws in the assumptions of the Black–Scholes world, I must emphasize how well the model has done in practice, how widespread is its use and how much impact it has had on financial markets, not to mention a Nobel Prize for two of its three creators. The model is used by everyone working in derivatives, whether they are salesmen, traders or quants. It is used confidently in situations for which it was not designed, usually successfully. The value of vanilla options are often not quoted in monetary terms, but in volatility terms with the understanding that the price of a contract is its Black–Scholes value using the quoted volatility. You use a model other than Black–Scholes only with extreme caution, and you will have to be pretty convincing to persuade your colleagues that you know what you are doing. The ideas of delta hedging and risk-neutral pricing have taken a formidable grip on the minds of academics and practitioners alike. In many ways, especially with regards to commercial success, the Black–Scholes model is remarkably robust.

Nevertheless, there is room for improvement. Certainly, we can find models that better describe the underlying; what is not clear is whether these models make us more money. In the next few sections I introduce the ideas and models that will be expanded upon in the rest of the book. Some of these ideas are classical, in the sense that they are academically respectable, require no great leaps of imagination and are thoroughly harmless. Other ideas are too new to say what the future has in store for them. Yet there is at least a grain of truth in all that follows.

Each section corresponds to a chapter. First I give the relevant Black–Scholes assumption, and then I say why it is wrong and how, later, I try to relax the assumption. All the details will be found in the relevant chapters.



46.2 DISCRETE HEDGING

Black–Scholes assumes

- Delta hedging is continuous

Consider first, the continuous-time world of stochastic calculus. When we derived the Black–Scholes equation we used the continuous-time Itô's lemma. The delta hedging that was necessary for risk elimination also had to take place continuously. If there is a finite time between rehedges then there is risk that has not been eliminated.

In Chapter 47 we consider the effect of hedging at discrete intervals of time, taking real expectations over finite time periods rather than applying continuous-time calculus.

46.3 TRANSACTION COSTS

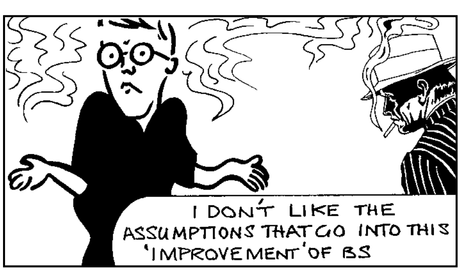
Black–Scholes assumes

- Delta hedging is continuous
- There are no costs in delta hedging

But not only must we worry about hedging discretely, we must also worry about how much it costs us to rehedge. The buying and selling of assets exposes us to bid-offer spreads. In some markets this is insignificant; then we rehedge as often as we can. In other markets, the cost can be so great that we cannot afford to hedge as often as we would like. These issues, and several models, are described in Chapter 48.

46.4 OVERVIEW OF VOLATILITY MODELING

Chapter 49 summarizes the key points of how to (and why to) model volatility. After all, volatility is *the* single most important determinant of an option's value that we don't accurately know. This chapter sets the scene and puts later material into context.



46.5 DETERMINISTIC VOLATILITY SURFACES

Black–Scholes assumes

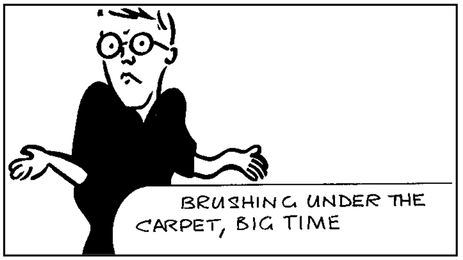
- Volatility is a known constant

If volatility is not a simple constant then perhaps it is a more complicated function of time and/or the underlying. If it is a function of time alone then we can find explicit solutions as described in Chapter 8. But what if it is a function of both time and the underlying? Perhaps the market even knows what this function is. Perhaps the market is so clever that it prices all traded contracts consistently with this volatility function and all we have to do is deduce from these traded prices what the volatility function is.

The link between the prices of vanilla option in the market and a deterministic volatility is the subject of Chapter 50. We see how to back out from market prices the 'implied volatility surface' and the resulting 'local volatility surface.'

This technique is popular for pricing exotic options, yielding prices that are 'consistent' with traded prices of similar contracts.

46.6 STOCHASTIC VOLATILITY



Black–Scholes assumes

- Volatility is a known constant (or a known deterministic function)

The Black–Scholes *formulae* require the volatility of the underlying to be a known deterministic function of time. The Black–Scholes *equation* requires the volatility to be a known function of time and the asset value. Neither of these is true. All volatility time series show volatility to be a highly unstable quantity. It is very variable and unpredictable. It is therefore natural to represent volatility itself as a random variable. I show the theory behind this in Chapter 51.

Stochastic volatility models are currently popular for the pricing of contracts that are very sensitive to the behavior of volatility. Barrier options are the most obvious example.

46.7 UNCERTAIN PARAMETERS

Black–Scholes assumes

- Volatility, interest rates and dividends are known constants (or known deterministic functions)

So volatility is not constant. Nor, actually, is it a deterministic function of time and the underlying. It's definitely unpredictable. Worse still it may not even be measurable.

Volatility cannot be directly observed and its measurement is very difficult. How then can we hope to model it? Maybe we should not attempt to model something we can't even observe. What we should do is to make as few statements about its behavior as possible; we will not say what volatility currently is or even what probability distribution it has. We shall content ourselves with placing a bound on its value, restricting it to lie within a given range. The probability distribution of the volatility within this range will not be prescribed. If it so desires, the volatility can jump from one extreme to the other as often as it wishes. This 'model' is then used to price contracts in a 'worst-case scenario' (see Chapter 52).

The idea of ranges for parameters is extended to allow the short-term interest rate and dividends to be uncertain, and to lie within specified ranges.

46.8 EMPIRICAL ANALYSIS OF VOLATILITY

Black–Scholes assumes

- Volatility is a known constant

In Chapter 53 I show how to estimate such quantities as the volatility of volatility, and how to deduce the drift rate of the volatility by analyzing its distribution.

Having determined a plausible stochastic differential equation model for the volatility from data, I suggest ways in which it can be used. It can be used directly in a stochastic volatility model, or indirectly in an uncertain volatility model to determine the likelihood of volatility ranges being breached.

46.9 STOCHASTIC VOLATILITY AND MEAN-VARIANCE ANALYSIS

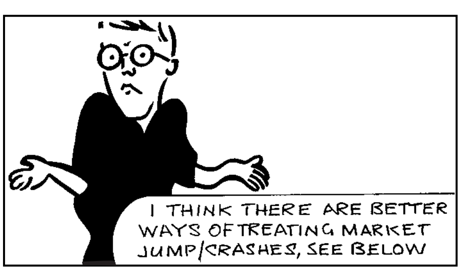
Black–Scholes assumes

- Volatility is a known deterministic function of asset value and time

When volatility is itself stochastic we can derive a theory (Chapter 51) that is consistent but requires knowledge of a new function, the market price of risk. This function is only observable via option prices themselves and so we find ourselves with a circular argument; we can price options if we know their market values. This is not entirely satisfactory and so in Chapter 54 we explore the possibility of valuing options when we know that there is some unhedged volatility risk.

46.10 ASYMPTOTIC ANALYSIS OF VOLATILITY

There is often a conflict between whether to use a scientifically accurate model, that is infuriatingly slow to crunch, or a fast but not so good model. It turns out that you can sometimes have your cake and eat it (perhaps during your free lunch) if you exploit the relative largeness or smallness of parameters in a mathematical model. In Chapter 55 we see how to take advantage of the typically fast speed of mean reversion and the large volatility of volatility that you see in equity markets to solve for option prices under very general stochastic volatility models.



46.11 JUMP DIFFUSION

Black–Scholes assumes

- The underlying asset path is continuous

It is common experience that markets are discontinuous, from time to time they 'jump,' usually downwards. This is not incorporated in the lognormal asset price model, for which all paths are continuous.

When I say 'jump' I mean two things. First, that the sudden moves are not contained in the lognormal model; they are too large, occurring too frequently, to be from a Normally distributed returns model. Second, they are unhedgeable; the moves are too sudden for continuous hedging through to the bottom of the jump.

The jump-diffusion model described in Chapter 57 is an attempt to incorporate discontinuities into the price path. These discontinuities are not modeled by the lognormal random walk that we have been using so far. Jump-diffusion is an improvement on the model of the underlying but introduces some unsatisfactory elements: Risk elimination is no longer possible and we must price in an 'expected' sense.

46.12 CRASH MODELING

Black–Scholes assumes

- The underlying asset path is continuous

If risk elimination is not possible can we consider worst-case scenarios? That is, assume that the worst does happen and then allow for it in the pricing. But what exactly is 'the worst'? The worst outcome will be different for different portfolios.

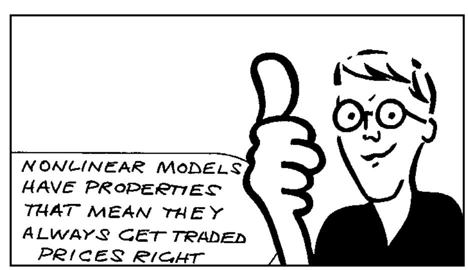
In Chapter 58 I show how to model the worst-case scenario and price options accordingly.

46.13 SPECULATING WITH OPTIONS

In Chapter 59 I show how to 'value' options when one is *not* hedging; rather, when one is speculating with derivatives because one has a good idea where the underlying is going. If one has a view on the underlying, then it is natural to invest in options because of their gearing. But how can this view be quantified? One way is to estimate real expected returns from an unhedged position. This together with an estimate of the risk in a position enables one to choose which option gives the best risk/reward profile for the given market view.

This idea can be extended to consider many types of model for the underlying, each one representing a different view of the behavior of the market. Furthermore, many trading strategies can be modeled. For example, how can one model the optimal closure of an option position? When should one sell back an option? Should you at times hedge, and at others speculate? Is there a best way to choose between the two?

46.14 OPTIMAL STATIC HEDGING



Many of the non-Black–Scholes models of this part of the book are non linear. This includes the models of Chapters 48, 52, 58 and 59. If the governing equation for pricing is non linear then the value of a portfolio of contracts is *not* the same as the sum of the values of each component on its own. One aspect of this is that the value of a contract depends on what it is hedged with. As an extreme example, consider the contract whose cashflows can be hedged exactly with traded instruments: To price this contract we do not even need a model. In fact, to use a model would be suicidal.

The beauty of the non-linear equation is that fitting parameters to traded prices (as in the implied volatility surfaces in Chapter 50 or in 'yield curve fitting' of Chapter 31) becomes redundant. Traded prices may be right or wrong, we don't much care which. All we care about is that if we want to put them into our portfolio then we know how much they will cost.

Imagine that we have a contract called 'contract,' and it has a value that we can write as

$$V_{NL}(\text{contract}),$$

where V_{NL} means the value of the contract using whatever is our non-linear pricing equation (together with relevant boundary and final conditions). Now imagine we want to hedge 'contract' with another contract called 'hedge.' And suppose that it costs 'cost' to buy or sell this second contract in the market. Suppose that we buy λ of these hedging contracts and put them in our portfolio; then the *marginal value* of our original 'contract' is

$$V_{NL}(\text{contract} + \lambda \text{ hedge}) - \lambda \text{ cost}. \tag{46.1}$$

In this expression 'contract + λ hedge' should be read as the portfolio made up of the 'union' of the original contract and λ of the hedging contract. Since V_{NL} is non linear, this marginal value is not the value of the contract on its own. We have hedged 'contract' statically, we may hold 'hedge' until expiry of 'contract.' We can go one step further and hedge *optimally*. Since the quantity λ can be chosen, let us choose it to maximize the marginal value of 'contract.' That is, choose λ to maximize (46.1). This is optimal static hedging. We can, of course, have as many traded contracts for hedging as we want, and we can easily incorporate bid-offer spread.

In the event that 'contract' can have all its cashflows hedged away by one or more 'hedge' contracts, we find that we are using our non-linear equation to value an empty portfolio and that the contract value is model independent (see Chapter 60).

46.15 THE FEEDBACK EFFECT OF HEDGING IN ILLIQUID MARKETS

Black–Scholes assumes

- The underlying asset is unaffected by trade in the option

The buying and selling of assets move their prices. A large trade will move prices more than a small trade. In the Black–Scholes model it is assumed that moves in the underlying are exogenous; that some cosmic random number generator tells us the prices of all 'underlyings.' In reality, a large trade in the underlying will move the price in a fairly predictable fashion. For example, it is not unheard of for unscrupulous people to move prices deliberately to ensure that a barrier option is triggered. In that case, a small move in the underlying could have a very big effect on the payoff of an option.

In Chapter 61 I will try to quantify this effect, introducing the idea that a trade in the underlying initiated by the need to delta hedge can move the price of the underlying. Thus it is no longer the case that the underlying moves and the option price follows, now it is more of a chicken-and-egg scenario. We will see that close to expiry of an option, when the gamma is large, the underlying can move in a very dramatic way.

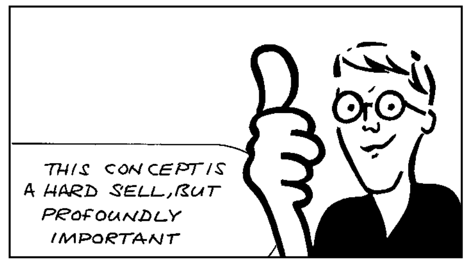
46.16 UTILITY THEORY

Black–Scholes assumes

- Delta hedging eliminates all risk

If, for any reason, we cannot perfectly delta hedge then we are left with some residual risk. We will therefore need a framework for valuing this risk. In Chapter 62 I introduce some ideas in this direction, which we will build on in later chapters.

46.17 MORE ABOUT AMERICAN OPTIONS AND RELATED MATTERS



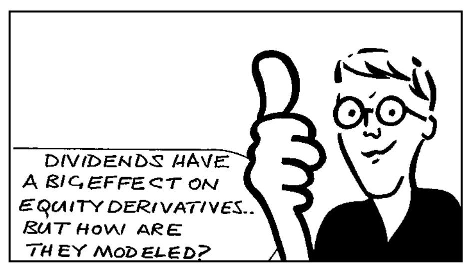
Black–Scholes assumes

- The American option should/will be exercised at the optimal time

But what does optimal mean? Does it mean the same thing to both the writer and the holder of the American option?

In Chapter 63 we examine the exercise strategy from the position of the option holder and I explain why he might want to exercise at an unexpected time. Such apparently 'non-optimal' exercise may actually be quite rational and will have a large impact on the writer of the option.

46.18 ADVANCED DIVIDEND MODELING



Black–Scholes assumes

- The underlying asset has deterministic dividends

Dividend payment is a subtle subject for modeling, and can have a significant effect on the prices of derivatives. In Chapter 64 we look at various forms of dividend models, including uncertain amount and payment dates, and stochastic dividends.

46.19 SERIAL AUTOCORRELATION IN RETURNS

Black–Scholes assumes

- There is no serial autocorrelation in returns

Each day is a new beginning, grasshopper. Today's return is random, and independent from what happened yesterday or any time in the past. That's the usual assumption. In Chapter 65 we take a brief look at what the data say, and then model a serially autocorrelated random walk.

46.20 SUMMARY

There are many faults with the Black–Scholes assumptions. Some of these are addressed in the next few chapters. Although it is easy to come up with any number of models that improve on Black–Scholes from a technical and mathematical point of view, it is nearly impossible to improve on its commercial success.

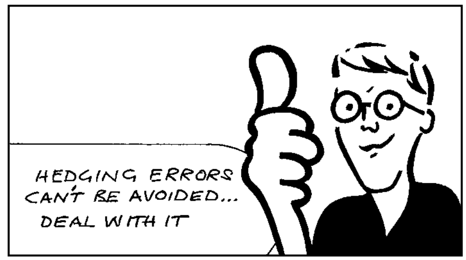
FURTHER READING

- For a discussion of some real-world modeling issues see the collection of papers edited by Kelly, Howison & Wilmott (1995).
- There are many other ways to improve on the lognormal model for the underlying. One which I don't have space for is a model of asset returns using hyperbolic, instead of Normal, distributions: See Eberlein & Keller (1995).

CHAPTER 47
discrete hedging

In this Chapter...

- the effect of hedging at discrete times
- hedging error
- the real distribution of profit and loss



47.1 INTRODUCTION

In this chapter we concentrate on one of the erroneous assumptions in the Black–Scholes model, that of continuous hedging. The Black–Scholes analysis requires the continuous rebalancing of a hedged portfolio according to a delta-neutral strategy. This strategy is, in practice, impossible.

The structure of this chapter is as follows. We begin by examining the concept of delta hedging in a discrete-time framework. We will see that taking expectations leads to the Black–Scholes equation without any need for stochastic calculus. I then show how this can be extended to a higher-order approximation, valid when the hedging period is not infinitesimal. We then discuss the nature of the hedging error, the error between the expected change in portfolio and the actual. This quantity is commonly ignored (perhaps because it averages out to zero) but is important, especially when one examines the real distribution of returns.

47.2 MOTIVATING EXAMPLE: THE TRINOMIAL MODEL

We've seen how perfectly we are able to eliminate risk in the binomial and Black–Scholes worlds. Now, what is the simplest generalization to these models that will show us how well delta hedging works in other situations? The simplest such model is the trinomial, as shown in Figure 47.1.

In this example the stock starts at 100 and can rise to 101, stay the same at 100, or fall to 99. Suppose that a call option with a strike of 100 simultaneously expires. The payoffs will then be either 1, if the stock rises, and zero in the other two cases. How can we hedge this payoff?

Introduce Δ as the quantity of stock we sell in order to hedge. Since the hedged portfolio value can be written as

$$V - \Delta S,$$

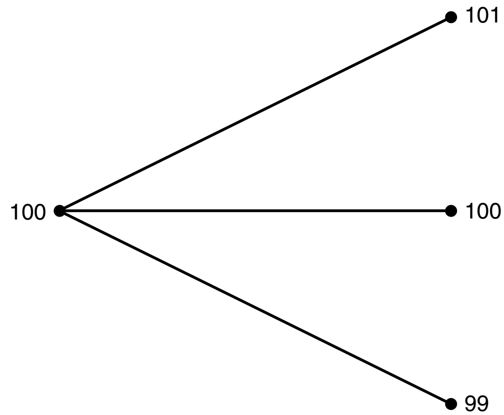


Figure 47.1 The trinomial random walk.

the portfolio value at expiration will be one of

$$1 - \Delta \times 101, \quad 0 - \Delta \times 100 \quad \text{or} \quad 0 - \Delta \times 99.$$

By hedging we are trying to make the portfolio have the same value whatever state the stock ends up in. So we are trying to find Δ such that

$$1 - \Delta \times 101 = 0 - \Delta \times 100 = 0 - \Delta \times 99.$$

Can't be done. Two equations, one unknown.

Conclusion, even in this rather trivial extension to the binomial world we are not able to hedge.

47.3 A MODEL FOR A DISCRETELY HEDGED POSITION

Not only does delta hedging fail in the trinomial world, it even fails in the Black–Scholes world if we can't hedge infinitely often. The Black–Scholes analysis requires *continuous* hedging, which is possible in theory but impossible, and even undesirable, in practice. Hence one hedges in some discrete way.

Our first step in analyzing the discrete-hedging problem is to choose a hedging strategy. If there are no transaction costs then there is no penalty for continuous rehedging and so the 'optimal' strategy is simply the Black–Scholes strategy, and the option value is the Black–Scholes value. This is the 'mathematical solution.' However, this strategy is clearly impractical. Unfortunately, it may be difficult to associate a 'cost' to the inconvenience of continuous rehedging. (This contrasts with the later problems when we do include transaction costs and optimal strategies are found.) For this reason a common assumption is that rehedging takes place regularly at times separated by a constant interval, the hedging period, here denoted by δt. This is a strategy commonly used in practice with δt ranging from half a day to a couple of weeks.

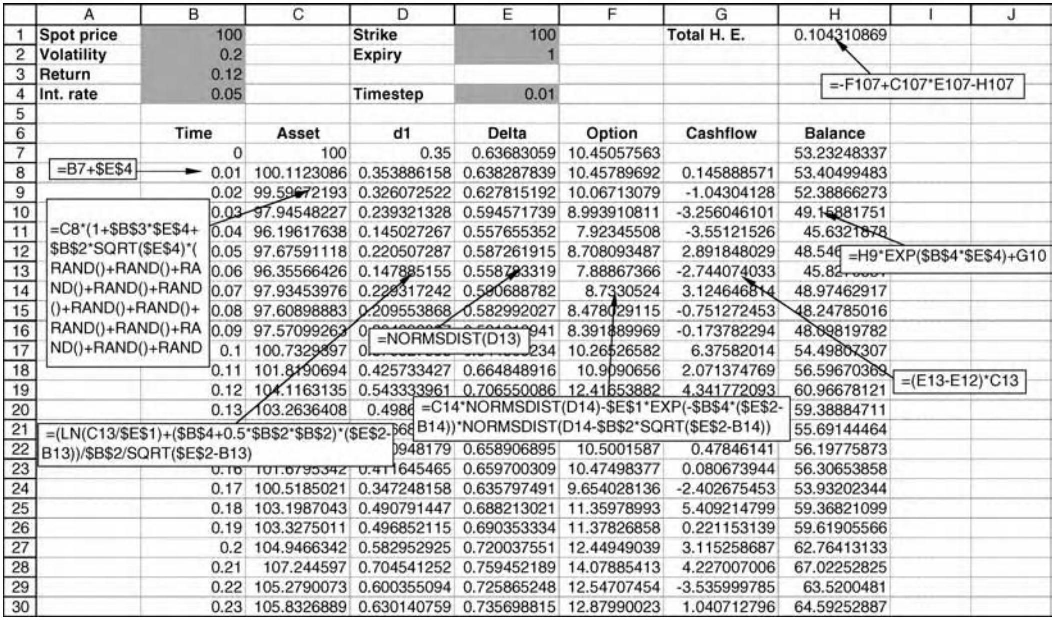


Figure 47.2 Spreadsheet for simulating hedging during the life of an option.

Often positions will be rebalanced to delta neutral just before the market closes each day. Note the use of $\delta\cdot$ to denote a discrete change in a quantity, this is to make a distinction with $d\cdot$, the earlier continuous changes.

The first work in this area was by Boyle and Emanuel who examined the discrepancy, the hedging error, between the Black–Scholes strategy of continuous rehedging and discrete hedging. They find that in each interval the hedging error is a random variable, from a chi-squared distribution, and proportional to the option's gamma. We will see why this is so shortly.

The spreadsheet in Figure 47.2 shows how to simulate delta hedging in the Black–Scholes fashion, but in discrete time. In this example the option is a call, there are no dividends on the underlying and there are no transaction costs. All of these can easily be incorporated. The spreadsheet uses a simple approximation to the Normal distribution. Note that the real drift of the asset is used in the simulation.

Results of this simulation are shown in Figures 47.3 and 47.4. In the first figure is the time series of the running **total hedging error** for a single realization of the underlying asset. This is the difference between the actual and theoretical values of the net option position as time evolves from the start to the end of the contract's life. The final value of -0.282 means that if you sold the option for the Black–Scholes value and hedged in the Black–Scholes manner until expiry to eliminate risk, you would have lost 0.282. Notice how the changes in the P&L seem to be very asymmetrical; there are a lot of small down moves and a few large up moves. We'll see the reason for this shortly.

In Figure 47.4 are the results of option replication over many realizations of the underlying asset. Each dot represents the final asset price and accumulated profit and loss after hedging for

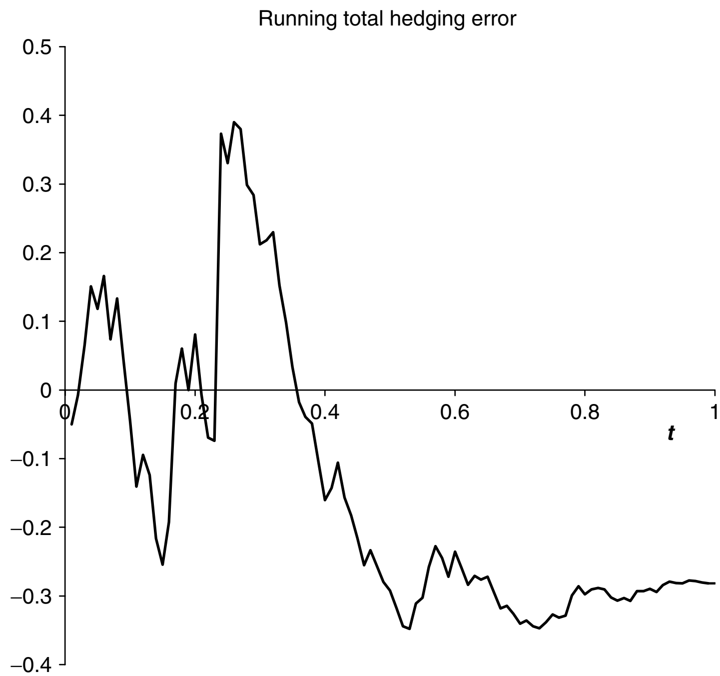


Figure 47.3 The running total hedging error.

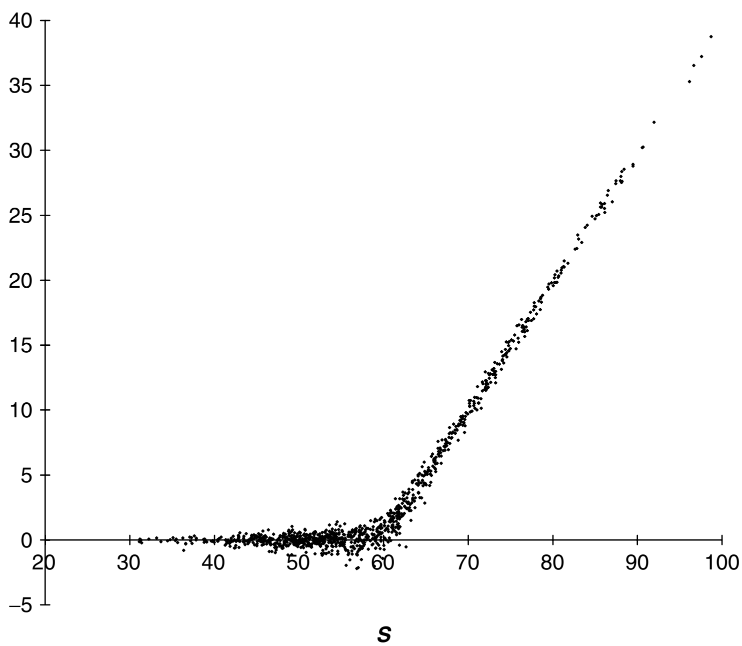


Figure 47.4 The results of discrete hedging.

the life of the contract. If hedging had been perfect each dot would lie on the payoff function, in this case $\max(S - 60, 0)$.

47.4 A HIGHER-ORDER ANALYSIS

Now we take this analysis one stage further. We use simple Taylor series expansions first to find a *better* hedge than the Black–Scholes and second to find an adjusted value for the option. The better hedge comes from hedging the option with the underlying using the number of shares that minimizes the variance of the hedged portfolio over the next time step.

The first step towards valuing an option is to choose a good model for the underlying in discrete time. A sensible choice is

$$S = e^x \tag{47.1}$$

where

$$\delta x = \left(\mu - \frac{\sigma^2}{2}\right)\delta t + \sigma\phi\,\delta t^{1/2}. \tag{47.2}$$

This is a discrete-time version of the earlier continuous-time stochastic differential equation for S.[1] Here ϕ is a random variable drawn from a standardized normal distribution and the term $\phi\ \delta t^{1/2}$ replaces the earlier Wiener process. In principle, the ideas that we describe do not depend on the random walk being lognormal and many models for S could be examined. In particular, ϕ need not be Normal but could even be measured empirically. If historic volatility is to be used then it should be measured at the same frequency as the rehedging takes place, i.e., using data at intervals of δt.

As in the Black–Scholes analysis we construct a hedged portfolio

$$\Pi = V - \Delta S, \tag{47.3}$$

with Δ to be chosen. As in Black–Scholes we first choose the hedge and then use it to derive an equation for the option value. (I have called this a 'hedged portfolio' but we will see that it is not perfectly hedged.)

We no longer have Itô's lemma, since we are in discrete time, but we still have the Taylor series expansion. Thus, it is a very simple matter to derive $\delta\Pi$ as a power series in δt and δx. On substituting for δx from (47.2) this expression becomes

$$\delta\Pi = \delta t^{1/2} A_1(\phi, \Delta) + \delta t\ A_2(\phi, \Delta) + \delta t^{3/2} A_3(\phi, \Delta) + \delta t^2 A_4(\phi, \Delta) + \cdots. \tag{47.4}$$

Actually, V and its derivatives appear in each A_i, but only the dependence on ϕ and Δ is shown. For example, the first term A_1 is given by

$$A_1(\phi, \Delta) = \sigma\phi S\left(\frac{\partial V}{\partial S} - \Delta\right),$$

[1] Why have I not chosen $\delta S = \mu S\,\delta t + \sigma S\phi\,\delta t^{1/2}$? Because that would only be an approximation to the continuous-time stochastic differential equation. We are going to look at high order terms in the Taylor series expansion of V and the above is exact.

and the second term by

$$A_2(\phi, \Delta) = \frac{\partial V}{\partial t} + S\left(\frac{\partial V}{\partial S} - \Delta\right)\left(\mu + \tfrac{1}{2}\sigma^2(\phi^2 - 1)\right) + \tfrac{1}{2}\sigma^2\phi^2 S^2 \frac{\partial^2 V}{\partial S^2}.$$

This expansion in powers of $\delta t^{1/2}$ can be continued indefinitely; we stop at the order shown above because it is at this order that we find results that differ from Black–Scholes. Since time is measured in units of a year, δt is small but not zero. I have not put all of the algebraic details here, this is definitely big-picture time. They are not hard to derive, just take up a lot of space.

Now I can state the very simple hedging strategy and valuation policy.

- Choose Δ to minimize the variance of $\delta\Pi$
- Value the option by setting the *expected* return on Π equal to the risk-free rate

The first of these, the hedging strategy, is easy to justify: The portfolio is, after all, hedged so as to reduce risk. But how can we justify the second, the valuation policy, since the portfolio is not riskless? The argument for the latter is that since options are in practice valued according to Black–Scholes yet necessarily discretely hedged, the second assumption is already being used by the market, but with an inferior choice for Δ.

Because we cannot totally eliminate risk I could also argue for a pricing equation that depends on risk preferences. I won't pursue that here, just note that even though we'll get an option 'value' that is different from the Black–Scholes value this does not mean that there is an arbitrage opportunity.

47.4.1 Choosing the Best Δ

The variance of $\delta\Pi$ is easily calculated from (47.4) since

$$\mathrm{var}[\delta\Pi] = E[\delta\Pi^2] - (E[\delta\Pi])^2. \tag{47.5}$$

In taking the expectations of $\delta\Pi$ and $\delta\Pi^2$ to calculate (47.5), all of the ϕ terms are integrated out leaving the variance of $\delta\Pi$ as a function of V, its derivatives, and, most importantly, Δ. Then, to minimize the variance we find the value of Δ for which

$$\frac{\partial}{\partial\Delta}\mathrm{var}[\delta\Pi] = 0.$$

The result is that the optimal Δ is given by

$$\Delta = \frac{\partial V}{\partial S} + \delta t\,(\cdots)\,. \tag{47.6}$$

The first term will be recognized as the Black–Scholes delta. The second term, which I give explicitly in a moment, is the correction to the Black–Scholes delta that gives a better reduction in the variance of $\delta\Pi$, and thus a reduction in risk. This term contains V and its derivatives.

47.4.2 The Hedging Error

The leading-order random term in the 'hedged' portfolio is, with this choice for Δ,

$$\tfrac{1}{2}\sigma^2\phi^2 S^2 \frac{\partial^2 V}{\partial S^2}.$$

We can write this as

$$\tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \tfrac{1}{2}(\phi^2 - 1)\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}.$$

The first term will, in the next section, be part of our pricing equation (and is part of the Black–Scholes equation). The second, which has a mean value of zero because ϕ is drawn from a standardized Normal distribution, is the **hedging error**. This term is random. The distribution of the square of a standardized Normal variable is called the **chi-squared distribution** (with a single degree of freedom). This chi-squared distribution is plotted in Figure 47.5. It has a mean of 1.

This distribution is asymmetrical about its mean, to say the least.[2] The average value is 1 but 68% of the time the variable is less than this, and only 32% above. If one is long gamma,

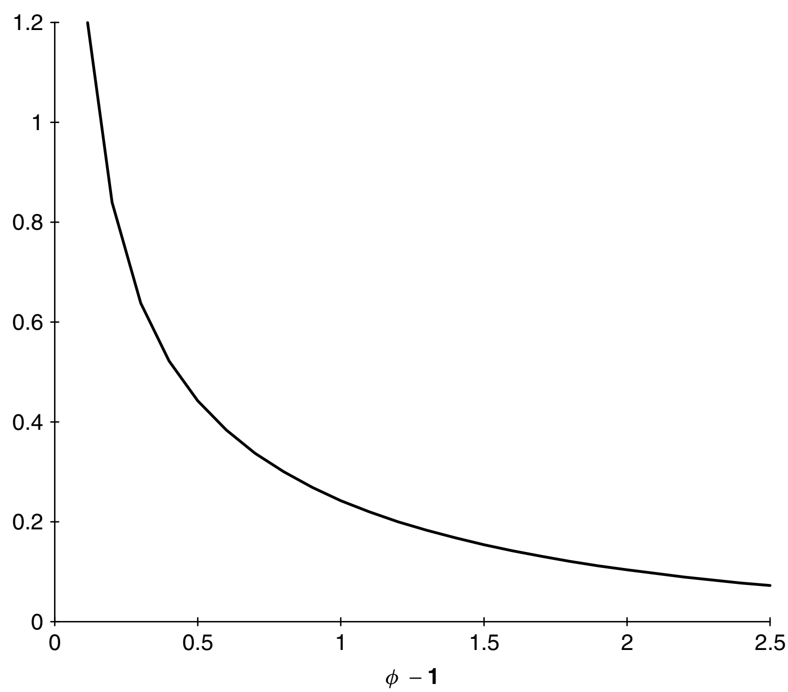


Figure 47.5 The chi-squared distribution.

[2] Remember my comments about Figure 47.3. There are a lot of small down moves and a small number of large up moves.

most of the time one loses money on the hedge (these are the small moves in the underlying), but 32% of the time you gain (and the size of that gain is on average larger than the small losses). The net position is zero. This result demonstrates the path dependency of hedging errors: Ideally, if long gamma, one would like large moves when gamma is large and small moves when gamma is small.

During the life of the hedged option the hedging errors at each rehedge add up to give the total hedging error. This is the final discrepancy between the theoretical profit and loss on the position and the actual.

Since the hedging error over one interval has a mean of zero then the total hedging error also has a mean of zero. The hedging error over one interval has a standard deviation of size δt; if you add up lots of these and they are all independent from each other, what size will be the standard deviation of the total hedging error? Since there are $T/\delta t$ rehedges, and therefore hedging errors, between now and expiration, the standard deviation of the total hedging error comes from estimating the variance of the sum of $T/\delta t$ independent random variables, each of size δt. The end result is a total hedging error of size $O(\delta t^{1/2})$.

The actual outcome for the total hedging error is highly path dependent. In Figure 47.6 is shown the distribution of the *total* hedging error for a call option hedged at different intervals. Note how the spread of the hedging error is much greater for the more infrequently hedged option. We will return to the distribution of the hedging error in theory and practice later.

In Figures 47.7 and 47.8 are shown two examples of realized stock price and the resulting P&L. The path dependency can be clearly seen. In the former the stock moves well away from

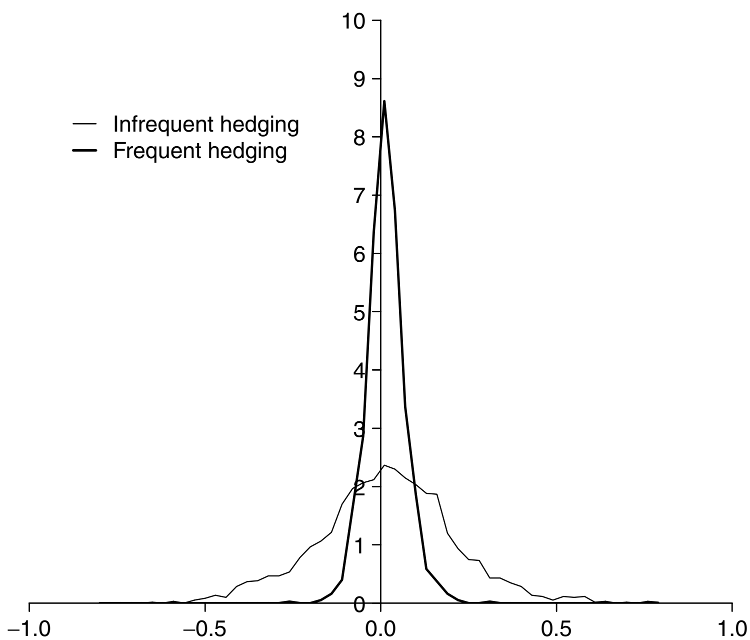


Figure 47.6 The distribution of the total hedging error for a call hedged frequently and infrequently.

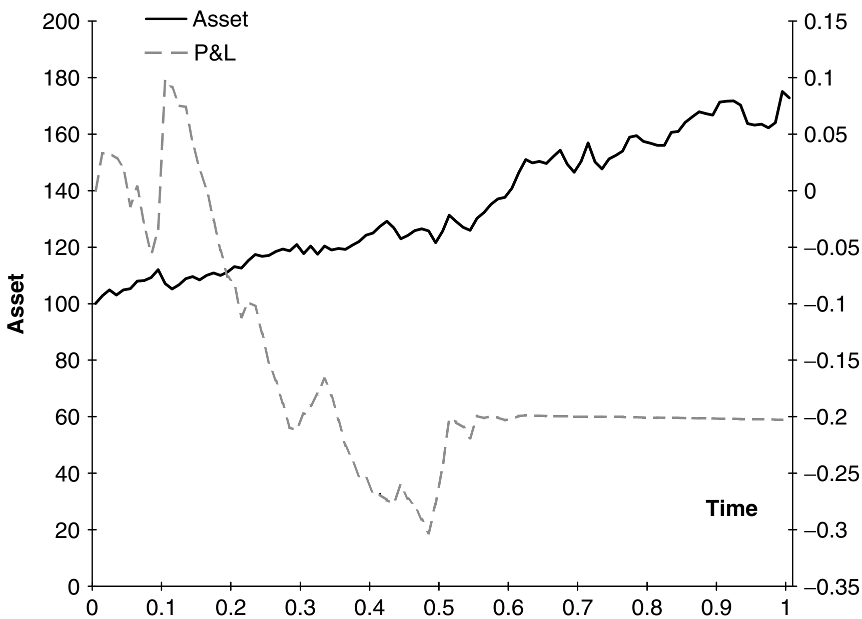


Figure 47.7 Asset price and the running total hedging error; stock price moves away from strike.

Figure 47.8 Asset price and the running total hedging error; stock price stays close to strike.

the strike, the gamma thus becomes small and the hedging error settles down to a constant level. In the second case the stock hovers around the strike so that gamma is large, and gets larger towards expiry, so that the P&L is always fluctuating.

In passing we note that the commonly held belief that 'time decay is the expected profit and loss (P&L) from a position' is wrong. The expected P&L from a self-financed and delta-hedged option position is zero.

47.4.3 Observations about the Hedging Error

Hedging error is one of the main reasons why the Black–Scholes model goes wrong. So it is very important to understand and appreciate how hedging error behaves. Here is a summary of the key features of hedging error.

- The hedging error should be zero on average for correctly priced and hedged options.
- Hedging error is asymmetric, drawn from a chi-squared distribution if the returns are Normal.
- The hedging error at each rebalancing is proportional to gamma.
- The hedging error at each rebalancing is proportional to the time between rehedges.
- Only the distribution of the square of returns matters (if we are delta hedging an option).
- The standard deviation of the real distribution of the hedging error is greater in practice than in theory because of fat tails.
- The total hedging error up to expiration is highly path-dependent.
- The total hedging error has a standard deviation which is of the order of the square root of the rebalancing interval.

Because the total hedging error does disappear as the time between rehedges shrinks we can easily argue for the validity of the Black–Scholes equation, but only in that limit.

47.4.4 Pricing the Option

Having chosen the best Δ, we now derive the pricing equation. The option should not be valued at the Black–Scholes value since that assumes perfect hedging and no risk: The fair value to an imperfectly hedged investor may be different. We find that the option value to the investor is equal to the Black–Scholes value plus a correction.[3]

The pricing policy that we have adopted has been stated as equating the expected return on the discretely hedged portfolio with the risk-free rate. This may be written as

$$E[\delta\Pi] = \left(r\ \delta t + \tfrac{1}{2}r^2\delta t^2 + \cdots\right)\Pi. \qquad (47.7)$$

[3] Note that I will be working in terms of real expectations. I am not pricing in the risk, simply looking at what happens on average but to a higher order than usual.

This is slightly different from the usual right-hand side, but simply represents a consistent higher-order correction to exponential growth: $e^{r\delta t} = 1 + r\ \delta t + r^2\delta t^2/2 + \cdots$. Now substitute (47.3) and (47.4) into (47.7) to get the equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \delta t\,(\cdots) = 0. \tag{47.8}$$

Again the first term is that derived by Black and Scholes and the second term is a correction to allow for the imperfect hedge; it contains V and its derivatives. I give the second term shortly.

47.4.5 The Adjusted Δ and Option Value

The as yet undisclosed terms in parentheses in (47.6) and (47.8) contain V and its derivatives, up to the second derivative with respect to t and up to the fourth with respect to S. However, since the adjusted option price is clearly close in value to the Black–Scholes price we can put the Black–Scholes value into the terms in parentheses without any reduction in accuracy. This amounts to solving (47.6) and (47.8) iteratively.

The result[4] is that the adjusted option price satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \tfrac{1}{2}\delta t(\mu - r)\left(r - \mu - \sigma^2\right) S^2 \frac{\partial^2 V}{\partial S^2} = 0 \tag{47.9}$$

and the better Δ is given by

$$\Delta = \frac{\partial V}{\partial S} + \delta t\left(\mu - r + \tfrac{1}{2}\sigma^2\right) S \frac{\partial^2 V}{\partial S^2}.$$

The important point to note about the above results is that the growth rate of the asset μ appears explicitly. This is a very important contrast with the Black–Scholes result. The Black–Scholes formulae do not contain μ. Thus any ideas about 'risk-neutral valuation' must be used with great care. There is no such thing as 'perfect hedging' in the real world. In practice the investor is *necessarily* exposed to risk in the underlying, and this manifests itself in the appearance of the drift of the asset price.

Notice how the second derivative terms in (47.9) are both of the form constant $\times\ S^2$. Therefore the correction to the option price can very easily be achieved by adjusting the volatility and using the value σ^* where

$$\sigma^* = \sigma\left(1 + \frac{\delta t}{2\sigma^2}(\mu - r)\left(r - \mu - \sigma^2\right)\right).$$

There is a similar volatility adjustment when there are transaction costs (see Chapter 48). The correction is symmetric for long and short positions.

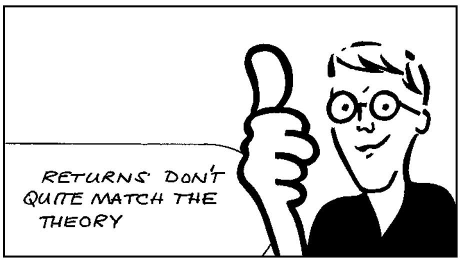
Is this volatility effect important? Fortunately, in most cases it is not. With typical values for the parameters and daily rehedging there is a volatility correction of one or two percent. In trending markets, however, when large μ can be experienced, this correction can reach five or ten percent, a value that cannot be ignored.

More importantly, in trending markets the corrected Δ will give a better risk reduction since it is in effect an anticipatory hedge: The variance is minimized over the time horizon until the next rehedge. This has been called **hedging with a view**.

[4] They were published in *Risk* magazine in 1994 but with an algebraic error. Oops!

The effect can also become important close to the strike near expiration when the gamma is very large. You should then use an adjustment for the delta of the form recommended above. Incidentally, close to strike and expiration, this correction is really just like estimating the delta not at the current time, but running the clock forward to the *next* time you will be delta hedging.

Since the option should be valued with a modified volatility, the difference between the adjusted option value and the Black–Scholes value is proportional to the option vega, its derivative with respect to σ.



47.5 THE REAL DISTRIBUTION OF RETURNS AND THE HEDGING ERROR

All of the above assumes that the return on the underlying is Normally distributed with a known volatility. In reality the distribution is close to, but certainly not identical to, Normal. In Figure 47.9 is shown the distribution of daily returns for the Dow Jones index from 1977 until 1996, scaled to have a mean of zero and a standard deviation of one, and the standardized Normal distribution. The empirical distribution has a higher peak and fatter tails than the Normal. This is typical of the randomness in financial variables, whether they are stocks, commodities, currencies or even interest rates. How does this distribution affect our hedging arguments and the hedging error?

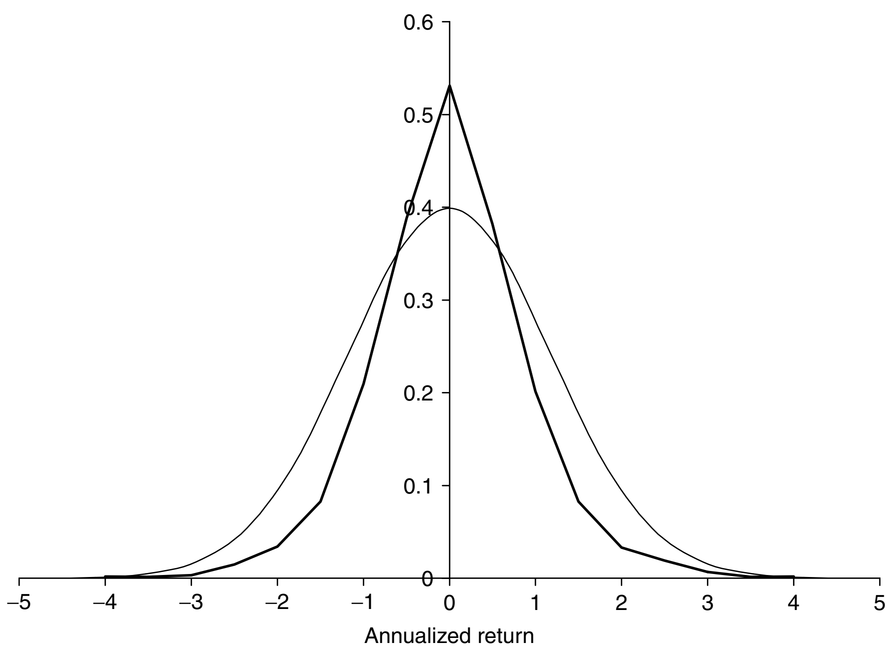


Figure 47.9 The distribution of returns on the Dow Jones index and the Normal distribution.

Let us assume that the return on the underlying in a time δt is given by

$$\frac{\delta S}{S} = \mu\ \delta t + \sigma\psi\ \delta t^{1/2},$$

where ψ is a random variable (with distribution determined empirically). This distribution is almost certainly not going to be Normal. It will have a mean of zero and a standard deviation of one, since those two parameters have been scaled out by using μ and σ.

We will examine the change in the value of a hedged portfolio assuming that the option component satisfies the Black–Scholes equation with a volatility of σ.

Our hedged portfolio has a random return in excess of the risk-free rate given by

$$\delta\Pi - r\Pi\ \delta t = S\left(\Delta - \frac{\partial V}{\partial S}\right)(r\ \delta t - \mu\ \delta t - \sigma\psi\ \delta t^{1/2}) + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}(\psi^2 - 1)\ \delta t + \cdots. \tag{47.10}$$

If we delta hedge in the Black–Scholes fashion then we are left with

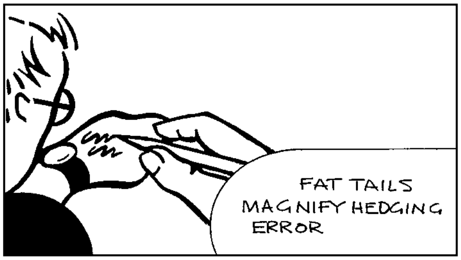
$$\delta\Pi - r\Pi\ \delta t = \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}(\psi^2 - 1)\ \delta t,$$

to leading order. Obviously, we are interested in the distribution of the hedging error, and from the above that means we are really interested in the distribution of ψ^2. Actually, if we are delta hedging options then we don't really care about the distribution of the returns, ψ, at all, only in the square of returns.

In Figure 47.10 is shown the normalized distribution of the square of returns and the chi-squared distribution, translated to have zero means. Both of these distributions have a mean of zero, but the empirical distribution is considerably higher at small values of the square and has a fatter tail (too far out to fit on the figure). Assuming that the standard deviation of the distribution of returns is $\sigma\ \delta t^{1/2}$ and that the market is using this σ to give the price (i.e. the actual and implied volatilities are the same), then the Black–Scholes value for a discretely-hedged position will be correct in the sense of *expectations*.

The fact that tails are fat doesn't affect the validity of the Black–Scholes model as long as the time between rehedges is small. There is one technical caveat to that, however. I have explicitly assumed that the standard deviation of the empirical distribution of returns does exist. If we find ourselves with empirical distributions that are so fat tailed as to have infinite standard deviation then the above argument fails, and all bets are off as far as the validity of the Black–Scholes equation is concerned.

Clearly, on *average* the portfolio earns the risk-free rate, but this is achieved via a particularly extreme distribution. Most of the time the stock price moves less than the theory would have, but makes up for this by the occasional large movement.



If the distribution were Normally distributed then 68% of the time the return would be less than the standard deviation and 32% greater. With a long gamma position you would lose a small amount 68% of the time but regain that loss on average due to the rarer but larger moves. However, this situation is exaggerated with the real distribution. From the Dow Jones data the higher peak and fatter tails mean that 78% of the time the move is less than the standard deviation and only 22% of the time is it greater. If you have a long gamma position then

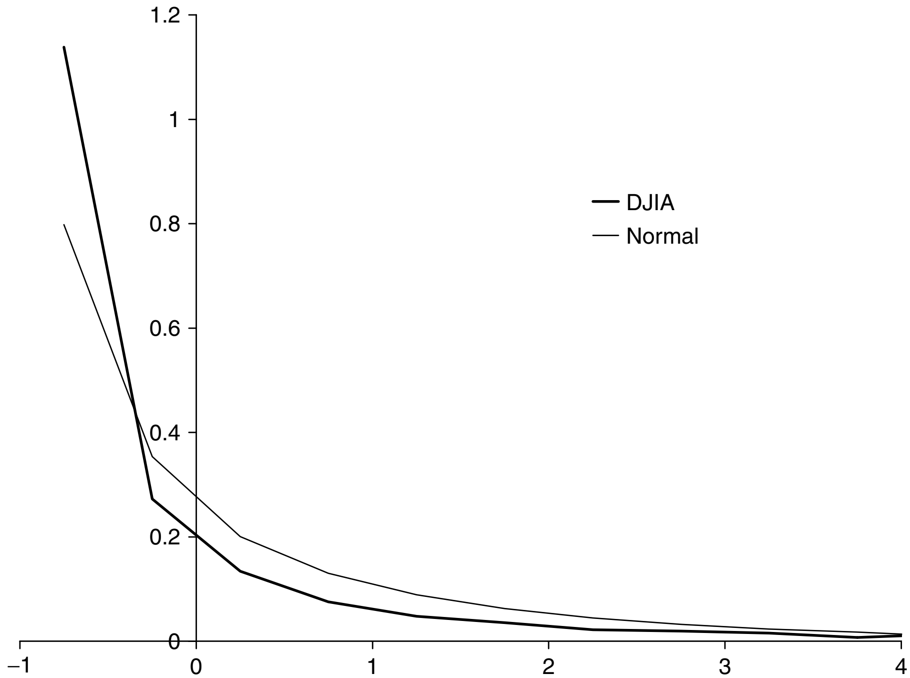


Figure 47.10 The distribution of the square of returns on the Dow Jones index and the translated chi-squared distribution.

approximately one fifth of the asset moves will lose you money. But again this is recovered on average with the rare large asset moves.

47.6 TOTAL HEDGING ERROR FOR THE REAL DISTRIBUTION OF RETURNS

We've looked at the 'real' hedging error on each rehedge, and these add up to give a total hedging error by the expiry of the option. This quantity is highly path dependent since each local hedging error depends on S. The order of magnitude of the total hedging error is $\delta t^{1/2}$, as we've seen, since the errors should not be serially autocorrelated. This is the size of the standard deviation of the accumulated errors. But what is the difference between the total hedging error when returns are Normal and the total hedging error for the real returns?

The mean of each local error is

$$E[\psi^2 - 1]\tfrac{1}{2}\sigma^2 S^2 \Gamma\ \delta t = 0.$$

The standard deviation of each hedging error is

$$\tfrac{1}{2}\sigma^2 S^2 \Gamma\ \delta t \sqrt{E[(\psi^2 - 1)^2]}.$$

For a Normal distribution the expectation term in this is 2. This need not be true for the actual distribution of returns. In Table 47.1 is shown the square root of the ratio of this expectation to its

Table 47.1 Total hedging error factors.

Asset	Factor
Ajinomoto	1.9
Bank of Scotland	1.7
Bass	1.3
Dover	1.8
Eastman Kodak	1.5
Guinness	1.6
Kanebo	2.3
Mitsui Mng & Smelt.	2.2
Nepool	7.3
Nippon Yakin Kogyo	1.9
Tesco	1.3

Normal equivalent for a number of assets. These numbers, which are scaled so that theoretically they should be 1, give a better idea of the size of the total hedging error that can be expected. All of the numbers are significantly greater than 1, and so total hedging error would in practice be larger than the theory says. One of the numbers is enormous; guess what asset that is.

The above factors tell you how much worse the hedging error will be in practice, than the Normal distribution theory suggests. So that for Ajinomoto, for example, the standard deviation of total hedging error will be worse than theory by a factor of 1.9. These factors are related to kurtosis.

47.7 WHICH MODELS ALLOW PERFECT DELTA HEDGING

Table 47.2 When is perfect hedging possible?

Worlds in which perfect hedging *is* possible	Worlds in which perfect hedging *is not* possible
Binomial	Trinomial etc.
Black–Scholes	Discrete time hedging + lognormal
	Stochastic volatility
	Jump diffusion
	Fat-tailed returns (infinite std dev.)

47.8 SUMMARY

Hedging error is often overlooked in the pricing of a contract. In each time step the order of magnitude of the change in P&L is the same order as the growth due to interest, i.e. of the order of the time step. Just because this averages out to zero does not mean it should be ignored. In this chapter I showed some of the effects of hedging error on an option value and how real returns differ from the theoretical Normal in such a way as to make the hedging error distribution even worse than theoretical. The asset with the worst hedging error factor was Nepool, the price of electricity. We will see why this is such a strange asset in Chapter 72.

FURTHER READING

- Boyle & Emanuel (1980) explain some of the problems associated with the *discrete* rehedging of an option portfolio.
- Wilmott (1994) derives the 'better' hedge.
- Leland (1985), Henrotte (1993) and Lacoste (1996) derive analytical results for the statistical properties of the tracking error.
- Mercurio & Vorst (1996) discuss the implications of discrete hedging strategies.
- See Hua (1997) for more details and examples of hedging error.

APPENDIX 1: THE SIMPLEST POSSIBLE DERIVATION OF THE BLACK–SCHOLES EQUATION ... SHOWING WHERE IT GOES WRONG

The Black–Scholes analysis requires *continuous* hedging, which is possible in theory but impossible, and even undesirable, in practice. Hence one hedges in some discrete way. Our first step in analyzing the discrete hedging problem is to choose a hedging strategy. We will make the simple assumption that rehedging takes place regularly at times separated by a constant interval, the hedging period, here denoted by δt.

Let us consider the **value changes** associated with a delta-hedged option.

- We start with zero cash.
- We buy an option.
- We sell some stock short.
- Any cash left (positive or negative) is put into a risk-free account.

We start by buying the option, as shown in Figure 47.11.

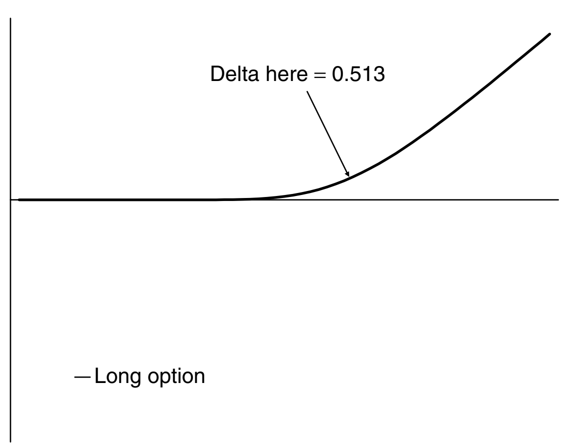


Figure 47.11 First, buy your option.

This option has a delta, and so we sell delta of the underlying stock in order to hedge (see Figure 47.12). This gives us a portfolio whose dependence on S is shown in Figure 47.13.

We are only concerned with small movements in the stock over a small time period, so zoom in on the region shown in Figure 47.14.

Locally the curve is approximately a parabola (see Figure 47.15).

There are three reasons for our total wealth to change from today to tomorrow.

1. The option price curve changes
2. There is an interest payment
3. The stock moves

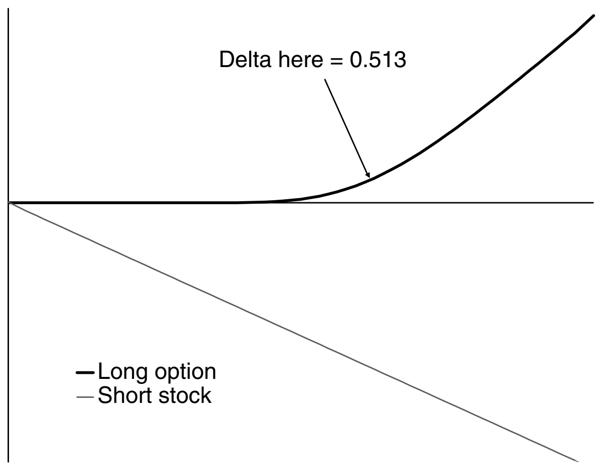


Figure 47.12 Now hedge it.

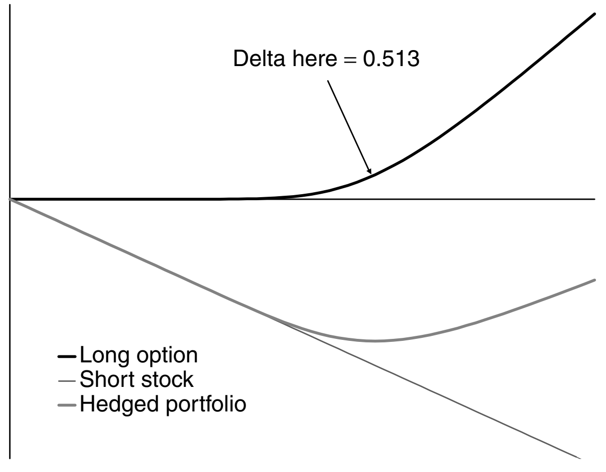


Figure 47.13 How our portfolio depends on S.

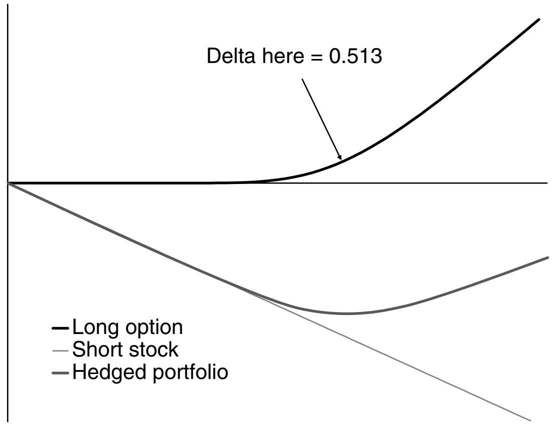


Figure 47.14 Zoom in here.

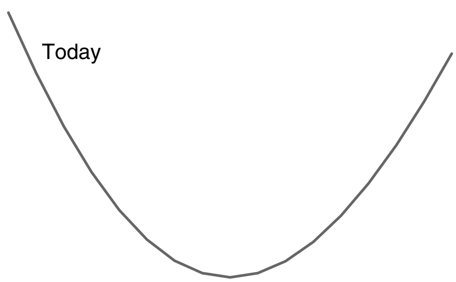


Figure 47.15 The curve is approximately quadratic.

Time Decay

The option curve falls by the time value, the theta multiplied by the time step (one day, in this case):

$$\Theta \times \delta t.$$

This is shown in Figure 47.16.

Interest

How much interest is received? How much money did we put in the bank?

$$\Delta \times S$$

from the stock sale and

$$-V$$

from the option purchase. Therefore the interest we receive is

$$r(S\Delta - V)\,\delta t.$$

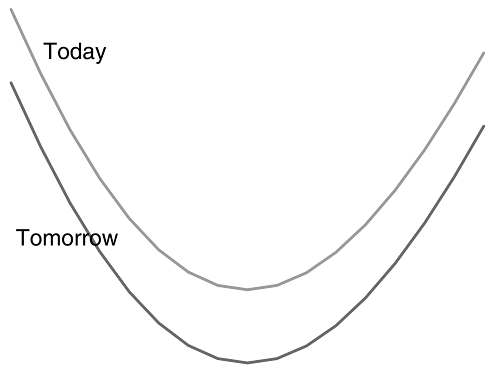


Figure 47.16 The curve falls by theta multiplied by the time step.

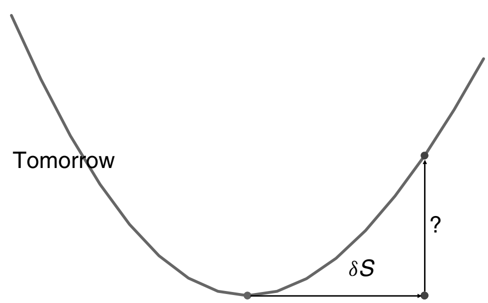


Figure 47.17 The profit from the stock move is proportional to the square of the stock move.

Profit from the Stock Move

Since gamma is positive, any stock price move is good for us. The larger the move the better (see Figure 47.17).

The option increases in value by

$$\tfrac{1}{2}\delta S^2\ \Gamma.$$

But this is random. What is its *expected value*? Since

$$\delta S = \mu S\ \delta t + \sigma S\ \phi\sqrt{\delta t}$$

then

$$\delta S^2 \approx \sigma^2 S^2 \delta t\ \phi^2$$

and so

$$E\left[\delta S^2\right] \approx \sigma^2 S^2 \delta t.$$

And the Sum is...

Put these three value changes together (ignoring the δt term which multiplies all of them):

$$\Theta + \tfrac{1}{2}\sigma^2 S^2 \Gamma + r(S\Delta - V) = 0.$$

And this is the Black–Scholes equation. In words, the total change in the value of a delta-hedged portfolio is equal to zero ... *on average*.

BS in Practice

This derivation is closer to what happens in practice than what you would think from either the binomial model or the Black–Scholes, continuous delta-hedging, model. This shows that Black–Scholes only works on average if you hedge discretely. Delta hedging works only *on average*. But if you go to the limit $\delta t = 0$ then the total hedging error vanishes like $\sqrt{\delta t}$ and this justifies the Black–Scholes model, but only in that limit.

APPENDIX 2: THE PROBABILITY DENSITY FUNCTION FOR THE CHI-SQUARED DISTRIBUTION

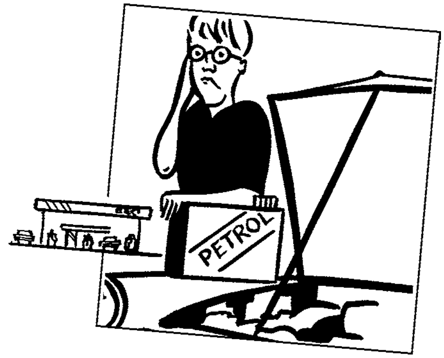
The cumulative distribution function for the chi-squared random variable $x = \phi^2$ is

$$\text{Prob}(X < x) = \frac{1}{\sqrt{2\pi}} \int_{-x^{1/2}}^{-x^{1/2}} e^{-\phi^2/2}\, d\phi = \sqrt{\frac{2}{\pi}} \int_0^{-x^{1/2}} e^{-\phi^2/2}\, d\phi.$$

The probability density function is then simply the derivative of this expression with respect to x:

$$\frac{1}{\sqrt{2\pi x}} e^{-x/2} \quad \text{for} \quad x > 0$$

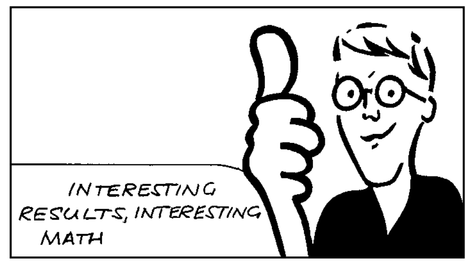
and zero otherwise.



CHAPTER 48
transaction costs

In this Chapter...

- how to allow for transaction costs in option prices
- how economies of scale work
- a variety of hedging strategies and their effects on option prices



48.1 INTRODUCTION

Transaction costs are the costs incurred in the buying and selling of the underlying, related to the bid-offer spread. The Black–Scholes analysis requires the continuous rebalancing of a hedged portfolio, and no 'friction' such as transaction costs in such rebalancing. In practice, this assumption is incorrect. Depending on the underlying market in question, costs may or may not be important. In a market with high transaction costs, stocks in emerging markets for example, it will be too costly to rehedge frequently. In more liquid markets, government bonds in first-world countries for example, costs are low and portfolios can be hedged often. If costs are important then this will be an important factor in the bid-offer spread in option prices.

The modeling of transaction costs was initiated by Hayne Leland. I will describe his model and then a simple model for transaction costs, the Hoggard–Whalley–Wilmott model for non-vanilla options and option portfolios, which is based on Leland's hedging strategy and model. These models will give us some insight into the effect of costs on the prices of option. Due to the mathematical complexity of this subject, the latter part of this chapter is necessarily just a review of key results with little of the underlying mathematical theory. The interested reader is referred to the original papers for further details.

48.2 THE EFFECT OF COSTS

In the Black–Scholes analysis we assumed that hedging took place continuously. In a sense, we take the limit as the time between rehedges goes to zero, $\delta t \to 0$. We therefore find ourselves rehedging an infinite number of times until expiry. With a bid-offer spread on the underlying the cost of rehedging leads to infinite total transaction costs. Clearly, even with discrete hedging the consequences of these costs associated with rehedging are important. Different people have different levels of transaction costs; there are economies of scale, so that the larger the amount that a person trades, the less significant are his costs. In contrast with the basic Black–Scholes model, we may expect that there is no unique option value. Instead, the value of the option depends on the investor. This is a whole new concept that we haven't seen in this book so far.

Not only do we expect different investors to have different values for contracts, we also expect an investor *if they are hedging* to have different values for long and short positions in the same contract. Why is this? It is because transaction costs are always a sink of money for hedgers, they always lose out on the bid-offer spread on the underlying. Thus we expect a hedger to value a long position at less than the Black–Scholes value and a short position at more than the Black–Scholes value; whether the position is long or short, some estimate of hedging costs must be taken away from the value of the option. Since the 'sign' of the payoff is now important, it is natural to think of a long position as having a positive payoff and a short position a negative payoff.

48.3 THE MODEL OF LELAND (1985)

The groundwork of modeling the effects of transaction costs was done by Leland (1985). He adopted the hedging strategy of rehedging at *every* time step. That is, every δt the portfolio is rebalanced, whether or not this is optimal in any sense. He assumes that the cost of trading ν assets costs an amount $\kappa\nu S$ for both buying and selling; this models bid-offer spread, the cost is proportional to the value traded.

In the main the Leland assumptions are those mentioned in Chapter 5 for the Black–Scholes model but with the following exceptions:

- The portfolio is revised every δt where δt is a finite and fixed, small time step.
- The random walk is given in discrete time by[1]

  $$\delta S = \mu S\ \delta t + \sigma S\phi\ \delta t^{1/2}$$

  where ϕ is drawn from a standardized Normal distribution.
- Transaction costs are proportional to the *value* of the transaction in the underlying. Thus if ν shares are bought ($\nu > 0$) or sold ($\nu < 0$) at a price S, then the cost incurred is $\kappa|\nu|S$, where κ is a constant. The value of κ will depend on the individual investor. For example, suppose that the stock is valued at \$100, we have to buy 2000 of them for hedging, and we lose 0.1% in transaction costs, then $S = 100$, $\nu = 2000$ and $\kappa = 0.001$. Total amount lost due to costs will be $0.001 \times 2000 \times \$100 = \$200$. A more complex cost structure can be incorporated into the model with only a small amount of effort, see later. We will then also see economies of scale appearing.
- The hedged portfolio has an *expected* return equal to that from a risk-free bank deposit. This is exactly the same valuation policy as earlier on discrete hedging with no transaction costs.

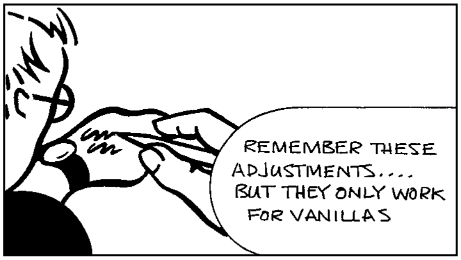
[1] We don't need the more accurate discrete lognormal model of (47.1) and (47.2) since we are not going to a high order of accuracy.

Leland allows for the cost of trading in valuing his hedged portfolio and by equating the *expected* return on the portfolio with the risk-free rate, he finds that long call and put positions should be valued with an adjusted volatility of $\check{\sigma}$ where

$$\check{\sigma} = \sigma\left(1 - \sqrt{\left(\frac{8}{\pi\,\delta t}\right)}\frac{\kappa}{\sigma}\right)^{1/2}.$$

Similarly short positions should be valued using $\hat{\sigma}$ where

$$\hat{\sigma} = \sigma\left(1 + \sqrt{\left(\frac{8}{\pi\,\delta t}\right)}\frac{\kappa}{\sigma}\right)^{1/2}.$$



As I have mentioned, long and short positions have different values.

Although the Leland concept is sound it is, in this form, only applicable to vanilla calls and puts, or any contract having a gamma of the same sign for all S and t. In the next section the Leland idea is extended to arbitrary option payoffs or portfolios, and the Leland result is derived along the way.

48.4 THE MODEL OF HOGGARD, WHALLEY & WILMOTT (1992)

The Leland strategy can be applied to arbitrary payoffs and to portfolios of options but the final result is not as simple as an adjustment to the volatility in a Black–Scholes formula. Instead, we will arrive at a non-linear equation for the value of an option, derived by Hoggard, Whalley & Wilmott, and one of the first non-linear models in derivatives theory.

Let us suppose we are going to hedge and value a portfolio of European options and allow for transaction costs. We can still follow the Black–Scholes analysis but we must allow for the cost of the transaction. If Π denotes the value of the hedged portfolio and $\delta\Pi$ the change in the portfolio over the time step δt, then we must subtract the cost of any transaction from the equation for $\delta\Pi$ at each time step. Note that we are not going to the limit $\delta t = 0$. Full details can be found in the appendix at the end of this chapter.

After a time step the change in the value of the hedged portfolio is now given by

$$\begin{aligned}\delta\Pi = {} & \sigma S\left(\frac{\partial V}{\partial S} - \Delta\right)\phi\ \delta t^{1/2} \\ & + \left(\tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\phi^2 + \mu S\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} - \mu\Delta S\right)\delta t - \kappa S|\nu|.\end{aligned} \tag{48.1}$$

You will recognize this as similar to the Black–Scholes expression but it now contains a transaction cost term. Transaction costs, $\kappa S|\nu|$, have been subtracted from the change in the value of the portfolio. Because these costs are always positive, there is a modulus sign, $|\cdot|$, in the above.

We will follow the same hedging strategy as before, choosing $\Delta = \partial V/\partial S$. However, the portfolio is only rehedged at discrete intervals. The number of assets held short is therefore

$$\Delta = \frac{\partial V}{\partial S}(S, t)$$

where this has been evaluated at time t and asset value S. I have not given the details, but this choice minimizes the risk of the portfolio, as measured by the variance, to leading order. After a time step δt and rehedging, the number of assets we hold is

$$\frac{\partial V}{\partial S}(S+\delta S, t+\delta t).$$

Note that this is evaluated at the new time and asset price. We can subtract the former from the latter to find the number of assets we have traded to maintain a 'hedged' position. The number of asset traded is therefore

$$\nu = \frac{\partial V}{\partial S}(S+\delta S, t+\delta t) - \frac{\partial V}{\partial S}(S, t).$$

Since the time step and the asset move are both small we can apply Taylor's theorem to expand the first term on the right-hand side:

$$\frac{\partial V}{\partial S}(S+\delta S, t+\delta t) = \frac{\partial V}{\partial S}(S, t) + \delta S\frac{\partial^2 V}{\partial S^2}(S, t) + \delta t\frac{\partial^2 V}{\partial S \partial t}(S, t) + \cdots.$$

Since $\delta S = \sigma S\phi\ \delta t^{1/2} + O(\delta t)$, the dominant term is that which is proportional to δS; this term is $O(\delta t^{1/2})$ and the other terms are $O(\delta t)$. To leading order the number of assets bought or sold is

$$\nu \approx \frac{\partial^2 V}{\partial S^2}(S, t)\ \delta S \approx \frac{\partial^2 V}{\partial S^2}\sigma S\phi\ \delta t^{1/2}.$$

We don't know beforehand how many shares will be traded, but we can calculate the *expected* number, and hence the expected transaction costs. The expected transaction cost over a time step is

$$E\left[\kappa S|\nu|\right] = \sqrt{\frac{2}{\pi}}\kappa\sigma S^2 \left|\frac{\partial^2 V}{\partial S^2}\right| \delta t^{1/2}, \tag{48.2}$$

where the factor $\sqrt{2/\pi}$ is the expected value of $|\phi|$. We can now calculate the expected change in the value of our portfolio from (48.1), including the usual Black–Scholes terms and also the new cost term:

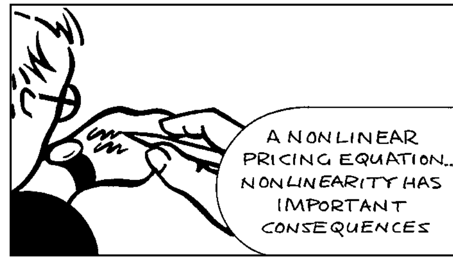
$$E[\delta\Pi] = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \kappa\sigma S^2 \sqrt{\frac{2}{\pi\ \delta t}} \left|\frac{\partial^2 V}{\partial S^2}\right|\right)\delta t. \tag{48.3}$$

Except for the modulus sign, the new, non-Black–Scholes, term in the above is of the same form as the second S derivative that has appeared before; it is a gamma term, multiplied by the square of the asset price, multiplied by a constant.

Now assuming that the holder of the option *expects* to make as much from his portfolio as if he had put the money in the bank, then we can replace the $E[\delta\Pi]$ in (48.3) with $r(V - S(\partial V/\partial S))\ \delta t$

as before to yield an equation for the value of the option:

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}-\kappa\sigma S^2\sqrt{\frac{2}{\pi\,\delta t}}\left|\frac{\partial^2 V}{\partial S^2}\right| + rS\frac{\partial V}{\partial S}-rV=0. \tag{48.4}$$



There is a nice financial interpretation of the term that is not present in the usual Black–Scholes equation. The second derivative of the option price with respect to the asset price, the gamma, $\Gamma = \partial^2V/\partial S^2$, is a measure of the degree of mishedging of the hedged portfolio. The leading-order component of randomness is proportional to δS and this has been eliminated by delta-hedging. But this delta hedging leaves behind a small component of risk proportional to the gamma. The gamma of an option or portfolio of options is related to the amount of rehedging that is expected to take place at the next rehedge and hence to the expected transaction costs.

The equation is a *non-linear* parabolic partial differential equation, one of the first such in finance. It is obviously also valid for a portfolio of derivative products. This is one of the first times in this book that we distinguish between single options and a portfolio of options. But for much of the rest of the book this distinction will be important. In the presence of transaction costs, the value of a portfolio is not the same as the sum of the values of the individual components. We can best see this by taking a very extreme example.

We have positions in two European call options with the same strike price and the same expiry date and on the same underlying asset. One of these options is held long and the other short. Our net position is therefore exactly zero because the two positions exactly cancel each other out. But suppose that we do not notice the cancelation effect of the two opposite positions and decide to hedge each of the options separately. Because of transaction costs we lose money at each rehedge on both options. At expiry the two payoffs will still cancel, but we have a negative net balance due to the accumulated costs of all the rehedges in the meantime. This contrasts greatly with our net balance at expiry if we realize beforehand that our positions cancel. In the latter case we never bother to rehedge, leaving us with no transaction costs and a net balance of zero at expiry.

Now consider the effect of costs on a single vanilla option held long. We know that

$$\frac{\partial^2 V}{\partial S^2} > 0$$

for a single call or put held long in the absence of transaction costs. Postulate that this is true for a single call or put when transaction costs are included. If this is the case then we can drop the modulus sign from (48.4). Using the notation

$$\check{\sigma}^2 = \sigma^2 - 2\kappa\sigma\sqrt{\frac{2}{\pi\,\delta t}}. \tag{48.5}$$

the equation for the value of the option is identical to the Black–Scholes value with the exception that the actual variance σ^2 is replaced by the modified variance $\breve{\sigma}^2$. Thus our assumption that $\partial^2 V/\partial S^2 > 0$ is true for a single vanilla option even in the presence of transaction costs. The modified volatility will be recognized as the Leland volatility correction mentioned at the start of this chapter.

For a short call or put option position we simply change all the signs in the above analysis with the exception of the transaction cost term, which must always be a drain on the portfolio. We then find that the call or put is valued using the new variance

$$\hat{\sigma}^2 = \sigma^2 + 2\kappa\sigma\sqrt{\frac{2}{\pi\ \delta t}}. \tag{48.6}$$

Again this is the Leland volatility correction.

The results (48.5) and (48.6) show that a long position in a single call or put with costs incorporated has an apparent volatility that is less than the actual volatility. When the asset price rises the owner of the option must sell some assets to remain delta hedged. But then the effect of the bid-offer spread on the underlying is to reduce the price at which the asset is sold. The effective increase in the asset price is therefore less than the actual increase, being seen as a reduced volatility. The converse is true for a short option position.

The above volatility adjustments are applicable when you have an option or a portfolio of options having a gamma of one sign. If the gamma is always and everywhere positive use the lower volatility value for a long position, and the higher value for a short position. If gamma is always and everywhere negative, swap these values around.

For a single vanilla call or put, we can get some idea of the total transaction costs associated with the above strategy by examining the difference between the value of an option with the cost-modified volatility and that with the usual volatility; that is, the difference between the Black–Scholes value and the value of the option taking into account the costs. Consider

$$V(S,t) - \hat{V}(S,t),$$

where the hatted function is Black–Scholes with the modified volatility. Expanding this expression for small κ we find that it becomes

$$(\sigma - \hat{\sigma})\frac{\partial V}{\partial \sigma} + \cdots.$$

This is proportional to the vega of the option. We know the formula for a European call option and therefore we find the expected spread to be

$$\frac{2\kappa S N'(d_1)\sqrt{(T-t)}}{\sqrt{2\pi\ \delta t}},$$

where $N'(d_1)$ has its usual meaning.

The most important quantity appearing in the model is

$$K = \frac{\kappa}{\sigma\,\delta t^{1/2}}. \tag{48.7}$$

K is a non-dimensional quantity, meaning that it takes the same value whatever units are used for measuring the parameters. If this parameter is very large, we write $K \gg 1$, then the transaction costs term is much greater than the underlying volatility. This means that costs are high and that the chosen δt is too small. The portfolio is being rehedged too frequently. If the transaction costs are very large or the portfolio is rehedged very often then it is possible to have $\kappa > \sigma/\sqrt{8\,\pi\,\delta t}$. In this case the equation becomes forward parabolic for a long option position. Since we are still prescribing final data, the equation is ill-posed. Although the asset price may have risen, its effective value due to the addition of the costs will have actually dropped. I discuss such ill posedness later.

If the parameter K is very small, we write $K \ll 1$, then the costs have only a small effect on the option value. Hence δt could be decreased to minimize risk. The portfolio is being rehedged too infrequently.

We can see how to use this result in practice if we have data for the bid-offer spread, volatility and time between rehedges for a variety of stocks. Plot the parameter κ/σ against the quantity $1/\delta t^{1/2}$ for each stock. An example of this for a real portfolio is shown in Figure 48.1. In this figure are also shown lines on which K is constant. To be consistent in our attitude towards transaction costs across all stocks we might decide that a value of $K = K'$ is ideal. If this

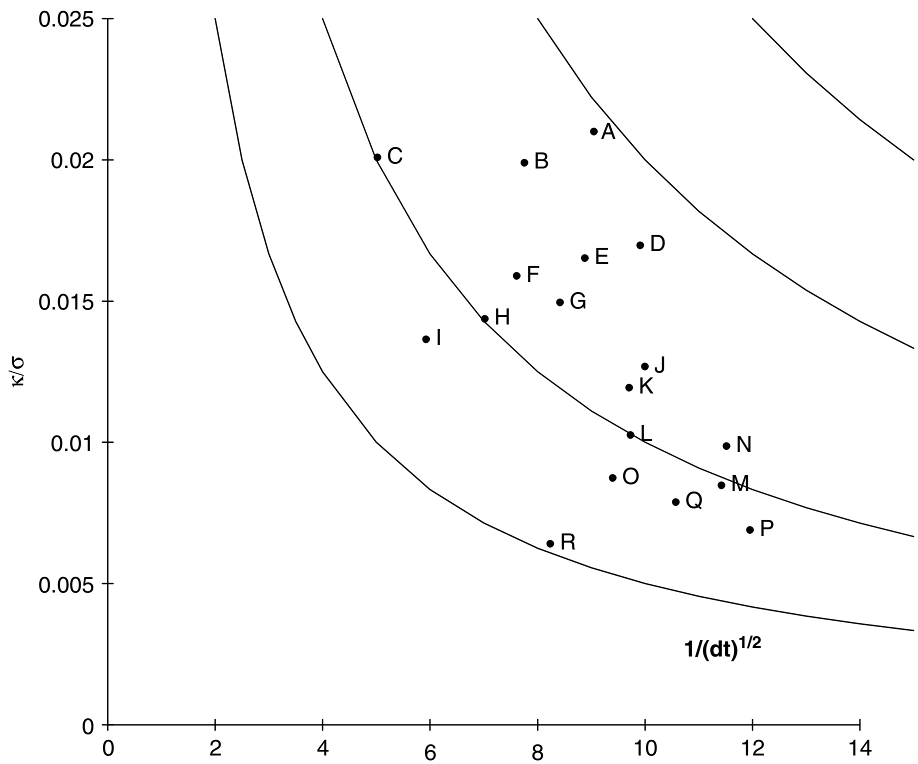


Figure 48.1 The parameter κ/σ against $1/\delta t^{1/2}$ for a selection of stocks. On each curve the transaction cost parameter K is the same.

is the bold line in the figure then options on those stocks above the line, such as A, are too infrequently hedged, while those below, such as R, are hedged too often. Of course, this is a very simple approach to optimizing a hedging strategy. A more sophisticated approach would also take into account the advantage of increased hedging: The reduction of risk.

48.5 NON-SINGLE-SIGNED GAMMA

For an arbitrary portfolio of options, the gamma, $\partial^2 V/\partial S^2$, is not of one sign. If this is the case then we cannot drop the modulus sign. Since the Hoggard–Whalley–Wilmott equation is non linear we must in general solve it numerically.

In Figures 48.2 and 48.3 is shown the value of a long bull spread consisting of one long call with $E = 45$ and one short call with $E = 55$ and the delta at six months before expiry for the two cases, with and without transaction costs. The volatility is 20% and the interest rate 10%. In this example $K = 0.25$. The bold curve shows the values in the presence of transaction costs and the other curve in the absence of transaction costs. The latter is simply the Black–Scholes value for the combination of the two options. The bold line approaches the other line as the transaction costs decrease.

In Figures 48.4 and 48.5 is shown the value of a long butterfly spread and its delta, before and at expiry. In this example the portfolio contains one long call with $E = 45$, two short calls with $E = 55$ and another long call with $E = 65$. The results are with one month until expiry for the two cases, with and without transaction costs. The volatility, the interest rate and K are as in the previous example.

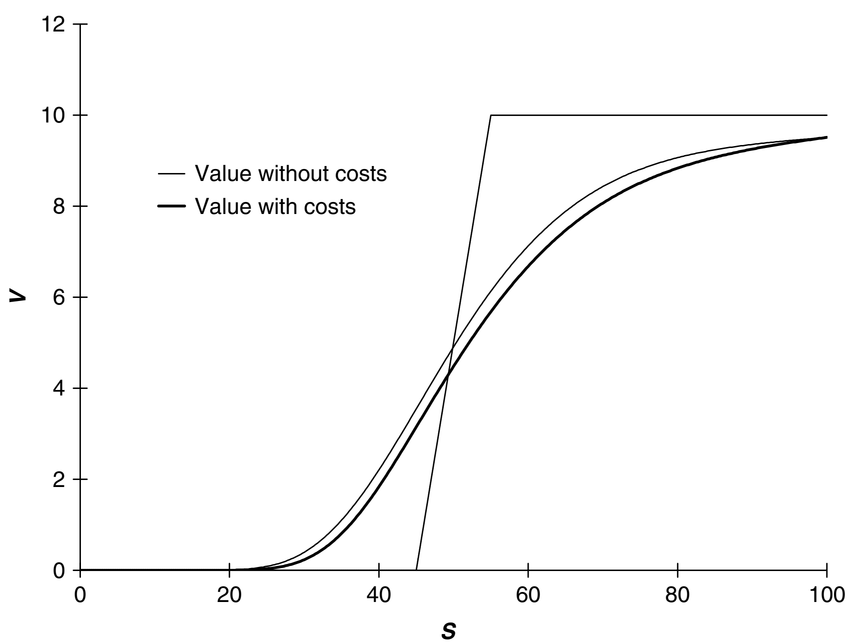


Figure 48.2 The value of a bull spread with (bold) and without transaction costs. The payoff is also shown.

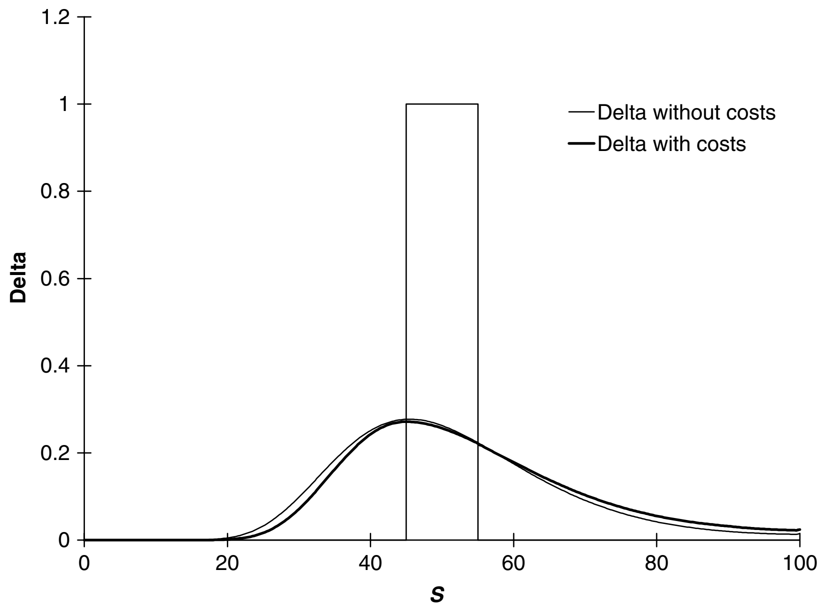


Figure 48.3 The delta for a bull spread prior to and at expiry with (bold) and without transaction costs.

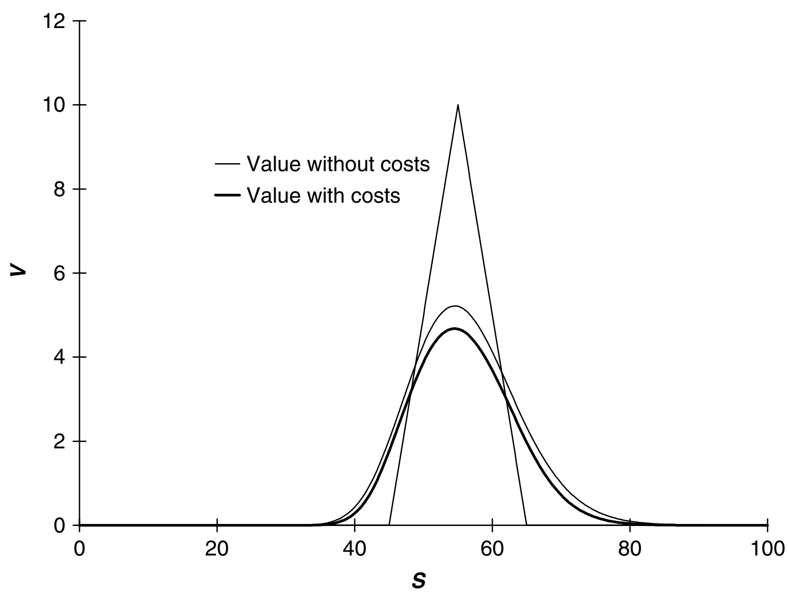


Figure 48.4 The value of a butterfly spread with (bold) and without transaction costs.

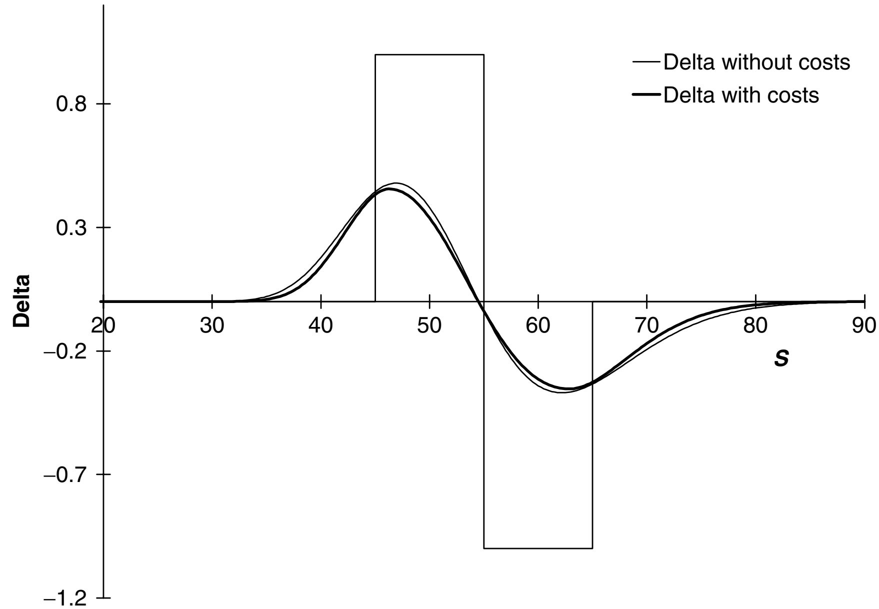


Figure 48.5 The delta for a butterfly spread with (bold) and without transaction costs.

48.6 THE MARGINAL EFFECT OF TRANSACTION COSTS

Suppose we hold a portfolio of options, let's call its value $P(S,t)$, and we want to add another option to this portfolio. What will be the effect of transaction costs? Call the value of the new, larger, portfolio $P+V$: What equation is satisfied by the marginal value V? This is not just the cost equation applied to V because of the nonlinearity of the problem. We can write

$$\frac{\partial P}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2P}{\partial S^2}-\kappa\sigma S^2\sqrt{\frac{2}{\pi\ \delta t}}\left|\frac{\partial^2P}{\partial S^2}\right|+rS\frac{\partial P}{\partial S}-rP=0$$

since the original portfolio must satisfy the costs equation. But now the new portfolio also satisfies this equation:

$$\frac{\partial(P+V)}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2(P+V)}{\partial S^2}-\kappa\sigma S^2\sqrt{\frac{2}{\pi\ \delta t}}\left|\frac{\partial^2(P+V)}{\partial S^2}\right|$$
$$+rS\frac{\partial(P+V)}{\partial S}-r(P+V)=0.$$

Both of these equations are non linear. But if the size of the new option is much less than the original portfolio we can linearize the latter equation to examine the equation satisfied by the marginal value V. We find that

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}-\operatorname{sgn}\left(\frac{\partial^2P}{\partial S^2}\right)\kappa\sigma S^2\sqrt{\frac{2}{\pi\ \delta t}}\frac{\partial^2V}{\partial S^2}+rS\frac{\partial V}{\partial S}-rV=0.$$

This equation is now linear. The important point to note is that the volatility correction only depends on the sign of the gamma of the original portfolio, P. If the gamma of the original portfolio is positive then the addition of another contract with positive gamma only makes the cost situation worse. However, the addition of a new contract with negative gamma will reduce the level of transaction costs. The benefits of reducing gamma may even make it worthwhile to buy/sell a contract for more/less than the Black–Scholes value, theoretically.

48.7 OTHER COST STRUCTURES

The above model can be extended to accommodate other transaction cost structures. Suppose that the cost of buying or selling ν assets is $\kappa(\nu, S)$. We follow the analysis of Section 48.4 up to the point where we take expectations of the transaction cost of hedging. The number of assets traded is still proportional to the gamma, and the expected cost of trading is

$$E\left[\kappa\left(\sigma S\,\delta t^{1/2}\Gamma\phi, S\right)\right].$$

The option pricing equation then becomes

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = \frac{1}{\delta t}E\left[\kappa\left(\sigma S\,\delta t^{1/2}\Gamma\phi, S\right)\right].$$

For example, suppose that trading in shares costs $k_1 + k_2\nu + k_3\nu S$, where k_1, k_2 and k_3 are constants. This cost structure contains fixed costs (k_1), a cost proportional to volume traded ($k_2\nu$), and a cost proportional to the value traded ($k_3\nu S$). The option value satisfies the non-linear diffusion equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = \frac{k_1}{\delta t} + \sqrt{\frac{2}{\pi\,\delta t}}\sigma S(k_2 + k_3 S)\left|\frac{\partial^2 V}{\partial S^2}\right|. \tag{48.8}$$

48.8 HEDGING TO A BANDWIDTH: THE MODEL OF WHALLEY & WILMOTT (1993) AND HENROTTE (1993)

We have so far seen how to model option prices when hedging takes place at fixed intervals of time. Another commonly used strategy is to rehedge whenever the position becomes too far out of line with the perfect hedge position. Prices are therefore monitored continuously but hedging still has to take place discretely.

Due to the complexity of this problem and those that follow, I only give a brief sketch of the ideas and results.

With $V(S, t)$ as the option value, the perfect Black–Scholes hedge is given by

$$\Delta = \frac{\partial V}{\partial S}.$$

Suppose, however, that we are not perfectly hedged, that we hold $-D$ of the underlying asset but do not want to accept the extra cost of buying or selling to reposition our hedge. The risk, as measured by the variance over a time step δt of this imperfectly hedged position is, to leading order,

$$\sigma^2 S^2 \left(D - \frac{\partial V}{\partial S} \right)^2 \delta t.$$

I can make two observations about this expression. The first is simply to confirm that when $D = \partial V / \partial S$ this variance is zero. The second observation is that a natural hedging strategy is to bound the variance within a given tolerance and that this strategy is equivalent to restricting D so that

$$\sigma S \left| D - \frac{\partial V}{\partial S} \right| \leq H_0. \tag{48.9}$$

The parameter H_0 is now a measure of the maximum expected risk in the portfolio. When the perfect hedge ($\partial V / \partial S$) and the current hedge (D) move out of line so that (48.9) is violated, then the position should be rebalanced. Equation (48.9) defines the **bandwidth** of the hedging position.

The model of Whalley & Wilmott (1993) and Henrotte (1993) takes this as the hedging strategy: The investor prescribes H_0 and on rehedging rebalances to $D = \partial V / \partial S$.

We find that the option value satisfies the non-linear diffusion equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = \frac{\sigma^2 S^4 \Gamma^2}{H_0} \left(k_1 + (k_2 + k_3 S)\frac{H_0^{\frac{1}{2}}}{S} \right), \tag{48.10}$$

where Γ is the option's gamma and the parameters k_1, k_2 and k_3 are the cost parameters for the cost structure introduced in Section 48.7. Note that again there is a non-linear correction to the Black–Scholes equation that depends on the gamma.

48.9 UTILITY-BASED MODELS

48.9.1 The Model of Hodges & Neuberger (1989)

All of the above models for transaction costs take the hedging strategy as exogenously given. That is, the investor chooses his strategy and then prices his option afterwards. Strategies like this have been called **local-in-time** because they only worry about the state of an option at the present moment in time. An alternative, first examined by Hodges & Neuberger (1989), is to find a strategy that is in some sense *optimal*. These have been called **global-in-time models** because they are concerned with what may happen over the rest of the life of the option.

The seminal work in this area, combining both utility theory and transaction costs, was by Hodges & Neuberger (HN), with Davis, Panas & Zariphopoulou (DPZ) making improvements to the underlying philosophy. HN explain that they assume that a financial agent holds a portfolio that is already optimal in some sense but then has the opportunity to issue an option and hedge the risk using the underlying. However, since rehedging is costly, they must define their strategy in terms of a 'loss function.' They thus aim to maximize expected utility. This entails the investor specifying a 'utility function.' The case considered in most detail by HN and DPZ is of the exponential utility function. This has the nice property of constant risk aversion. Mathematically, such a problem is one of stochastic control and the differential equations involved are very similar to the Black–Scholes equation.

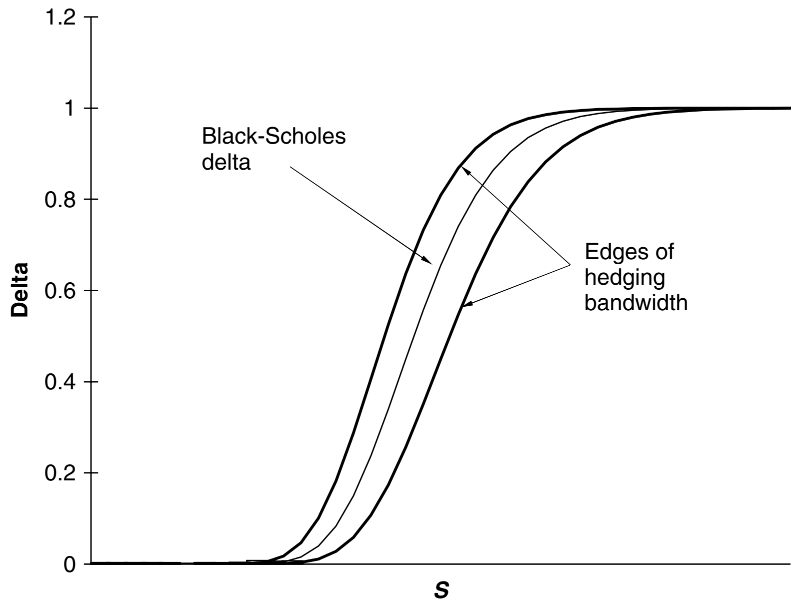


Figure 48.6 The optimal hedging strategy with proportional costs.

48.9.2 The Model of Davis, Panas & Zariphopoulou (1993)

The ideas of HN were modified by DPZ. Instead of valuing an option on its own, they embed the option valuation problem within a more general portfolio management approach. They then consider the effect on a portfolio of adding the constraint that at a certain date, expiry, the portfolio has an element of obligation due to the option contract. They introduce the investor's utility function, in particular, they assume it to be exponential. They only consider costs proportional to the value of the transaction ($\kappa \nu S$), in which case they find that the optimal hedging strategy is not to rehedge until the position moves out of line by a certain amount. Then, the position is rehedged as little as possible to keep the delta at the edge of this hedging bandwidth. This result is shown schematically in Figure 48.6. Here we see the Black–Scholes delta position and the hedging bandwidth.

In HN and DPZ the value of the option and, most importantly, the hedging strategy are given in terms of the solution of a three-dimensional free boundary problem. The variables in the problem are asset price S, time t, as always, and also D, the number of shares held in the hedged portfolio.

48.9.3 The Asymptotic Analysis of Whalley & Wilmott (1993)

The models of HN and DPZ are unwieldy because they are time consuming to compute. As such it is difficult to gain any insight into the optimal hedging strategy. Whalley & Wilmott did an asymptotic analysis of the DPZ model assuming that transaction costs are small, which is, of course, the case in practice. This analysis shows that the option price is given by the solution of an inhomogeneous diffusion equation, similar to the Black–Scholes equation.

This asymptotic analysis also shows that the HN optimal hedging bandwidth is symmetric about the Black–Scholes delta so that

$$\left| D - \frac{\partial V}{\partial S} \right| \leq \left(\frac{3k_3 S e^{-r(T-t)} F(S,t,\Gamma)^2}{2\gamma} \right)^{1/3},$$

where

$$F(S,t,\Gamma) = \left| \Gamma - \frac{e^{-r(T-t)}(\mu - r)}{\gamma S^2 \sigma^2} \right|.$$

The parameter γ is the index of risk aversion in the utility function.

These results are important in that they bring together all the local-in-time models mentioned above and the global-in-time models of HN and DPZ into the same diffusion equation framework.

This hedging bandwidth has been tested using Monte Carlo simulations by Mohamed (1994) and found to be the most successful strategy that he tested. The model has been extended by Whalley & Wilmott (1994) to an arbitrary cost structure, which is described below.

48.9.4 Arbitrary Cost Structure

The above description concentrates on the proportional cost case. If there is a fixed cost component then shares are traded to position the number of shares to be at some **optimal rebalance point**. This is illustrated schematically in Figure 48.7.

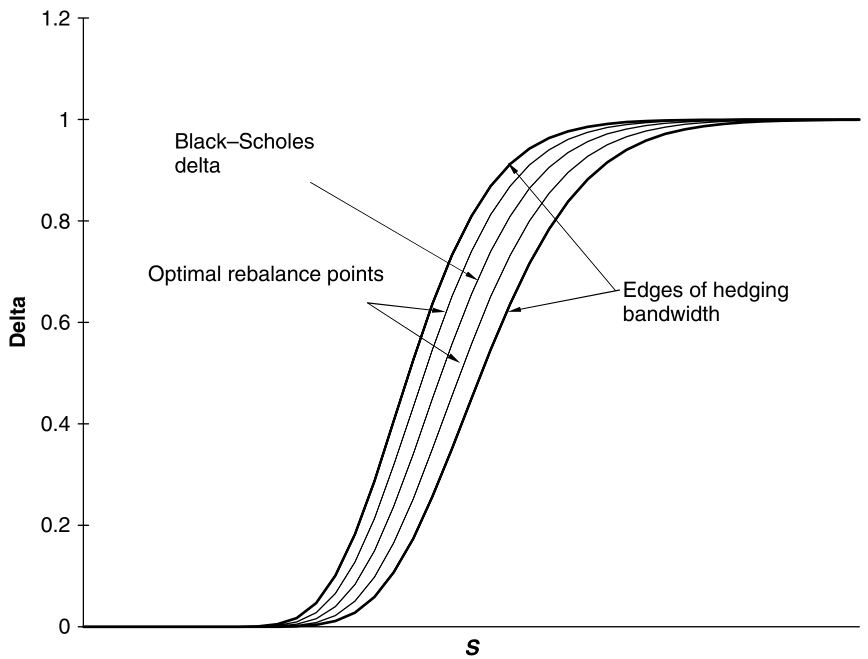


Figure 48.7 The optimal hedging strategy with arbitrary cost structure.

I do not give any of the details but note that the algorithm for finding the optimal rebalance point and the hedging bandwidth is as follows.

Assume that costs take the form $K(S, \nu)$, and that this is symmetric for buying and selling. The bandwidth is given by

$$\Delta - A(S,t) \leq D \leq \Delta + A(S,t)$$

where Δ is the Black–Scholes delta. The optimal rebalance points are given by

$$D = \Delta \pm B(S,t).$$

A and B come from solving

$$\gamma AB(A+B) = 3e^{-r(T-t)}\Gamma^2 \left.\frac{\partial K}{\partial \nu}\right|_{\nu=A-B}$$

and

$$\gamma (A+B)^3(A-B) = 12e^{-r(T-t)}\Gamma^2 K(S, A-B),$$

where Γ is the Black–Scholes gamma and γ is the index of risk aversion.

48.10 INTERPRETATION OF THE MODELS

Non-linear and inhomogeneous diffusion equations appear throughout the physical science literature. Thus there is a ready-made source of theoretical results and insights. I describe some of these in this section.

48.10.1 Nonlinearity

The effect of the nonlinearity on the valuation equations is that the sum of two or more solutions is not necessarily a solution itself. As I have said, a portfolio consisting of an equal number of the same options but held long and short (which has value identically zero), is not equal to the sum of the values of the two sub-portfolios of all the long and short options. This makes sense because in valuing each sub-portfolio separately we are assuming that each would be hedged separately, with attendant transaction costs to be taken into account. Upon recombining the two, the intrinsic values cancel, but the two sets of costs remain, giving a negative net value. The importance of nonlinearity extends far beyond this however.

Consider the following. Transaction cost models are non linear. The value of a portfolio of options is generally not the same as the sum of the values of the individual components. We can add contracts to our portfolio by paying the market prices, but the marginal value of these contracts may be greater or less than the amount that we pay for them. Is it possible to optimize the value of our portfolio by adding traded contracts until we give our portfolio its best value? This question is answered (in the affirmative) in Chapter 60. In a sense, the optimization amounts to finding the cheapest way to reduce the gamma of the portfolio globally, since the costs of hedging are directly related to the gamma.

48.10.2 Negative Option Prices

The transaction cost models above can result in negative option prices for some asset values depending on the hedging strategy implied by the model. So for example in the Hoggard–Whalley–Wilmott model with fixed transaction costs, $k_1 > 0$, option prices can become negative if they are sufficiently far out of the money. This model assumes that we rehedge at the end of *every* time step, irrespective of the level of risk associated with our position and also irrespective of the option value. Thus there is some element of obligation in our position, and the strategy should be amended so that we do not rehedge if this would make the option value go negative. In the case of a call therefore, there may be an asset price below which we would cease to rehedge and in this case we would regard the option as worthless.

Note that this is not equivalent to discarding the option; if the asset price were subsequently to rise above the appropriate level (which will change over time), we would begin to hedge again and the option would once more have a positive value. So we introduce the additional conditions for a moving boundary:

$$V(S_s(t), t) = 0$$

and

$$\frac{\partial V}{\partial S}(S_s(t), t) = 0.$$

The value $S_s(t)$ is to be found as part of the solution. This problem is now a 'free boundary problem,' similar mathematically to the American option valuation problem. In our transaction cost problem for a call option, we must find the boundary, $S_s(t)$, below which we stop hedging. This is illustrated in Figure 48.8.

In this figure we are valuing a long vanilla call with fixed costs at each rehedge. The top curve is the Black–Scholes option value as a function of S at some time before expiry. The bottom curve allows for the cost of rehedging but with the obligation to hedge at each time step. The option value is thus negative far out of the money. The middle curve also incorporates costs but without the obligation to rehedge. It thus always has a positive value and is, of course, also below the Black–Scholes option value.

48.10.3 Existence of Solutions

Linear diffusion equations have many nice properties, as we discussed in Chapter 6. The solution to a 'sensible' problem exists and is unique. This need not be the case for non-linear equations. The form of the equation and the final data $V(S, T)$ for the equation, (the payoff at expiry), may result in the solution 'blowing up,' that is, becoming infinite and thus financially unrealistic.

This can occur in some models even if transaction costs are small because of the effect of the option's gamma, which in those models is raised to some power greater than one in the extra transaction cost term. So wherever the gamma is large this term can dominate. For example, near the exercise price for a vanilla call or put option, $\partial^2 V/\partial S^2(S, 0)$ is infinite. We consider the case for the model of Equation (48.10).

The governing equation in this case has a transaction cost term proportional to Γ^2. Close to expiry and near the exercise price, E, we write $t = T - \tau$ and $S = E + s$ where $|s|/E \ll 1$

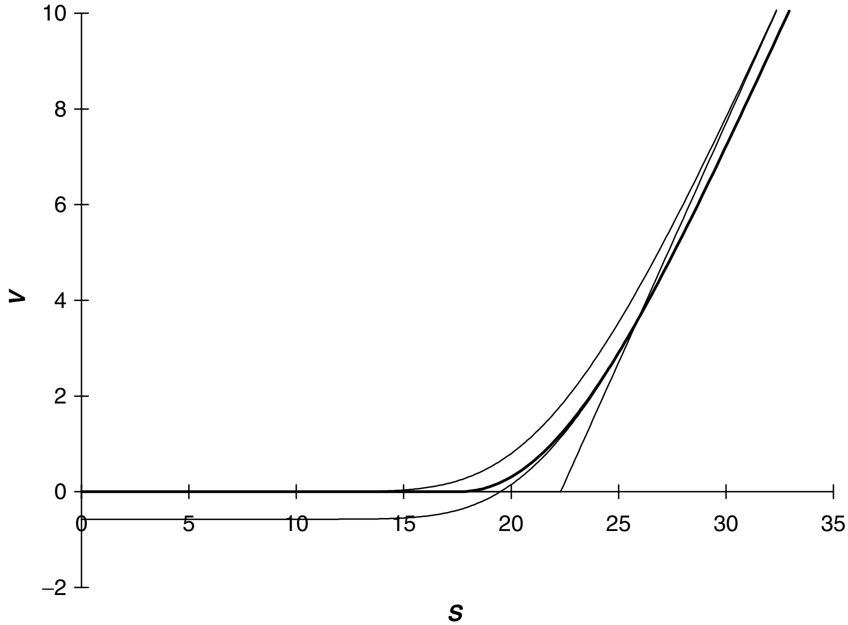


Figure 48.8 When to stop hedging if there are fixed costs.

and then the equation can be approximated by

$$\frac{\partial V}{\partial \tau} = \beta \left(\frac{\partial^2 V}{\partial s^2} \right)^2, \tag{48.11}$$

where

$$\beta = \frac{\sigma^2 E^4}{H_0} \left(k_1 + H_0^{\frac{1}{2}} \left(\frac{k_2}{E} + k_3 \right) \right). \tag{48.12}$$

Taking H_0 to be a constant, which is equivalent to a fixed bandwidth for the delta, it can be shown that equation (48.11) is ill-posed, that it is has no solution, if $\Gamma(S, 0) > 0$.[2]

So a long vanilla call or put hedged under this strategy has no finite value. Note that for short vanilla options, $\Gamma(S, 0) < 0$ and a solution does exist, so they *can* be valued under such a strategy, and can be hedged with a constant level of risk throughout the life of the option. However, returning to the case of payoffs with positive gamma, as the option approaches expiry, the number of hedging transactions required, and hence the cost of maintaining the hedging strategy, increases unboundedly unless the level of risk allowed (H_0) is itself allowed to become unbounded.

[2] We can see this intuitively as follows. If $\Gamma(S, 0) > 0$ then the right-hand side of (48.11) is positive. Thus V increases in time, increasing fastest where the gamma is largest. This in turn further increases the gamma, making the growth in V even faster. The result is a blow up. The linear diffusion equation also behaves in this way but the increase in V does not get out of control, since the gamma is raised only to the power one.

48.11 NON-NORMAL RETURNS

We know that returns are not Normally distributed but is this important?

Suppose that returns are given by

$$\frac{\delta S}{S} = \mu \, \delta t + \sigma \, \psi \, \delta t^{1/2}$$

for some random variable ψ of empirically determined distribution. What matters as far as expected transaction costs are concerned is not the mean of ψ, nor its standard deviation. What is most important is the mean of the absolute value of ψ i.e. the average value of $|\psi|$. We can examine the data to see if this number is greater or less than the theory says, the ratio to the theoretical value giving us a transaction cost factor.

In Table 48.1 are given the transaction cost factors for a selection of stocks, scaled with time step and volatility so that all numbers would be one if the underlying distribution were Normal. You can see that they are all less than one, but not by an enormous amount. So costs are going to be slightly less important than you might think. As we saw in Chapter 47 electricity prices stand out as being the furthest from the theory.

Table 48.1 Transaction cost factors.

Asset	Factor
Asahi Breweries	0.93
Asda Group	0.93
Cable & Wireless	0.91
Std Chartered	0.68
Equifax	0.88
Fleetwood Ents	0.92
Ford	0.96
Nepool	0.50
Sumitomo Bank	0.88
Toshiba	0.88

48.12 EMPIRICAL TESTING

In this section we look at empirical results for transaction costs and hedging error using various hedging strategies described above. We will look at the following four strategies:

- Basic Black–Scholes strategy, delta hedging at fixed intervals
- Leland volatility-modified delta, hedging at fixed intervals
- Delta tolerance, hedging to the Black–Scholes delta when the difference between quantity of the underlying held and ideal delta move too far out of line
- The asymptotic version of the utility model

We will use stock price data that is generated randomly, with known and constant volatility, and we will also use real data. Many stock path realizations will be used so that we can examine the statistical properties of the total costs and hedging errors.

Finally, we examine

- Average total transaction costs
- Average price (i.e. Black–Scholes value plus average costs)
- Standard deviation of price
- Price of 95th percentile

The third of these includes the hedging error that would be present even if there were no costs at all. The last simply means the price at which the contract must be sold to ensure that 95% of the time we do not lose money.

When we come to look at real data we also examine for each stock price time series which of the four strategies is the winner. Here 'winner' means the strategy that gives the lowest total cost plus hedging error for that particular realized asset path.

To understand how the random simulations were done, see Chapter 80.

To start with we value and hedge an at-the-money call with a volatility of 20%, risk-free rate of 5%. The underlying is currently at 100 and there is one year to expiry. The values for k_1, k_2 and k_3 are all 0.01.

48.12.1 Black–Scholes and Leland Hedging

This is a straight fight between using the Black–Scholes delta or the Leland volatility-adjusted delta.

Figure 48.9 shows the average amount of costs incurred and the standard deviation of the option price when the time between hedging varies. For each hedging period 5000 simulations were used. The number of days between rehedging varied from 1 up to 25.

Not surprisingly, as the time gap increases the amount of transaction costs paid decreases.

The figure also shows the standard deviation of the price. This includes the hedging error that would still be present in the absence of costs. It is very clear that the level of risk increases

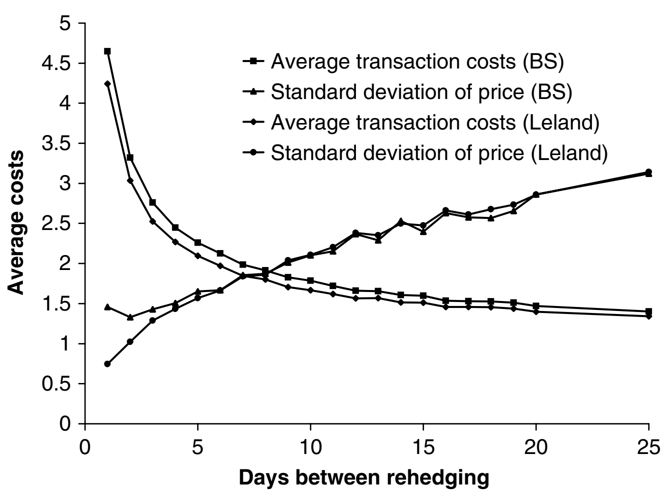


Figure 48.9 Average costs and standard deviation of price under the Black–Scholes and Leland strategies.

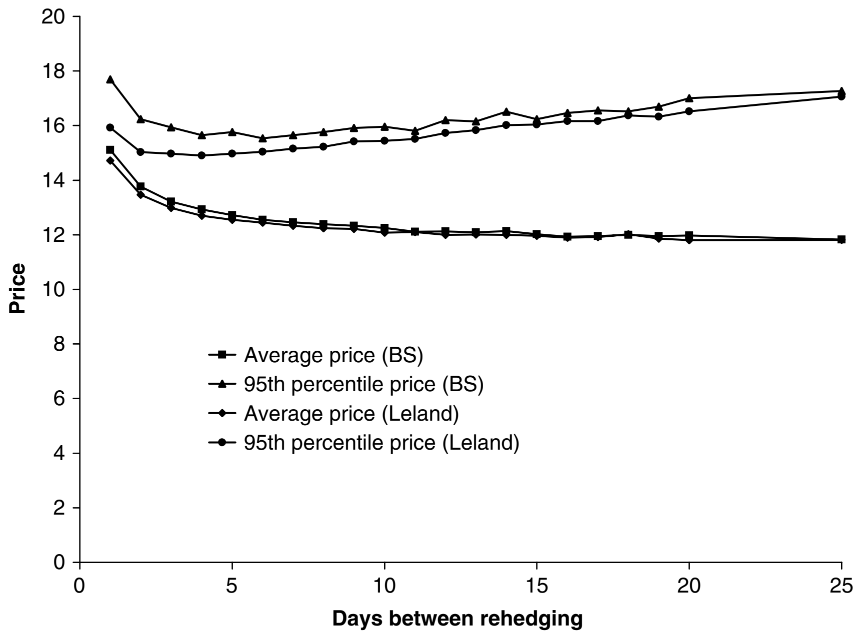


Figure 48.10 Average price and 95th percentile of price under the Black–Scholes and Leland strategies.

when the number of trades decreases. The kinks in the graph between 13 and 20 days reflect the fact that as the number of days increases, the gap between the last trade and the penultimate trade sometimes decreases, or sometimes increases, which affects the level of risk taken on.

Figure 48.10 shows the average total price of hedging the option. This is a straight average of the 5000 prices accumulated from the simulations. This graph suggests that the fewer trades the better. The Leland strategy produces a lower price than the Black–Scholes.

Figure 48.10 also shows the 95th percentile price. This is the price we must charge for the option for us to make a profit 95% of the time. This is the most informative picture from a risk management point of view.

Frequent hedging is good for risk control but bad for pricing, infrequent hedging is the opposite. The 95th percentile price is a compromise between taking on risk and incurring transaction costs. With this as our option 'value' the figure shows that the Leland model outperforms the Black–Scholes method.

The optimal number of days between rehedging for the Leland method was 4, giving a 95th percentile of 14.90.

Now we repeat all of this analysis and plotting for the delta-tolerance strategy.

48.12.2 Market Movement or Delta-tolerance Strategy

In this model the risk in the hedged position is restricted by the parameter H_0, see Equation (48.9).

Figure 48.11 shows the number of trades required on average as H_0 varies.

The next figure, 48.12, shows how much the strategy affects the average total transaction costs. Observe that instead of plotting the costs against the bandwidth I have plotted costs against

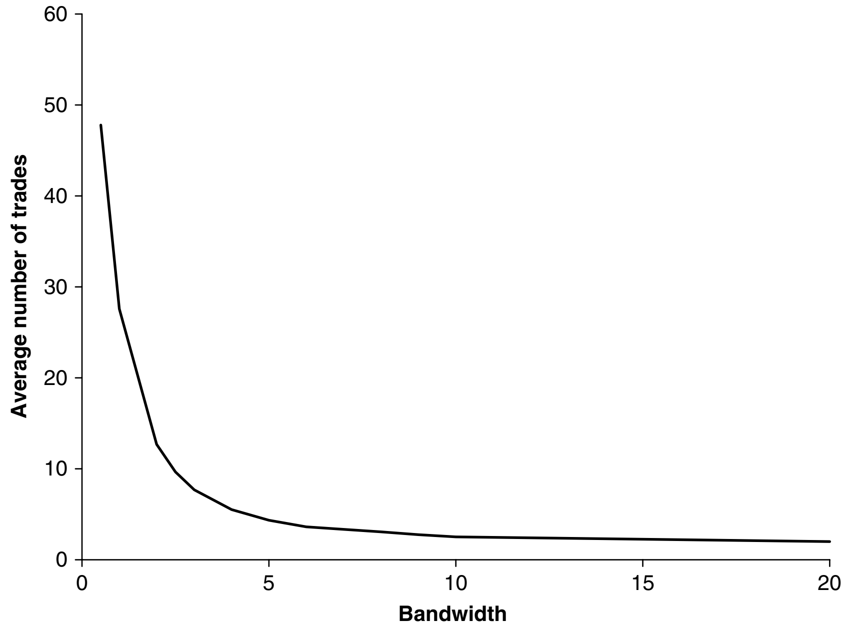


Figure 48.11 Average number of trades for the delta-tolerance strategy.

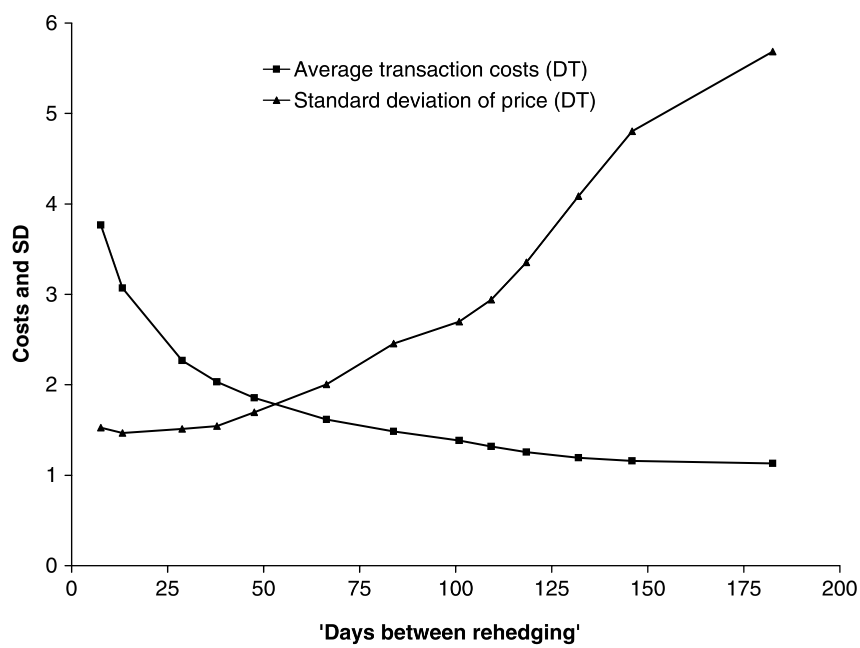


Figure 48.12 Average costs and standard deviation of price for the delta-tolerance strategy.

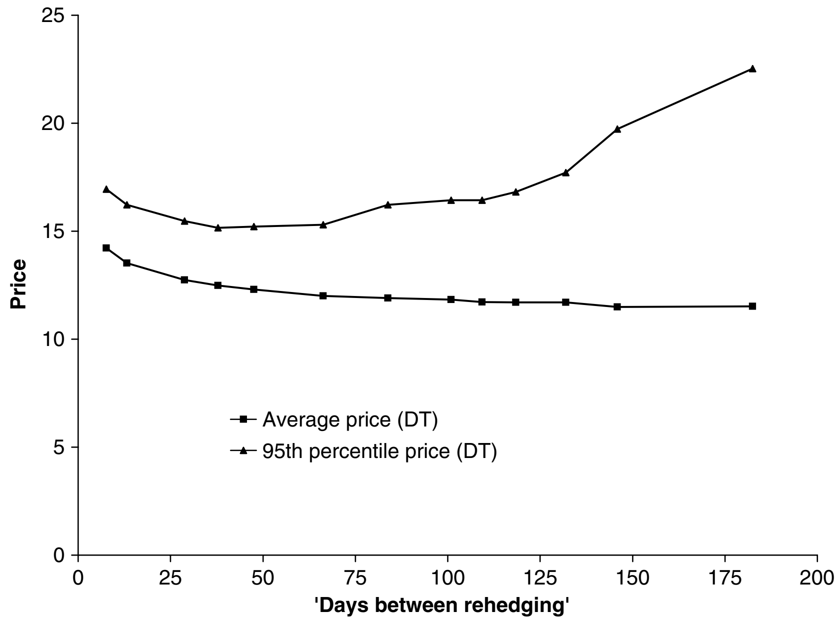


Figure 48.13 Average price and 95th percentile of price for the delta-tolerance strategy.

the inverse of the average number of trades. This is not quite the same as average number of days between rehedges, hence the inverted commas. As before restricting the number of trades restricts the amount of costs, but increases the standard deviation of price, the risk.

The average contract price is shown in Figure 48.13 along with the 95th percentile of price. As in the Leland strategy, there is a compromise point. With the 95th percentile 'value' determining this point we find that we get an option 'value' of 15.15. This is slightly worse than the Leland strategy but better than normal Black–Scholes hedging.

48.12.3 The Utility Strategy

In this strategy the parameter to be varied is γ, the level of risk aversion.

Figure 48.14 shows the average number of trades versus the risk aversion parameter. Again, this is used to convert from risk aversion to a measure of the number of days between rehedges, so that all plots can be better compared across strategies.

From the 95th percentile plot the best option 'value' is 15.03. Thus this method turned out to be better than the delta-tolerance method, but still not as good as the Leland fixed-time step hedging strategy. However over the whole range of values considered for γ this utility method produced a far lower 95th percentile price than the ranges produced from the other strategies; the 95th percentile seems to be quite insensitive to the risk aversion parameter. See Figures 48.15 and 48.16.

There are a couple of points to note about the use of the utility strategy. First, we have only used the asymptotic version since the computational time necessary for the solution of the full partial differential equation would be prohibitively large. Second, we looked at a fairly general cost model, not just proportional costs. The addition of the extra transaction costs, fixed and

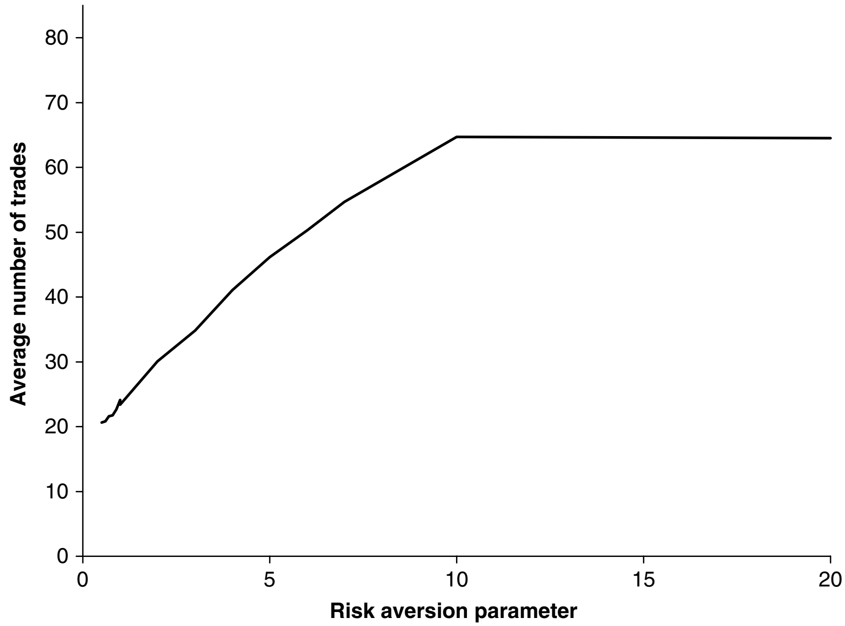


Figure 48.14 Average number of trades for the utility strategy.

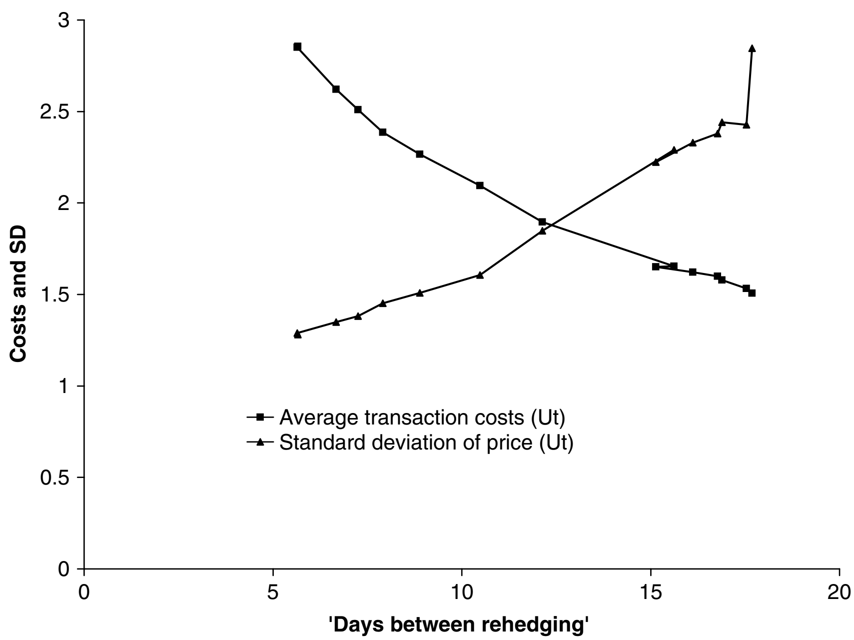


Figure 48.15 Average costs and standard deviation of price for the utility strategy.

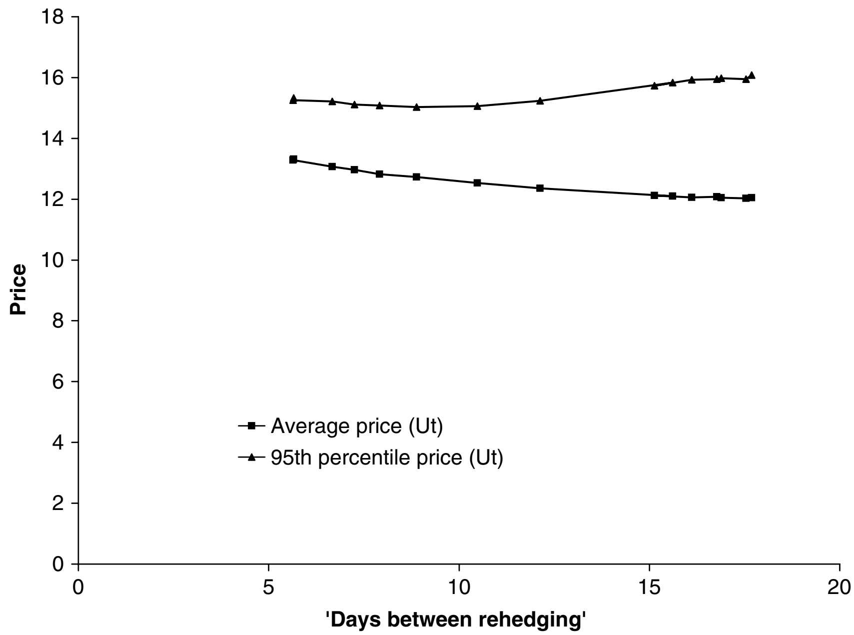


Figure 48.16 Average price and 95th percentile of price for the utility strategy.

proportional to volume, made the simulations much more general and more realistic. However, solving for the edges of the bandwidth and the rebalance point did take quite a lot of time. This wouldn't matter during the real hedging of a position, but slows down the back-testing.

48.12.4 Using the Real Data

In addition to incorporating the full cost structure into the testing framework, the strategies have been used on real data. For nearly 40 stocks, using daily prices from 1985 to 1996, one-year options were set up, commencing on 1st January, and expiring on the 31st December. The options were at-the-money when struck.

This is a more realistic test than the basic Monte Carlo simulations used above, since it involves a large range of different volatilities and starting stock prices.

48.12.5 And the Winner is. . .

Although the Leland strategy seemed to be marginally better than the others for the simulations, the results were quite different when real data were used. Admittedly, a slightly different criterion was used when deciding exactly what was meant by 'best.' Because of the relative lack of data, compared with the random simulations, the criterion used was based on which strategy gave the lowest hedging cost for each realized asset path.

The delta-tolerance method proved to be easily the best method. The delta-tolerance method won in nearly 70% of cases. The second place goes to the utility method with 23%. The Leland method produces the lowest price in only 7% of stock price paths.

Given the different results for the simulation and for the real data, it's not at all clear which is the best strategy to use.

48.13 TRANSACTION COSTS AND DISCRETE HEDGING PUT TOGETHER

In this chapter we have seen models for transaction costs and in Chapter 47 we saw a model for the effects of discrete hedging. Let's now put the two together to see what the combined effect is on our delta-hedged options. We'll work with the models based on hedging at fixed interval for simplicity.

Recall that the effect of transaction costs on our P&L is of the order of the inverse square root of the hedging period. This is an expected loss, and let's write it as

$$\frac{\alpha}{\sqrt{\delta t}}.$$

Of course, we could be more precise in what α is; we have an expression for it involving σ etc. Anyway, let's keep the following as uncluttered as possible.

The effect of discrete hedging is to add a random (approximately Normal, thanks to the CLT) component to the P&L. This is very path-dependent and therefore trickier to quantify. But at least we know that its mean is zero and its standard deviation is proportional to the square root of the hedging period. Let's write this as

$$\beta\sqrt{\delta t}.$$

Now suppose that we want to know how much delta hedging will cost us with a specified degree of confidence. In the following we will work with 95% confidence. We can say that hedging will cost us less than

$$-\frac{\alpha}{\sqrt{\delta t}} - 1.64 \times \beta\sqrt{\delta t}.$$

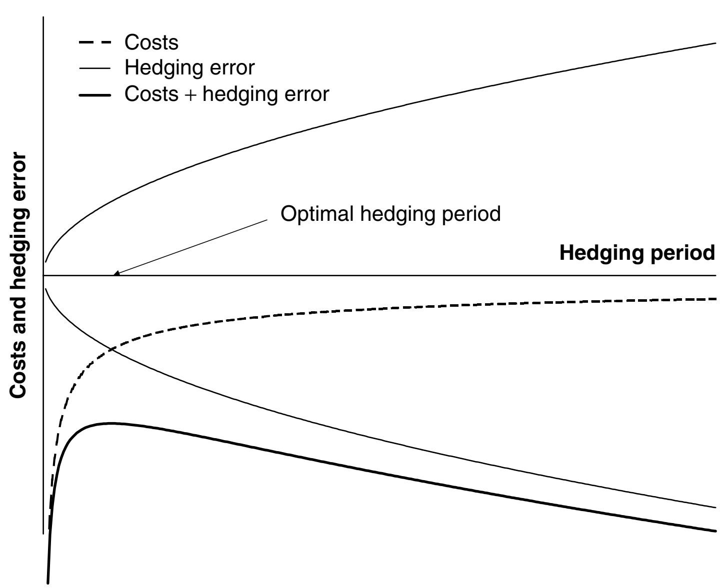


Figure 48.17 Combined effects of costs and hedging error.

The first term is the loss due to costs and the second is the risk due to hedging error, with the 1.64 factor coming from the 95% confidence for a Normal distribution. This is plotted (along with some other useful information) in Figure 48.17 as a function of δt. There is clearly an optimal hedging period. If you hedge more frequently than that your costs will get too large, but if you hedge less frequently the random hedging error will get too big.

A simple analysis such as this can show you what sort of drag you need to overcome before you will be able to make any profit from delta hedging an option position, should you have a different estimate of the future volatility than the market, for example. Note that α will be different from trader to trader, depending on the costs he incurs, but the β is the same for everyone.

48.14 SUMMARY

For equity derivatives and derivatives on emerging market underlyings transaction costs are large and important. In these markets pricing and hedging must take into account costs. In other markets the underlying may be so liquid that costs are irrelevant, then you would hedge as often as you could. When transaction costs are important the problem of pricing and hedging is non-linear; if you don't price and hedge contracts together then you will miss out on cancelation effects and economies of scale. This is interesting from a mathematical point of view, but cumbersome from a practical point of view.

FURTHER READING

- Some of the material of this chapter is based on the model of Leland (1985) as extended to portfolios of options by Hoggard, Whalley & Wilmott (1994) and Whalley & Wilmott (1993a).
- Gemmill (1992) gives an example taken from practice of the effect of transaction costs on a hedged portfolio.
- The Leland model was put into a binomial setting by Boyle & Vorst (1992). They find a similar volatility adjustment to that of Leland.
- Whalley & Wilmott (1993b, 1994b, 1996) and Henrotte (1993) discuss various hedging strategies and derive more non-linear equations using ideas similar to those in this chapter.
- For alternative approaches involving 'optimal strategies' see Hodges & Neuberger (1989), Davis & Norman (1990) and Davis, Panas & Zariphopoulou (1993), and the asymptotic analyses of Whalley & Wilmott (1994a,b, 1995, 1997) for small levels of transaction costs.
- Dewynne, Whalley & Wilmott (1994, 1995) discuss the pricing of exotic options in the presence of costs.
- Avellaneda & Parás (1994) discuss the fixed-time step equation when transaction costs are large.
- Ahn, Dayal, Grannan & Swindle (1998) discuss the variance of replication error when there are transaction costs.
- For a model of a Poisson process for stocks with transaction costs, see Neuberger (1994).

- Jefferies (1999) examines the various hedging strategies using simulated and real data. His results are those presented above.

APPENDIX: DERIVATION OF THE HOGGARD–WHALLEY–WILMOTT EQUATION

In discrete time we approximate the stochastic differential equation

$$dS = \mu S\, dt + \sigma S\, dX$$

by

$$\delta S = \mu S\, \delta t + \sigma S \phi\, \delta t^{1/2}.$$

We still set up a 'hedged' portfolio

$$\Pi = V(S, t) - \Delta S,$$

but we must accept that this cannot be perfect. After a time δt the portfolio value becomes

$$\Pi + \delta\Pi = V(S + \delta S, t + \delta t) - \Delta(S + \delta S).$$

It follows that

$$\delta\Pi = V(S + \delta S, t + \delta t) - \Delta(S + \delta S) - V(S, t) + \Delta S.$$

Expanding this in Taylor series gives

$$\begin{aligned}\delta\Pi &= \frac{\partial V}{\partial t}\delta t + \frac{\partial V}{\partial S}\delta S + \tfrac{1}{2}\frac{\partial^2 V}{\partial S^2}(\delta S)^2 + \cdots - \Delta\, \delta S \\ &= \delta t^{1/2}\sigma S\phi\left(\frac{\partial V}{\partial S} - \Delta\right) + \delta t\left(\frac{\partial V}{\partial t} + \mu S\left(\frac{\partial V}{\partial S} - \Delta\right) + \tfrac{1}{2}\sigma^2 S^2\phi^2\frac{\partial^2 V}{\partial S^2}\right) + \cdots.\end{aligned}$$

But this has not accounted for the inevitable transaction costs that will be incurred on rehedging. The costs are

$$\kappa S|\nu|.$$

The quantity ν of the underlying asset that are bought is given by the change in the delta from one time step to the next:

$$\nu = \frac{\partial V}{\partial S}(S + \delta S, t + \delta t) - \frac{\partial V}{\partial S}(S, t).$$

This can be approximated by

$$\nu = \frac{\partial V}{\partial S} + \frac{\partial^2 V}{\partial S^2}\delta S + \frac{\partial^2 V}{\partial S\,\partial t}\delta t + \cdots - \frac{\partial V}{\partial S}$$

where all derivatives are now evaluated at (S, t). Two terms in this cancel, leaving the leading-order approximation

$$\nu \approx \frac{\partial^2 V}{\partial S^2}\delta S \approx \frac{\partial^2 V}{\partial S^2}\sigma S\phi\ \delta t^{1/2}.$$

Subtracting the costs from the change in the portfolio value gives a total change of

$$\begin{aligned}\delta\Pi = {} & \delta t^{1/2}\sigma S\phi\left(\frac{\partial V}{\partial S} - \Delta\right) + \delta t\left(\frac{\partial V}{\partial t} + \mu S\left(\frac{\partial V}{\partial S} - \Delta\right) + \tfrac{1}{2}\sigma^2 S^2\phi^2\frac{\partial^2 V}{\partial S^2}\right)\\ & - \kappa\sigma S^2|\phi|\ \delta t^{1/2}\left|\frac{\partial^2 V}{\partial S^2}\right| + \cdots.\end{aligned}$$

The mean of this is

$$E[\delta\Pi] = \delta t\left(\frac{\partial V}{\partial t} + \mu S\left(\frac{\partial V}{\partial S} - \Delta\right) + \tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right) - \kappa\sigma S^2\sqrt{\frac{2}{\pi}}\delta t^{1/2}\left|\frac{\partial^2 V}{\partial S^2}\right| + \cdots.$$

This uses

$$E[\phi] = 0, \quad E[\phi^2] = 1 \quad \text{and} \quad E[|\phi|] = \sqrt{\frac{2}{\pi}}.$$

We also find that

$$\begin{aligned}E[(\delta\Pi)^2] &= \delta t\ E\left[\sigma^2 S^2\phi^2\left(\frac{\partial V}{\partial S} - \Delta\right)^2 - 2\kappa\sigma S^2\left|\frac{\partial^2 V}{\partial S^2}\right|\sigma S\left(\frac{\partial V}{\partial S} - \Delta\right)\phi|\phi|\right.\\ &\qquad \left. + \kappa^2\sigma^2 S^4\left(\frac{\partial^2 V}{\partial S^2}\right)^2\phi^2\right] + \cdots\\ &= \delta t\ \left(\sigma^2 S^2\left(\frac{\partial V}{\partial S} - \Delta\right)^2 + \kappa^2\sigma^2 S^4\left(\frac{\partial^2 V}{\partial S^2}\right)^2\right) + \cdots\end{aligned}$$

since

$$E[\phi|\phi|] = 0.$$

The variance of the portfolio change is therefore

$$\begin{aligned}\mathrm{var}[\delta\Pi] &= E[(\delta\Pi)^2] - (E[\delta\Pi])^2\\ &= \delta t\ \left(\sigma^2 S^2\left(\frac{\partial V}{\partial S} - \Delta\right)^2 + \left(1 - \frac{2}{\pi}\right)\kappa^2\sigma^2 S^4\left(\frac{\partial^2 V}{\partial S^2}\right)^2\right)\end{aligned}$$

to leading order. For finite hedging period δt and finite costs κ this cannot generally be made to vanish. However, the variance, or risk, can be minimized by choosing

$$\Delta = \frac{\partial V}{\partial S}.$$

With this choice,

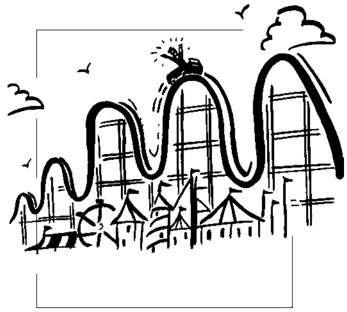
$$E[\delta\Pi] = \delta t\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) - \kappa\sigma S^2\sqrt{\frac{2}{\pi}}\delta t^{1/2}\left|\frac{\partial^2 V}{\partial S^2}\right|$$

to leading order. This quantity is an *expectation*, allowing for the *expected* amount of transaction costs. I am now going to set this quantity equal to the amount that would have been earned by a risk-free account:

$$\begin{aligned}&\delta t\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) - \kappa\sigma S^2\sqrt{\frac{2}{\pi}}\delta t^{1/2}\left|\frac{\partial^2 V}{\partial S^2}\right| \\ &\quad = r\Pi\ \delta t = r\left(V - S\frac{\partial V}{\partial S}\right)\delta t.\end{aligned}$$

On dividing by δt and rearranging gives the Hoggard–Whalley–Wilmott equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \kappa\sigma S^2\sqrt{\frac{2}{\pi\ \delta t}}\left|\frac{\partial^2 V}{\partial S^2}\right| + rS\frac{\partial V}{\partial S} - rV = 0.$$



CHAPTER 49
overview of volatility modeling

In this Chapter...

- a look at the volatility joys to come

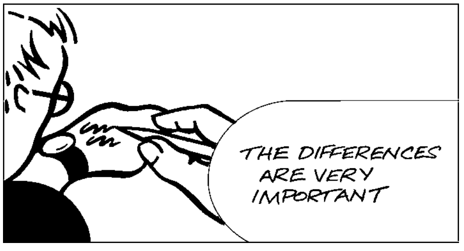
49.1 INTRODUCTION

The purpose of this chapter is to orientate you for most of the rest of this part of the book.

In the main, the Black–Scholes model is very robust and does a decent job of pricing derivatives, including exotics. The two most important flaws in the model concern the discreteness of hedging, already discussed, and the behavior of volatility. Quite frankly, we do not know what volatility currently is, never mind what it may be in the future. And the correct pricing of derivatives requires us to know what the volatility is going to be.

For this reason, volatility analysis and modeling takes a prominent role in the working life of a quant.

49.2 THE DIFFERENT TYPES OF VOLATILITY



We can't *see* volatility in the same way we can see stock prices or interest rates. The best we can hope to do is to measure it statistically. But such a measure is necessarily backwards looking, and we really want to know what volatility is going to be in the future.[1] For this reason people talk about different volatilities, as proxies for the real thing. Here are a few adjectives you can put in front of 'volatility.'

- Actual
- Historical/realized
- Implied
- Forward

[1] You can see I am already assuming that volatility is not the nice constant that earlier chapters may have led you to believe.

49.2.1 Actual Volatility

This is the measure of the amount of randomness in an asset return at any particular time. It is very difficult to measure, but is supposed to be an input into all option pricing models. In particular, the actual (or 'local') volatility goes into the Black–Scholes *Equation*.

- There is no 'timescale' associated with actual volatility; it is a quantity that exists at each instant, possibly varying from moment to moment.

Example: The actual volatility is now 20%... now it is 22%... now it is 24%...

49.2.2 Historical or Realized Volatility

This is a measure of the amount of randomness over some period in the past. The period is always specified, and so is the mathematical method for its calculation. Sometimes this backward-looking measure is used as an estimate for what volatility will be in the future.

- There are two 'timescales' associated with historical or realized volatility: one short, and one long.

Example: The 60-day volatility using daily returns. Perhaps of interest if you are pricing a 60-day option, which you are hedging daily.

In pricing an option we are making an estimate of what actual volatility will be over the lifetime of the option. After the option has expired we can go back and calculate what the volatility actually was over the life of the option. This is the realized volatility.

Example: I sold a 30-day option for a 30% volatility, I hedged it every day. Did I make money?

49.2.3 Implied Volatility

The implied volatility is the volatility which when input into the Black–Scholes option pricing formulae gives the market price of the option. It is often described as the market's view of the future actual volatility over the lifetime of the particular option.

However, it is also influenced by other effects such as supply and demand.

- There is one 'timescale' associated with implied volatility: expiration.

Example: A stock is at 100, a call has strike 100, expiration in one year, interest rates 5% and the option market price is $10.45. What volatility are traders using?

49.2.4 Forward Volatility

The adjective 'Forward' can be applied to many forms of volatility, and refers to the volatility (whether actual or implied) over some period in the future.

- Forward volatility is associated with either a time period, or a future instant.

49.3 VOLATILITY ESTIMATION BY STATISTICAL MEANS

49.3.1 The Simplest Volatility Estimate: Constant Volatility/Moving Window

If we believe that volatility is constant or slowly varying, and we have N days' data, we can use

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} R_i^2$$

where

$$R_i = \frac{S_i - S_{i-1}}{S_{i-1}}$$

and is the return on the ith day. Note that I haven't annualized quantities here.

There are obvious major problems associated with this volatility measure; because the returns are equally weighted you will get a plateauing affect associated with a large return. If there is a large one-day return it will increase the volatility instantaneously, but the estimate of volatility will stay raised until N days later when that return drops out of the sample. This is a totally spurious effect. This effect can be seen in Figure 49.1.

49.3.2 Incorporating Mean Reversion

Now let's consider time-varying volatility. We don't just have one σ but must consider σ_n, our estimate of the volatility on the nth day, using data available up to that point. If we believe

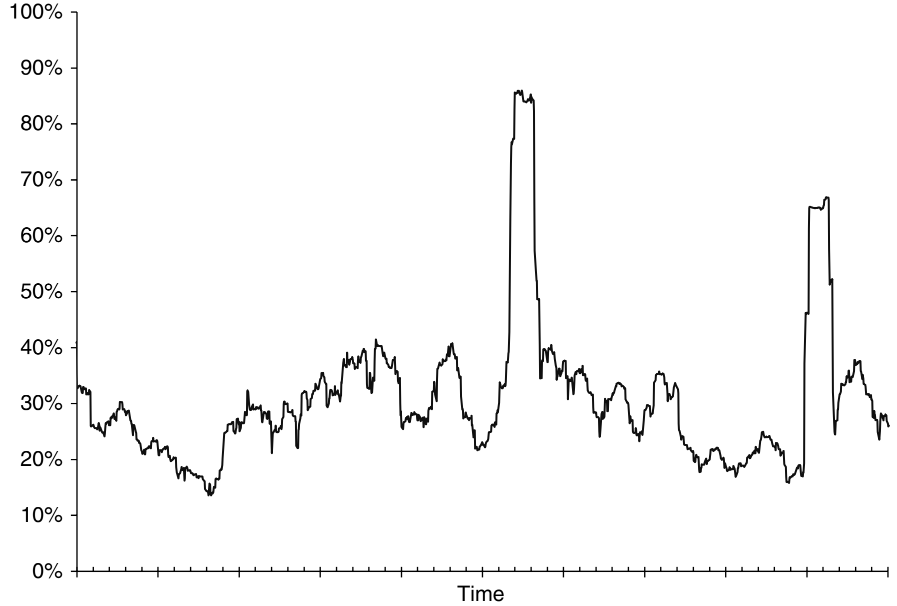


Figure 49.1 Moving-window volatility; observe the plateauing.

that volatility tends to vary about a long-term mean $\overline{\sigma}$, then we could use

$$\sigma_n^2 = \alpha\overline{\sigma}^2 + (1-\alpha)\frac{1}{n}\sum_{i=1}^{n} R_i^2.$$

Here there is a weighting assigned to each of the long-run volatility estimate and the current estimate based on the last n returns. This is called an **ARCH** model, for **Autoregressive Conditional Heteroscedasticity**.

But why should each of the last n returns be equally important?

49.3.3 Exponentially Weighted Moving Average

Consider this estimate for volatility:

$$\sigma_n^2 = (1-\lambda)\sum_{i=1}^{\infty} \lambda^{i-1} R_{n-i+1}^2.$$

The parameter λ must be greater than zero and less than one. This is an example of an **exponentially weighted moving average estimate**. The more recent the return, the more weight is attached. The sum extends back to the beginning of time.

The coefficient of $1-\lambda$ ensures that the weights all add to one.

This expression can be simplified to

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)R_n^2.$$

Note that this uses the most recent return *and* the previous estimate of the volatility. *This is the RiskMetrics volatility measure.*

Figure 49.2 uses the same stock price data as in Figure 49.1 but now there is no plateauing.[2]

49.3.4 A Simple GARCH Model

Put the preceding models together to get

$$\sigma_n^2 = \alpha\overline{\sigma}^2 + (1-\alpha)\left(\lambda\sigma_{n-1}^2 + (1-\lambda)R_n^2\right).$$

This is a **Generalized Autoregressive Conditional Heteroscedasticity** or **GARCH** model.

49.3.5 Expected Future Volatility

We are currently at day n and we want to estimate the volatility k days into the future i.e. on day $n+k$.

Expected future volatility: EWMA

Recall

$$\sigma_{n+k}^2 = \lambda\sigma_{n+k-1}^2 + (1-\lambda)R_{n+k}^2.$$

[2] You could say that the obvious exponential decay in the volatility estimate is just as spurious as the discontinuity in the equally weighted case.

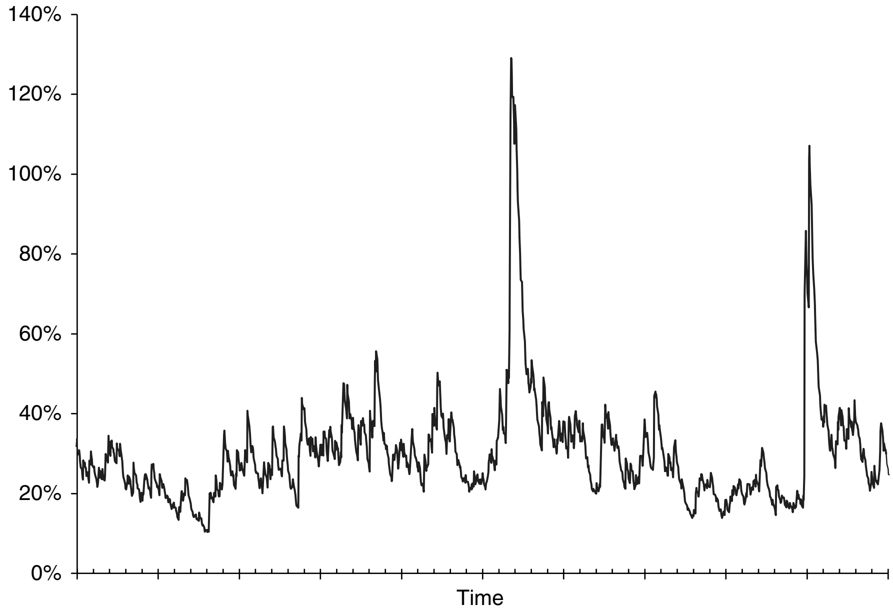


Figure 49.2 Exponentially weighted volatility; no plateauing.

Take expectations of this...

$$E[\sigma^2_{n+k}] = \lambda E[\sigma^2_{n+k-1}] + (1-\lambda)E[R^2_{n+k}].$$

But, of course,

$$E[R^2_{n+k}] = \sigma^2_{n+k}.$$

Therefore

$$E[\sigma^2_{n+k}] = \lambda E[\sigma^2_{n+k-1}] + (1-\lambda)E[\sigma^2_{n+k}]$$

or, on rearranging,

$$E[\sigma^2_{n+k}] = E[\sigma^2_{n+k-1}].$$

In other words, the expected future value of the variance is the same as the previous day's value. Working backwards to the present, the expected future variance is the same as today's.

Expected future volatility: GARCH

Recall

$$\sigma^2_{n+k} = \alpha\overline{\sigma}^2 + (1-\alpha)\left(\lambda\sigma^2_{n+k-1} + (1-\lambda)R^2_{n+k}\right).$$

Taking expectations of this results in

$$E[\sigma^2_{n+k}] = \alpha\overline{\sigma}^2 + (1-\alpha)\left(\lambda E[\sigma^2_{n+k-1}] + (1-\lambda)E[\sigma^2_{n+k}]\right).$$

On rearranging, we get

$$E[\sigma^2_{n+k}] = \frac{\alpha\overline{\sigma}^2}{1-(1-\alpha)(1-\lambda)} + \frac{\lambda(1-\alpha)}{1-(1-\alpha)(1-\lambda)} E[\sigma^2_{n+k-1}].$$

The next expected value of the variance is a weighting between the most recent value and the long-term mean. Looking further into the future

$$E[\sigma^2_{n+k}] = \overline{\sigma}^2 + \left(E[\sigma^2_n] - \overline{\sigma}^2\right)(1-\nu)^k$$

where

$$\nu = \frac{\alpha}{1-(1-\alpha)(1-\lambda)}.$$

The path of expected future volatility[3] is shown in Figure 49.3.

49.3.6 Beyond Close-Close Estimators: Range-based Estimation of Volatility

The problem with estimating volatility is that you need lots and lots of data to avoid sampling-error problems.

But then if you use too many days' worth of data you will be trying to estimate a parameter during a period when that parameter is almost certainly varying.

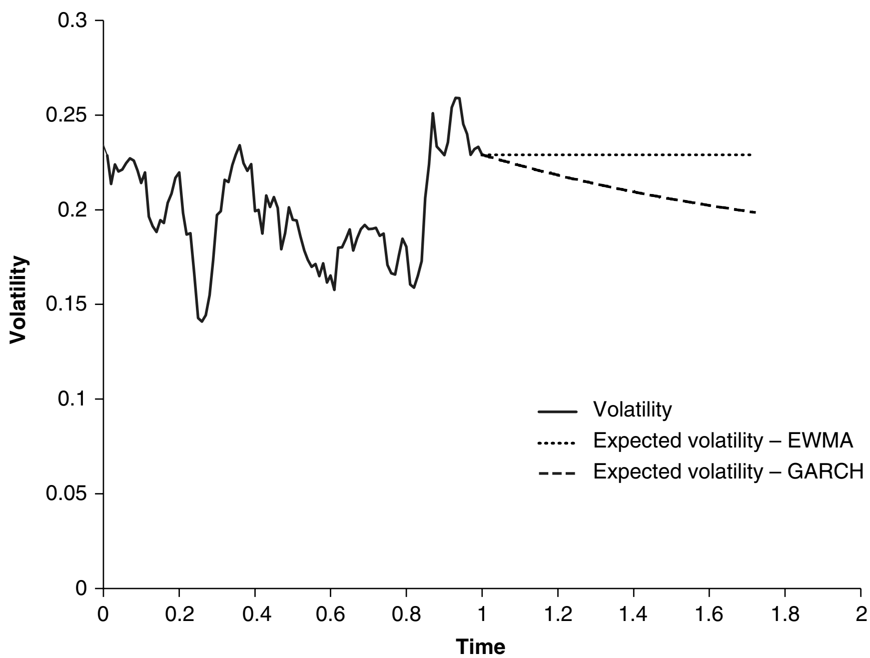


Figure 49.3 Expected future volatility for EWMA and GARCH models.

[3] Actually variance, of course, which is more important than volatility anyway.

A Catch 22 situation.

Or is it?

Why not use more information contained within a single day? That is, go down to finer timescales for the data. The problem with that is the behavior of returns over very short timescales, such as minutes, does not appear to be Normally distributed... there is even some evidence that the returns do not have a finite standard deviation.

Setting aside such worries(!), let's look at very simple ways of better estimating volatility using readily available price data, and not just closing prices.

Traditional close-to-close measure:

When drift is small

$$\sigma_{cc}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\log\left(\frac{C_i}{C_{i-1}}\right)\right)^2.$$

Here there is a slight change of notation from before; C_i is the closing price on the ith day. Note also that we are looking at logarithms.

To adjust this for the drift take

$$\sigma_{acc}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\left(\log\left(\frac{C_i}{C_{i-1}}\right)\right)^2 - \frac{\left(\log\left(\frac{C_n}{C_0}\right)\right)^2}{n(n-1)}\right).$$

(Don't forget to annualize: multiply by square root of number of trading days in a year.)

Parkinson (1980)

This estimator uses extreme value, the highs H and the lows L during the day.

$$\sigma_p^2 = \frac{1}{4n\log(2)}\sum_{i=1}^{n}\left(\log\left(\frac{H_i}{L_i}\right)\right)^2.$$

This is five times more efficient than the close-to-close estimate. (That means, for the same amount of data the variance of the data is one fifth that of the close-to-close measure.)

Garman & Klass (1980)

At 7.4 times more efficient than close-to-close, we have

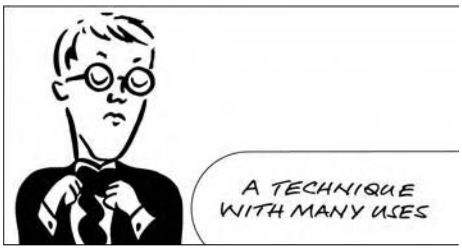
$$\sigma_{gk}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(0.511\left(\log\left(\frac{H_i}{L_i}\right)\right)^2 \right.$$
$$\left. -0.019\log\left(\frac{C_i}{O_i}\right)\log\left(\frac{H_iL_i}{O_i^2}\right) - 2\log\left(\frac{H_i}{O_i}\right)\log\left(\frac{L_i}{O_i}\right)\right).$$

Here O_i is the opening price.

Rogers & Satchell (1991)

Parkinson and Garman & Klass are not independent of the drift. Our final simple volatility estimate is:

$$\sigma_{rs}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\log\left(\frac{H_i}{C_i}\right)\log\left(\frac{H_i}{O_i}\right) + \log\left(\frac{L_i}{C_i}\right)\log\left(\frac{L_i}{O_i}\right)\right).$$



49.4 MAXIMUM LIKELIHOOD ESTIMATION

A large part of statistical modeling concerns finding model parameters. Two popular ways of doing this are regression and Maximum Likelihood Estimation (MLE). We look at the second method here.

49.4.1 A Simple Motivating Example: Taxi Numbers

You are attending a maths conference. You arrive by train at the city hosting the event. You take a taxi from the train station to the conference venue. The taxi number is 20,922. How many taxis are there in the city?

This is a parameter estimation problem. Getting into a specific taxi is a probabilistic event. Estimating the number of taxis in the city from that event is a question of assumptions and statistical methodology.

The assumptions

1. Taxi numbers are strictly positive integers
2. Numbering starts at one
3. No number is repeated
4. No number is skipped

The statistical methodology

We will look at the probability of getting into taxi number 20,922 when there are N taxis in the city. This couldn't be simpler, the probability of getting into any specific taxi is

$$\frac{1}{N}.$$

This is shown in Figure 49.4.

Which N maximizes the probability of getting into taxi number 20,922?

$$N = 20{,}922.$$

This example explains the concept of MLE: *Choose parameters that maximize the probability of the outcome actually happening.*

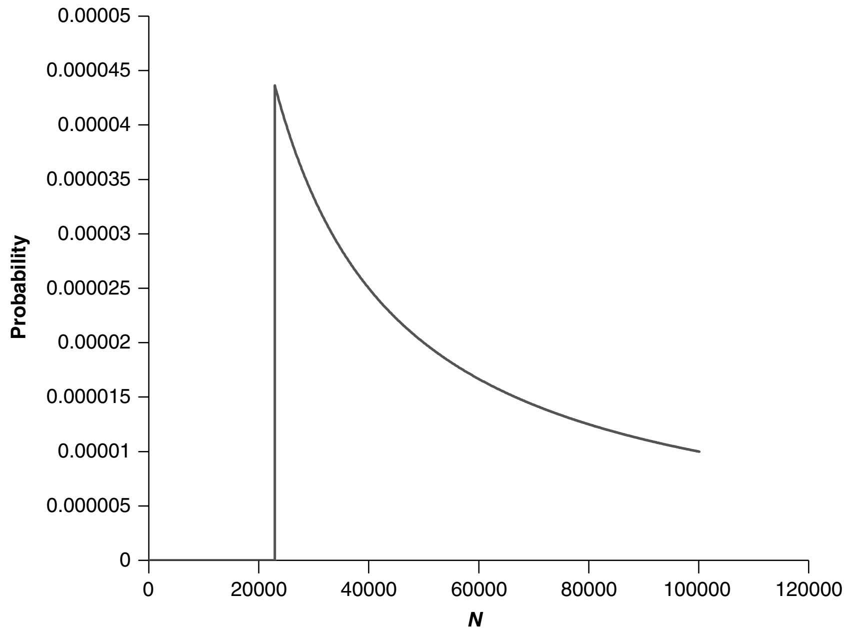


Figure 49.4 Probability of getting into the *N*th taxi.

49.4.2 Three Hats

You have three hats containing Normally distributed random numbers. One hat's numbers have mean of zero and standard deviation 0.1. This is hat A. Another hat's numbers have mean of zero and standard deviation 1. This is hat B. The final hat's numbers have mean of zero and standard deviation 10. This is hat C.

You don't know which hat is which.

You pick a number out of one hat, it is -2.6. Which hat do you think it came from?

The 'probability' of picking the number -2.6 from hat A (having a mean of zero and a standard deviation of 0.1) is

$$\frac{1}{\sqrt{2\pi}\ 0.1}\exp\left(-\frac{2.6^2}{2\times 0.1^2}\right) = 6\ \ 10^{-147}.$$

Very, very unlikely! (Note: The word 'probability' is in inverted commas to emphasize the fact that this is the value of the probability density function, not the actual probability. The probability of picking exactly -2.6 is, of course, zero.)

The 'probability' of picking the number -2.6 from hat B (having a mean of zero and a standard deviation of 1) is

$$\frac{1}{\sqrt{2\pi}\ 1}\exp\left(-\frac{2.6^2}{2\times 1^2}\right) = 0.014,$$

and from hat C (having a mean of zero and a standard deviation of 10)

$$\frac{1}{\sqrt{2\pi}\ 10}\exp\left(-\frac{2.6^2}{2\times 10^2}\right) = 0.039.$$

We would conclude that hat C is the most likely, since it has the highest probability for picking the number -2.6.

We now pick a second number from the same hat, it is 0.37. This looks more likely to have come from hat B. We get the following table of probabilities:

Hat	-2.6	0.37	Joint
A	$6\ \ 10^{-147}$	0.004	$2\ \ 10^{-149}$
B	0.014	0.372	0.005
C	0.039	0.040	0.002

The second column represents the probability of drawing the number -2.6 from each of the hats, the third column represents the probability of drawing 0.37 from each of the hats, and the final column is the joint probability, that is, the probability of drawing both numbers from each of the hats.

Using the information about *both* draws, we can see that the most likely hat is now B.

49.4.3 The Maths Behind This: Find the Standard Deviation

You have one hat containing Normally distributed random numbers, with a mean of zero and a standard deviation of σ which is unknown. You draw N numbers ϕ_i from this hat. Estimate σ.

Q. What is the 'probability' of drawing ϕ_i from a Normal distribution with mean zero and standard deviation σ?

A. It is

$$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{\phi_i^2}{2\sigma^2}}.$$

Q. What is the 'probability' of drawing all of the numbers $\phi_1, \phi_2, \ldots, \phi_N$ from independent Normal distributions with mean zero and standard deviation σ?

A. It is

$$\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{\phi_i^2}{2\sigma^2}}.$$

... choose the σ that maximizes this quantity. This is easy...

Differentiate with respect to σ (take logarithms first) and set result equal to zero:

$$\frac{d}{d\sigma}\left(-N\log(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\phi_i^2\right) = 0.$$

(A multiplicative factor has been ignored here.) That is

$$-\frac{N}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{N}\phi_i^2 = 0.$$

Therefore our best guess for σ is given by

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\phi_i^2.$$

This should be a familiar expression!

49.4.4 Quants' Salaries

In Figure 49.5 are the results of a 2004 survey on `www.wilmott.com` concerning the salaries of quants using the Forum.

This distribution looks vaguely lognormal, with distribution

$$\frac{1}{\sqrt{2\pi}\sigma E}\exp\left(-\frac{(\log E - \log E_0)^2}{2\sigma^2}\right),$$

where E is annual earnings, σ is the standard deviation and E_0 the mean. Using MLE find σ and E_0.

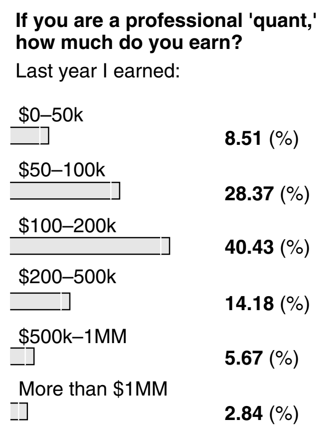


Figure 49.5 Distribution of quants' salaries.

Figure 49.6 Distribution of quants' salaries and MLE fit.

The MLE solution is shown in Figure 49.6.
The mean $E_0 = \$133{,}284$, with $\sigma = 0.833$.

49.5 SKEWS AND SMILES

We have briefly already mentioned skews and smiles, but here is a reminder of what they are about.

For a series of options that all expire at the same date, plot the value of implied volatility against strike. *If* actual volatility were constant, and *if* the Black–Scholes model were correct and *if* people priced options correctly then that plot would be flat; all options would have the same implied volatility. Of course, none of those assumptions is correct, and so there is plenty of scope for that plot to be curved, or skewed. If there is an appreciable slant to that curve, for example if it goes from top left to bottom right, then we have what is called a skew. A negative skew is downward sloping and a positive skew upward sloping. If there is curvature so that the curve has a minimum in the middle then we have a smile. Skews and smiles are the market's way of telling us that either they don't believe in the Black–Scholes model or its assumptions or they don't care.

'They don't care'? If out-of-the-money puts are expensive, does it matter? Perhaps not if they still cost only pennies and they are the easiest or cheapest way of getting downside protection. To not buy needed downside protection because implied volatility seems a bit high would be a foolish economy. 'Penny wise, pound foolish.'

So smiles and skews may give us information about future volatility or they may give us information about people's expectation of future volatility or they may give us an idea of how desperately they need to hedge. All very useful stuff.

You can speculate on what you think implied volatility might be going to do in the future by buying or selling specific portfolios of vanilla options. If you want to speculate on level of implied volatility you may buy or sell a straddle, and if you want to speculate on skew then you would consider a risk reversal. Let's see how this works.

49.5.1 Sensitivity of the Straddle to Skews and Smiles

The straddle is a portfolio made up of a long position in a call and a put, both having the same strike and expiration. Because of put-call parity the market will price these two contracts with the same volatility (bid-offer spread aside) and so there is really only one implied volatility to consider. Figure 49.7 shows the Black–Scholes value of this simple portfolio using three different levels of volatility. The strike is 100, the interest rate is 5% and there is one year to expiration. The Black–Scholes value of this portfolio is monotonic in the volatility; increase the volatility and the value rises. Such a portfolio is therefore ideal when it comes to speculating on implied volatility.

49.5.2 Sensitivity of the Risk Reversal to Skews and Smiles

The risk reversal is made up of a long call and a short put, the call having a higher strike than the put. Now that there are two strikes to consider we can see that the implied volatility skew is important. Figure 49.8 shows how the value of the risk reversal varies with the skew. In the case of the negative skew we have used a 25% volatility for the put, having strike 80, and a 15% volatility for the call, having a strike of 120. In the positive skew example these

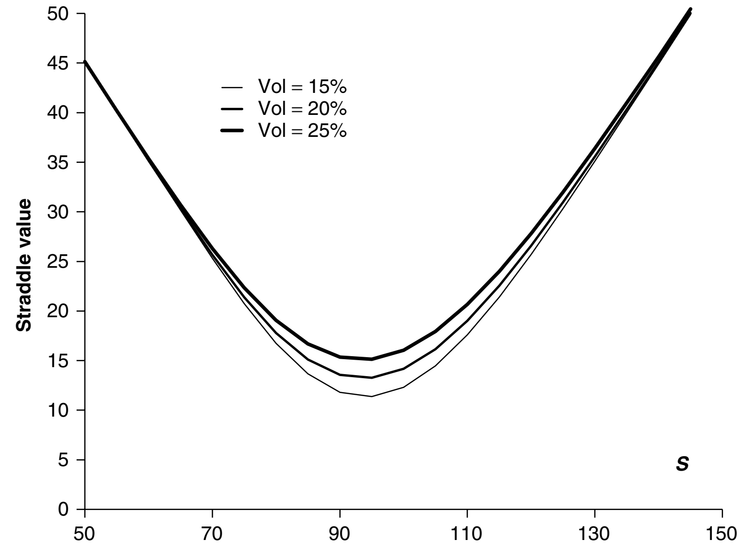


Figure 49.7 How the value of a straddle varies with level of volatility.

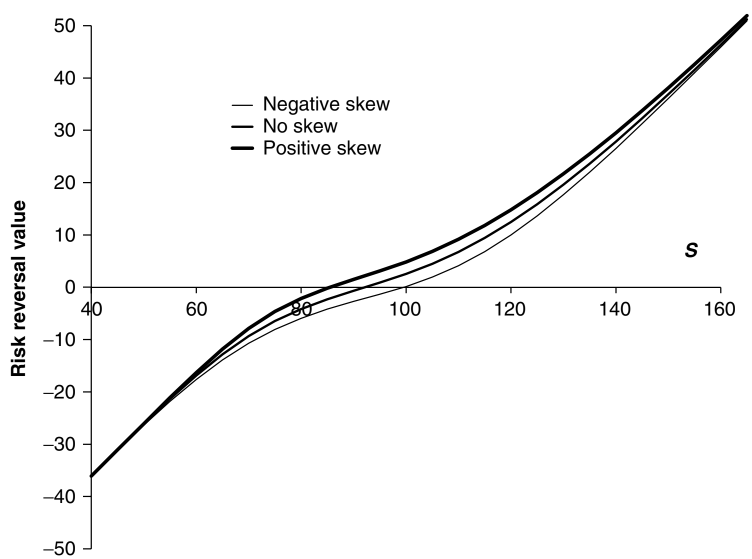


Figure 49.8 How the value of a risk reversal varies with volatility skew.

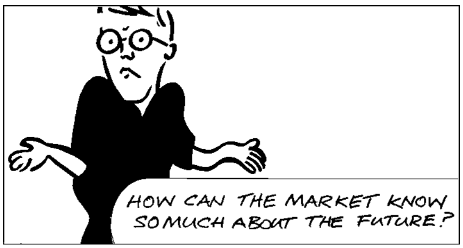
quantities are reversed and with the no-skew example both options have a strike of 20%. The risk reversal's sensitivity to skew makes it perfect for speculation on that skew.[4]

If you want to speculate on the smile what contract would you buy? Simple, you need a payoff with three kinks in it, a butterfly spread.

[4] At a meeting of the Committee on September 15th 2005 it was unanimously voted that sensitivity to skew be henceforward called 'Xena.' Present were Dominic Connor, Mike Weber and the author.

49.6 DIFFERENT APPROACHES TO MODELING VOLATILITY

Much of the rest of this part of the book focuses on sophisticated ways to approach the modeling of volatility. For the rest of this chapter I want to provide a brief overview of what is to come, and put it all into context.



49.6.1 To Calibrate or Not?

Perhaps the biggest question to face is that of how much information should we take from the prices of options in the markets. Given that we can't see what volatility is at any instant, and given that forecasting volatility is not easy (see just a small subset of possibilities above), we might be tempted to use implied volatility as an estimate of future volatility. What we might say is that implied volatility is the market's best estimate of what volatility will do in the future. In this sense, the implied volatility of a three-month option can be thought of as containing information about actual volatility over the next three months.

Well, yes and no. In some sense, perhaps. But I personally don't believe that the market is that statistically sophisticated. If people want to spend too much buying an option or charge too little when selling it that is their business, it doesn't mean that their prices are right. It is no different from the price of a liter of milk in your corner shop. Has the shopkeeper done some detailed analysis of the utility of individual customers, pricing in the opportunity cost of going to the corner shop versus the cheaper supermarket, allowing for the future price of petrol or bus fares etc.? No, he knows what it costs to buy the milk and he just tries to sell it for as much above that as he possibly can. Out-of-the-money puts are expensive, due to demand, otherwise known as 'fear'. Out-of-the-money calls are cheap, due to supply, otherwise known as 'greed'. Yes, there is arbitrage. Every business is based on arbitrage, so it is to be expected that arbitrage (or mispricing as theoreticians call it, or earning a living, as sensible people call it) exists. So I wouldn't be too keen on accepting that there is a great deal of information about volatility within option prices.

Anyway, whether to accept prices as containing information is part of the subject of calibration. If you have an option pricing model, should it output theoretical prices that are exactly the same as quoted market prices?

The subject of calibration will crop up a lot from now on.

49.6.2 Deterministic Volatility Surfaces

In Chapter 50 we look at the simplest model that is consistent with the market prices of options. In other words, that chapter is entirely about calibration and the assumption, in a sense, that market prices are correct. We work in the classical Black–Scholes world, with the one exception that actual volatility is not a constant but a function of stock price and time, $\sigma(S, t)$.

Stock- and time-dependent actual volatility are still entirely consistent with the Black–Scholes partial differential equation, you just have $\sigma(S, t)$ in the coefficient of the gamma term. The big difference from the constant-vol world is just that we cannot generally find closed-form solutions for option prices anymore.

The way that this idea works is that you have enough freedom in the function $\sigma(S, t)$ to make the theoretical option values (found numerically, say) consistent with all option values in the market, or, totally equivalently, consistent with all implied vols. Now normally you would specify the actual vol, work out the option prices and so the implied vols. This is the natural forward problem. Not so with this form of calibration. Here we specify the implied vols and ask what actual volatilities are consistent with them. This is what is known to mathematicians as an inverse problem.

Inverse problems are notoriously tricky to solve. One reason is the sensitivity of the result (actual volatilities) to the input (implied volatilities). If you specify actual volatilities and work out implied, then this is a smoothing operation; implied volatilities are averages of actuals. But going the other way does the opposite of smoothing. Make a very small change to implied volatilities and this can have an enormous effect on $\sigma(S, t)$.

If our problem were a diffusion equation from a nice physical problem, this might not matter. Having faith in the diffusion model, finding the parameters (actual volatilities), is a matter of 'regularization.' Unfortunately, we cannot have faith in our model, it is financial not physical, and so the sensitivity to initial input is something we have to live with.

This calibration idea has been around since 1993 (references are given at the end of Chapter 50). I have hated it with every atom of my being since the day I first read about it.[5]

One reason is that common sense says it cannot be right, the financial world is too complicated to obey the rather simplistic $\sigma(S, t)$ model. A second reason is that it is very, very easy to test scientifically. Do the following. Find $\sigma(S, t)$ using option data one day. Then come back a week later, when stock prices and option prices have all changed. Now recalibrate to find $\sigma(S, t)$ again. If the model is right then the first and second calculations will give exactly the same actual volatilities. Do they? No. Never.[6]

49.6.3 Stochastic Volatility

In Chapter 51 we move away from deterministic volatility to examine the idea of stochastic volatility; actual volatility that is itself random. Without a doubt it is impossible to forecast accurately future actual volatility so it makes sense to treat that quantity as being random. After all, modeling stock prices as random was the great breakthrough in pricing derivatives based on stocks, so perhaps modeling volatility as random is also a great breakthrough since options are really all about future volatility. Expect to read about concepts such as volatility of volatility in this chapter.

Although stochastic volatility models are commonly used in practice they do suffer from two major problems. First, what model should we choose? What is the volatility of volatility? Given that we can't even measure volatility at any moment particularly accurately, how on earth can we measure the volatility of that immeasurable volatility? Well, we can't. You'll see in this chapter how there are many such stochastic volatility models, named after very famous quants. You'll also see that the criteria for choosing which model to use are not at all the criteria you might expect.

The second problem concerns hedging. When only the stock price is random we have only one source of randomness. This randomness can be hedged away from an option by using a position in the underlying stock. One source of randomness, one traded quantity for hedging that

[5] Sometimes I feel like James Stewart in *No Highway*, except that I don't have the good looks and charisma, obviously.

[6] Phlogiston theory: Hypothesis regarding combustion. The theory, advanced by J.J. Becher late in the 17th century and extended and popularized by G.E. Stahl, postulates that in all flammable materials there is present phlogiston, a substance without color, odor, taste, or weight that is given off in burning. 'Phlogisticated' substances are those that contain phlogiston and, on being burned, are 'dephlogisticated.' The ash of the burned material is held to be the true material. Source: *The Columbia Electronic Encyclopedia*, 6th ed., 2004. Many financial theories seem very phlogiston-like to me. It is time for some debunking!

randomness, end result no randomness at all. Now, when volatility is stochastic we have two sources of randomness, the stock and volatility. But we still only have the one traded quantity to hedge with, the stock. We can't hedge with volatility to remove volatility risk because that isn't traded. We can, sort of, get rid of the volatility risk in one option by hedging with another option, an exchange-traded vanilla. The end result is a hedged portfolio, but now we have one equation for two unknowns, the value of the original option and the value of the option used for hedging. Oh dear. How we get out of that one is revealed in Chapter 51; it requires the introduction of a concept known as the market price of volatility risk.

49.6.4 Uncertain Parameters

First deterministic volatility, then random volatility. That just about covers the possibilities, no? No. There is also 'uncertain' volatility. This is subtly, but importantly, different from random. Random is when you have a probabilistic description, perhaps even a probability density function, for the random variable. Uncertainty is when you have no such concept. Models using uncertainty are therefore far more vague than models using randomness (and I mean 'vague' in a good way.)

In Chapter 52 we see a model in which volatility is treated as uncertain. The simplest such model allows volatility to lie within a range. But there is no mention made of how likely the volatility is to be at any point in that range; it is a genuine model of uncertainty not of randomness. Given a range of possible values for volatility, we find that there is a range of possible values for an option. Furthermore, long and short positions take different values, and the whole business of pricing derivatives becomes non linear. It turns out that the governing partial differential equation is exactly the Hoggard–Whalley–Wilmott equation we saw in Chapter 48.

49.6.5 Empirical Analysis of Volatility

We can't talk about stochastic volatility models without suggesting a way of analyzing data to derive a decent such model. So in Chapter 53 we look at how to analyze data to back out the volatility of volatility and the drift of volatility. The technique we use here is one that I find very useful for determining models for all sorts of financial data, and we'll see it again when we look at interest rate modeling.

49.6.6 Static Hedging

Hedging is used to reduce or, if possible, to eliminate risk. In options theory and practice we dynamically delta hedge to eliminate the stock-price risk in an option. We have seen such theory several times in this book. But there are other, perhaps more obscure, risks that we are only just starting to worry about here. One of those risks is caused by volatility. The plain and simple fact is that we don't know what volatility is going to be in the future therefore when we price an option we will be exposed to the input volatility being wrong; we therefore have volatility risk.

In Chapter 51 is discussed how to hedge one option with another dynamically, an exchange-traded vanilla, to eliminate volatility risk. As mentioned above this has associated conceptual problems. It also has the practical problem that you can't dynamically hedge with options; it would be far, far too costly since bid-offer spread on options is typically very large.

In practice, instead of dynamically hedging with vanilla options, we statically hedge. To see how this works, imagine that we want to price and hedge some exotic contract. Suppose that the payoff from that exotic contract can somehow be closely matched by the payoff from a portfolio of vanillas. If it can, then you hedge with those vanillas leaving only a small residual payoff which needs delta hedging. The original contract may have been very difficult to delta hedge, and so very exposed to your volatility input. But after statically hedging with the vanillas the delta-hedging task is much easier. And, of course, you know exactly what the portfolio is worth (the contracts are exchange traded) so you know almost exactly what the exotic should be worth. In practice, you won't be able to get a perfect match between the exotic and the vanillas because if you could then the exotic wasn't exotic at all.

In Chapter 60 we explain various ways of going about the process of static hedging in theory and in practice.

49.6.7 Stochastic Volatility and Mean-variance Analysis

As mentioned above, one of the drawbacks with stochastic volatility models is that when you build up the governing equation you assume that you can dynamically hedge with options. And the equation you end up with contains the market price of volatility risk, a wonderful concept in theory but rather too unreliable for practice. There are ways of tidying this all up without these two problems, and that is to accept that not all risk can be dynamically hedged away. So in Chapter 54 we see how to build up a perfectly consistent pricing—and a dynamic and static hedging framework provided one only dynamically hedges with stock—and statically hedge with vanillas.

49.6.8 Asymptotic Analysis of Volatility

When deciding which stochastic volatility model to choose we often face the dilemma of whether to pick an easy-to-compute model that might not be so accurate, or a slower but more scientifically precise model. Practitioners almost always go for fast and inaccurate. However, by exploiting a little known applied math concept known as asymptotic analysis[7] you can actually get the best of both worlds. Asymptotic analysis is all about finding approximate solutions to differential equations, models, whatever, by exploiting the relative largeness or smallness of a parameter in the model. If some terms in a complicated equation are multiplied by a small number then perhaps those terms don't matter; so we can ignore them, leaving us with a simpler equation that we may be able to solve. Of course, it is nowhere near as simple as that, but it gives you the flavor of the technique.

In Chapter 55 we look at applying asymptotic techniques to find closed-form solutions to stochastic volatility pricing models.

49.6.9 Volatility Case Study: The Cliquet Option

Our examination of volatility ends with a case study, the cliquet option, in Chapter 56. This is a particularly popular exotic option but one with which several banks have lost a lot of money. The way in which volatility impacts on the price of this contract is subtle, to say the least, so

[7] Little known in finance, at least. To applied mathematicians asymptotic analysis is a very commonly used tool in their toolbox. Prior to succumbing to the lure of finance almost every mathematical model I worked on used asymptotic analysis at some stage, mainly because they would otherwise have been far too complicated to solve or understand.

it makes the perfect subject for an in-depth study which I hope will reveal how important it can be to get your volatility model right.

49.7 THE CHOICES OF VOLATILITY MODELS

Model	Math	Popularity
Constant vol.	$\sigma =$ constant, Black–Scholes formulae	Very, especially for vanillas
Deterministic vol.	$\sigma(S, t)$, Black–Scholes pde	Very, for exotics
Stochastic vol.	$d\sigma = \cdots$, higher dimensions, transforms	Very, for exotics
Jump diffusion	Poisson processes, jumps in stock and/or vol.	Increasing
Uncertain vol.	$\sigma^- \le \sigma \le \sigma^+$, non-linear pde	Not at all, unfortunately
Stoch vol. and mean-variance	$d\sigma = \cdots$, higher dimensions, nonlinearity	Not at all, unfortunately

49.8 SUMMARY

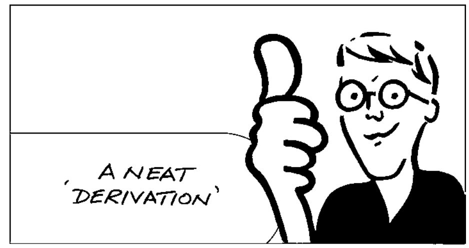
Derivatives are all about volatility. You can't price or hedge derivatives without a decent model for volatility. And if your volatility model is better than the market's you could make money by speculating on volatility. In the chapters to come we will be seeing many interesting volatility models, some better than others, some more practical than others. One point to watch out for is how much information we back out from vanilla option prices into our volatility model. Never forget that volatility is a property of the stock and would still be present even if there weren't any derivatives! Indeed, how on earth can markets possibly have perfect knowledge of the future, and the future volatility of the stock? Of course they can't. Option prices may be governed to some extent by what people expect to happen in the future in some rational sense, but they are also governed by fear and greed as interpreted by option prices through simple supply and demand. Option prices, and hence volatility, will rise when people panic and rush to buy those OTM puts, regardless of whether this panic is rational. And if you do have to pay 10 cents for an option that may only be worth 7 cents, well so what? It's still pennies after all. However, translate those pennies into implied volatility numbers and suddenly it looks like the market is expecting volatility to rise.

FURTHER READING

- Natenberg (1994), of course.
- Shu and Zhang (2003) look at data for SP500 index option volatility and compare with realized volatility.
- For something more (much more) quanty, see Rebonato (2004)
- Javaheri (2005) on volatility arbitrage is an essential read.

- A survey of stochastic volatility models, their calibration and how very differently they price the same contracts, is given in Schoutens, Simons & Tistaert (2004).

APPENDIX: HOW TO DERIVE BS PDE, MINIMUM FUSS



Here is how to derive the Black–Scholes equation, with the minimum fuss and the minimum (explicit) assumptions. This derivation also shows the special, almost arbitrary role, played by volatility.

Working with any Brownian-motion type model for stock price dynamics will give you a linear diffusion equation for pretty much everything, including option values. If $V(S, t)$ is the option's value then the general linear diffusion equation (and homogeneous, since the value of an empty portfolio should be zero) is

$$\frac{\partial V}{\partial t} + a\frac{\partial^2 V}{\partial S^2} + b\frac{\partial V}{\partial S} + cV = 0$$

where a, b and c are, for the moment, arbitrary.

A solution of this equation must be cash, i.e. $e^{-r(T-t)}$. Plug this function into the partial differential equation to find that

$$c = -r.$$

Another solution is the stock, S. Plug this function in and you will find

$$b = rS.$$

This gives you the risk-neutral pricing equation

$$\frac{\partial V}{\partial t} + a\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

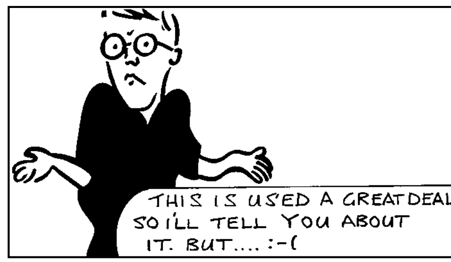
The only remaining 'real' or 'arbitrary' parameter left to determine or fudge/calibrate is a. And we know (from what we've done before) that this is related to volatility. So, this shows us a couple of things:

- Basic considerations pin down the coefficients of most of the terms in the pricing equation, and are clearly risk neutral as well.
- To determine or model the diffusion coefficient you will need to go to greater lengths. Or more simply, just use this function as the only one you can arguably calibrate.

CHAPTER 50
deterministic volatility surfaces

In this Chapter...

- volatility smiles and skews
- the implied volatility surface
- calibration: how to determine the local volatility surface that gives prices of options that are consistent with the market



50.1 INTRODUCTION

One of the erroneous assumptions of the Black–Scholes world is that the volatility of the underlying is constant. The fluctuations of volatility can be seen in any statistical examination of time-series data for assets, regardless of the sophistication of the analysis. Take a look at Figure 50.1 to see how volatility appears to change with time. This varying volatility is also observed *indirectly* through the market prices of traded contracts. In this chapter we are going to examine the relationship between the volatility of the underlying asset and the prices of derivative products. Since the volatility is not directly observable, and is certainly not predictable, we will try to exploit the relationship between prices and volatility to determine the volatility *from* the market prices. This is the exact inverse of what we have done so far. Previously we measured/modeled the volatility and then found the price, now we take the price and deduce the volatility.

50.2 IMPLIED VOLATILITY

In the Black–Scholes world of constant volatility, the value of a European call option is simply

$$V(S, t; \sigma, r; E, T) = SN(d_1) - Ee^{-r(T-t)}N(d_2),$$

where

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

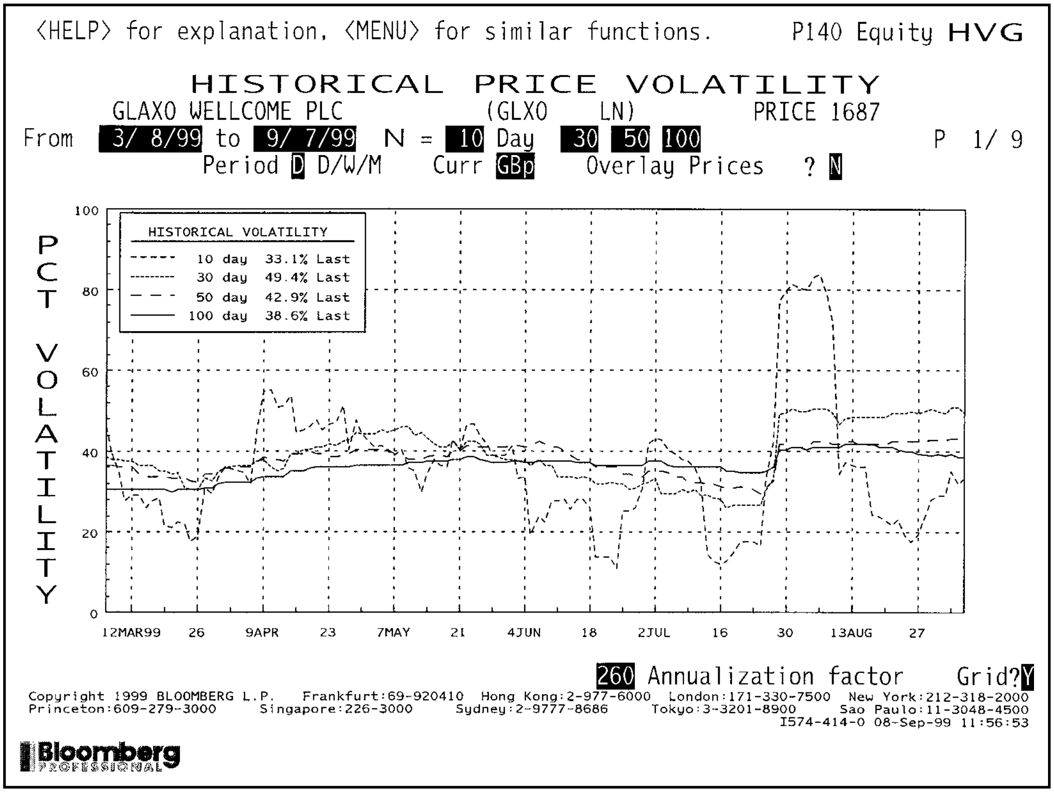


Figure 50.1 Historical volatility using several time periods.

and

$$d_2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

I have given the function V six arguments, the first two are the independent variables, the second two are parameters of the asset and the financial world, and the last two are specific to the contract in question. All but σ are easy to measure (r may be a bit inaccurate but the price is typically not too sensitive to this). If we know σ then we can calculate the option price. Conversely, if we know the option price V then we can calculate σ. We can do this because the value of a call option is monotonic in the volatility. Provided that the market value of the call option is greater than $\max(S - E^{-r(T-t)}, 0)$ and less than S there is one, and only one, value for σ that makes the theoretical option value and the market price the same. This is called the **implied volatility**. One is usually taught to think of the implied volatility as the market's view of the future value of volatility. Yes and no. If the 'market' does have a view on the future of volatility then it will be seen in the implied volatility. But the market also has views on the direction of the underlying, and also responds to supply and demand. Let me give examples.

In one month's time there is to be an election; it is not clear who will win. If the right wing party are elected markets will rise, if the left wing are successful, markets will fall. Before the election the market assumes the middle ground, splitting the difference. In fact, little trading

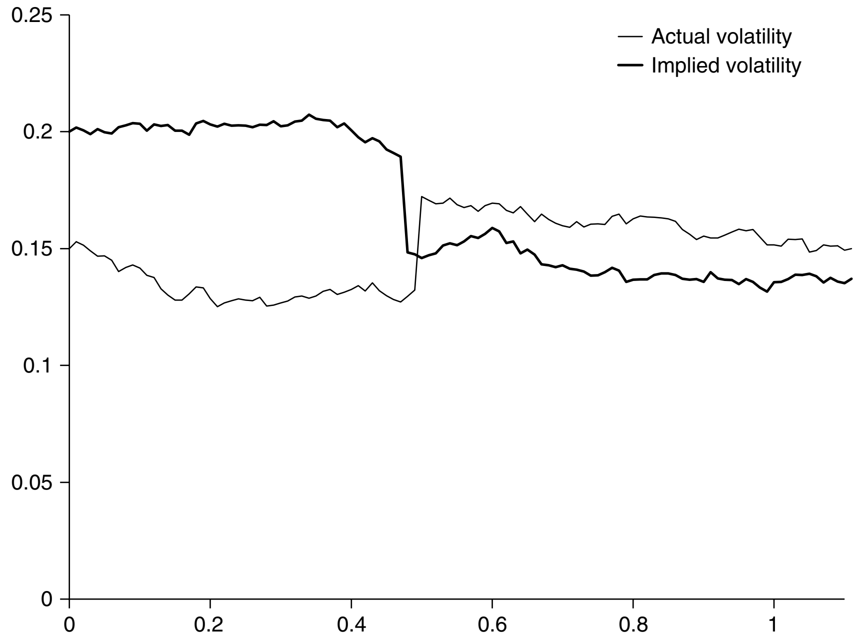


Figure 50.2 Actual and implied volatilities just before and just after an anticipated major news item.

occurs and markets have a very low volatility. But option traders know that after the election there will be a lot of movement one way or the other. Prices of both calls and puts are therefore high. If we back out implied volatilities from these option prices we see very high values. Actual and implied volatilities are shown in Figure 50.2 for this scenario.

Traders may increase option prices to reflect the expected sudden moves but if we are only observing implied volatilities then we are getting the underlying story very wrong indeed. This illustrates the fact that if you want to play around with prices there is only one parameter you can fudge: The volatility. As long as it is not too out of line compared with implied volatilities of other products no-one will disbelieve it.

Regardless of a market maker's view of future events, he is at the mercy of supply and demand. If everyone wants calls, then it is only natural for him to increase their prices. As long as he doesn't violate put-call parity (either with himself or with another market maker) who's to know that it is supply and demand driving the volatility?

50.3 TIME-DEPENDENT VOLATILITY

In Table 50.1 are the market prices of European call options[1] with one, three and seven months until expiry. All have strike prices of 105 and the underlying asset is currently 106.25. The short-term interest rate over this period is about 5.6%.

[1] I'm only using call options because put option prices should follow from the prices of the calls by put-call parity, therefore there will not be any more volatility information in the prices of puts than is already present in the prices of calls. Also the following analysis only really works if the options are European. Since most equity options are American we cannot use put prices because the values of American and European puts are different. In the absence of dividends the two call prices are the same.

Table 50.1 Market prices of European call options, see text for details.

Expiry	Value
1 month	3.50
3 months	5.76
7 months	7.97

Table 50.2 Market prices of European call options, and the implied volatilities.

Expiry	Value	Implied vol.
1 month	3.50	21.2%
3 months	5.76	20.5%
7 months	7.97	19.4%

As can easily be confirmed by substitution into the Black–Scholes call formula, these prices are consistent with volatilities of 21.2%, 20.5% and 19.4% for the one-, three- and seven-month options respectively, as shown in Table 50.2. Clearly these prices cannot be correct if the volatility is constant for the whole seven months. What is to be done?

The simplest adjustment we can make to the Black–Scholes world to accommodate these prices (without any serious effect on the theoretical framework) is to assume a time-dependent, deterministic volatility. Let's assume that volatility is a function of time:

$$\sigma(t).$$

As explained in Chapter 8 the Black–Scholes formulae are still valid when volatility is time-dependent provided we use

$$\sqrt{\frac{1}{T-t}\int_t^T \sigma(\tau)^2\,d\tau} \tag{50.1}$$

in place of σ, that is, now use

$$d_1 = \frac{\log(S/E) + r(T-t) + \frac{1}{2}\displaystyle\int_t^T \sigma(\tau)^2\,d\tau}{\sqrt{\displaystyle\int_t^T \sigma(\tau)^2\,d\tau}}$$

and

$$d_2 = \frac{\log(S/E) + r(T-t) - \frac{1}{2}\displaystyle\int_t^T \sigma(\tau)^2\,d\tau}{\sqrt{\displaystyle\int_t^T \sigma(\tau)^2\,d\tau}}.$$

Now normally we would say that we know $\sigma(t)$ and so we could find the values of options. We would then be able to figure out their implied volatilities. In other words, we usually work from actual volatility towards implied volatility:

$$\sigma(t) \Rightarrow \sigma_{\text{imp}}(t^*, T).$$

Here I have used $\sigma_{\text{imp}}(t^*, T)$ to mean the implied volatility measured at time t^* of a European option expiring at time T.

In what follows we are going to turn this on its head. Since we can observe σ_{imp} but not $\sigma(t)$ we shall works backwards to deduce what function $\sigma(t)$ is consistent with the implied volatilities, i.e.

$$\sigma_{\text{imp}}(t^*, T) \Rightarrow \sigma(t).$$

In our time-dependent example, all we need to do to ensure consistent pricing is to make

$$\sqrt{\frac{1}{T-t}\int_t^T \sigma(\tau)^2\, d\tau} = \text{implied volatilities.}$$

In words, this just means that the implied volatility is the square root of the average variance. This makes sense if you recall that variances add up, not standard deviations.

We do this 'fitting' at time t^*, and then

$$\boxed{\sigma(t) = \sqrt{\sigma_{\text{imp}}(t^*, t)^2 + 2(t - t^*)\sigma_{\text{imp}}(t^*, t)\frac{\partial \sigma_{\text{imp}}(t^*, t)}{\partial t}}.} \tag{50.2}$$

See Figure 50.3.

Practically speaking, we do not have a continuous (and differentiable) implied volatility curve. We have a discrete set of points (three in the above example). We must therefore make some assumption about the term structure of volatility between the data points. Usually one assumes that the function is piecewise constant or linear.

Example: One-month implied volatility is 30%. Two-month implied volatility is 25%. Three-month implied volatility is 26%. Construct an actual volatility function of time that is consistent with these.

First of all, note that there is no unique solution. We must make some assumptions about the time dependence of σ. For simplicity let us assume that $\sigma(t)$ is piecewise constant.

For t less than one month the answer is simple:

$$\sigma(t) = 0.3.$$

In other words, if actual volatility is assumed to be constant then that constant must be the same as the implied volatility. Use σ_1 to denote this constant value over the first period.

For t greater than one month but less than two months, this is harder. The actual volatility must be chosen to be consistent with two-month implied volatility. But two-month implied volatility depends on $\sigma(t)$ from t zero up to two months.

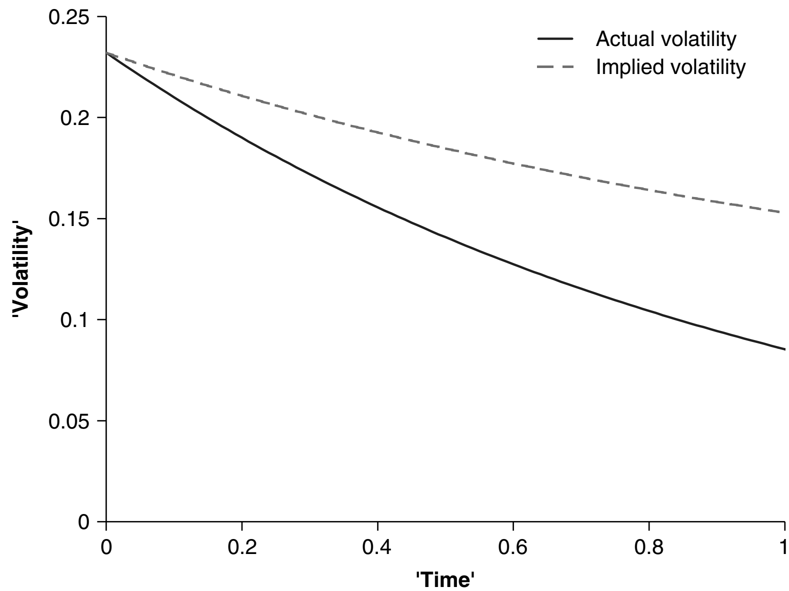


Figure 50.3 Implied volatilities and actual volatility structure that is consistent.

We know that the implied variance is the time-weighted average of actual variance:

$$\frac{2}{12} \times 0.25^2 = \frac{1}{12} \times 0.3^2 + \frac{1}{12} \times \sigma_2^2.$$

Solve this to get

$$\sigma_2 = 0.187.$$

Finally, to get σ_3, which is the actual volatility during the third month, we must have

$$\frac{3}{12} \times 0.26^2 = \frac{1}{12} \times 0.3^2 + \frac{1}{12} \times 0.187^2 + \frac{1}{12} \times \sigma_3^2.$$

The solution is just

$$\sigma_3 = 0.279.$$

See Figure 50.4.

If we have implied volatility for expiries T_i and we assume the volatility curve to be piecewise constant then

$$\boxed{\sigma(t) = \sqrt{\frac{(T_i - t^*)\sigma_{\text{imp}}(t^*, T_i)^2 - (T_{i-1} - t^*)\sigma_{\text{imp}}(t^*, T_{i-1})^2}{T_i - T_{i-1}}} \quad \text{for} \quad T_{i-1} < t < T_i}$$

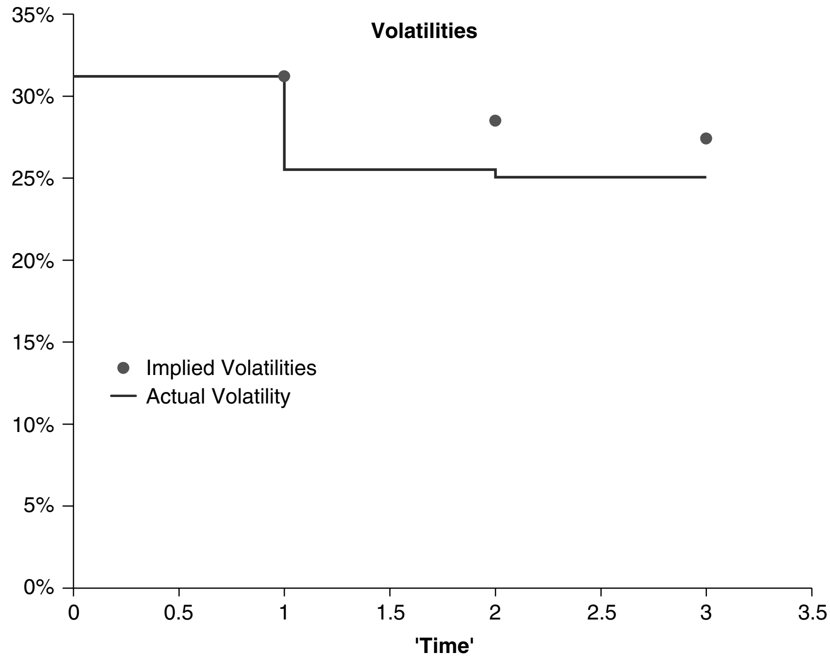


Figure 50.4 Implied volatilities and piecewise actual volatility structure that is consistent.

50.4 VOLATILITY SMILES AND SKEWS

Now let me throw the cat among the pigeons. Continuing with the example above, suppose that there is also a European call option struck at 100 with an expiry of seven months and a price of 11.48. This corresponds to a volatility of 20.8% in the Black–Scholes equation. Now we have two conflicting volatilities up to the seven-month expiry, 19.4% and 20.8%. Clearly we cannot adjust the time dependence of the volatility in any way that is consistent with *both* of these values. What else can we do? Before I answer this, we'll look at a few more examples. Concentrating on the same example, suppose that there are call options traded with an expiry of seven months and strikes of 90, 92.5, 95, 97.5, 100, 102.5, 105, 107.5 and 110. In Figure 50.5 I plot the implied volatility of these options against the strike price (the actual option prices do not add anything to our insight so I haven't given them). This is the function $\sigma_{\text{imp}}(t^*, E, T)$, the implied volatility at time t^*, as a function of E and T.

The shape of this implied volatility versus strike curve is called the **smile**. In some markets it shows considerable asymmetry, a **skew**, and sometimes it is upside down in a **frown**. The general shape tends to persist for a long time in each underlying.

If we managed to accommodate implied volatility that varied with expiry by making the volatility time-dependent perhaps we can accommodate implied volatility that varies with strike by making the volatility asset price-dependent. This is exactly what we'll do. Unfortunately, it's much harder to make analytical progress except in special cases; if we have $\sigma(S)$ then rarely can we solve the Black–Scholes equation to get nice closed-form solutions for the values of derivative products. In fact, we may as well go all the way and assume that volatility is a function of *both* the asset and time, $\sigma(S, t)$.

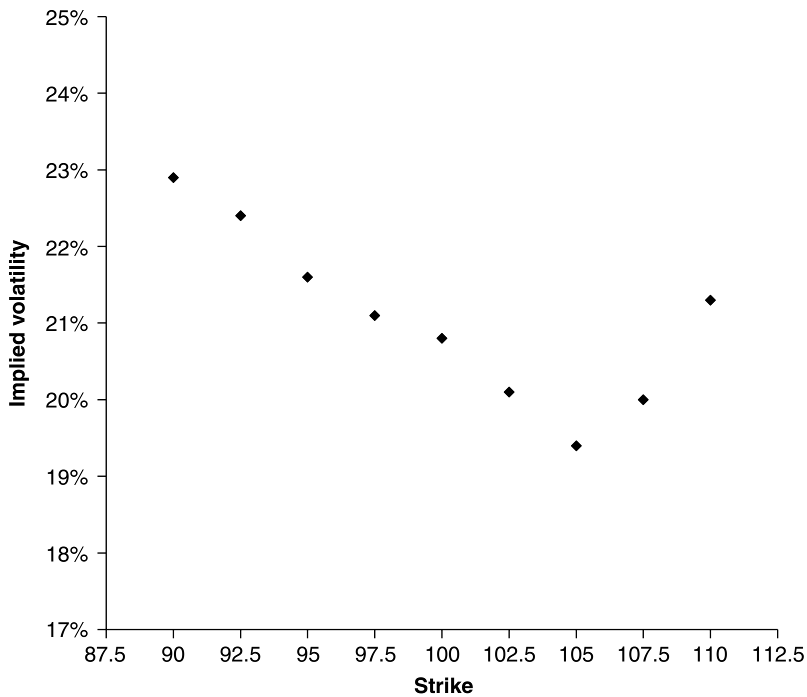


Figure 50.5 Implied volatilities against strike price.

50.5 VOLATILITY SURFACES

We can show implied volatility against both maturity and strike in a three-dimensional plot. One is shown in Figure 50.6.

This implied volatility surface represents the constant value of volatility that gives each traded option a theoretical value equal to the market value.

Again, we would normally want to work from actual volatility towards implied volatility:

$$\sigma(S, t) \Rightarrow \sigma_{\text{imp}}(t^*, E, T).$$

However, above we saw how the time dependence in implied volatility could be turned into a volatility of the underlying that was time-dependent i.e. we deduced $\sigma(t)$ from $\sigma_{\text{imp}}(t^*, T)$. Can we similarly deduce $\sigma(S, t)$ from $\sigma_{\text{imp}}(t^*, E, T)$, the implied volatility at time t^*? If we could, then we might want to call it the **local volatility surface** $\sigma(S, t)$. Thus we are going to work backwards:

$$\sigma_{\text{imp}}(t^*, E, T) \Rightarrow \sigma(S, t).$$

This local volatility surface can be thought of as the market's view of the future value of volatility when the asset price is S at time t (which, of course, may not even be realized). The local volatility is also called the **forward volatility** or the **forward forward volatility**.

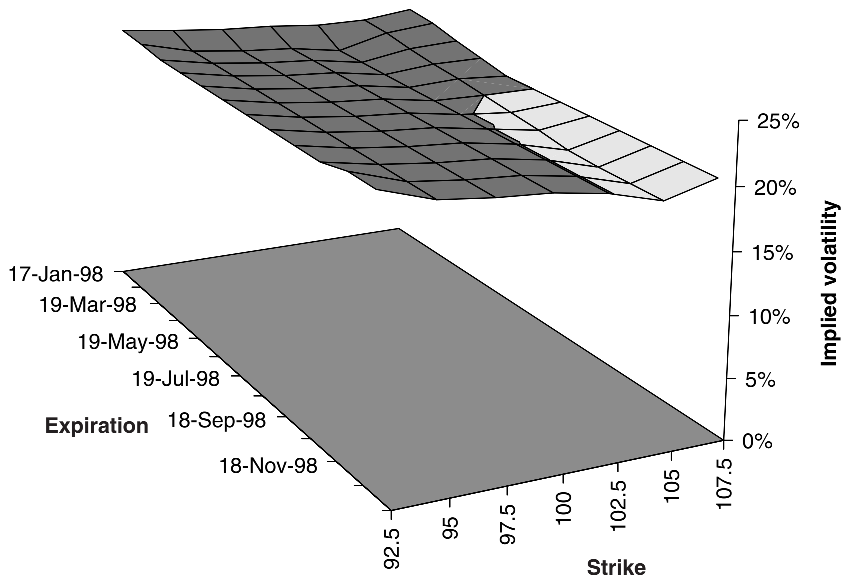


Figure 50.6 Implied volatilities against expiry and strike price.

50.6 BACKING OUT THE LOCAL VOLATILITY SURFACE FROM EUROPEAN CALL OPTION PRICES

Market prices of traded vanilla options are never, in practice, consistent with the constant volatility assumed by Black–Scholes. Nor are they consistent with either a time-dependent or an asset price-dependent local volatility. To match the theoretical prices of traded options to their market prices always requires a volatility structure that is a function of both the asset price, S, and time, t i.e. $\sigma(S,t)$. To back out the local volatility surface from the prices of market traded instruments I am going to assume that we have a distribution of European call prices of all strikes and maturities. This is not a realistic assumption but it gets the ball rolling. These prices will be denoted by $V(E,T)$. I could use puts but these can be converted to call prices by put-call parity. This notation is vastly different from before. Previously, we had the option value as a function of the underlying and time. Now the asset and time are fixed at S^* and t^*, today's values. I will use the dependence of the market prices on strike and expiry to calculate the volatility structure.

I will assume that the risk-neutral random walk for S is

$$dS = rS\,dt + \sigma(S,t)S\,dX.$$

This is our usual one-factor model for which all the building blocks of delta hedging and arbitrage-free pricing hold. The only novelty is that the volatility is dependent on the level of the asset and time.

In the following, I am going to rely heavily on the transition probability density function $p(S^*,t^*;S,T)$ for the risk-neutral random walk. Note that the backward variables are fixed at today's values and the forward time variable is T. Recalling that the value of an option is the present value of the expected payoff, I can write

$$V(E,T) = e^{-r(T-t^*)}\int_0^\infty \max(S-E,0)\,p(S^*,t^*;S,T)\,dS$$
$$= e^{-r(T-t^*)}\int_E^\infty (S-E)\,p(S^*,t^*;S,T)\,dS. \qquad (50.3)$$

We are *very* lucky that the payoff is the maximum function so that after differentiating with respect to E we get

$$\frac{\partial V}{\partial E} = -e^{-r(T-t^*)}\int_E^\infty p(S^*,t^*;S,T)\,dS.$$

And after another differentiation, we arrive at

$$\boxed{p(S^*,t^*;E,T) = e^{r(T-t^*)}\frac{\partial^2 V}{\partial E^2}} \qquad (50.4)$$

Before even calculating volatilities we can find the transition probability density function. In a sense, this is the market's view of the future distribution. But it's the market view of the risk-neutral distribution and not the real one. An example is plotted in Figure 50.7.

The next step is to use the forward equation for the transition probability density function, the Fokker–Planck equation,

$$\frac{\partial p}{\partial T} = \tfrac{1}{2}\frac{\partial^2}{\partial S^2}(\sigma^2 S^2 p) - \frac{\partial}{\partial S}(rSp). \qquad (50.5)$$

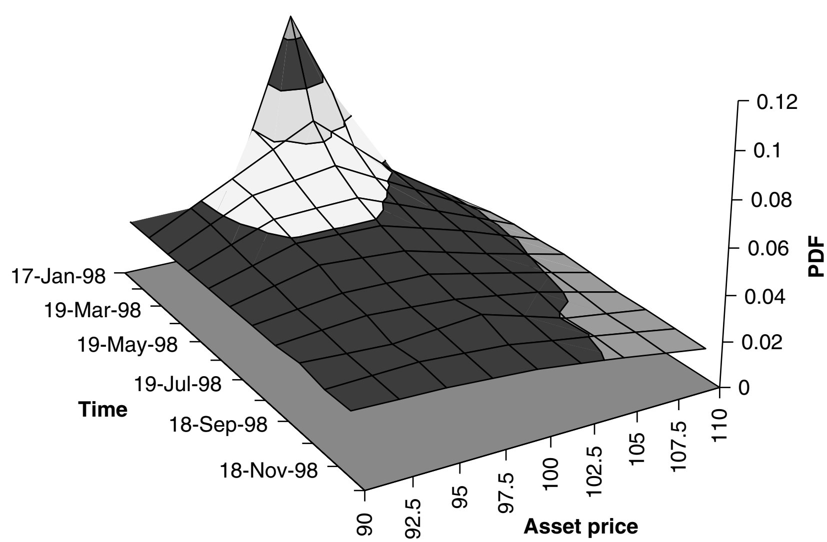


Figure 50.7 Risk-neutral transition probability density function calculated from European call prices.

Here σ is our, still unknown, function of S and t. However, *in this equation* $\sigma(S,t)$ *is evaluated at* $t = T$.

From (50.3) we have

$$\frac{\partial V}{\partial T} = -rV + e^{-r(T-t^*)}\int_E^\infty (S-E)\frac{\partial p}{\partial T}\,dS.$$

This can be written as

$$\frac{\partial V}{\partial T} = -rV + e^{-r(T-t^*)}\int_E^\infty \left(\frac{1}{2}\frac{\partial^2(\sigma^2 S^2 p)}{\partial S^2} - \frac{\partial(rSp)}{\partial S}\right)(S-E)\,dS.$$

using the forward equation (50.5). Integrating this by parts twice, assuming that p and its first S derivative tend to zero sufficiently fast as S goes to infinity, we get

$$\frac{\partial V}{\partial T} = -rV + \tfrac{1}{2}e^{-r(T-t^*)}\sigma^2 E^2 p + re^{-r(T-t^*)}\int_E^\infty Sp\,dS. \tag{50.6}$$

In this expression $\sigma(S,t)$ has $S = E$ and $t = T$. Writing

$$\int_E^\infty Sp\,dS = \int_E^\infty (S-E)p\,dS + E\int_E^\infty p\,dS$$

and collecting terms, we get

$$\frac{\partial V}{\partial T} = \tfrac{1}{2}\sigma^2 E^2\frac{\partial^2 V}{\partial E^2} - rE\frac{\partial V}{\partial E}.$$

Rearranging this we find that

$$\sigma = \sqrt{\frac{\dfrac{\partial V}{\partial T} + rE\dfrac{\partial V}{\partial E}}{\tfrac{1}{2}E^2\dfrac{\partial^2 V}{\partial E^2}}}.$$

This gives us $\sigma(E,T)$ and hence, by *relabeling the variables*, $\sigma(S,t)$.

This calculation of the volatility surface from option prices worked because of the particular form of the payoff, the call payoff, which allowed us to derive the very simple relationship between derivatives of the option price and the transition probability density function.

When there is a constant and continuous dividend yield on the underlying the relationship between call prices and the local volatility is

$$\boxed{\sigma = \sqrt{\frac{\dfrac{\partial V}{\partial T} + (r-D)E\dfrac{\partial V}{\partial E} + DV}{\tfrac{1}{2}E^2\dfrac{\partial^2 V}{\partial E^2}}}} \tag{50.7}$$

There is no change in this expression when the interest rate and dividend yield are time-dependent, just use the relevant forward rates.

One of the problems with this expression concerns data far in or far out of the money. Unless we are close to at the money both the numerator and denominator of (50.7) are small, leading to inaccuracies when we divide one small number by another. One way of avoiding this is to relate the local volatility surface to the implied volatility surface as I now show.

In the same way that we found a relationship between the local volatility and the implied volatility in the purely time-dependent case, Equation (50.2), we can find a relationship in the general case of asset- and time-dependent local volatility. This relationship is obviously quite complicated and I omit the details of the derivation. The result is

$$\sigma(E,T) = \sqrt{\frac{\sigma_{\mathrm{imp}}^2 + 2(T-t^*)\sigma_{\mathrm{imp}}\dfrac{\partial\sigma_{\mathrm{imp}}}{\partial T} + 2(r-D)E(T-t^*)\sigma_{\mathrm{imp}}\dfrac{\partial\sigma_{\mathrm{imp}}}{\partial E}}{\left(1 + Ed_1\sqrt{T-t^*}\dfrac{\partial\sigma_{\mathrm{imp}}}{\partial E}\right)^2 + E^2(T-t^*)\sigma_{\mathrm{imp}}\left(\dfrac{\partial^2\sigma_{\mathrm{imp}}}{\partial E^2} - d_1\left(\dfrac{\partial\sigma_{\mathrm{imp}}}{\partial E}\right)^2\sqrt{T-t^*}\right)}}$$

$$d_1 = \frac{\log(S^*/E) + (r - D + \frac{1}{2}\sigma_{\mathrm{imp}}^2)(T-t^*)}{\sigma_{\mathrm{imp}}\sqrt{T-t^*}}. \tag{50.8}$$

In terms of the implied volatility the implied risk-neutral probability density function is

$$p(S^*, t^*; E, T) = \frac{1}{E\sigma_{\mathrm{imp}}\sqrt{2\pi(T-t^*)}}e^{-\frac{1}{2}d_2^2}$$

$$\left(\left(1 + Ed_1\sqrt{T-t^*}\frac{\partial\sigma_{\mathrm{imp}}}{\partial E}\right)^2 + E^2(T-t^*)\sigma_{\mathrm{imp}}\left(\frac{\partial^2\sigma_{\mathrm{imp}}}{\partial E^2} - d_1\left(\frac{\partial\sigma_{\mathrm{imp}}}{\partial E}\right)^2\sqrt{T-t^*}\right)\right).$$

One of the advantages of writing the local volatility and probability density function in terms of the implied volatility surface is that if you put in a flat implied volatility surface you get out a flat local surface and a lognormal distribution.

In practice there only exists a finite, discretely-spaced set of call prices. To deduce a local volatility surface from these data requires some interpolation and extrapolation. This can be done in a number of ways and there is no correct way. One of the problems with these approaches is that the final result depends sensitively on the form of the interpolation. The problem is actually 'ill-posed,' meaning that a small change in the input can lead to a large change in the output. There are many ways to get around this ill-posedness, coming under the general heading of 'regularization.' Several suggestions for further reading in this area are given at the end of the chapter.

An example of a local volatility surface is plotted in Figure 50.8.

Once this local volatility surface has been found it must be set in stone. If the model is correct then when you come back a week later to recalibrate, even though time has moved on and stock and option prices have all changed, you will still get *exactly* the same local volatility surface.

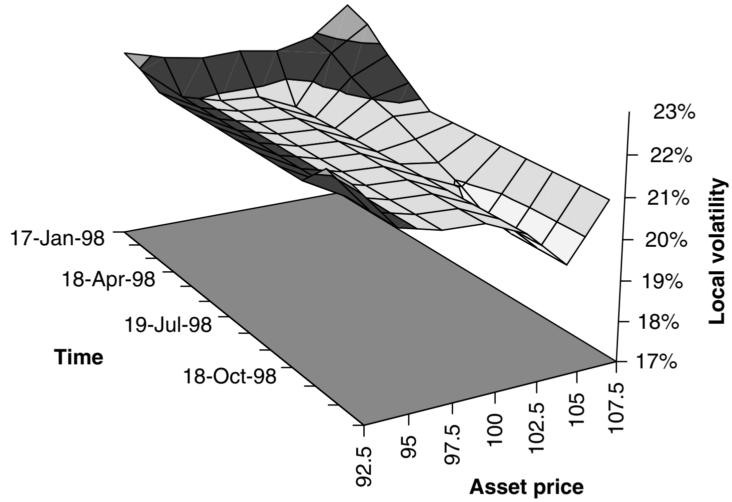


Figure 50.8 Local volatility surface calculated from European call prices.

In practice do you think that happens? That the local volatility surface remains unchanged after one week? You are correct.[2]

50.7 A SIMPLE VOLATILITY SURFACE PARAMETERIZATION

Backing out the deterministic local volatility surface from the market prices of traded instruments has become increasingly popular since the introduction of the idea in the early 1990s. But it is clear is that the practical implementation of the method suffers from serious numerical instabilities and can be slow to compute. In this section we examine an approximation to the local volatility surface that is simple and fast to implement.

There is a major problem with using the results of the previous section, and that concerns the loss of accuracy due to the discreteness of the implied volatility data: we only know σ_{imp} at a finite number of points (E, T). The resulting local volatility surface often looks unrealistic, and is very sensitive to the input implied volatilities. There is also some evidence (Dumas, Fleming & Whaley, 1998) that the information contained in the market prices is lost if the local volatility surface is over-fitted. In other words, a parsimonious representation of volatility surfaces is best.

In this section I am going to assume that the implied volatilities are of the simple form

$$\sigma_{\text{imp}}(E, T) = a(T)(E - S^*) + b(T). \tag{50.9}$$

The motivation behind assumption (50.9) is that the most important information in market prices concerns the at-the-money volatility and the volatility skew, both of which may be time-dependent.

Later we will see how to calculate $a(T)$ and $b(T)$ from the market prices of an at-the-money straddle and a risk reversal. First, however, we will find the local volatility surface assuming (50.9).

[2] It doesn't.

50.8 AN APPROXIMATE SOLUTION

Substituting (50.9) into (50.8) we find that

$$\sigma(E,T)=\sqrt{\frac{\begin{array}{c}(a(E-S^*)+b)^2+2(T-t^*)(a(E-S^*)+b)(a'(E-S^*)+b')\\ +2(r-D)E(T-t^*)(a(E-S^*)+b)a\end{array}}{\left(1+Ed_1\sqrt{T-t^*}a\right)^2-E^2(T-t^*)^{3/2}(a(E-S^*)+b)d_1a^2}},$$

where

$$d_1=\frac{\log(S^*/E)+(r-D+\frac{1}{2}(a(E-S^*)+b)^2)(T-t^*)}{(a(E-S^*)+b)\sqrt{T-t^*}}.$$

Here $'$ denotes d/dT. Rewriting σ in terms of its more natural arguments, S and t, we have

$$\sigma(S,t)=\sqrt{\frac{\begin{array}{c}(a(S-S^*)+b)^2+2(t-t^*)(a(S-S^*)+b)(a'(S-S^*)+b')\\ +2(r-D)S(t-t^*)(a(S-S^*)+b)a\end{array}}{\left(1+Sd_1\sqrt{t-t^*}a\right)^2-S^2(t-t^*)^{3/2}(a(S-S^*)+b)d_1a^2}},$$

where

$$d_1=\frac{\log(S^*/S)+(r-D+\frac{1}{2}(a(S-S^*)+b)^2)(t-t^*)}{(a(S-S^*)+b)\sqrt{t-t^*}}.$$

If we can find a and b then our simple volatility surface parameterization is complete.

50.9 VOLATILITY INFORMATION CONTAINED IN AN AT-THE-MONEY STRADDLE

The straddle position is made up of a long call and a long put with the same strikes and expiries. Practitioners use the market prices of at-the-money straddles to deduce the at-the-money volatility. The Black–Scholes value of a straddle is given by

$$V_S=C+P=C+C-Se^{-D(T-t)}+Ee^{-r(T-t)}=2C-Se^{-D(T-t)}+Ee^{-r(T-t)},$$

where C and P are the values of the call and the put respectively and we have used put-call parity. We can therefore deduce the price of a single call from

$$C=\tfrac{1}{2}(V_S+Se^{-D(T-t)}-Ee^{-r(T-t)})$$

and hence the implied volatility. Since the straddle is at-the-money we have $S=E=S^*$ and, of course, $t=t^*$.

From our assumed form (50.9) for the implied volatility we have

$$e^{-D(T-t^*)}N(d_1)-e^{-r(T-t^*)}N(d_2)=\frac{1}{2}\left(\frac{V_S}{S^*}+e^{-D(T-t^*)}-e^{-r(T-t^*)}\right),$$

where

$$d_1=\frac{(r-D+\frac{1}{2}b(T)^2)\sqrt{T-t^*}}{b(T)}\quad\text{and}\quad d_2=d_1-b(T)\sqrt{T-t^*}.$$

Since V_S is known from the market we can calculate $b(T)$.

50.10 VOLATILITY INFORMATION CONTAINED IN A RISK-REVERSAL

The risk-reversal is a long call, with strike above the current spot, and a short put with a strike below the current spot. Both have the same expiry. Practitioners use the market price of the risk-reversal to deduce the volatility skew.

In the following we will assume that the strikes of the call and the put are a short distance ϵ away from the current spot: the strike of the call is thus $S^* + \epsilon$ and the strike of the put is $S^* - \epsilon$; this can easily be generalized, but this assumption allows us to go a little further with our analysis. I will shortly be expanding various quantities in Taylor series in ϵ.

I will now need a slightly more informative notation. I will include the strike and implied volatility of each option as arguments e.g. $C(E, \sigma_{\text{imp}})$ means a call with strike E and implied volatility σ_{imp}, similarly for puts.

The Black–Scholes value of the risk-reversal is given by

$$\begin{aligned} V_{\text{RR}} &= C(S^* + \epsilon, \sigma_{\text{imp}}(S^* + \epsilon, T)) - P(S^* - \epsilon, \sigma_{\text{imp}}(S^* - \epsilon, T)) \\ &= C(S^* + \epsilon, \sigma_{\text{imp}}(S^* + \epsilon, T)) - C(S^* - \epsilon, \sigma_{\text{imp}}(S^* - \epsilon, T)) \\ &\quad + S^* e^{-D(T-t^*)} - (S^* - \epsilon)e^{-r(T-t^*)}. \end{aligned}$$

I have directly written everything in terms of the current spot and time, and again used put-call parity.

If ϵ is small, that is if the two strikes are close together, which is usually the case, we can expand the above. We find that

$$V_{\text{RR}} - S^*(e^{-D(T-t^*)} - e^{-r(T-t^*)}) = \epsilon \left(e^{-r(T-t^*)} + 2\frac{\partial C}{\partial E}(S^*, b) + 2\frac{\partial C}{\partial \sigma_{\text{imp}}}(S^*, b)a \right).$$

Using the formulae for a call we get

$$\begin{aligned} &V_{\text{RR}} - S^* \left(e^{-D(T-t^*)} - e^{-r(T-t^*)} \right) \\ &\quad = \epsilon \left(e^{-r(T-t^*)} - 2e^{-r(T-t^*)} N(d_2) + 2aS^* \sqrt{T - t^*} N'(d_1) e^{-D(T-t^*)} \right), \end{aligned}$$

where

$$d_1 = \frac{(r - D + \frac{1}{2}b(T)^2)\sqrt{T - t^*}}{b(T)} \quad \text{and} \quad d_2 = d_1 - b(T)\sqrt{T - t^*}.$$

If we have found b from the at-the-money straddle then finding a from the above is very easy:

$$a(T) = \frac{e^{D(T-t^*)}}{2\epsilon S^* \sqrt{T - t^*} N'(d_1)} \left(V_{\text{RR}} - S^* \left(e^{-D(T-t^*)} - e^{-r(T-t^*)} \right) \right) + \frac{e^{(D-r)(T-t^*)} N(d_2)}{S^* \sqrt{T - t^*} N'(d_1)}.$$

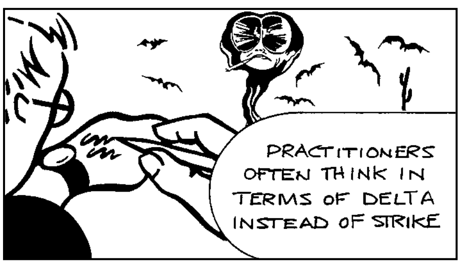
50.11 TIME DEPENDENCE

We have dealt with the discreteness of the data in the direction of the strike, but we must still deal with the discreteness in the expiry dates. In practice we only have a small number of expiries from which to deduce the term structure of the parameters a and b.

There are several approaches we can take depending on how much smoothness we want to insist upon. Fortunately, the time dependence of the local volatility surfaces is less extreme than the typical asset price dependence. The simplest interpolation between data points in the time direction is linear. This is adequate in most situations. If we know $a(T_i)$ and $b(T_i)$ we take the implied volatility parameters to be

$$\frac{a(T_i)(T_{i+1}-T)+a(T_{i+1})(T-T_i)}{T_{i+1}-T_i} \quad \text{and} \quad \frac{b(T_i)(T_{i+1}-T)+b(T_{i+1})(T-T_i)}{T_{i+1}-T_i}$$

for times T between the two expiries T_i and T_{i+1}.



50.12 A MARKET CONVENTION

Instead of quoting an option price by specifying its expiry, maturity and volatility, it is common to quote its expiry, *delta* and volatility. For instance, one might be told that the six-month 25 delta call has a volatility of 13%. This means that the option with an expiry in six months and a theoretical Black–Scholes delta of 0.25 will cost the theoretical Black–Scholes value using a volatility of 0.13. What seems to be missing from this? The strike of the option. But we can work this out from the rest of the information. Let's continue with this example. Recalling the formula for the delta of a call, we know that

$$e^{-D(T-t)}N(d_1)=0.25$$

where

$$d_1=\frac{\log(S/E)+(r-D+\frac{1}{2}0.13^2)(T-t)}{0.13\sqrt{(T-t)}}.$$

Also in this example

$$T-t=0.5,$$

and we presumably know the spot price S, the interest rate r and the dividend yield D. We now solve for E.

Another example might be that we are told the volatility of a specified maturity zero-delta straddle. Again, we can work backwards to find the location of the strike.[3]

The vol surface parameterization can straightforwardly be put into the form of this market convention.

[3] It will be close to, but not exactly at, at-the-money forward, because of the skew caused by lognormality.

50.13 HOW DO I USE THE LOCAL VOLATILITY SURFACE?

There are two ways to look at the local volatility surface. One is to say that it is the market's view of future volatility and that these predictions will come to pass. We then price other, more complex, products using this asset- and time-dependent volatility. This is a very naive belief. Not only do the predictions not come true but even if we come back a few days later to look at the 'prediction,' i.e. to refit the surface, we see that it has changed.

The other way of using the surface is to acknowledge that it is only a snapshot of the market's view and that tomorrow it may change. But it can be used to price non-traded contracts in a way that is consistent across all instruments. As long as we price our exotic contract consistently with the vanillas, that is, with the same volatility structure, *and simultaneously hedge with these vanillas* then we are reducing our exposure to model error. This approach is readily justifiable, although it is a bit difficult to estimate by how much we have reduced our model exposure. However, if you price using the calibrated volatility surface but only delta hedge, then you are asking for trouble. Suppose you price, and sell, a volatility-sensitive instrument such as an up-and-out call with a fitted volatility surface which increases with stock level. If it turns out that when the volatility is realized it is a downward-sloping function of the asset, then you are in big trouble.

As an example, calculate the local volatility surface using vanilla calls. Now price a barrier option using this volatility structure. This means solving the Black–Scholes partial differential equation with the asset- and time-dependent volatility and with the relevant boundary conditions. This must be done numerically, by the methods explained in Part Six, for example. Now statically hedge the barrier by buying and selling the vanilla contracts to mimic as closely as possible the payoff and boundary conditions of the barrier option. This is described fully in Chapter 60.

50.14 SUMMARY

Whether you believe in them or not, local volatility surfaces have taken the practitioner world by storm. Now that they are commonly used for pricing and hedging exotic contracts there is no way back to the world of constant volatility.

Personally I am in two minds about this issue. But as long as you hedge as much of the cashflows using traded vanilla options then you will be reducing your model exposure anyway. Once you have done this then you have done your best to reduce dependence on volatility. The danger is always going to be there if you never bother to statically hedge, only delta hedge. This is discussed in detail in Chapter 60.

FURTHER READING

- Merton (1973) was the first to find the explicit formulae for European options with time-dependent volatility.
- See Dupire (1993, 1994) and Derman & Kani (1994) for more details of fitting the local volatility surface.
- Rubinstein (1994) constructs an implied tree using an optimization approach, and this has been generalized by Jackwerth & Rubinstein (1996).

- To get around the problem of ill-posedness, Avellaneda, Friedman, Holmes & Samperi (1997) propose calibrating the local volatility surface by entropy minimization. Their article is also a very good source of references to the volatility surface literature.
- For trading strategies involving views on the direction of volatility see Connolly (1997).
- See Dewynne, Ehrlichman & Wilmott (1998) for more details of the simple volatility surface parameterization.
- I suspect that there is not much information about future volatility contained in the local volatility surface. This is demonstrated in Dumas, Fleming & Whaley (1998).
- For details of techniques for calibration you must read Rebonato (2004).
- For a critique of the whole smile problem see Ayache, Henrotte, Nassar & Wang (2004).

APPENDIX: CURVE FITTING 101

In Figure 50.9 we see a set of data points to be fitted. It doesn't matter what the independent and dependent variables are, maybe they are strike and implied volatility. Just eyeballing the dots shows there is a very clear relationship between the variables; generally speaking the larger the independent variable the larger the dependent. It's not a straight line, but then it's not too far off.

There are lots of ways of fitting a curve to this dataset so that we can interpolate and extrapolate. Figure 50.10, for example, shows a perfect fit using a polynomial. Perfect, yes?

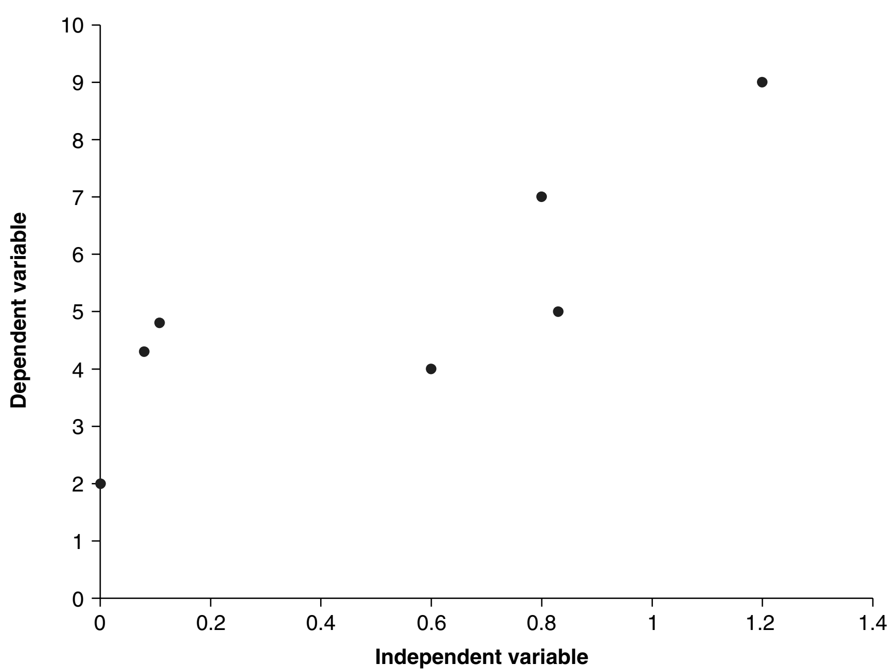


Figure 50.9 A set of data points to be fitted.

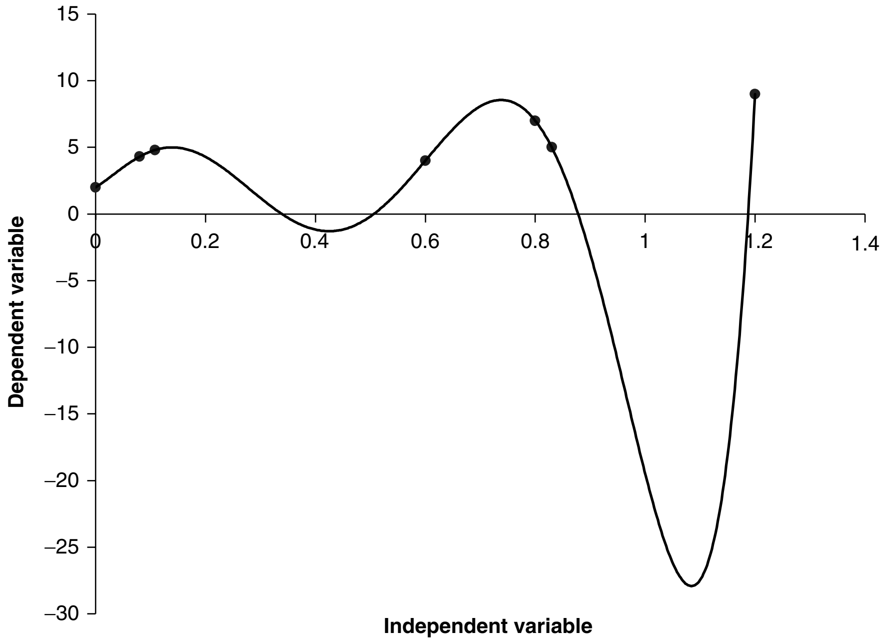


Figure 50.10 The perfect fit.

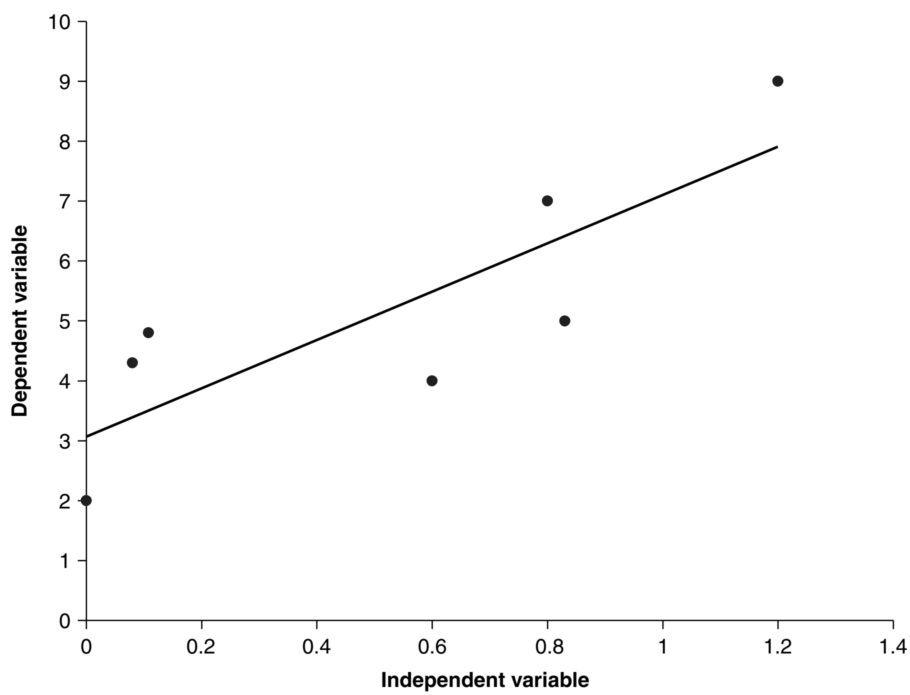


Figure 50.11 A better 'fit.'

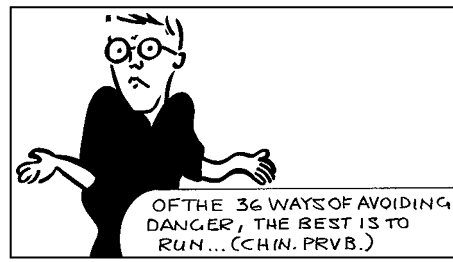
Well, somehow it doesn't look right. The apparent relationship between the variables has now been swamped by the wiggles caused by the polynomial. Would you prefer to rely on this curve for interpolation purposes, or perhaps something that just looks more plausible?

In Figure 50.11 the same data have been fitted by a simple straight line. Obviously nowhere is the fit exact, but then also nowhere does it look too bad. There is an obvious moral to this story: Think before you fit.

CHAPTER 51
stochastic volatility

In this Chapter...

- modeling volatility as a stochastic variable
- discontinuous volatility
- how to price contracts when volatility is stochastic



51.1 INTRODUCTION

Volatility does not behave how the Black–Scholes equation would like it to behave; it is not constant, it is not predictable, it is not even directly observable. This makes it a prime candidate for modeling as a random variable. There is plenty of evidence that returns on equities, currencies and commodities are not Normally distributed; they have higher peaks and fatter tails than predicted by a Normal distribution. We have seen this in several places in this book. This has also been cited as evidence for non-constant volatility.

51.2 RANDOM VOLATILITY

Before saying what we mean by stochastic volatility, let's be clear about what random volatility is not about.

If we draw a number at random from a distribution with either a 10% volatility, 20% volatility or 30% volatility with equal probability, then the resulting distribution would have the higher peak-fatter tails properties we have seen in the data. However, as long as the standard deviation of this distribution is finite (and scales with the square root of the time step) we can price options using the standard deviation in place of the volatility in Black–Scholes (see also Chapter 47). So this is *not* what we mean when we talk about stochastic volatility. The key point is timescale.

Pricing only becomes a problem with random volatility when the timescale of the evolution of volatility is of the same order as the evolution of the underlying. For example, if we have a stochastic differential equation model for the volatility then we must move beyond the Black–Scholes world.

51.3 A STOCHASTIC DIFFERENTIAL EQUATION FOR VOLATILITY

Figure 51.1 shows four estimates of volatility of the Dow Jones Index, using weighted averages of the daily changes in the index over the previous 10, 30, 50 and 100 days. First of all, this graph shows that volatility is not constant. Moreover, it is not at all clear how to measure it. Each estimate gives a different answer. Modeling volatility as a stochastic process requires there to be such a quantity as volatility ... but maybe observability isn't that important.

We continue to assume that S satisfies

$$dS = \mu S\,dt + \sigma S\,dX_1,$$

but we further assume that volatility satisfies

$$d\sigma = p(S, \sigma, t)\,dt + q(S, \sigma, t)\,dX_2. \tag{51.1}$$

The two increments dX_1 and dX_2 have a correlation of ρ. The choice of functions $p(S, \sigma, t)$ and $q(S, \sigma, t)$ is crucial to the evolution of the volatility, and thus to the pricing of derivatives. The choice of these functions is discussed later in this chapter and in Chapter 53. For the moment, I only comment that a mean-reverting process is a natural choice.

The value of an option with stochastic volatility is a function of three variables, $V(S, \sigma, t)$.

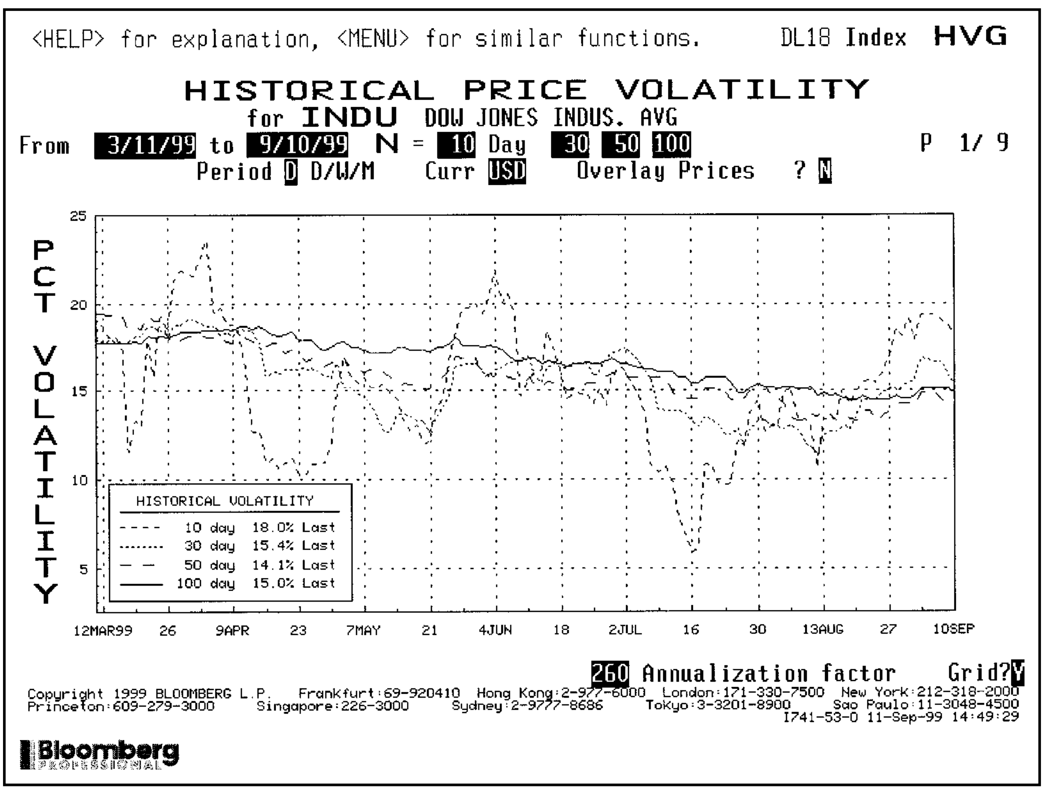


Figure 51.1 Volatility against time. Source: Bloomberg L.P.

51.4 THE PRICING EQUATION

The new stochastic quantity that we are modeling, the volatility, is not a traded asset. Thus, when volatility is stochastic we are faced with the problem of having a source of randomness that cannot be easily hedged away. Because we have two sources of randomness we must hedge our option with two other contracts, one being the underlying asset as usual, but now we also need another option to hedge the volatility risk. We therefore must set up a portfolio containing one option, with value denoted by $V(S, \sigma, t)$, a quantity $-\Delta$ of the asset and a quantity $-\Delta_1$ of another option with value $V_1(S, \sigma, t)$. We have

$$\Pi = V - \Delta S - \Delta_1 V_1. \tag{51.2}$$

The change in this portfolio in a time dt is given by

$$\begin{aligned} d\Pi = &\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 V}{\partial \sigma^2}\right) dt \\ &- \Delta_1 \left(\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 V_1}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 V_1}{\partial \sigma^2}\right) dt \\ &+ \left(\frac{\partial V}{\partial S} - \Delta_1 \frac{\partial V_1}{\partial S} - \Delta\right) dS + \left(\frac{\partial V}{\partial \sigma} - \Delta_1 \frac{\partial V_1}{\partial \sigma}\right) d\sigma. \end{aligned}$$

where I have used Itô's lemma on functions of S, σ and t.

To eliminate all randomness from the portfolio we must choose

$$\frac{\partial V}{\partial S} - \Delta - \Delta_1 \frac{\partial V_1}{\partial S} = 0,$$

to eliminate dS terms, and

$$\frac{\partial V}{\partial \sigma} - \Delta_1 \frac{\partial V_1}{\partial \sigma} = 0,$$

to eliminate $d\sigma$ terms.

So

$$\Delta_1 = \frac{\partial V / \partial \sigma}{\partial V_1 / \partial \sigma}$$

and

$$\Delta = \frac{\partial V}{\partial S} + \frac{\partial V / \partial \sigma}{\partial V_1 / \partial \sigma} \frac{\partial V_1}{\partial S}.$$

This leaves us with

$$\begin{aligned} d\Pi &= \frac{\partial V}{\partial t}\,dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\,dt + \rho\sigma S q \frac{\partial^2 V}{\partial S\,\partial \sigma}\,dt + \frac{1}{2}q^2 \frac{\partial^2 V}{\partial \sigma^2}\,dt \\ &\quad - \Delta_1 \left(\frac{\partial V_1}{\partial t}\,dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2}\,dt + \rho\sigma S q \frac{\partial^2 V_1}{\partial S\,\partial \sigma}\,dt + \frac{1}{2}q^2 \frac{\partial^2 V_1}{\partial \sigma^2}\,dt\right) \\ &= r\Pi\,dt = r(V - \Delta S - \Delta_1 V_1)\,dt, \end{aligned}$$

where I have used arbitrage arguments to set the return on the portfolio equal to the risk-free rate.

As it stands, this is *one* equation in the *two* unknowns, V and V_1. This contrasts with the earlier Black–Scholes case with one equation in the one unknown.

Collecting all V terms on the left-hand side and all V_1 terms on the right-hand side we find that

$$\frac{\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}+\rho\sigma Sq\frac{\partial^2V}{\partial S\,\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2V}{\partial\sigma^2}+rS\frac{\partial V}{\partial S}-rV}{\frac{\partial V}{\partial\sigma}}$$

$$=\frac{\frac{\partial V_1}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2V_1}{\partial S^2}+\rho\sigma Sq\frac{\partial^2V_1}{\partial S\,\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2V_1}{\partial\sigma^2}+rS\frac{\partial V_1}{\partial S}-rV_1}{\frac{\partial V_1}{\partial\sigma}}.$$

We are lucky that the left-hand side is a functional of V but not V_1 and the right-hand side is a function of V_1 but not V. Since the two options will typically have different payoffs, strikes or expiries, the only way for this to be possible is for both sides to be independent of the contract type. In other words, both sides must in some sense be equal to the same 'universal' constant. Except that really it is a universal function of all of the *independent* variables common to *all* options. Both sides can only be functions of the independent variables, S, σ and t. Thus we have

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}+\rho\sigma Sq\frac{\partial^2V}{\partial S\,\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2V}{\partial\sigma^2}+rS\frac{\partial V}{\partial S}-rV=-(p-\lambda q)\frac{\partial V}{\partial\sigma},$$

for some function $\lambda(S,\sigma,t)$.

Reordering this equation, we usually write

$$\boxed{\begin{aligned}&\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}+\rho\sigma Sq\frac{\partial^2V}{\partial S\,\partial\sigma}+\frac{1}{2}q^2\frac{\partial^2V}{\partial\sigma^2}+rS\frac{\partial V}{\partial S}\\&\quad+(p-\lambda q)\frac{\partial V}{\partial\sigma}-rV=0\end{aligned}}\tag{51.3}$$

The function $\lambda(S,\sigma,t)$ is called the **market price of (volatility) risk**.

51.5 THE MARKET PRICE OF VOLATILITY RISK

If we can solve Equation (51.3) then we have found the value of the option, and the hedge ratios. Generally, this must be done numerically. But note that we find *two* hedge ratios, $\partial V/\partial S$ and $\partial V/\partial\sigma$. We have two hedge ratios because we have two sources of randomness that we must hedge away.

Because one of the modeled quantities, the volatility, is not traded we find that the pricing equation contains a market price of risk term. What does this term mean? Suppose we hold one of the option with value V, and satisfying the pricing equation (51.3), delta hedged with the underlying asset only i.e. we have

$$\Pi = V - \Delta S.$$

The change in this portfolio value is

$$\begin{aligned} d\Pi = &\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma q S\frac{\partial^2 V}{\partial S\partial\sigma} + \frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2}\right) dt \\ &+ \left(\frac{\partial V}{\partial S} - \Delta\right) dS + \frac{\partial V}{\partial\sigma}\, d\sigma. \end{aligned}$$

Because we are delta hedging the coefficient of dS is zero. We find that

$$\begin{aligned} d\Pi - r\Pi\, dt &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma q S\frac{\partial^2 V}{\partial S\partial\sigma} + \frac{1}{2}q^2\frac{\partial^2 V}{\partial\sigma^2} + rS\frac{\partial V}{\partial S} - rV\right) dt + \frac{\partial V}{\partial\sigma}\, d\sigma \\ &= q\frac{\partial V}{\partial\sigma}(\lambda\, dt + dX_2). \end{aligned}$$

This has used both the pricing equation (51.3) and the stochastic differential equation for σ, (51.1). Observe that for every unit of volatility risk, represented by dX_2, there are λ units of extra return, represented by dt. Hence the name 'market price of risk.'

The quantity $p - \lambda q$ is called the **risk-neutral drift rate** of the volatility. Recall that the risk-neutral drift of the underlying asset is r and not μ. When it comes to pricing derivatives, it is the risk-neutral drift that matters and not the real drift, whether it is the drift of the asset or of the volatility.

51.5.1 Aside: The Market Price of Risk for Traded Assets

Let us return briefly to the Black–Scholes world of constant volatility. In Chapter 5 we derived the Black–Scholes equation for equities by constructing a portfolio consisting of one option and a number $-\Delta$ of the underlying asset. We were able to do this because the underlying asset, the equity, was traded. Suppose that instead we were to follow the analysis above and construct a portfolio of two *options* with different maturity dates (or different exercise prices, for that matter) instead of an option and the underlying. We would have

$$\Pi = V - \Delta_1 V_1.$$

Note that there is none of the underlying asset in this portfolio. The same argument as used above leads us to

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (\mu - \lambda_S\sigma)S\frac{\partial V}{\partial S} - rV = 0. \qquad (51.4)$$

What is special about the variable S? It is the value of a traded asset. This means that $V = S$ must itself be a solution of (51.4). Substituting $V = S$ into (51.4) we find that

$$(\mu - \lambda_S\sigma)S - rS = 0,$$

i.e.

$$\lambda_S = \frac{\mu - r}{\sigma};$$

this is the market price of risk for a traded asset. Now putting $\lambda_S = (\mu - r)/\sigma$ into (51.4) we arrive at

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

We are back at the Black–Scholes equation, which contains no mention of μ or λ_S.

51.6 THE VALUE AS AN EXPECTATION

Equation (51.3) is another backward parabolic partial differential equation, but in two 'space' variables S and σ. It has a nice, simple interpretation. It is the equation you solve not just to price an option, but the equation you solve when you want to find the present value of the expected payoff. Just as the classical Black–Scholes model tells us that an option value is the present value of an expectation, so this interpretation carries over into the stochastic volatility world.

In words, the option value can be written as the **present value of the expected payoff under a risk-neutral random walk**. Yes, it is risk-neutral random walks we have to consider again. Remember that risk-neutral random walks have different drifts from real random walks. Well, this is also true when volatility is stochastic.

In Table 51.1 we see the real and risk-neutral drifts for the two 'space' variables S and σ. Only the risk-neutral drifts appear in the pricing equation and so help determine the value of an option. Later we will be looking at Monte Carlo simulations used for pricing. We will see how it is necessary to simulate the risk-neutral random walks, not the real ones.

51.7 AN EXAMPLE

Figure 51.2 shows the theoretical value of a call option assuming stochastic volatility. The details of the volatility model and the code that produced the data are given in Chapter 82. One of the inputs into the model is the correlation between the movements in the underlying and the movements in the volatility. In Figure 51.3 are plots of call values for three different values of that correlation, for the same volatility. In this example the correlation effect is quite large for slightly out-of-the-money options. For in-the-money options the effect on the option's time value is also significant.

When the volatility model is independent of the stock price (p, q and λ are not functions of S) and there is zero correlation ($\rho = 0$) then the solution of the

Table 51.1 Real and risk-neutral drift functions.

Variable	Real drift	Risk-neutral drift
S	μS	r
σ	p	$p - \lambda q$

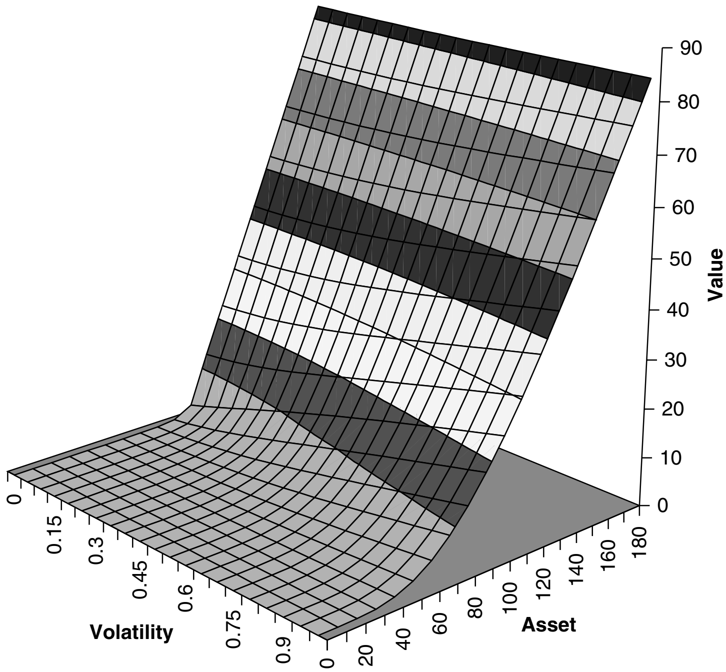


Figure 51.2 Call value when volatility is stochastic.

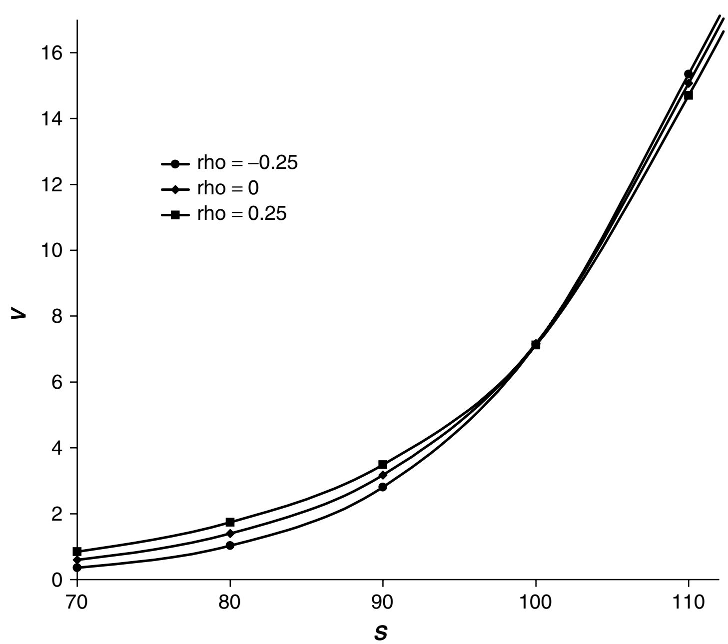


Figure 51.3 Call value when volatility is stochastic; three different correlations.

pricing equation has a nice interpretation. It turns out that you can then relate the option's value in the stochastic volatility world to the classical Black–Scholes formulae. The option price is just the weighted average of Black–Scholes formulae, where the weighting is the probability density function for the average variance from now to expiration. You will find that this generates a slight smile. But to get a realistic skew you will need to put in some correlation, and then you may have to crunch the equation numerically.

51.8 CHOOSING THE MODEL

Choosing the model means choosing the functional forms for p and q, or rather $p - \lambda q$ and q. This is not easy, principally because σ is not observable, so how can you model it? Strictly speaking, you ought to try and get p and q by looking at the statistics of the stock price S. Models such as ARCH, GARCH, REGARCH, mentioned below, try to do this. Then you would estimate λ from option prices, since λ is associated with how people value volatility risk, and that isn't observable in the stock price series.

None of that is easy. So what seems to be more common these days, although harder to justify than the statistical approach, is choosing $p - \lambda q$ and q so that the model correctly prices exchange-traded options. Often this means picking a model that is tractable, has closed-form formulae for vanillas, and has sufficient degrees of freedom (in terms of parameters) so that those vanillas can be priced exactly the same as the market. I'm not going to go into the details of calibration, you should look in the Further Reading section for pointers in that direction. Instead I will first explain what is meant by a model, and then mention a few of the popular ones.

Focus on the volatility of volatility function q first. This governs how much randomness there is in the volatility model. Suppose volatility is low. Would you expect changes in volatility to be small or large? If volatility is around 5%, will changes in that level be ball park 0.05% per day or 2% per day? (I'm not expecting you to give me an answer. Bear with me for a moment longer.) And if volatility is about 30%, will daily changes be 0.05% or 2%? The question is about how does q vary with the level of σ? Most people answer that the higher the value of volatility then the bigger the daily fluctuations in it. This seems reasonable and is borne out by research. But it is far from being sufficient information to pin down the functional form for q. It may be an increasing function of σ, but which increasing function?

The same applies to the drift function p. Most people would say that volatility is mean reverting, and this should be reflected in p. But again this means little more than p is negative when σ is large and positive when σ is small. More information is needed.

To model volatility as a stochastic process you need some statistics, or a simple model that you can calibrate. Some further ideas on the statistical approach are given in Chapter 53. Now let's look at the famous models.

51.9 NAMED/POPULAR MODELS

Hull & White (1987)

Hull & White considered both general and specific volatility modeling. The most important result of their analysis is that when the stock and the volatility are uncorrelated and the risk-neutral dynamics of the volatility are unaffected by the stock (i.e. $p - \lambda q$ and q are independent of S) then the fair value of an option is the average of the Black–Scholes values for the option, with the average taken over the distribution of σ^2.

One of the (risk-neutral) stochastic volatility models considered by Hull & White was

$$d(\sigma^2) = a(b - \sigma^2)\,dt + c\sigma^2\,dX_2.$$

Usually, the value of an option must be found numerically, but there are some simple approximations using Taylor series.

Square-root model/Heston (1993)

In Heston's model

$$dv = (a - bv)\,dt + c\sqrt{v}\,dX_2,$$

where $v = \sigma^2$. This has arbitrary correlation between the underlying and its volatility. This is popular because there are closed-form solutions for European options.

3/2 model

$$dv = (av - bv^2)\,dt + cv^{3/2}\,dX_2,$$

where $v = \sigma^2$. Again, this is popular because it has a closed-form solution.

GARCH-diffusion

Generalized autoregressive conditional heteroscedasticity, or GARCH for short, is a model for an asset and its associated volatility. The simplest such model is GARCH(1,1) which takes the form

$$\sigma_{t+1}^2 = \omega + \beta\sigma_t^2 + \alpha\epsilon_t^2,$$

where the ϵ_t are the asset price returns after removing the drift. That is,

$$S_{t+1} = S_t(1 + \mu + \sigma_t\epsilon_t).$$

Note how the notation is different from, but related to, that which we are used to. This is because, historically, GARCH was developed in an econometrical and not a financial environment. There should not be any confusion as this is the only place in the book where this notation is used. It can be shown that this simplest GARCH model becomes the same as the stochastic volatility model

$$dv = (a - bv)\,dt + cv\,dX_2.$$

as the time step tends to zero. Here $v = \sigma^2$. The parameters in the above are related to the parameters in the original GARCH specification and to the time step.

There are many references to GARCH in the Further Reading section.

Ornstein–Uhlenbeck process

With $y = \log\ v$, $v = \sigma^2$,

$$dy = (a - by)\,dt + c\,dX_2.$$

This is popular because it matches data well.[1] In the long run, volatility is lognormally distributed in this model.

51.9.1 The Heston Model

This is perhaps the most popular stochastic volatility model:

$$dv = (a - bv)\,dt + c\sqrt{v}\,dX_2,$$

$v = \sigma^2$. There are 'closed-form solutions' for simple options in this model, however, those formulae are not particularly nice. Here's how to find the formulae:

1. Change variables, $x = \log S$
2. Take Fourier transform in x
3. Solve transformed equation (now a diffusion equation in just time and v)
4. Invert the transform (numerical integration)

There are four parameters in the model: Speed of mean reversion; Level of mean reversion; Volatility of volatility; Correlation. The main four parameters can be chosen by matching data or by calibration. Experience suggests that calibrated parameters are unstable, and often unreasonable.

The Heston model with jumps

Increasingly popular is **stochastic volatility with jumps models**. Jump models (see Chapter 57) require a parameter to measure probability of a jump (a Poisson process) and a distribution for the jumps.

Pros: More parameters allow better fitting. The jump component of the model has most impact over short timescales. Therefore use longer-dated options to fit the stochastic volatility parameters and the shorter-dated options to fit the jump component.

Cons: Mathematics is slightly more complicated (and again we must work in the transform domain). Hedging is even harder when the underlying stock process is potentially discontinuous.

51.9.2 The REGARCH Model and its Diffusion Limit

REGARCH = Range-based Exponential GARCH. 'Range-based' refers to the use of the daily range, defined as the difference between the highest and lowest log asset price recorded throughout the day. 'Exponential' refers to modeling the logarithm of the variance (see Brandt & Jones, 2002).

Diffusion limits exist for GARCH-type of processes. That is, they can be expressed in continuous time using stochastic differential equations. This is achieved via 'moment matching.' The statistical properties of the discrete-time GARCH processes are recreated with the continuous-time sdes.

[1] I lied, again, this is not popular at all, even though it is a very good model.

The REGARCH model becomes the following three-factor model:

$$dS = \mu S\,dt + \sigma_1 S\,dX_0,$$

$$d(\log\sigma_1) = a_1(\log\sigma_2 - \log\sigma_1)\,dt + b_1\,dX_1,$$

$$d(\log\sigma_2) = a_2(c_2 - \log\sigma_2)\,dt + b_2\,dX_2.$$

This is a three-factor model, with two volatilities. σ_1 represents the actual volatility of the asset returns, which is stochastic. The σ_2 represents the level to which σ_1 reverts, and is itself stochastic.

For pricing options we must replace these sdes with the risk-neutral versions:

$$dS = rS\,dt + \sigma_1 S\,dX_0,$$

$$d(\log\sigma_1) = a_1(\log\sigma_2 - \log\sigma_1 - \lambda_1)\,dt + b_1\,dX_1,$$

$$d(\log\sigma_2) = a_2(c_2 - \log\sigma_2 - \lambda_2)\,dt + b_2\,dX_2.$$

The λs represent the market prices of risk.

The a and b coefficients and the correlations between the three sources of randomness give this system seven parameters. These parameters are related to the parameters of the original REGARCH model and can be estimated from asset data.

Example: The following numbers are not from an actual stock, but are typical orders of magnitude.

$$a_1 = 62.3, \quad b_1 = 1.22, \quad a_2 = 3.12, \quad b_2 = 0.401, \quad c_2 = -1.17.$$

The orders of magnitude are particularly instructive. The parameter a_1 is large; it represents the speed of mean reversion. The inverse has dimensions of time, and so mean reversion of σ_1 to σ_2 takes place over about one week. The volatility of σ_1, b_1, is quite large; this is the volatility of volatility. Mean reversion of σ_2 is much slower, over a period of about four months, with a lower volatility.

A realization of σ_1 and σ_2 is shown in Figure 51.4. Observe how σ_2 is less volatile and slower to mean revert than σ_1. And it is, of course, σ_1 that is the volatility of the stock itself.

Options can be priced either by Monte Carlo simulation or finite differences.

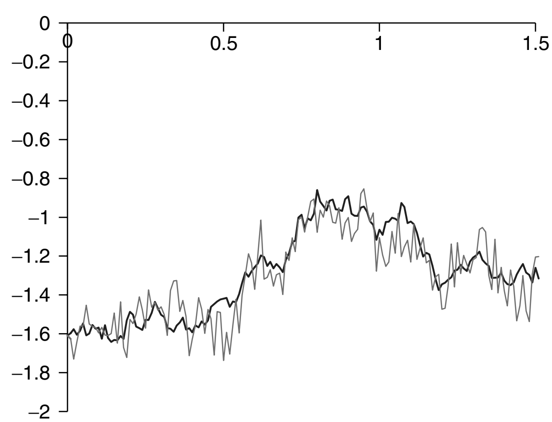


Figure 51.4 Simulation of the REGARCH model.

51.10 A NOTE ON BIASES

We aren't really trying to model or forecast volatility; we want to forecast variance, the square of volatility. That's what option values care about. We can see this in the formula for implied volatility in a deterministic volatility world, Equation (50.1). This is simply because variances add up, standard deviations don't. So suppose we have a volatility model; we should use that model to look at the statistics of variance, for example we may want to estimate the expected variance over the life of the option. Now what if our model outputs σ, or $\log\sigma$ etc. instead of σ^2; how easy is it to estimate properties of σ^2? This is the subject of 'biases.'

Suppose we model σ; what is the relationship between $E[\sigma]$ and $E[\sigma^2]$? And it is the latter we want to estimate.

Let's write

$$\sigma = \overline{\sigma} + \epsilon, \tag{51.5}$$

where $\overline{\sigma} = E[\sigma]$ so that $E[\epsilon] = 0$. We have

$$\sigma^2 = \overline{\sigma}^2 + 2\overline{\sigma}\epsilon + \epsilon^2,$$

so that

$$E[\sigma^2] = \overline{\sigma}^2 + \mathrm{Var}[\epsilon].$$

If we naively assumed that we could use $E[\sigma]^2$ instead of $E[\sigma^2]$ we would be underestimating the expected variance, underestimating option values, and making a beginner's mistake. We would have missed the **convexity adjustment**. We can do the general case, in which we model $f(\sigma)$, for example $\log\sigma$, and know the statistics of $f(\sigma)$, such as its mean and variance, and then try to deduce the stats for σ^2. To do this we make some assumptions; those consistent with us being able to perform a Taylor series and able to neglect all but the first few terms.

Using (51.5) and expanding $f(\sigma)$ in a Taylor series about $\overline{\sigma}$ we have

$$f(\sigma) = f(\overline{\sigma}) + \epsilon\, f'(\overline{\sigma}) + \tfrac{1}{2}\epsilon^2 f''(\overline{\sigma}) + \cdots,$$

where $'$ denotes $d/d\sigma$. So

$$E[f(\sigma)] = f(\overline{\sigma}) + \tfrac{1}{2}E[\epsilon^2]f''(\overline{\sigma}) + \cdots.$$

Similarly,

$$\begin{aligned} E[f(\sigma)^2] &= E\left[f(\overline{\sigma})^2 + 2\epsilon\, f(\overline{\sigma})\, f'(\overline{\sigma}) + \epsilon^2\left(f'(\overline{\sigma})^2 + f(\overline{\sigma})\, f''(\overline{\sigma})\right) + \cdots\right] \\ &= f(\overline{\sigma})^2 + E[\epsilon^2]\left(f'(\overline{\sigma})^2 + f(\overline{\sigma})\, f''(\overline{\sigma})\right) + \cdots. \end{aligned}$$

Here we have two equations for the two unknowns $\overline{\sigma}$ and $E[\epsilon^2]$ (since we know the mean and variance of $f(\sigma)$). The end result for the desired $E[\sigma^2]$ is

$$E[\sigma^2] \approx \overline{\sigma}^2 + \frac{2\left(E[f(\sigma)] - f(\overline{\sigma})\right)}{f''(\overline{\sigma})},$$

where we get $\overline{\sigma}$ from

$$f''(\overline{\sigma})\left(E[f(\sigma)^2] - f(\overline{\sigma})^2\right) - 2\left(f'(\overline{\sigma})^2 + f(\overline{\sigma})\, f''(\overline{\sigma})\right)\left(E[f(\sigma)] - f(\overline{\sigma})\right) = 0.$$

51.11 STOCHASTIC IMPLIED VOLATILITY: THE MODEL OF SCHÖNBUCHER

The problem with stochastic volatility models is that it is usually very hard to make the outputs consistent with the current prices of liquid instruments. In other words, calibrating or fitting is difficult. The reason that you might want to 'get right' the prices of liquid instruments is that you want to hedge your exotic contract with them. And it is no good hedging with something when you can't even get the price of that something right.

One way around this problem is to model instead the *implied* volatilities of the liquid instruments.

I am not going to go into the details, but will just give an overview of the approach. The details can be found in Schönbucher (1999).

The implied volatility, for each strike, E, and expiry, T, that we want to model is to satisfy some stochastic differential equation. In that stochastic differential equation there will be a deterministic drift term, a Brownian motion correlated with the underlying asset process and a part uncorrelated with that process. Let's just look at the single traded option case, so that we only have to worry about modeling one implied volatility.

The stochastic differential equation for the implied volatility $\hat{\sigma}$ will be

$$d\hat{\sigma} = u\,dt + \gamma\,dX_1 + v\,dX_2.$$

The dX_1 is the same as in the asset price process

$$dS = \mu S\,dt + \sigma S\,dX_1,$$

but the dX_2 is uncorrelated with that. The relationship between the actual spot volatility σ, and the other variables and functions will be made precise later on.

To be clear, $\hat{\sigma}$ is the implied volatility of a vanilla option with a strike E and an expiry T. Let's assume that it is a call option.

The value of this call option is given by

$$Se^{-D(T-t)}N(d_1) - Ee^{-r(T-t)}N(d_2)$$

in terms of the implied volatility etc., where

$$d_1 = \frac{\log(S/E) + (r - D + \frac{1}{2}\hat{\sigma}^2)(T-t)}{\hat{\sigma}\sqrt{T-t}}$$

and

$$d_2 = d_1 - \hat{\sigma}\sqrt{T-t}.$$

The process for the call price is given by

$$dC = \frac{\partial C}{\partial t}\,dt + \frac{\partial C}{\partial S}\,dS + \frac{\partial C}{\partial \hat{\sigma}}\,d\hat{\sigma} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\,dt + \sigma S\gamma\,\frac{\partial^2 C}{\partial S\partial\hat{\sigma}}\,dt + \tfrac{1}{2}(\gamma^2 + v^2)\frac{\partial^2 C}{\partial \hat{\sigma}^2}\,dt.$$

We could go through the usual business of setting up a risk-free portfolio etc. and applying the no-arb principle. Equivalently, we can say that the expected risk-adjusted drift rate of the call must be the same as the risk-free rate i.e.

$$E[dC] = rC\,dt.$$

Putting this all together we get

$$rC = \frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + u\frac{\partial C}{\partial \hat{\sigma}} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \sigma S\gamma \frac{\partial^2 C}{\partial S \partial \hat{\sigma}} + \tfrac{1}{2}(\gamma^2 + v^2)\frac{\partial^2 C}{\partial \hat{\sigma}^2}. \qquad (51.6)$$

Since we know that C must also satisfy the Black–Scholes equation with a volatility of $\hat{\sigma}$ we can simplify (51.6) to get

$$\tfrac{1}{2}(\sigma^2 - \hat{\sigma}^2)S^2 \frac{\partial^2 C}{\partial S^2} + u\frac{\partial C}{\partial \hat{\sigma}} + \sigma S\gamma \frac{\partial^2 C}{\partial S \partial \hat{\sigma}} + \tfrac{1}{2}(\gamma^2 + v^2)\frac{\partial^2 C}{\partial \hat{\sigma}^2} = 0.$$

We know the formula for C so we can write this as

$$\hat{\sigma} u = \frac{\hat{\sigma}^2 - \sigma^2}{2(T-t)} - \tfrac{1}{2} d_1 d_2\, (\gamma^2 + v^2) + \frac{d_2 \sigma \gamma}{\sqrt{T-t}}. \qquad (51.7)$$

Equation (51.7) is a relationship between the risk-adjusted drift of the implied volatility u, the implied volatility, the actual volatility and the two components of the volatility of the implied volatility. These cannot all be independent of each other. You have to make a decision about what you are going to model, and which will be deduced from the no-arbitrage condition (51.7).

It's natural to specify the two functions v and γ, which could be found from an examination of a time series of the volatility of the implied volatility. You can't get the drift u from the data because it is the *risk-neutral* drift of the implied volatility. So you are left with either specifying the risk-neutral drift, and from that deducing the process for the actual volatility σ, or vice versa.

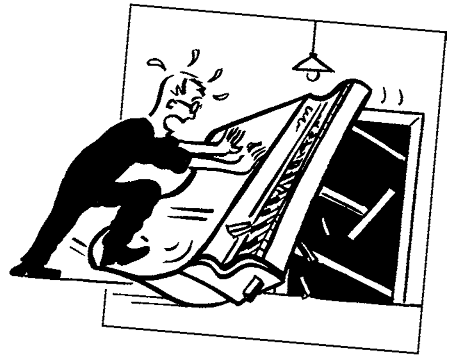
51.12 SUMMARY

Because of the profound importance of volatility in the pricing of options, and because volatility is hard to estimate, observe or predict, it is natural to model it as a random variable. For some contracts, most notably barriers, a constant volatility model is just too inaccurate.

FURTHER READING

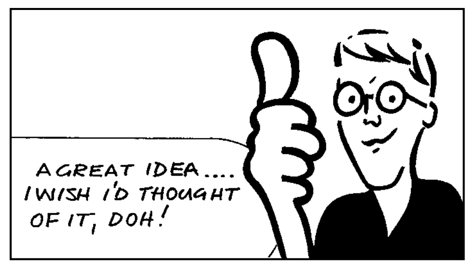
- See Hull & White (1987) and Heston (1993) for more discussion of pricing derivatives when volatility is stochastic.
- GARCH is explained in Bollerslev (1986).
- Nelson (1990) shows the relationship between GARCH models and diffusion processes.
- There are many articles in *Risk* magazine, starting in 1992, covering GARCH and its extensions in detail. Other developments are described in Engle & Mezrich (1996).
- For further work on GARCH see the papers by Engle (1982), Engle & Bollerslev (1987), Alexander (1995, 1996b, 1997b), Alexander & Riyait (1992), Alexander & Chibuma (1997) Alexander & Williams (1997), and the collection of papers edited by Engle (1995).
- Derman & Kani (1997) model the stochastic evolution of the whole local volatility surface.

- Ahn, Arkell, Choe, Holstad & Wilmott (1999) examine the risk involved in delta and static hedging under stochastic volatility.
- See Brandt & Jones (2002) for details of the REGARCH model.
- For details of the market model for stochastic *implied* volatility see Schönbucher (1999).
- The book by Rebonato (2004) is excellent for most aspects of volatility modeling.
- Mikhailov & Nögel (2003) give examples of the calibration of the Heston model.
- Lucic (2003) prices forward start options in a stochastic volatility framework.
- Fouque, Papanicolaou, Sircar & Solna (2004) show how to price path-dependent contracts in a stochastic volatility framework consistently with vanillas.
- See Swishchuk (2004) for stochastic volatility pricing of volatility swaps.



CHAPTER 52
uncertain parameters

In this Chapter...



- why the parameters in the Black–Scholes model are not reliable
- the difference between 'random' and 'uncertain'
- how to price contracts when volatility, interest rate and dividend are uncertain

52.1 INTRODUCTION

The Black–Scholes equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - D)S\frac{\partial V}{\partial S} - rV = 0$$

is a parabolic partial differential equation in two variables, S and t, with three parameters, σ, r and D, not to mention other parameters such as strike price, barrier levels etc. specific to the contract. Out of these variables and parameters, which ones are easily measurable?

- *Asset price*: The asset price is quoted and therefore easy to measure in theory. In practice, two prices are quoted, the bid and the ask prices; and even these prices will differ between market makers. This issue of transaction costs and their effect on option prices was discussed in Chapter 48.
- *Time to expiry*: Today's date and the expiry date are the easiest quantities to measure. (There is some question about how to treat weekends, but this is more a question of modeling asset price movements than of parameter estimation.)
- *Volatility*: There are two traditional ways of measuring volatility: Implied and historical. Whichever way is used, the result cannot be the future value of volatility; either it is the market's estimate of the future or an estimate of values in the past. The correct value of volatility to be used in an option calculation cannot be known until the option has expired. A time series plot of historical volatility, say, might look something like Figure 52.1, and it

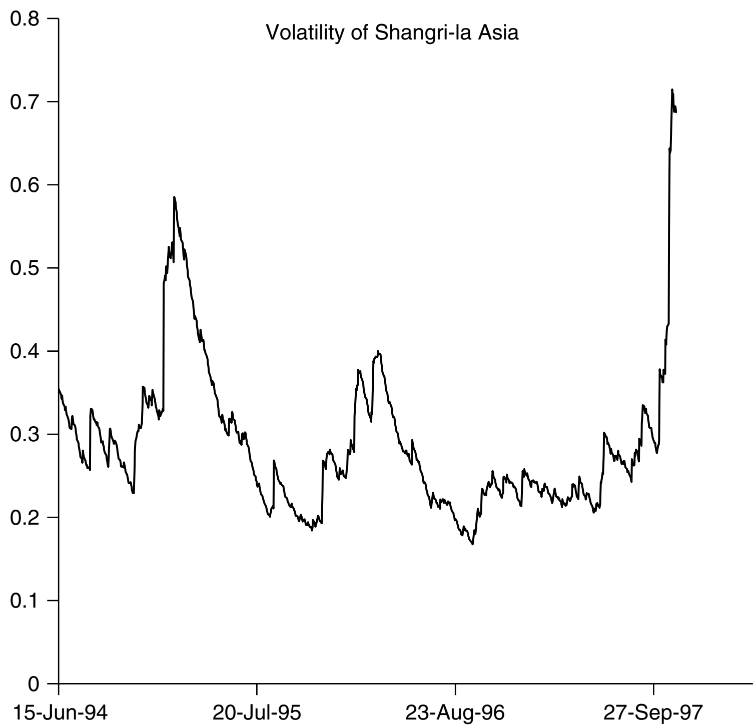


Figure 52.1 A typical time series for historical volatility; an implied volatility time series would look similar.

is certainly not constant as assumed in the simple Black–Scholes formulae. We can see in this figure that volatility for this stock typically ranges between 20 and 60%. The exception to this was during the October/November 1997 crash, for which jump/crash models are perhaps more relevant.

- *Risk-free interest rate*: Suppose we are valuing an option with a lifespan of six months. We can easily find the yield to maturity of a six-month bond and this could be our value for the risk-free rate. However, because the hedged portfolio earns the instantaneous spot rate, the Black–Scholes theory requires knowledge of the future behavior of the *spot* interest rate, and this is not the same as the six-month rate. We can, of course, couple an asset price model *and* an interest rate model, as in Chapter 33, but then we have even more problems with the accuracy of our interest rate model and estimating its parameters.
- *Dividends*: Dividends are declared a few months before they are paid. However, before then what value do we use in our option value calculation? Again, we have to make a guess at the dividend value, obviously using the past as a guide. See Chapters 8 and 64 for detailed discussions of modeling dividend structure.

In this chapter we address the problem of how to value options when parameter values are *uncertain*.

Uncertainty is different from randomness in a subtle but important way. Uncertainty means that you can't predict an outcome (just like with randomness) but it also means that you have no probabilistic description of what may happen, no probability density functions, for example. The easiest way to understand this is to think that in models of uncertainty all you know is what **cannot** happen, but you have no idea of the likelihood of what is left. Now since we do not have probabilities associated with events we clearly cannot talk about expectations, either real or risk neutral.

Rather than assuming that we know the precise value for a parameter, we assume that all we know about the parameters is that *they lie within specified ranges*. With this assumption, we do not find a *single* value for an option, instead we find that the option's value can also lie within a range: There is no such thing as *the* value. In fact, the correct value cannot be determined until the expiry of the option when we know the path taken by the parameters. Until then, there are many possible values any of which *might* turn out to be correct.

We will see that this problem is non linear, and thus an option valued in isolation has a different range of values from an option valued as part of a portfolio: We find that the range of possible option values depends on what we use to hedge the contract. If we put other options into the portfolio this will change the value of the original portfolio. This leads to the idea of incorporating traded options into an OTC portfolio in such a way as to maximize its value. This is called optimal static hedging, discussed in depth in Chapter 60. In that chapter, I also show how to apply the idea to path-dependent contracts such as barrier options.

Two of the advantages of the approach we adopt in this chapter are obvious. First, we can be more certain about the correctness of a range of values than a single value: We will be happier to say that the volatility of a stock lies within the range 20–30% over the next nine months than to say that the average volatility over this period will be 24%. Another advantage concerns the crucial matter of whether to believe market prices for a contract or the value given by some model. We have seen in Chapters 31 and 50 that it is common practice to 'adjust' a model so that it gives theoretical values for liquid contracts that exactly match the market values. Since the uncertain parameter model gives ranges for option values this means that we no longer have to choose between the correctness of a prediction and of a market price. Now they can differ; all we can say is that, according to the model, arbitrage is only certain if the market value lies outside our predicted range.

52.2 BEST AND WORST CASES

The first step in valuing options with uncertain parameters is to acknowledge that we can do no better than give ranges for the future values of the parameters. For volatility, for example, this range may be the range of past historical volatility, or implied volatilities, or encompass both of these. Then again, it may just be an educated guess. The range we choose represents our estimate of the upper and lower bounds for the parameter value for the life of the option or portfolio in question. These ranges for parameters lead to ranges for the option's value. Thus it is natural to think in terms of a lowest and highest possible option value; if you are long the option, then we can also call the lowest value the *worst* value and the highest the *best*. Work in this area was started by Avellaneda, Levy, Parás and Lyons.

We begin by considering uncertain volatility and shortly address the problem of uncertain interest rate (very similar) and then uncertain discretely paid dividends (slightly different).

52.2.1 Uncertain Volatility: The Model of Avellaneda, Levy & Parás (1995) and Lyons (1995)

Let us suppose that the volatility lies within the band

$$\sigma^- < \sigma < \sigma^+.$$

We will follow the Black–Scholes hedging and no-arbitrage argument as far as we can and see where it leads us.

Construct a portfolio of one option, with value $V(S, t)$, and hedge it with $-\Delta$ of the underlying asset. The value of this portfolio is thus

$$\Pi = V - \Delta S.$$

We still have

$$dS = \mu S\, dt + \sigma S\, dX,$$

even though σ is unknown, and so the change in the value of this portfolio is

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \left(\frac{\partial V}{\partial S} - \Delta\right) dS.$$

Even with the volatility unknown, the choice of $\Delta = \partial V/\partial S$ eliminates the risk:

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt.$$

At this stage we would normally say that if we know V then we know $d\Pi$. This is no longer the case since we do not know σ. The argument now deviates subtly from the vanilla Black–Scholes argument. What we will now say is that we will be pessimistic: We will assume that the volatility over the next time step is such that our portfolio increases by the least amount. If we have a long position in a call option, for example, we assume that the volatility is at the lower bound σ^-; for a short call we assume that the volatility is high. This amounts to considering the *minimum* return on the portfolio, where the minimum is taken over all possible values of the volatility within the given range. The return on this worst-case portfolio is then set equal to the risk-free rate:

$$\min_{\sigma^- < \sigma < \sigma^+} (d\Pi) = r\Pi\, dt.$$

Thus we set

$$\min_{\sigma^- < \sigma < \sigma^+} \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt = r\left(V - S\frac{\partial V}{\partial S}\right) dt.$$

Now observe that the volatility term in the above is multiplied by the option's gamma. Therefore the value of σ that will give this its minimum value depends on the sign of the gamma. When the gamma is positive we choose σ to be the lowest value σ^- and when it is negative we choose σ to be its highest value σ^+. We find that the worst-case value V^- satisfies

$$\frac{\partial V^-}{\partial t} + \tfrac{1}{2}\sigma(\Gamma)^2 S^2 \frac{\partial^2 V^-}{\partial S^2} + rS\frac{\partial V^-}{\partial S} - rV^- = 0 \qquad (52.1)$$

where

$$\Gamma = \frac{\partial^2 V^-}{\partial S^2}$$

and

$$\sigma(\Gamma) = \begin{cases} \sigma^+ & \text{if } \Gamma < 0 \\ \sigma^- & \text{if } \Gamma > 0. \end{cases}$$

This is all highly intuitive. If you hold a delta-hedged long volatility/gamma position what is the worst that could happen? The stock price doesn't move, volatility is low. If you hold a delta-hedged short volatility/gamma position what is the worst that could happen? The stock price moves a great deal, volatility is high.

We can find the best option value V^+, and hence the range of possible values by solving

$$\frac{\partial V^+}{\partial t} + \tfrac{1}{2}\sigma(\Gamma)^2 S^2 \frac{\partial^2 V^+}{\partial S^2} + rS\frac{\partial V^+}{\partial S} - rV^+ = 0$$

where

$$\Gamma = \frac{\partial^2 V^+}{\partial S^2}$$

but this time

$$\sigma(\Gamma) = \begin{cases} \sigma^+ & \text{if } \Gamma > 0 \\ \sigma^- & \text{if } \Gamma < 0. \end{cases}$$

We won't find much use for the problem for the best case in practice since it would be financially suicidal to assume the best outcome. (Note that just by changing the signs in Equation (52.1) we go from the worst-case equation to the best. In other words, the problem for the worst price for long and short positions in a particular contract is mathematically equivalent to valuing a long position only, but in worst and best cases. This distinction between long and short positions is an important consequence of the nonlinearity of the equation and we discuss this in depth shortly.)

Equation (52.1), derived by Avellaneda, Levy & Parás and Lyons, is exactly the same as the Hoggard–Whalley–Wilmott transaction cost model that we saw in Chapter 48. The partial differential equation may be the same, but the reason for it is completely different.

52.2.2 Example: An Up-and-out Call

Consider the following up-and-out call option. The strike is 100, the spot interest rate is 5%, the barrier is at 120 and there is one year to expiry. Let's suppose we are not sure what the volatility is, but it should be in the range 17–23%.

If we priced the option using first a 17% volatility and then a 23% volatility we would get two curves looking like those in Figure 52.2. If the asset value were at around 80 you might think that the option was insensitive to the volatility. But actually it is around that asset value where the option is very sensitive to the volatility ranging over 17–23%. Let's see why this is by solving the worst-case scenario problem with

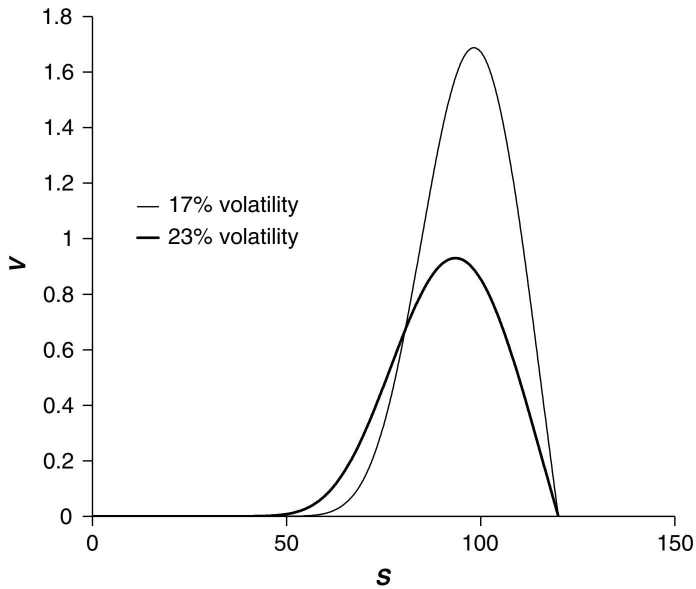


Figure 52.2 Up-and-out call values in a Black–Scholes world with volatilities of 17% and 23%.

the range 17–23%; after all the volatility could do anything within this range, it doesn't have to remain constant.

Equation (52.1) must in general be solved numerically, because it is non linear. In Figure 52.3 are shown the best and worst prices for an up-and-out call option. As I said above, the best and worst prices could be interpreted as worst-case prices for short and long positions respectively. The two bold lines give the worst-case long and short values assuming a volatility ranging over 17% to 23%.

We must solve the non-linear equation numerically because the gamma for this contract is not single signed. The problem is genuinely non linear, and we cannot just substitute each of 17% and 23% into a Black–Scholes formula.

Observe that the best/worst of the two curves (the 'envelope') in Figure 52.2 is not the same as the best/worst of Figure 52.3. If you were to price the option using the worst out of the two standard Black–Scholes prices, with constant volatility, you would significantly overestimate the value of the option in the worst-case scenario. For this reason it can be extremely dangerous to calculate a contract's vega when the contract has a gamma that changes sign. Having said that, practitioners do have ways of fudging their vega calculations to make some allowance for this effect.

Continuing with this barrier option example, let us look at implied volatilities. In Figure 52.4 is shown the Black–Scholes value of an up-and-out call option as a function of the volatility. The strike is 100, the stock is at 80, the spot interest rate is 5%, the barrier is at 120 and there is one year to expiry. This contract has a gamma that changes sign, and a price that is not monotonic in the volatility. This figure shows that there is a maximum option value of 0.68 when the volatility is about 20%. Now turn the problem around. Suppose that the market is pricing this contract at 0.55. From Figure 52.4 we can see that there are *two* volatilities that correspond to this market price. One has a positive vega and the other negative. Which, if either,

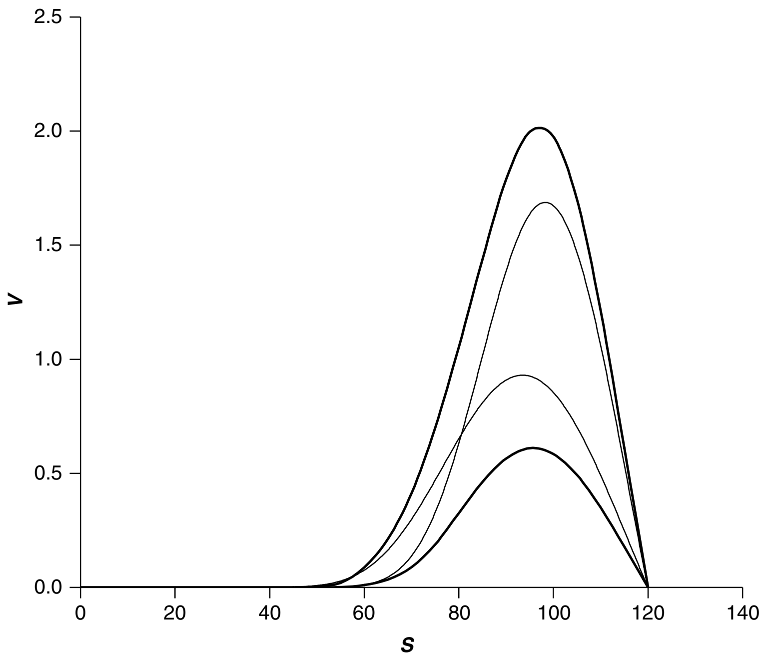


Figure 52.3 Up-and-out call value assuming a range for volatility, and two Black–Scholes prices assuming constant volatility.

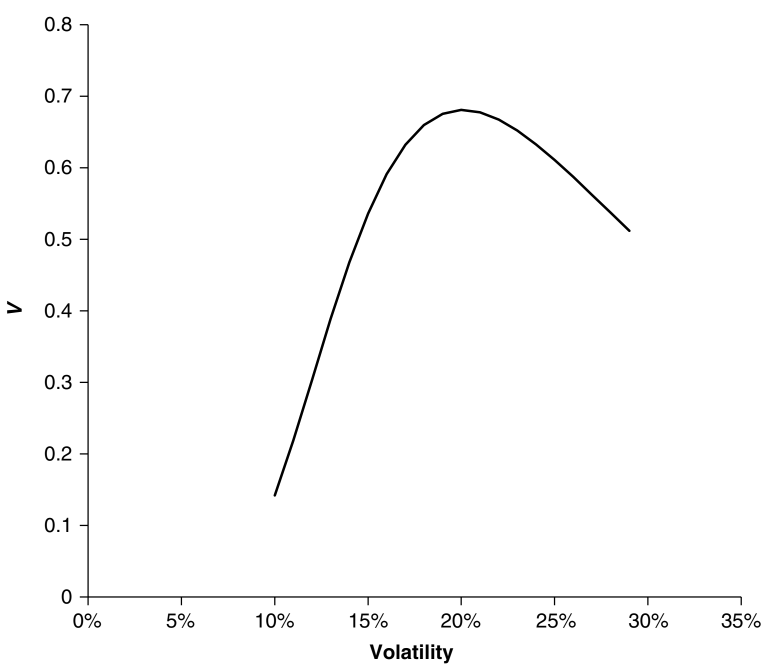


Figure 52.4 Black–Scholes values for an up-and-out call option against volatility.

is correct? The question is probably meaningless because of the non-single-signed gamma of this contract.

Take this example further. What if the market price is 0.72? This value cannot be reached by any single volatility. Does this mean that there are arbitrage opportunities? Not necessarily. This could be due to the market pricing with a non-constant volatility, either with a volatility surface, stochastic volatility or a volatility range. As we have seen from the best/worst prices for this contract, the uncertainty in the option value may be large enough to cover the market price of the option, and there may be no guaranteed arbitrage at all. The best value is greater than 1 for the volatility range 17–23%, as can be seen in Figure 52.3.

Whenever you see a non-monotonic plot of price versus a parameter it should set off alarm bells. Clearly there are competing effects going on here. Something makes the price increase with the parameter value, and something else makes it decrease. In the barrier example this is because of the area with positive gamma competing with the area with negative gamma. You must ask yourself what if the cause of the increase were to become ineffective; this would leave you with only the bad stuff. In the barrier example suppose that the benefits of the positive gamma area were not to apply because volatility didn't increase there. The net result would be far worse than you might otherwise have expected. So the rule is to start thinking more deeply about scenarios whenever you see such non monotonicity.

In Figure 52.5 is shown the envelope of all possible Black–Scholes up-and-out call option values assuming a *constant* volatility ranging from *zero to infinity*. Option values outside this

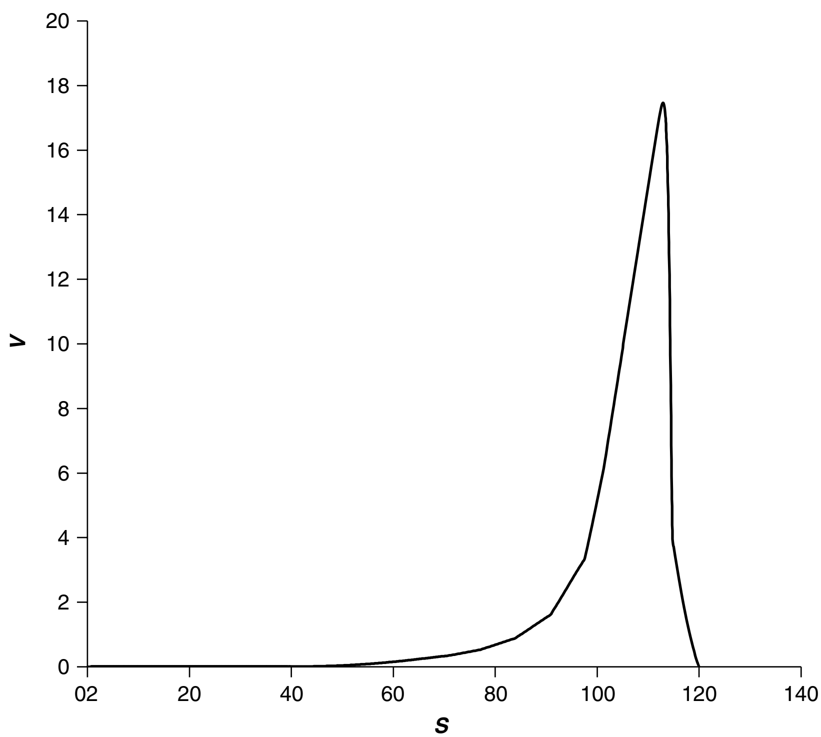


Figure 52.5 The envelope of all Black–Scholes values for an up-and-out call option.

envelope cannot be attained with constant volatility. With uncertain volatility, prices outside this envelope *can* easily be attained.

52.2.3 Uncertain Interest Rate

Exactly the same idea can be applied to the case of uncertain interest rate. Let us suppose that the risk-free interest rate lies within the band

$$r^- < r < r^+.$$

Continuing with the worst case, we assume that if our portfolio has a net positive value ($\Pi > 0$) then interest rates will be high, if it has a negative value then rates will be low. The reason for this is that if we have a positive amount invested in options then in the worst case we are failing to benefit from high interest rates. Thus, the interest rate we choose will depend on the sign of Π. As before, we set the return on the portfolio equal to the risk-free rate. With our special choice for this rate we arrive at the equation for the worst-case scenario:

$$\frac{\partial V^-}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V^-}{\partial S^2} + r(\Pi)\left(S\frac{\partial V^-}{\partial S} - V^-\right) = 0 \qquad (52.2)$$

where

$$\Pi = V^- - S\frac{\partial V^-}{\partial S}$$

and

$$r(\Pi) = \begin{cases} r^+ & \text{if } \Pi > 0 \\ r^- & \text{if } \Pi < 0. \end{cases}$$

The equation for the best case is obvious.

52.2.4 Uncertain Dividends

Consider a dividend yield that is independent of the asset price. There are two cases to consider here: Continuously paid and discretely paid. In the former case (and with FX options where the dividend yield is replaced by the foreign interest rate) we assume that the dividend yield lies between two values:

$$D^- \le D \le D^+.$$

Now, for the option's worst value we simply solve

$$\frac{\partial V^-}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V^-}{\partial S^2} + r\left(S\frac{\partial V^-}{\partial S} - V^-\right) - D(\Delta)S\frac{\partial V^-}{\partial S} = 0 \qquad (52.3)$$

where

$$D(\Delta) = \begin{cases} D^+ & \text{if } \Delta > 0 \\ D^- & \text{if } \Delta < 0. \end{cases}$$

The case of discretely paid dividends is slightly different. Consider what happens across a dividend date, t_i. Let us suppose that the discretely paid dividend yield lies within the band

$$D^- < D < D^+.$$

The jump condition across a discretely paid dividend date is

$$V(S, t_i^-) = V((1 - D)S, t_i^+)$$

as shown in Chapter 8. When this dividend yield is uncertain we simply minimize over the possible values of D and this gives the following jump condition

$$V^-(S, t_i^-) = \min_{D_- < D < D_+} V^-((1 - D)S, t_i^+). \tag{52.4}$$

There is a corresponding jump condition for the best case.

Pricing with an uncertain dividend, instead of dividend yield, is just the same. Figure 52.6 gives an example of a portfolio consisting of long two calls struck at 80 and short three struck at 110, all expire in six months. The risk-free rate is 7% and volatility 20%. A dividend is paid after four months, and this dividend lies in the range zero to five.

Of course, it is a simple matter to put together the results of the above sections to model uncertainty in all of volatility, interest rate and dividend.

I emphasize that from now on we only consider the pessimistic case. We assume the worst outcome and price contracts accordingly. By so doing we guarantee that we never lose money provided that our uncertain parameter ranges are not violated.

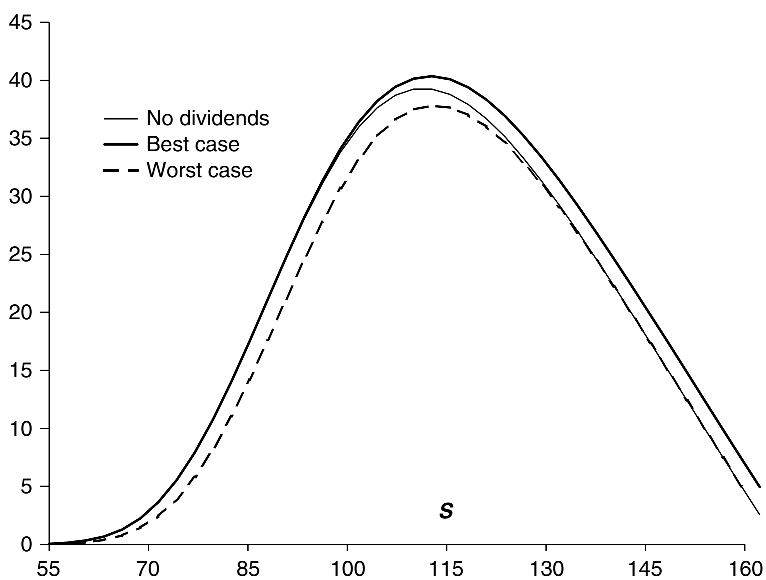


Figure 52.6 Portfolio value under uncertain dividends, best and worst cases.

52.3 UNCERTAIN CORRELATION

Exactly the same idea can be applied to multi-asset instruments which have a dependence on the correlation between the underlyings. Correlation is something that is particularly difficult to calculate or predict so there is obviously a role to be played by uncertainty. The downside is that the spread on correlation is likely to be so large that the spread on the option price will be unrealistic. And since there are so few correlation instruments with which to hedge it may not be possible to reduce this spread to anything tradeable.

52.4 NONLINEARITY

All of the new partial differential equations that we have derived in this chapter are non linear. In particular, the uncertain volatility model results in the same non-linear equation, Equation (52.1), as the Hoggard–Whalley–Wilmott model for pricing options in the presence of transaction costs, discussed in Chapter 48. Because of this nonlinearity, we must therefore distinguish between long and short positions, for example for a long call we have

$$V^-(S,T) = \max(S-E,0)$$

and for a short call

$$V^-(S,T) = -\max(S-E,0).$$

Explicit solutions exist only in special cases, and these cases are, depending on which of the volatility, interest rate and dividend are uncertain, for options for which the gamma (for uncertain volatility), the portfolio value (for uncertain interest rate) and the delta (for uncertain dividend) are single signed. Then the formulae are simply the Black–Scholes formulae with the parameters set to the appropriate extreme values.

Because Equations (52.1) and (52.2) are non linear the value of a portfolio of options is not necessarily the same as the sum of the values of the individual components. Long and short positions have different values. For example, a long call position has a lower value than a short call. In both cases we are being pessimistic: If we own the call we assume that it has a low value, if we are short the call and thus may have to pay out at expiry, we assume that the value of the option to its holder is higher. Note that here we mean a long (or short) position valued *in isolation*. Obviously, if we hold one of each simultaneously then they will cancel each other regardless of the behavior of any parameters. This is a very important point to understand: The value of a contract depends on what else is in the portfolio.

One of the really nice things about these non-linear models is that the value of a portfolio of options is no less than the value of the individual options. In symbols this can be written as

$$\text{Value}\left(\sum_{i=1}^{N}\text{Option}_i\right) \geq \sum_{i=1}^{N}\text{Option}_i$$

or, as I like to think of it,

$$\text{Value}(J+P+G+R) \geq \text{Value}(J) + \text{Value}(P) + \text{Value}(G) + \text{Value}(R).$$

Why have I used these symbols? Answer in the footnote.[1]

[1] The Beatles. The whole being greater than the sum of the parts. Ouch!

Unfortunately, the model as it stands predicts very wide spreads on options. For example, suppose that we have a European call, strike price \$100, today's asset price is \$100, there are six months to expiry, no dividends but a spot interest rate that we expect to lie between 5% and 6% and a volatility between 20% and 30%. We can calculate the values for long and short calls assuming these ranges for the parameters directly from the Black–Scholes formulae. This is because the gamma and the portfolio value are single-signed for a call. A long call position is worth \$6.85 (the Black–Scholes value using a volatility of 20% and an interest rate of 5%) and a short call is worth \$9.85 (the Black–Scholes value using a volatility of 30% and an interest rate of 6%). This spread is much larger than that in the market, and in this example, is mostly due to the uncertain volatility. The market prices may, for example, be based on an interest rate of 5.5% with a volatility between 24% and 26%. Unless the model can produce narrower spreads the model will be useless in practice.

The spreads *can* be tightened by 'static hedging.' This means the purchase and sale of traded option contracts so as to improve the marginal value of our original position. This only works because we have a *non-linear* governing equation: The price of a contract depends on what else is in the portfolio. This static hedge can be optimized so as to give the original contract its best value; we can squeeze even more value out of our contract with the best hedge.

The issues of static hedging and optimal static hedging are covered in detail in Chapter 60.

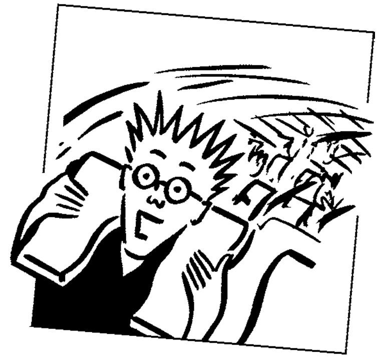
52.5 SUMMARY

Greeks such as vega and rho can be completely misleading if used carelessly and without understanding. This is because they are derivatives with respect to parameters rather than variables, and in the Black–Scholes formulae those parameters are constant. Differentiating with respect to a constant can be dangerous. The uncertain parameter model gives a consistent way to eliminate dependence of a price on a parameter, and to some extent reduce model dependence. Out of the three volatility models (deterministic calibrated smile, stochastic and uncertain), uncertain volatility is easily my favorite. As presented in this chapter the spreads on the best/worst prices are too large for the model to be of any practical use. Fortunately this situation will be remedied in Chapter 60.

FURTHER READING

- See Avellaneda, Levy & Parás (1995), Avellaneda & Parás (1996) and Lyons (1995) for more details about the modeling of uncertain volatility.
- See Bergman (1995) for the derivation of Equation (52.2) for a world in which there are different rates for borrowing and lending.
- See Oztukel (1996) for the uncertain parameter technique applied to correlation.
- Ahn, Muni & Swindle (1997) explain how to modify the final payoff to reduce the effect of volatility errors in the worst-case scenario.
- See Ahn, Muni & Swindle (1999) for a stochastic control approach to utility maximization in the worst-case model.

CHAPTER 53
empirical analysis of volatility



In this Chapter...

- how to analyze volatility data to determine the most suitable time-independent stochastic model
- how to determine the probability that the volatility will stay within any specified range
- how to assign a degree of confidence to your uncertain volatility model price

53.1 INTRODUCTION

In this chapter we examine real data in order to model the behavior of volatility. Our principal aim in this is to derive a good stochastic volatility model. This volatility model can then be used in a number of ways: In a two-factor option-pricing model; To examine the time evolution of volatility from an initial known value today to, in the long run, a steady-state distribution; To estimate the probability of our uncertain volatility model price being correct. The approach we adopt must be contrasted with the traditional approach to stochastic volatility which seems to be to write down something nice and tractable and then fit the parameters.

53.2 STOCHASTIC VOLATILITY AND UNCERTAIN PARAMETERS REVISITED

The classical way of dealing with random variables is to model them stochastically. This we do here for volatility, deriving a stochastic model such that drift and variance, and the steady-state mean and dispersion of volatility are *compatible with historical data*. We can then, among other things, determine the evolution of volatility from the known value today to the steady-state distribution in the long term.

Recall also, the approach of Avellaneda, Levy, Parás and Lyons for modeling volatility. In their model they allow volatility to do just about anything as long as it doesn't move outside some given range. With this model it is natural to calculate not a single option value but a worst price and a best price for the option. This results in a non-linear partial differential equation. This procedure yields a 'certainty interval' for the price of the option, driven by the input volatility band. Hence we know that, for example, if we could access the option in the market below our worst price then within the limits of our certainty band assumptions, we are guaranteed a profit. But what if we are not one hundred percent sure about the volatility range? Can we use our stochastic volatility model to see how likely volatility is to stay in the range?

53.3 DERIVING AN EMPIRICAL STOCHASTIC VOLATILITY MODEL

In this chapter we are going to work with daily Dow Jones Industrial Average spot data from January 1975 until August 1995. The method that I am going to demonstrate for determining the volatility model can be used for many financial time series; indeed in this book we also use the same method for modeling US short-term interest rates and inflation.

In Figure 53.1 is shown the calculated annualized 30-day volatility of daily returns for the DJIA over the period of interest. (Whether or not the volatility that has been calculated here is representative of the 'actual' volatility is debatable. Nevertheless, the techniques I describe below are applicable to whatever volatility series we have.)

Previous attempts at modeling volatility stochastically are numerous. However, since adding another stochastic variable complicates the problem of pricing options, the emphasis to date has been on deriving a tractable model. We are going to try to determine a stochastic model for volatility by fitting the drift and variance functions to empirical data, letting tractability take a secondary role. We model the volatility in isolation, assuming that it does not depend on the

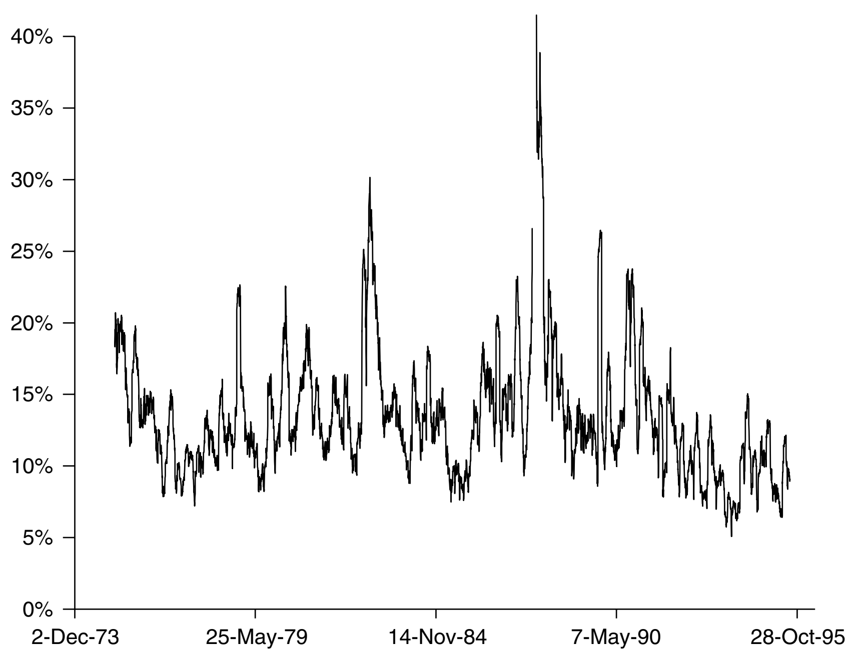


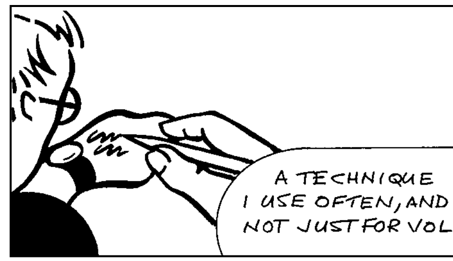
Figure 53.1 Thirty-day volatility of daily returns for the Dow Jones Industrial Average from 1975 to 1995.

level of the index. Thus we assume that the stochastic process for volatility is given by

$$d\sigma = \alpha(\sigma)\,dt + \beta(\sigma)\,dX. \tag{53.1}$$

We use twenty years of daily Dow Jones closing prices (in total over 5,000 observations) to calculate daily returns and 30-day volatility of these daily returns.

53.4 ESTIMATING THE VOLATILITY OF VOLATILITY



From (53.1), the square of the day-to-day change in volatility is given by

$$(\delta\sigma)^2 = \beta(\sigma)^2\phi^2\delta t \tag{53.2}$$

to leading order, where ϕ is a standardized Normal variable.

From our time series for volatility we can easily calculate the time series for $\delta\sigma$, and thus a time series for $(\delta\sigma)^2$. Then, in order to examine the dependence of the changes in volatility on the level of the volatility, split σ into buckets and calculate the mean of $(\delta\sigma)^2$ for each σ falling in a particular bucket. This gives us an estimate of the dependence of $E\left[(\delta\sigma)^2\right]$ on σ.

In Figure 53.2 is plotted $\log\left(E\left[(\delta\sigma)^2\right]\right)$ versus $\log(\sigma)$. Superimposed on this figure is a straight line, fitted to the dots by ordinary least squares.

The straight line suggests that

$$\log\left(E\left[(\delta\sigma)^2\right]\right) = 2\log(\beta(\sigma)) + \log\delta t = a + b\log(\sigma)$$

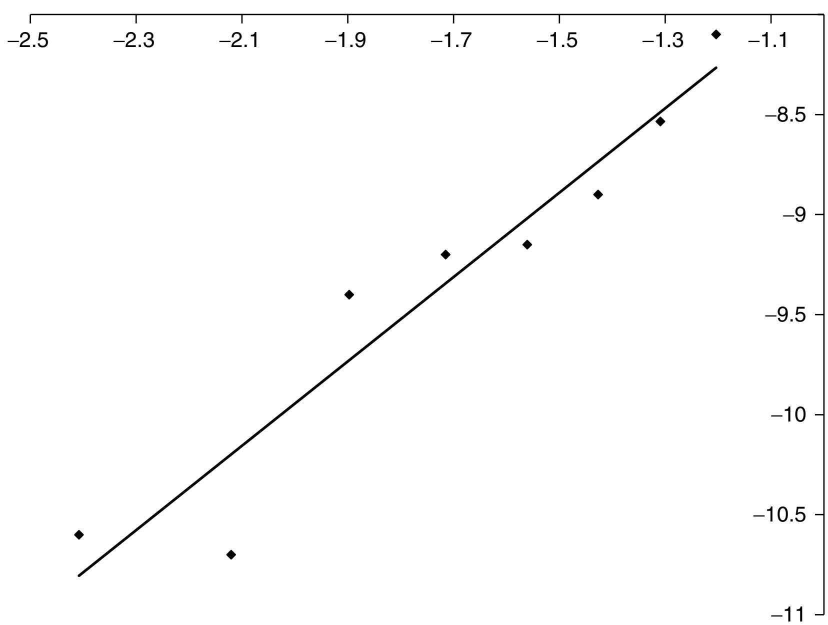


Figure 53.2 Plot of $\log\left(E\left[(\delta\sigma)^2\right]\right)$ versus $\log(\sigma)$ and the fitted line.

is a good approximation. From the data we can estimate the parameters a and b and so deduce that a good model is

$$\beta(\sigma) = \nu\sigma^{\gamma}. \tag{53.3}$$

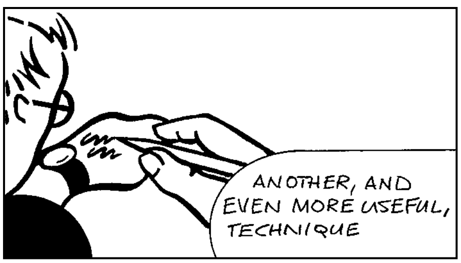
From the data we find that

$$\nu = 50.88 \quad \text{and} \quad \gamma = 1.05.$$

In our model we take

$$d\sigma = \alpha(\sigma)\,dt + \nu\sigma^{\gamma}\,dX.$$

This gives a model with volatility of volatility that is consistent with empirical data. To complete the modeling we still need to find the drift of the volatility $\alpha(\sigma)$.



53.5 ESTIMATING THE DRIFT OF VOLATILITY

To calculate $\alpha(\sigma)$ we need to examine the time series for σ in a slightly different way. In particular we are going to determine the steady-state probability density function for σ and from that deduce the drift function. The equation governing the probability density function for σ is of the form derived in Chapter 10; it is the Fokker–Planck equation (or forward Kolmogorov equation)

$$\frac{\partial p}{\partial t} = \tfrac{1}{2}\frac{\partial^2}{\partial \sigma^2}(\beta^2 p) - \frac{\partial}{\partial \sigma}(\alpha p) \tag{53.4}$$

where $p(\sigma, t)$ is the probability density function for σ and I am using σ and t to denote the forward variables. Suppose that somehow we know the steady-state distribution, $p_\infty(\sigma)$, for σ. This will satisfy

$$0 = \tfrac{1}{2}\frac{d^2}{d\sigma^2}(\beta^2 p_\infty) - \frac{d}{d\sigma}(\alpha p_\infty).$$

This is simply a steady-state version of (53.4). Integrating this once we get

$$\alpha(\sigma) = \frac{1}{2p_\infty}\frac{d}{d\sigma}(\beta^2 p_\infty). \tag{53.5}$$

The constant of integration is zero, as can be shown by examining the behavior of α for large and small σ.

From (53.5), we see that if we know p_∞ then we can find the drift term. But can we find p_∞? Yes, we can. If we assume that all parameters are independent of time, then we can determine the steady-state probability density function by using the 'ergodic property' of random walks; that is, the equivalence of ensemble and time averages. In layman's terms this means that the

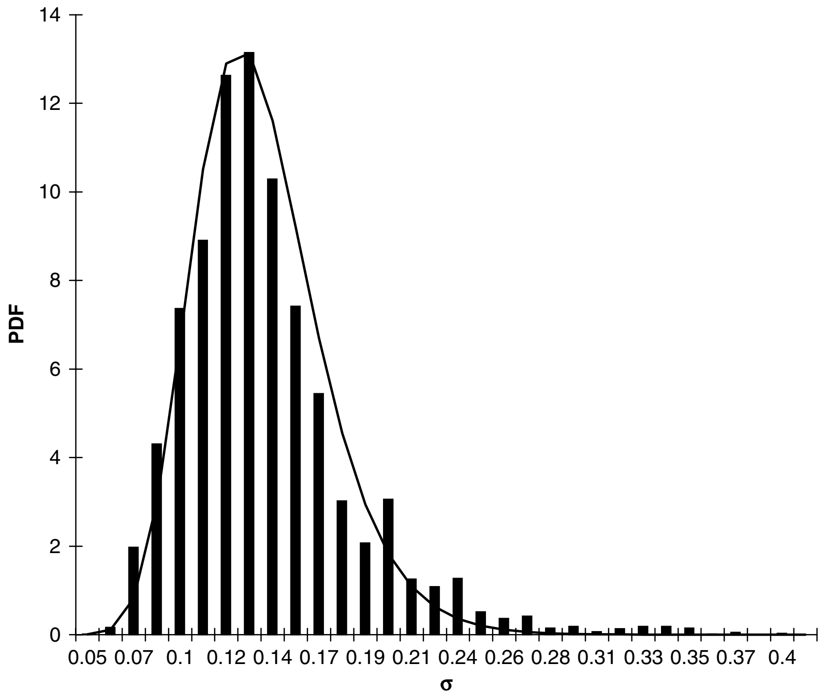


Figure 53.3 The steady-state distribution of σ and the fitted lognormal curve.

distribution of the volatility that you get from a single, very long, realization will be the same as what you get from the last data point of many separate realizations.

We find p_∞ from the data by plotting the frequency distribution of σ versus buckets of σ; that is, how many observations fall into each bucket. The empirical distribution, shown in Figure 53.3, closely resembles a lognormal curve. Since we need to differentiate this function to find the drift according to (53.5) we fit a curve to the distribution. We shall assume p_∞ to be of the form of a lognormal distribution:

$$p_\infty = \frac{1}{\sqrt{2\pi}a\sigma} e^{-(1/2a^2)(\log(\sigma/\overline{\sigma}))^2}$$

where $\log \overline{\sigma}$ represents the mean of the distribution of $\log \sigma$ and a describes the dispersion of the distribution about the mean.

This graph and fitted curve are shown in Figure 53.3.

We find from the data that,

$$a = 0.33 \quad \text{and} \quad \overline{\sigma} = 13.4\%.$$

We now have that

$$\alpha(\sigma) = \nu^2 \sigma^{2\gamma - 1} \left(\gamma - \frac{1}{2} - \frac{1}{2a^2} \log(\sigma/\overline{\sigma}) \right).$$

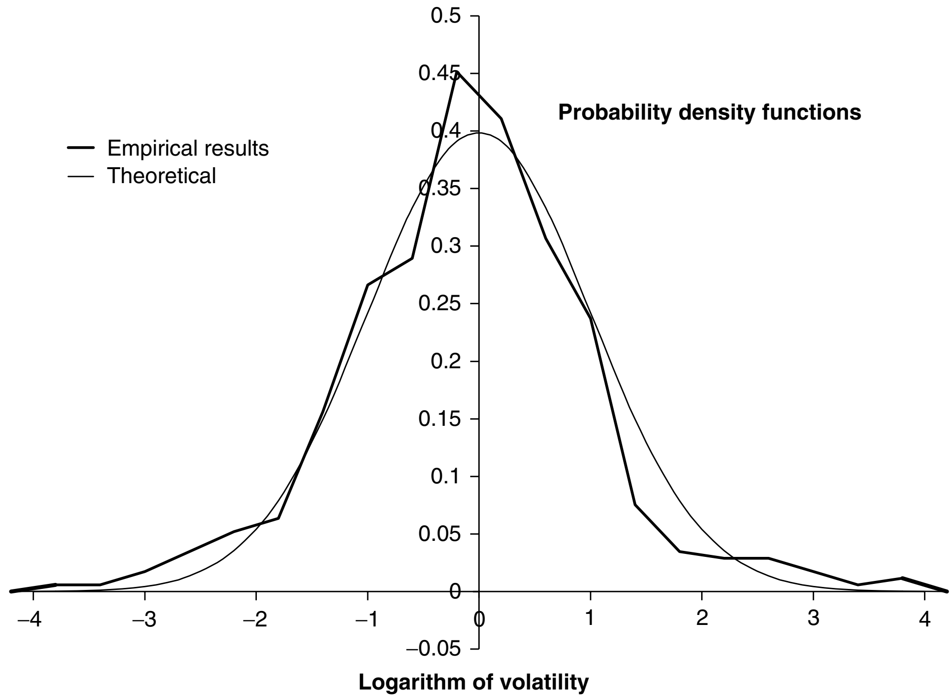


Figure 53.4 The theoretical volatility distribution and the empirical out of sample.

53.6 OUT-OF-SAMPLE RESULTS

Figure 53.4 shows the results of this model using results from approximately 150 stocks and indices. To each of these underlyings is fitted a model as described above, based on past data. This model is then used to forecast volatility ranges, with a certain degree of confidence, over the next two months for each stock/index. After two months we come back and look at what happened to volatility. As can be seen from the graph, the distribution of forecast and actual volatilities are very similar.

53.7 HOW TO USE THE MODEL

We can use our model in a number of ways, here I just make a few suggestions. The three possibilities that I will describe are straightforward option pricing, determining the future distribution for volatility and estimating our confidence in volatility ranges.

53.7.1 Option Pricing with Stochastic Volatility

The first way we can use the model is in a two-factor model for pricing options. The two factors would be the underlying asset and its volatility. The only problem with this is that we have estimated from data the *real* drift rate and not the *risk-neutral* rate that we need for option pricing. Whether this matters or not is debatable.

It is possible to estimate or model the market price of volatility risk but that would require looking at the market price of options. The real drift of volatility and the volatility of volatility exist even in the absence of options, but to measure the market price of volatility risk we need that option data. We won't be going that extra step here, but in Chapter 36 you saw how we can do a similar analysis in the slightly simpler world of stochastic interest rates.

53.7.2 The Time Evolution of Stochastic Volatility

From the stochastic volatility model we can derive the probabilistic evolution of volatility from an initial value. We turn once again to the Fokker–Planck equation, given by (53.4), which describes the evolution over time of the probability density function of a random variable defined by a stochastic differential equation. We can use the functions α and β derived above to find the probability density function of volatility over a specified time horizon. This way we can observe the evolution of the distribution of volatility.

As the initial condition for Equation (53.4) we would apply a delta function—meaning that we know the value of today's volatility with certainty—and solve for the resulting probability density function as time evolves.

What we would see would be the density function starting out with a delta function spike, and gradually smoothing out until it stabilized at the input steady-state probability density function. Because the volatility has a sensible distribution in the steady state we would not see any unreasonable behavior from the model. Other models, having nice but otherwise arbitrary drift rate and volatility of volatility might not be so well behaved.

53.7.3 Stochastic Volatility, Certainty Bands and Confidence Limits

The third way we could use the model is in conjunction with the uncertain volatility model described in Chapter 52. From our stochastic volatility model we can estimate the probability of our chosen volatility range being breached. We can therefore assign a probability to our uncertain volatility model price.

The likelihood of the volatility staying in any range can be mathematically represented by the function $C(\sigma, t)$ satisfying

$$\frac{\partial C}{\partial t} + \tfrac{1}{2}\beta(\sigma)^2\frac{\partial^2 C}{\partial \sigma^2} + \alpha(\sigma)\frac{\partial C}{\partial \sigma} = 0,$$

see Chapter 10. The final condition is given by

$$C(\sigma, T) = 0 \quad \text{for all } \sigma$$

since at $t = T$ the likelihood of σ breaching the barriers is zero. The boundary conditions are given by

$$C(\sigma^+, t) = C(\sigma^-, t) = 1 \quad \text{for all } t.$$

Here σ^- and σ^+ are the lower and upper barriers respectively. These conditions are obvious: If $\sigma = \sigma^-$ or $\sigma = \sigma^+$ then the barrier is certain to be breached. The final and boundary conditions are shown schematically in Figure 53.5.

This third use of the model is quite a nice compromise. One reason I like it is that it allows the volatility to be random, satisfying a given model, yet gives an option price that doesn't depend on the market price of risk.

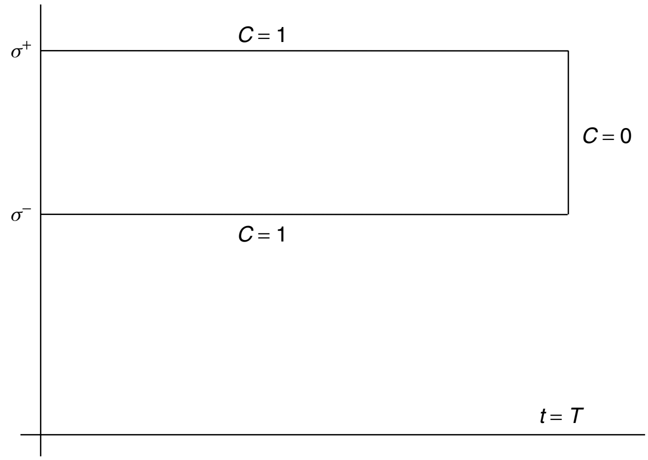


Figure 53.5 The final and boundary conditions for the exit probability.

53.8 SUMMARY

Generally speaking I don't like to accept a model because it looks nice. If at all possible I prefer to look at the data to try and find the best model. In this chapter we examined volatility data to find the best model having time-independent functions describing the drift and volatility of volatility. Crucially we modeled the volatility of volatility by examining the daily changes of volatility. After all, the randomness in the level of the volatility is seen over the shortest timescale. However, to find the best drift we looked over a longer timescale, and we estimated the steady-state distribution for the volatility. The drift only really comes into its own over longer timescales. To examine the short timescale to estimate the drift would at best not make much sense and at worst result in a model with poor long-term properties.

The resulting stochastic differential equation model for the volatility can be used in several ways. One is just to accept it as the model for volatility in a two-factor option model. Of course, then you have to estimate the market price of risk; not nice. Another use for the model, which gets around the market price of risk problem, is in conjunction with an uncertain volatility model, such as described in Chapter 52. Prescribe a range for volatility and then use the stochastic differential equation to estimate the probability of the volatility staying within this range.

FURTHER READING

- The subject of bounds for option values when volatility is stochastic was first addressed by El Karoui, Jeanblanc-Picque & Viswanathan (1991).
- See Oztukel (1996) and Oztukel & Wilmott (1998) for further details of the analysis of volatility.

CHAPTER 54
stochastic volatility and mean-variance analysis

In this Chapter...

- how to analyze risk when volatility is stochastic
- mean-variance analysis

54.1 INTRODUCTION

I tend not to like any model that requires the input of a market price of risk. The main reason is that this quantity is not directly observable. At best it can be deduced from the prices of derivatives, so called 'fitting.' But this is far from adequate, since the fitting will only work if those who set the prices of derivatives are using the same model and they are consistent in that the fitted market price of risk does not change when the model is refitted a few days later. To see what I mean, think back to Chapter 50. In that chapter I showed how to find the local deterministic volatility surface that is consistent with all quoted prices of vanilla options. If a few days later we find that this surface has changed, still being chosen to match market prices, then the model was wrong. Exactly the same problem occurs when we have stochastic volatility and we have fitted to find the market price of volatility risk.[1]

Whether we have a deterministic volatility surface or a stochastic volatility model with prescribed or fitted market price of risk, we will always be faced with how to interpret refitting. Was the market wrong before but is now right, or was the market correct initially and now there are arbitrage opportunities? We won't be faced with awkward questions like this if we don't expect our model, whatever it may be, to give unique and correct values. In this chapter we'll see how to estimate probabilities for prices being correct. We do this by only delta hedging and not dynamically vega hedging. Instead we look at means and variances for option values.

54.2 THE MODEL FOR THE ASSET AND ITS VOLATILITY

We are going to work with the very classical model

$$dS = \mu S\,dt + \sigma S\,dX_1$$

[1] I call this the market-price-of-risk risk.

and

$$d\sigma = p(S,\sigma)\,dt + q(S,\sigma)\,dX_2$$

with a correlation of $\rho(S,\sigma)$. We'll only consider a non-dividend-paying asset; the modifications needed to allow for dividends are the usual.

We are going to examine the statistical properties of a portfolio that tries to replicate as closely as possible the original option position. We will not hedge the portfolio dynamically with other options so our portfolio will not be risk free. Instead we will examine the mean and variance of the value of our portfolio as it varies through time.

With Π representing the value of the position $-\Delta$ in the asset,

$$\Pi = -\Delta S.$$

Thus

$$d\Pi = -\Delta\,dS,$$

and

$$d\Pi - r\Pi\,dt = -\Delta\,dS + r\Delta S\,dt = -(\mu - r)S\Delta\,dt - \sigma S\Delta\,dX_1.$$

54.3 ANALYSIS OF THE MEAN

Define the mean m of the portfolio value at any time by

$$m(S,\sigma,t) = E\left[\Pi(S,\sigma,t)|\text{state of the world at time } t\right].$$

Since

$$E[dm - rm\,dt] = E[d\Pi - r\Pi\,dt] = -(\mu - r)S\Delta\,dt,$$

we have

$$\frac{\partial m}{\partial t} + \tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 m}{\partial S^2} + \rho\sigma Sq\frac{\partial^2 m}{\partial S\partial\sigma} + \tfrac{1}{2}q^2\frac{\partial^2 m}{\partial\sigma^2} + \mu S\frac{\partial m}{\partial S} + p\frac{\partial m}{\partial\sigma} - rm = -(\mu - r)S\Delta. \quad (54.1)$$

We still have to decide on Δ. We will choose it to minimize the variance locally, so we can't choose it until we've analyzed the variance in the next section. Note also that the final condition for (54.1) will be the payoff for our original option that we are trying to replicate.

This equation for m was easy to derive, the equation for the variance is a bit harder.

54.4 ANALYSIS OF THE VARIANCE

Define the variance $v(S,\sigma,t)$ by

$$v(S,\sigma,t) = E\left[(\Pi(S,\sigma,t) - m(S,\sigma,t))^2\,|\text{state of the world at time } t\right].$$

Thus

$$
\begin{aligned}
v(S+dS,\sigma+d\sigma,t+dt) &= E\left[(\Pi(S+dS,\sigma+d\sigma,t+dt)-m(S+dS,\sigma+d\sigma,t+dt))^2 \,|\right.\\
&\qquad\left.\text{state of the world at time } t+dt\right]\\
&= E\left[(\Pi(S,\sigma,t)+d\Pi-m(S,\sigma,t)-dm)^2 \,|\right.\\
&\qquad\left.\text{state of the world at time } t+dt\right].
\end{aligned}
$$

Taking expectations over what happens from t to $t+dt$

$$
\begin{aligned}
&E[v(S,t,t)+dv]\\
&\quad = E\left[(\Pi(S,\sigma,t)-m(S,\sigma,t))^2\right]+E\left[(\Pi(S,\sigma,t)-m(S,\sigma,t))\,(d\Pi-dm)\right]\\
&\qquad + E\left[(d\Pi-dm)^2\right].
\end{aligned}
$$

I've been a bit loose with my notation here, and I've also neglected any discounting. I'll put the latter back in a moment.

The middle term on the right-hand side of the above is zero from the definition of m. The third term becomes

$$
E\left[\left(-\sigma S\Delta dX_1-\frac{\partial m}{\partial\sigma}q\,dX_2-\frac{\partial m}{\partial S}\sigma S\,dX_1\right)^2\right].
$$

I'll leave this for you to simplify, but the end result is, for an arbitrary Δ,

$$
\begin{aligned}
&\frac{\partial v}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 v}{\partial S^2}+\rho\sigma Sq\frac{\partial^2 v}{\partial S\,\partial\sigma}+\tfrac{1}{2}q^2\frac{\partial^2 v}{\partial\sigma^2}+\mu S\frac{\partial v}{\partial S}+p\frac{\partial v}{\partial\sigma}-2rv\\
&\quad+\sigma^2S^2\left(\frac{\partial m}{\partial S}\right)^2+q^2\left(\frac{\partial m}{\partial\sigma}\right)^2+2\rho\sigma Sq\frac{\partial m}{\partial S}\frac{\partial m}{\partial\sigma}\\
&\quad+\sigma^2S^2\Delta^2+2\Delta\left(\sigma^2S^2\frac{\partial m}{\partial S}+\rho\sigma Sq\frac{\partial m}{\partial\sigma}\right)=0. \qquad (54.2)
\end{aligned}
$$

Observe that I've put the discounting back in.

54.5 CHOOSING Δ TO MINIMIZE THE VARIANCE

Only the last two terms in (54.2) contain Δ. We therefore choose Δ to minimize this quantity, to ensure that the variance in our portfolio is as small as possible. This gives

$$
\Delta=-\frac{\partial m}{\partial S}-\frac{\rho q}{\sigma S}\frac{\partial m}{\partial\sigma}. \qquad (54.3)
$$

54.6 THE MEAN AND VARIANCE EQUATIONS

Substituting (54.3) into (54.1) and (54.2) we get

$$
\frac{\partial m}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 m}{\partial S^2}+\rho\sigma Sq\frac{\partial^2 m}{\partial S\partial\sigma}+\tfrac{1}{2}q^2\frac{\partial^2 m}{\partial\sigma^2}+rS\frac{\partial m}{\partial S}+\left(p-(\mu-r)\frac{\rho q}{\sigma}\right)\frac{\partial m}{\partial\sigma}-rm=0 \qquad (54.4)
$$

and

$$\frac{\partial v}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 v}{\partial S^2} + \rho\sigma S q \frac{\partial^2 v}{\partial S \partial \sigma} + \tfrac{1}{2} q^2 \frac{\partial^2 v}{\partial \sigma^2}$$
$$+ \mu S \frac{\partial v}{\partial S} + p \frac{\partial v}{\partial \sigma} + q^2(1-\rho^2)\left(\frac{\partial m}{\partial \sigma}\right)^2 - 2rv = 0. \quad (54.5)$$

The final conditions for these are obviously the payoff, for $m(S,\sigma,T)$, and zero for $v(S,\sigma,T)$. If the portfolio contains options with different maturities, the equations must satisfy the corresponding jump conditions as well.

Since the final condition for v is zero and the only 'forcing term' in (54.5) is $(\partial m/\partial\sigma)^2$, equation (54.5) shows that the only way we can have a perfect hedge is for either q to be zero, i.e. deterministic volatility, or to have $\rho = \pm 1$. In the latter case the asset and volatility (changes) are perfectly correlated. The solution of (54.4) is then different from the Black–Scholes solution.

Equation (54.4) is very much like the pricing equation for stochastic volatility in a risk-neutral setting. It's rather like having a market price of volatility risk of $(\mu - r)\rho/\sigma$. But, of course, the reasoning and model are completely different in our case.

The system of equations is non linear (actually two linear equations, coupled by a non-linear forcing term). We are going to exploit this fact shortly.

54.7 HOW TO INTERPRET AND USE THE MEAN AND VARIANCE

Take an option position in a world with stochastic volatility, and delta hedge as proposed above. Because we cannot eliminate all the risk we cannot be certain how accurate our hedging will be. Think of the final value of the portfolio together with accumulated hedging as being the 'outcome.' The distribution of the outcome will generally not be Normal. The shape will depend very much on the option position we are hedging. But we have calculated both the mean and the variance of the hedged portfolio. If we made the assumption that the distribution was not too far from Normal then the mean and the variance are sufficient to describe the probabilities of any outcome. If we wanted to be 95% certain that we would make money then we would have to sell the option for

$$m + 1.644853 v^{1/2}$$

or buy it for

$$m - 1.644853 v^{1/2}.$$

The 1.644853 comes from the position of the 95th percentile assuming a Normal distribution.

We'll use this idea below, but with a requirement that we are within one standard deviation of the mean, i.e. we make money 84% of the time.

More generally we would price at

$$m \pm \xi\, v^{1/2},$$

where the ξ is a personal choice.

Clearly the larger ξ the greater the potential for profit from a single trade (see Figure 54.1).

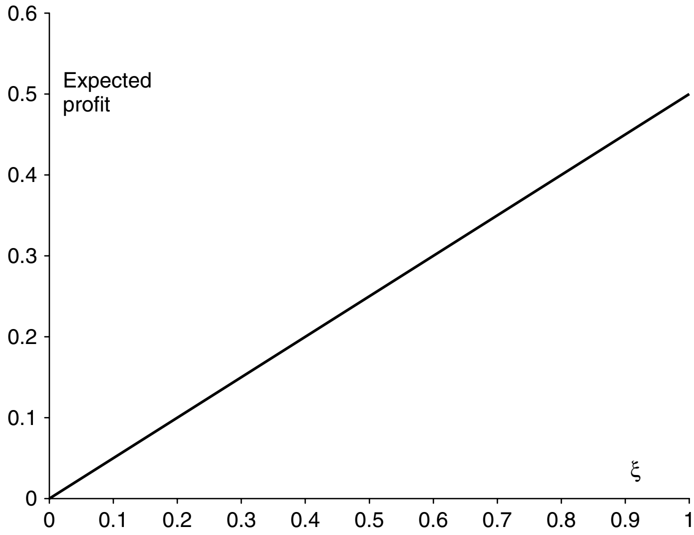


Figure 54.1 Expected profit from a single trade versus ξ.

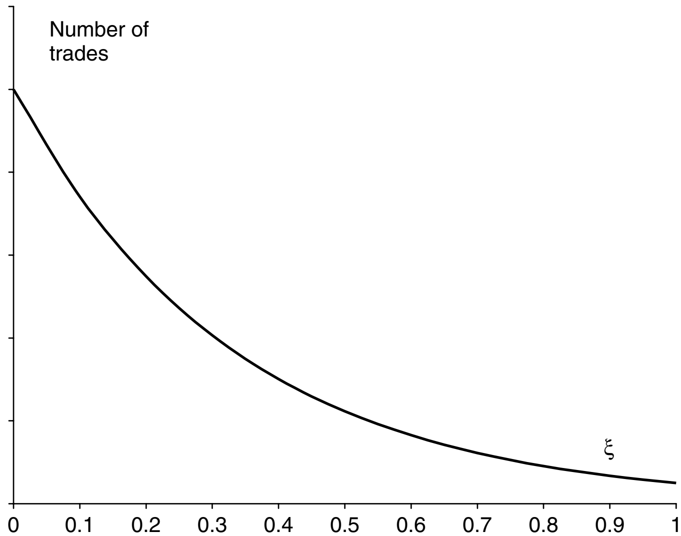


Figure 54.2 Number of trades versus ξ.

However, the larger ξ the fewer trades (see Figure 54.2).

The net result is that the total profit potential, being a product of the number of trades and the profit from each trade, is of the form shown in Figure 54.3. Don't be too greedy or too generous.

We'll use this idea in the example below, but we will insist that we are within one standard deviation of the mean so that $\xi = 1$. This is simply so that we have fewer parameters to carry around.

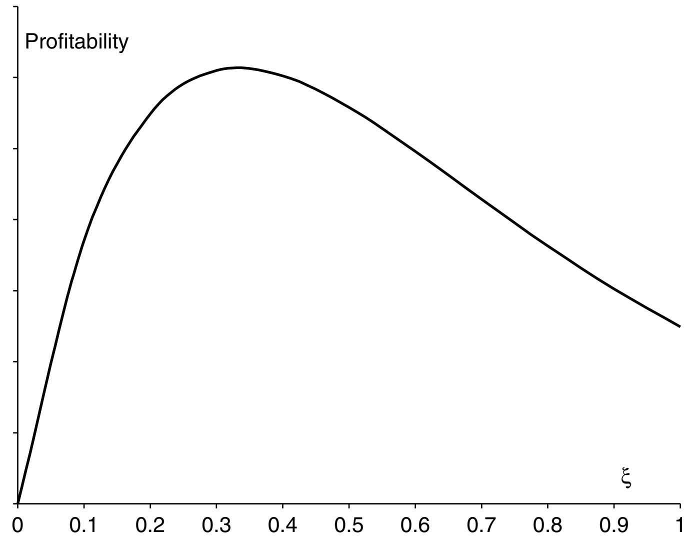


Figure 54.3 Total profit potential versus ξ.

54.8 STATIC HEDGING AND PORTFOLIO OPTIMIZATION

If we use as our option (portfolio) 'price' the following

$$\text{mean} - (\text{variance})^{1/2} = m - v^{1/2}$$

then we have a non-linear model. Everything that will be said in Chapter 60 about non-linear pricing models applies here, in particular the possibility of optimal static hedging.

54.9 EXAMPLE: VALUING AND HEDGING AN UP-AND-OUT CALL

In this section, we consider the pricing and hedging of a short up-and-out call. Throughout this section, our choice of mean-variance combination is:

$$m - v^{1/2}. \tag{54.6}$$

First consider a single up-and-out call with barrier located at S_u. In this case, we solve the equations (54.4) and (54.5) subject to:

(a) $m(S_u, \sigma, t) = v(S_u, \sigma, t) = 0$ for each $(\sigma, t) \in (0, \infty) \times (0, T)$ where T is maturity;
(b) $m(S, \sigma, T) = -\max(S - E, 0)$ for each $(S, \sigma) \in (0, X) \times (0, \infty)$ where E is the strike;
(c) $v(S, \sigma, T) = 0$ for each (S, σ).

The discontinuity of the payoff at the knock-out barrier makes this position particularly difficult to hedge. In fact this can be easily seen from our equations. Figures 54.4 and 54.5 are the pictures of calculated mean and variance respectively with the following specifications:

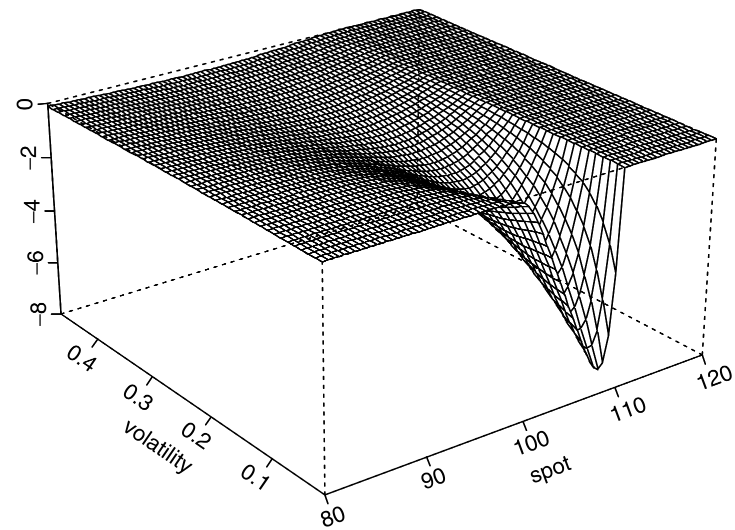


Figure 54.4 Mean for a single up-and-out call.

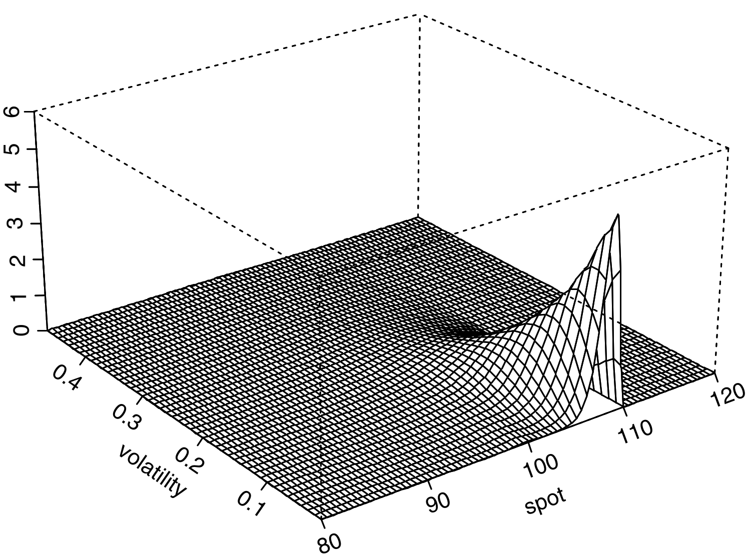


Figure 54.5 Variance for a single up-and-out call.

- Strike at 100, barrier at 110, and expiry 30 days;
- $p(\sigma) = 0.8(\sigma^{-1} - 0.2)$ and $q(\sigma) = 0.5$.

Near the barrier, $(\partial m/\partial\sigma)^2$ is huge (see Figure 54.4) and this feeds the variance, being the source term in (54.5). If the spot S is 100, and the current spot volatility σ is 20% per annum, the mean is -1.1101 and the variance is 0.3290. Thus if there is no other instrument available in the market, one would price this option at \$1.6836 to match with Equation (54.6).

These results are shown in the following table:

	Mean (m)	Var. (v)	Value
Unhedged	-1.1101	0.329	1.6836

54.10 STATIC HEDGING

Suppose that there are six 30-day vanilla call options available in the market with the following specifications:

Option	1	2	3	4	5	6
Strike	96.62	100.00	104.17	108.70	112.36	116.96
Bid Price	4.6186	2.6774	1.1895	0.4302	0.1770	0.0557
Ask Price	4.6650	2.7043	1.2014	0.4345	0.1788	0.0562

Aside

These hypothetical market prices were generated by computing the mean of each call option, with

$$d\sigma = \left(\frac{1}{\sigma} - 0.2\right)dt + 0.5\,dX \qquad (54.7)$$

where X is a Brownian motion with respect to the risk-neutral measure. Then 0.5% bid-ask spread was added.

How can we use these vanillas or the information contained within these prices? We are going to see how to incorporate quantities of the vanillas into a portfolio along with the barrier option to construct a portfolio that has a better 'value' than the barrier on its own, unhedged as above. Here 'value' means the theoretical value of the portfolio under this non-linear model after the cost of the static hedge, the vanillas, has been subtracted off. We saw this idea in Chapter 52 and will see it again in Chapter 60.

Suppose we trade $(q_1, \ldots, q_6)$ of the above instruments and let E_i be the strikes among the payoffs. Furthermore, let $(m^{(0)}, v^{(0)})$ be the mean variance pair *after* knock out and $(m^{(1)}, v^{(1)})$ be that *before* knock out. Then $(m^{(i)}, v^{(i)})$, $i = 0, 1$, satisfy the equations (54.4) and (54.5) subject to:

(a) $m^{(1)}(110, \sigma, t) = m^{(0)}(110, \sigma, t)$ and $v^{(1)}(110, \sigma, t) = v^{(0)}(110, \sigma, t)$ for each (σ, t) in $(0, \infty) \times [0, T)$;

(b) $m^{(0)}(S, \sigma, T) = \sum_{i=1}^{6} q_i \max(S - E_i, 0)$ for each $(S, \sigma) \in (0, \infty) \times (0, \infty)$;

(c) $m^{(1)}(S, \sigma, T) = \sum_{i=1}^{6} q_i \max(S - E_i, 0) - \max(S - 100, 0)$ for each (S, σ) in $(0, 110) \times (0, \infty)$;

(d) $v^{(1)}(S, \sigma, T) = v^{(0)}(S, \sigma, T) = 0$ for each (S, σ) in $(0, \infty) \times (0, \infty)$.

Thus $m^{(1)}(S, \sigma, 0)$ stands for the mean of the cashflows excluding the upfront premium.

We then find a $(q_1, \ldots, q_6)$ that maximizes

$$m^{(1)}(S, \sigma, 0) - \sum_{i=1}^{6} p(q_i) - \sqrt{v^{(1)}(S, \sigma, 0)}$$

where $p(q_i)$ is the market price of trading q_i shares of strike E_i, allowing for bid-offer prices.

We find that in the case of $S = 100$ and $\sigma = 0.2$, our optimal choice for this 'vega hedge' is given by:

Option	1	2	3	4	5	6
Strike	96.62	100.00	104.17	108.70	112.36	116.96
Quantity	0.0000	−1.1688	1.0207	3.1674	−3.6186	0.8035

The cost of this hedge position is \$1.1863. Figures 54.6 and 54.7 are the pictures of $m^{(1)}$ and $v^{(1)}$ after the optimal static vega hedge. After the optimal static vega hedge, the mean is 0.0398 and the variance is reduced to 0.0522. Thus the price for the up-and-out call that matches with our mean-variance combination (54.6) is \$1.3752 $(1.1863 - 0.0398 + \sqrt{0.0522})$. In the risk-neutral set-up (54.7), the price for this up-and-out call is \$1.1256. The difference mainly comes from the standard deviation term (variance$^{1/2}$) in (54.6) which is $\sqrt{0.0522} = 0.2286$.

These results are shown in the following table.

	Mean (m)	Var. (v)	Hedge	Value
Unhedged	−1.1101	0.329		1.6836
Hedged	0.0398	0.0522	1.1863	1.3752

By statically hedging we have reduced the price at which we can safely sell the option, from 1.6836 to 1.3752, while still making money 84% of the time. Alternatively, we can still sell the option for 1.6836 and make even more profit.

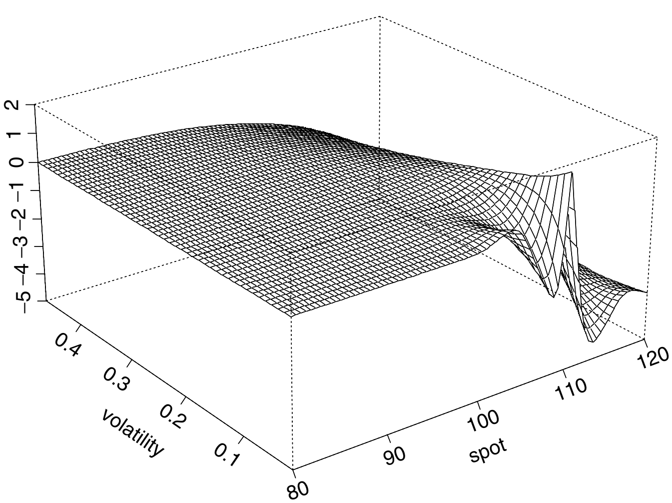


Figure 54.6 Mean of portfolio after optimal static vega hedging.

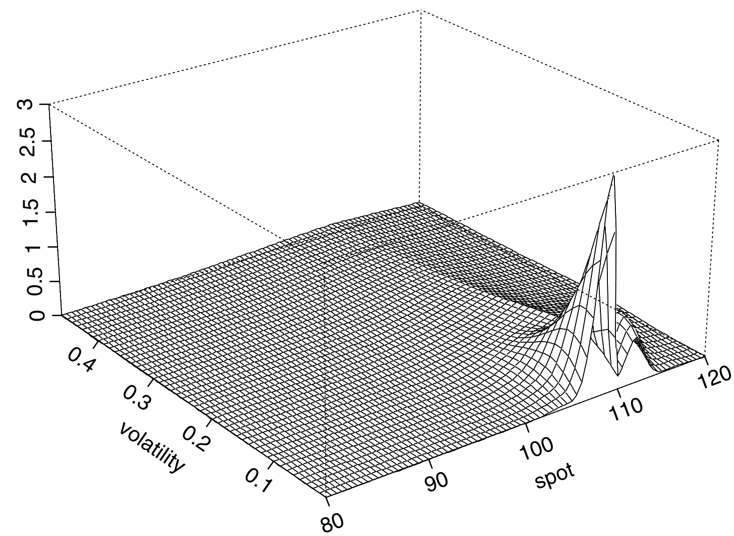


Figure 54.7 Variance of portfolio after optimal static vega hedging.

54.11 OTHER DEFINITIONS OF 'VALUE'

In the above example we have statically hedged so as to find the best value according to our definition of value. This is by no means the only static hedging strategy. One can readily imagine different players having different criteria.

Obvious strategies that spring to mind are as follows.

- Minimize variance, that is minimize the function v. This has the effect of reducing model risk as much as possible using all available instruments (the underlying and all traded options). This may be a strategy adopted by the sell side.
- Maximize the return-risk ratio. This is perhaps more of a buy-side strategy, for maximizing Sharpe ratio, for example.

54.12 SUMMARY

Constructing a risk-neutral model to fit the market prices of exchange traded options consistently over a reasonable time period is a difficult task. Putting aside the fundamental question of whether the axiomatic risk-neutral model for stochastic volatility is legitimate or not, we must face potential financial losses due to re-calibration. In this chapter we take another approach. We first evaluate the mean and variance of the discounted future cashflow and then find market instruments that reduce the volatility risk optimally.

I've set this problem up in a mean-variance framework but it could easily be extended to a more general utility theory approach.

FURTHER READING

- For all the details of the above model see Ahn, Arkell, Choe, Holstad & Wilmott (1999).
- The above was published in *Wilmott* magazine, Ahn & Wilmott (2003).

CHAPTER 55
asymptotic analysis of volatility

In this Chapter...

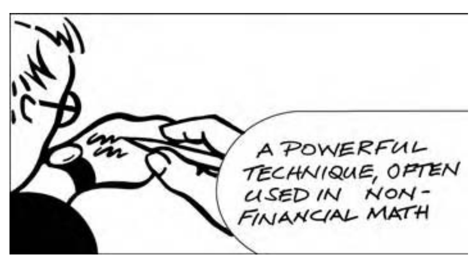
- asymptotic analysis and small or large parameters
- a series solution for implied volatility under arbitrary stochastic volatility

55.1 INTRODUCTION

Asymptotic analysis is a systematic method for exploiting the largeness or smallness of a parameter in some equation. In our context the equation is that for an option under stochastic volatility and the parameters measure the speed of mean reversion of volatility and the volatility of volatility. The mean reversion is fast (large parameter) and the volatility of volatility is large (another large parameter). Since we'll be looking for asymptotic solutions we'll see how the precise specification of the model is irrelevant as far as finding closed-form solutions is concerned. In other words, we don't have to sacrifice accuracy for tractability anymore.

This chapter is heavily based on a paper by myself and Henrik Rasmussen, Rasmussen & Wilmott (2002).

55.2 FAST MEAN REVERSION AND HIGH VOLATILITY OF VOLATILITY



We consider the pricing of options when the underlying asset value S and its volatility σ are described by the stochastic differential equations,

$$dS/S = r\,dt + \sigma\,dX$$

$$d\sigma = A\,dt + B\,dY$$

$$dX \cdot dY = \rho\,dt$$

where X and Y are Brownian motions, r is the short rate, and the coefficients A and B are functions of only σ. When calibrating such a stochastic volatility model to market prices, one usually finds that the volatility of volatility B/σ is greater than the volatility σ of the underlying. For instance (Wiggins, 1987),

$$\sigma \propto 0.2$$

$$B \propto 0.2,$$

in which case the ratio between volatility and the volatility of volatility becomes

$$\frac{B}{\sigma^2} \propto 5.$$

Accordingly, we introduce a small parameter ϵ

$$B = \frac{\beta}{\sqrt{\epsilon}},$$

such that (for typical values of σ)

$$\frac{\beta}{\sigma^2} \propto 1.0$$

which implies

$$\epsilon \propto 0.04.$$

The parameter ϵ can be interpreted as a ratio of time scales for fluctuations in volatility σ and in the asset price S. If T is a characteristic timescale for S, then $\epsilon\, T$ is a characteristic timescale for the volatility process σ.

Note that we are writing large parameters as inverses of a small parameter. That's because we will shortly be looking for a series expansion in powers of ϵ.

Since we would like to think of ϵ as a ratio of characteristic timescales, we let the drift term A scale like

$$A = \alpha/\epsilon,$$

and thereby arrive at the equations

$$\begin{aligned} dS/S &= r\, dt + \sigma\, dX \\ d\sigma &= \frac{\alpha}{\epsilon}\, dt + \frac{\beta}{\sqrt{\epsilon}}\, dY \\ dX \cdot dY &= \rho\, dt. \end{aligned}$$

Let V denote the value of a European option on S. It satisfies

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} + r\, S \frac{\partial V}{\partial S} - r\, V + \rho \frac{S\, \sigma\, \beta}{\sqrt{\epsilon}} \frac{\partial^2 V}{\partial S\, \partial \sigma} + \frac{\beta^2}{2\, \epsilon} \frac{\partial^2 V}{\partial \sigma^2} + \frac{\alpha}{\epsilon} \frac{\partial V}{\partial \sigma} = 0. \tag{55.1}$$

In the following, we first seek a regular perturbation series solution of this equation up to and including order ϵ^2,

$$V(S, \sigma, t) \sim \sum_{n=0}^{N} \epsilon^{n/2}\, V_n(S, \sigma, t), \qquad N = 4.$$

We next estimate the implied volatility for a vanilla option (these are used for calibration). Our approximation to the implied volatility is a fifth-order polynomial in moneyness with coefficients depending on the time to expiry and on the current volatility.

These results extend those of Fouque *et al.* (1998–2000) and Sircar & Papanicolaou (1998) in being valid for many more models and in going to higher orders in ϵ (where model differences become important). The motivation for going to higher orders in ϵ is to better fit the 'smiles' observed in the market. Specifically, Fouque *et al.* (2000) solve the equation to order $\epsilon^{1/2}$ for the case of a volatility being a deterministic function of an Ornstein–Uhlenbeck process. But the corresponding approximation to the implied volatility only grows linearly with log-moneyness (log of forward over strike), whereas the implied volatility observed in the market often has significant curvature with respect to log-moneyness (the 'smile'). To capture curvature in the implied volatility, we need an approximation that is at least quadratic in log-moneyness.

55.3 CONDITIONS ON THE MODELS

The volatility process should be autonomous in the sense that α and β are functions of only σ,

$$\alpha = \alpha(\sigma)$$
$$\beta = \beta(\sigma)$$

In addition, the risk-neutral probability density of volatility should have a stationary limit and this limit should be well-behaved. Let p denote the stationary limit of the risk-neutral probability density of volatility. It is a function of σ only and it satisfies (if it exists) the stationary version of the Kolmogorov forward equation for the volatility process,

$$\tfrac{1}{2}\frac{\partial^2}{\partial\sigma^2}\left[\beta^2 p\right] - \frac{\partial}{\partial\sigma}\left[\alpha p\right] = 0,$$

A stationary density exists if the solution of this equation is normalizable,

$$p(\sigma) = \frac{C}{\beta^2(\sigma)} \exp\left\{\int_{\sigma_0}^{\sigma} 2\,\alpha(x)/\beta^2(x)\,dx\right\} \in L^1$$

Here, C is a constant needed for normalization,

$$\|p\|_1 = 1.$$

This density p should be well-behaved in the sense that certain integrals must be finite. In particular,

$$\hat{\sigma}^2 = \int_0^\infty \sigma^2\, p(\sigma)\, d\sigma < \infty,$$

i.e., the risk-neutral volatility process has a finite root-mean-square value with respect to the stationary density. When this condition is satisfied, $\hat{\sigma}$ becomes an effective long-term volatility, around which the real volatility σ fluctuates with a characteristic timescale of order ϵ.

55.4 EXAMPLES OF MODELS

Many of the common models satisfy the above conditions. An important exception is the lognormal model (without mean reversion), for which no stationary probability density exists. The following table (where the C_i are constants) gives some examples of stochastic volatility models.

Author(s)	Model
Hull & White, Zhu & Avellaneda	$d\sigma^2 = C_1\,\sigma^2\,dt + C_2\,\sigma^2\,dY$
Scott, Taylor & Xu	$d\log\sigma = C_1\,(C_2 - \log\sigma)\,dt + C_3\,dY$
Stein & Stein	$\sigma = \lvert Z\rvert$, where $dZ = C_1\,(C_2 - Z)\,dt + C_3\,dY$
Ball & Roma, Heston	$d\sigma^2 = C_1\,(C_2 - \sigma^2)\,dt + C_3\,\sigma\,dY$
Ravanelli	$d\sigma^2 = C_1\,(C_2 - \sigma^2)\,dt + C_3\,\sigma^2\,dY$

55.4.1 Scott's Model

In this model, $\log(\sigma_t)$ is a mean-reverting Ornstein–Uhlenbeck process,

$$d\log(\sigma) = C_1\,(C_2 - \log(\sigma))\,dt + C_3\,dY$$

(assuming volatility is non-dimensionalized). A stationary density exists for $C_1 > 0$ and is given by

$$p(\sigma) = C\,C_3^2\,\sigma^{-1+2\,C_1C_2/C_3^2}\,\exp\bigl[-C_1\,\log^2(\sigma)/C_3^2\,\bigr].$$

$C_1 > 0$ ensures that the log of volatility is mean reverting, at rate C_1 and towards the level C_2/C_1.

55.4.2 The Heston/Ball–Roma Model

This model is given by

$$d\sigma^2 = C_1\,(C_2 - \sigma^2)\,dt + C_3\,\sigma\,dY$$

The stationary density is given by

$$p(\sigma) = \frac{4\,C}{C_3^2}\,\sigma^{4C_1(C_2-C_3^2/4)/C_3^2}\,\exp(-2\,C_1\,\sigma^2/C_3^2),$$

and it exists for $C_1 > 0$ and $C_2 - C_3^2/4 > 0$. The Heston model is perhaps the most popular model in applications, since it allows semi-closed-form solutions for option prices.

55.5 NOTATION

To simplify the notation, we introduce the differential operators,

$$
\begin{aligned}
D_0^\sigma &= \frac{\partial}{\partial t} + \tfrac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2}{\partial S^2} + r\, S\, \frac{\partial}{\partial S} - r \\
D_1 &= S\, \frac{\partial}{\partial S} \\
D_2 &= \sigma\, \frac{\partial}{\partial \sigma} \\
D_3 &= \tfrac{1}{2}\,\beta^2\, \frac{\partial^2}{\partial \sigma^2} + \alpha\, \frac{\partial}{\partial \sigma} \\
D_4 &= S^2\, \frac{\partial^2}{\partial S^2}.
\end{aligned}
$$

It is easily shown that

$$D_4\, D_0^\sigma = D_0^\sigma\, D_4 \tag{55.2}$$

$$D_4\, D_1 = D_1\, D_4 \tag{55.3}$$

$$D_0^\sigma = D_0^{\hat\sigma} + \tfrac{1}{2}\left(\sigma^2 - \hat\sigma^2\right) D_4 \tag{55.4}$$

$$D_3^{-1}(f) = -\int_\sigma^\infty \frac{1}{\beta^2(y)\, p(y)} \int_0^y f(x)\, p(x)\, dx\, dy. \tag{55.5}$$

To prove Equations (55.2) and (55.3), we simply note that

$$x = \log(S/S_0),$$

where S_0 is a constant with the same dimension as S, transforms D_0^σ and D_4 into operators with constant coefficients. Equation (55.4) follows immediately from the definitions, while Equation (55.5) can be proved by first applying D_3 on both sides, and then simplifying using the definition of the stationary density p.

Finally, we define

$$
\begin{aligned}
F_1(\sigma) &= \tfrac{1}{2} \int_\sigma^\infty \frac{1}{\beta^2(y)\, p(y)} \int_0^y p(x)\,(\,\hat\sigma^2 - x^2\,)\, dx\, dy \\
\gamma &= \int_0^\infty p(x)\, \beta(x)\, D_2(F_1)[x]\, dx \\
F_2(\sigma) &= \int_\sigma^\infty \frac{1}{\beta^2(y)\, p(y)} \int_0^y p(x)\left(\gamma - \beta(x)\right) D_2(F_1)[x]\, dx\, dy \\
F_3(\sigma) &= \tfrac{1}{2} \int_\sigma^\infty \frac{1}{\beta^2(y)\, p(y)} \int_0^y p(x)\,(\,\hat\sigma^2 - x^2\,)\, F_1(x)\, dx\, dy \\
F_4(\sigma) &= \int_\sigma^\infty \frac{1}{\beta^2(y)\, p(y)} \int_0^y p(x)\, \beta(x)\, D_2[F_1](x)\, dx\, dy \\
F_5(\sigma) &= \int_\sigma^\infty \frac{1}{\beta^2(y)\, p(y)} \int_0^y p(x)\, \beta(x)\, F_2(x)\, dx\, dy.
\end{aligned}
$$

These functions are all straightforward to evaluate, either in closed form, or by analytical or numerical approximation.

55.6 ASYMPTOTIC ANALYSIS

Using the above notation, Equation (55.1) becomes

$$\epsilon\, D_0^{\sigma}(V) + \sqrt{\epsilon}\,\rho\,\beta\, D_1 D_2(V) + D_3(V) = 0. \tag{55.6}$$

Inserting the perturbation series Equation (55.2) in to Equation (55.6), we find that

$$D_3(V_0) = 0 \tag{55.7}$$

$$\rho\,\beta\, D_1 D_2(V_0) + D_3(V_1) = 0 \tag{55.8}$$

$$\forall n \geq 2: \qquad D_0^{\sigma}(V_{n-2}) + \rho\,\beta\, D_1 D_2(V_{n-1}) + D_3(V_n) = 0. \tag{55.9}$$

To satisfy the expiry condition at time $t = T$, we need the expiry conditions

$$V_0(S, \sigma, T) = V(S, T)$$

$$V_n(S, \sigma, T) = 0 \qquad \text{for} \quad n \geq 1.$$

The solution has a thin expiry layer for t near T. However, we are only interested in the solution outside this expiry layer, so we need only impose the expiry condition for $n = 0$.

Equations (55.7) and (55.8) together imply that V_0 and V_1 are independent of σ. It then follows from the definition of D_2 above that the second term in Equation (55.9) vanishes for $n = 2$, so that

$$D_0^{\sigma}(V_0) + D_3(V_2) = 0. \tag{55.10}$$

After multiplying by the stationary density $p(\sigma)$ and then integrating with respect to σ over $[0, \infty]$, the contribution from the second term is going to be zero. To see this, integrate by parts and note that p satisfies the Kolmogorov forward equation. Since V_0 is independent of σ, we obtain

$$D_0^{\hat{\sigma}}(V_0) = 0. \tag{55.11}$$

This is just the Black–Scholes equation with volatility equal to the root-mean-square volatility $\hat{\sigma}$,

$$\frac{\partial V_0}{\partial t} + \tfrac{1}{2}\,\hat{\sigma}^2\, S^2\, \frac{\partial^2 V_0}{\partial S^2} + r\, S\, \frac{\partial V_0}{\partial S} - r\, V_0 = 0,$$

so that

$$V_0(S, t) = V_{BS}(S, \hat{\sigma}, t), \tag{55.12}$$

where $V_{BS}(S, \sigma, t)$ denotes a Black–Scholes price at volatility σ. The expiry condition for V_0 is the original expiry condition $V(S, T)$.

Since V_0 is known, we may solve for V_2 in Equation (55.10). We first subtract Equation (55.11) from Equation (55.10), and then invert D_3,

$$V_2 = \tfrac{1}{2}\left[\, D_3^{-1}(\,\hat{\sigma}^2 - \sigma^2\,)\,\right] D_4(V_B S).$$

By definition of F_1,

$$V_2(S, \sigma, t) = F_1(\sigma)\, D_4(V_{BS}). \tag{55.13}$$

Note that V_2 does not satisfy the expiry condition $V_2 = 0$ at $t = T$. The reason is that V has a thin expiry layer, and our series in Equation (55.2) converges only outside this layer.

Consider now Equation (55.9) for $n = 3$,

$$D_0^{\sigma}(V_1) + \rho\,\beta\, D_1 D_2(V_2) + D_3(V_3) = 0. \tag{55.14}$$

As before, we multiply by p and integrate with respect to σ. Since V_1 too is independent of σ,

$$D_0^{\hat{\sigma}}(V_1) + \rho \int_0^{\infty} p(\sigma)\,\beta(\sigma)\, D_1 D_2(V_2)[\sigma]\, d\sigma = 0.$$

It then follows from Equation (55.13) that

$$D_0^{\hat{\sigma}}(V_1) + \rho \left\{ \int_0^{\infty} p(\sigma)\,\beta(\sigma)\, D_2(F_1)[\sigma]\, d\sigma \right\} D_1 D_4(V_{BS}) = 0.$$

By the definition of γ,

$$D_0^{\hat{\sigma}}(V_1) + \rho\,\gamma\, D_1 D_4(V_{BS}). = 0. \tag{55.15}$$

To solve this equation, we first note that $D_1(f)$ and $D_4(f)$ are both solutions of the Black–Scholes equation if f is a solution. Therefore, since V_{BS} satisfies the Black–Scholes equation with volatility $\hat{\sigma}$,

$$D_0^{\hat{\sigma}} D_1 D_4(V_{BS}) = 0.$$

Using this fact, it is easily checked that

$$V_1(S, t) = \rho\,\gamma\,(T - t)\, D_1 D_4(V_{BS}) \tag{55.16}$$

is the solution of Equation (55.15). In the next section, we determine $D_1 D_4(V_{BS})$ explicitly for the case of a European vanilla option, and we shall see that the expiry condition $V_1 = 0$ is satisfied. In contrast to V_2, V_0 and V_1 do not depend on the current volatility σ, but only on current stock price S and date t. Note also that, to the leading order, the deviation from a Black–Scholes price increases linearly with the correlation coefficient ρ.

To get an even better approximation, we calculate two more terms in Equation (55.2). The method is the same as before. To determine V_3, subtract Equation (55.15) from Equation (55.14),

$$D_0^{\sigma}(V_1) - D_0^{\hat{\sigma}}(V_1) + \rho\,\beta\, D_1 D_2(V_2) - \rho\,\gamma\, D_1 D_4(V_{BS}) + D_3(V_3) = 0. \tag{55.17}$$

Using Equation (55.4) for the first and second term, and Equation (55.13) to simplify the third term, we obtain

$$\tfrac{1}{2}\,(\,\sigma^2 - \hat{\sigma}^2\,)\, D_4(V_1) + \rho\,\big(\beta\, D_2(F_1) - \gamma\big)\, D_1 D_4(V_{BS}) + D_3(V_3) = 0. \tag{55.18}$$

Insert now V_1 from Equation (55.16) and invert D_3,

$$V_3(S,\sigma,t) = \rho\,\gamma\,(T-t)\,F_1(\sigma)\,D_1 D_4^2(V_{BS}) + \rho\,F_2(\sigma)\,D_1 D_4(V_{BS}). \tag{55.19}$$

Here we used Equation (55.3) to change the order of D_1 and D_4 and simplified the inverse of D_3 by means of the above definitions of F_1 and F_2. We see that V_3, like V_1, vanishes when $\rho = 0$. In fact, it is true in general that odd order terms vanish for $\rho = 0$. Finally, by a similar argument as used for V_3, we find that

$$V_4(S,\sigma,t) = F_3(\sigma)\,D_4^2(V_{BS}) + \rho^2\,\gamma\,(T-t)\,F_4(\sigma)\,D_1^2 D_4^2(V_{BS}) - \rho^2\,F_5(\sigma)\,D_1^2 D_4(V_{BS}). \tag{55.20}$$

In this section, we have shown that

$$\begin{aligned}
V(S,\sigma,t) = {} & V_{BS}(S,\overline{\sigma},t)\\
& + \epsilon^{1/2}\,\rho\,\gamma\,(T-t)\,D_1 D_4(V_{BS})\\
& + \epsilon\,F_1(\sigma)\,D_4(V_{BS})\\
& + \epsilon^{3/2}\,\rho\left\{\gamma\,(T-t)\,F_1(\sigma)\,D_1 D_4^2(V_{BS}) + F_2(\sigma)\,D_1 D_4(V_{BS})\right\}\\
& + \epsilon^2\left\{F_3(\sigma)\,D_4^2(V_{BS}) + \rho^2\,\gamma\,(T-t)\,F_4(\sigma)\,D_1^2 D_4^2(V_{BS})\right.\\
& \qquad \left. - \rho^2\,F_5(\sigma)\,D_1^2 D_4(V_{BS})\right\}\\
& + O(\epsilon^{5/2}).
\end{aligned}$$

Note that the first term is a Black–Scholes value with volatility equal to the risk-neutral root-mean-square (rms) $\overline{\sigma}$ of volatility. This rms value is also the effective volatility for the asset price S in the limit of an infinitesimal characteristic timescale for the volatility process. See Skorokhod (1987) and Khasminskii & Yin (1996) for discussions of 'homogenization' of a slow process in the presence of a fast process.

55.7 VANILLA OPTIONS: ASYMPTOTICS FOR VALUES

Stochastic volatility models are often calibrated to vanilla options, so let us consider these in detail, i.e. a European call or put.

Since gamma of such options is given by

$$\begin{aligned}
\Gamma(S,\hat{\sigma},t) &= \frac{1}{S\,\hat{\sigma}\,\sqrt{T-t}}\,n(d_1)\\
d_1 &= \frac{\log(S/K) + (r + \hat{\sigma}^2/2)(T-t)}{\hat{\sigma}\,\sqrt{T-t}}\\
n(x) &= \frac{1}{\sqrt{2\pi}}\,e^{-x^2/2},
\end{aligned}$$

we get

$$D_4(V_{BS}) = \frac{S}{\hat{\sigma}\sqrt{T-t}}\, n(d_1)$$

where V_{BS} is the Black–Scholes price with volatility $\hat{\sigma}$. To simplify the notation, let

$$\begin{aligned} x &= d_2/\sqrt{2} \\ d(t) &= \frac{K}{\sqrt{2\pi}} \exp\left[-r\,(T-t) + \hat{\sigma}^2\,(T-t)/2\right] \\ \tau &= 2\,\hat{\sigma}^2\,(T-t), \end{aligned}$$

where

$$d_2 = \frac{\log(S/K) + (r - \hat{\sigma}^2/2)(T-t)}{\hat{\sigma}\sqrt{T-t}}.$$

Then

$$D_4(V_{BS}) = d(t)\, e^{-x^2}$$

and

$$\begin{aligned} D_1(f) &= \tau^{-1/2}\,\frac{\partial f}{\partial x} \\ D_1^2(f) &= \tau^{-1}\,\frac{\partial^2 f}{\partial x^2} \\ D_4(f) &= \tau^{-1}\,\frac{\partial^2 f}{\partial x^2} - \tau^{-1/2}\,\frac{\partial f}{\partial x} \\ D_1 D_4(f) &= \tau^{-3/2}\,\frac{\partial^3 f}{\partial x^3} - \tau^{-1}\,\frac{\partial^2 f}{\partial x^2} \\ D_1^2 D_4(f) &= \tau^{-2}\,\frac{\partial^4 f}{\partial x^4} - \tau^{-3/2}\,\frac{\partial^3 f}{\partial x^3} \end{aligned}$$

Using

$$\frac{d^n}{dx^n} e^{-x^2} = (-1)^n\, H_n(x)\, e^{-x^2},$$

where H_n is the nth Hermite polynomial, and the fact that

$$\begin{aligned} H_0 &= 1 \\ H_1 &= 2\,x \\ H_2 &= 4\,x^2 - 2 \\ H_3 &= 8\,x^3 - 12\,x \\ H_4 &= 16\,x^4 - 48\,x^2 + 12, \end{aligned}$$

we get

$$D_1 D_4(V_{BS}) = -\tau^{-1/2} H_1(x) D_4(V_{BS})$$
$$D_1^2 D_4(V_{BS}) = -\tau^{-1} H_2(x) D_4(V_{BS})$$
$$D_4^2(V_{BS}) = \tau^{-1} \left[H_2(x) + \tau^{1/2} H_1(x) \right] D_4(V_{BS})$$
$$D_1 D_4^2(V_{BS}) = -\tau^{-3/2} \left[H_3(x) + \tau^{1/2} H_2(x) \right] D_4(V_{BS})$$
$$D_1^2 D_4^2(V_{BS}) = \tau^{-2} \left[H_4(x) + \tau^{1/2} H_3(x) \right] D_4(V_{BS}).$$

Then, inserting into the formulae from the previous section, we obtain the correction terms to the Black–Scholes prices V_0,

$$V_1(S,\sigma,t) = -\tfrac{1}{2}\rho\gamma\hat{\sigma}^{-2}\tau^{1/2} H_1(x) S^2 \Gamma$$
$$V_2(S,\sigma,t) = F_1(\sigma) S^2 \Gamma$$
$$V_3(S,\sigma,t) = -\tfrac{1}{2}\rho\gamma\hat{\sigma}^{-2}\tau^{-1/2} F_1(\sigma)\left[H_3(x) + \tau^{1/2} H_2(x) \right] S^2 \Gamma$$
$$+ \tau^{-1/2} F_2(\sigma) H_1(x) S^2 \Gamma$$
$$V_4(S,\sigma,t) = \tau^{-1} F_3(\sigma)\left[H_2(x) + \tau^{1/2} H_1(x) \right] S^2 \Gamma$$
$$+ \tfrac{1}{2}\rho^2\gamma\hat{\sigma}^{-2}\tau^{-1} F_4(\sigma)\left[H_4(x) + \tau^{1/2} H_3(x) \right] S^2 \Gamma$$
$$- \rho^2\tau^{-1} F_5(\sigma) H_2(x) S^2 \Gamma,$$

where the volatility used in Γ is $\hat{\sigma}$. From the definitions of the first Hermite polynomial H_1, of x and of Γ, it follows that V_1 vanishes at expiry, i.e. both V_0 and V_1 satisfy their corresponding expiry conditions. But this is not the case for terms V_n with $n \geq 2$.

55.8 VANILLA OPTIONS: IMPLIED VOLATILITIES

In this section, we look for an asymptotic expansion of the implied volatility

$$I = \hat{\sigma} + \epsilon^{1/2} I_1 + \epsilon I_2 + \epsilon^{3/2} I_3 + \epsilon^2 I_4 + O(\epsilon^{5/2}).$$

Since the implied volatility I satisfies

$$V_{BS}(S, I, t) = V(S, \sigma, t),$$

we determine $\{I_1, I_2, I_3, I_4\}$ by Taylor expanding on the LHS, inserting previous results on the right-hand side, and then equating terms of equal powers in ϵ. On the left-hand side, we use

$$V_{BS}(S, I, t) = \sum_{n=0}^{4} \left(\epsilon^{1/2} I_1 + \epsilon I_2 + \epsilon^{3/2} I_3 + \epsilon^2 I_4 \right)^n \frac{\partial^n V_{BS}}{\partial \hat{\sigma}^n} + O(\epsilon^{5/2}). \tag{55.21}$$

Since vega for a European Vanilla option is given by

$$\frac{\partial V_{BS}}{\partial \hat{\sigma}} = S\sqrt{T-t}\,N(d_1), \tag{55.22}$$

and since

$$\frac{\partial d_1}{\partial \hat{\sigma}} = -\hat{\sigma}^{-1} d_2$$
$$\frac{\partial d_2}{\partial \hat{\sigma}} = -\hat{\sigma}^{-1} d_1,$$

we get

$$\frac{\partial^2 V_{BS}}{\partial \hat{\sigma}^2} = \eta_2 \frac{\partial V_{BS}}{\partial \hat{\sigma}}$$
$$\frac{\partial^3 V_{BS}}{\partial \hat{\sigma}^3} = \eta_3 \frac{\partial V_{BS}}{\partial \hat{\sigma}}$$
$$\frac{\partial^4 V_{BS}}{\partial \hat{\sigma}^4} = \eta_4 \frac{\partial V_{BS}}{\partial \hat{\sigma}},$$

where we have used the 'vega coefficients,'

$$\begin{aligned}
\eta_1 &= 1\\
\eta_2 &= \hat{\sigma}^{-1} d_1 d_2\\
\eta_3 &= -\hat{\sigma}^{-2}\big(1 + d_1 + d_2 - d_1 d_2\big)\\
\eta_4 &= \hat{\sigma}^{-3} d_1 d_2 \big([d_1 d_2 - 1]^2 - d_1^2 d_2 - d_1 d_2^2 - d_2^2 - d_1^2 + 2d_2 + 2 d_1\big).
\end{aligned}$$

Equation (55.21) now becomes

$$V_{BS}(S, I, t) = V_{BS}(S, \hat{\sigma}, t) + \left[\sum_{n=1}^{4}\big(\epsilon^{1/2} I_1 + \epsilon I_2 + \epsilon^{3/2} I_3 + \epsilon^2 I_4\big)^n \eta_n\right]\frac{\partial V_{BS}}{\partial \hat{\sigma}} + O(\epsilon^{5/2}).$$

Since

$$\frac{\partial V_{BS}}{\partial \hat{\sigma}} = \hat{\sigma}\,(T-t)\,S^2\,\Gamma,$$

and after collecting terms of equal order in ϵ, we get

$$\begin{aligned}
V_{BS}(S, I, t) = V_{BS}(S, \hat{\sigma}, t)&\\
&+ \epsilon^{1/2}\,\hat{\sigma}\,(T-t)\,I_1\,S^2\,\Gamma\\
&+ \epsilon\,\hat{\sigma}\,(T-t)\,(I_2 + I_1^2\,\eta_2)\,S^2\,\Gamma\\
&+ \epsilon^{3/2}\,\hat{\sigma}\,(T-t)\,(I_3\,\eta_1 + 2\,I_1\,I_2\,\eta_2 + I_1^3\,\eta_3)\,S^2\,\Gamma\\
&+ \epsilon^2\,\hat{\sigma}\,(T-t)\,(I_1^4\,\eta_4 + 3\,I_1^2\,I_2\,\eta_3 + 2\,I_1\,I_3\,\eta_2 + I_2^2\,\eta_2 + I_4\,\eta_4)\,S^2\,\Gamma\\
&+ O(\epsilon^{5/2}).
\end{aligned}$$

Setting this equal to

$$\begin{aligned}V_{BS}(S, I, t) &= V(S, \hat{\sigma}, t)\\ &= V_0 + \epsilon^{1/2} V_1 + \epsilon V_2 + \epsilon^{3/2} V_3 + \epsilon^2 V_4 + O(\epsilon^{5/2}),\end{aligned}$$

it is straightforward to determine $\{I_1, I_2, I_3, I_4\}$ in terms of $\{V_1, V_2, V_3, V_4\}$. In particular, using the results for V_1 and V_2, we obtain for I_1 and I_2,

$$\begin{aligned}I_1 &= -\rho\,\gamma\,\hat{\sigma}^{-2}\,(T-t)^{-1/2}\,d_2\\ I_2 &= \frac{1}{2\,\hat{\sigma}\,(T-t)}\int_{\sigma}^{\infty}\frac{1}{\beta^2(y)\,p(y)}\int_0^y p(x)\,(\hat{\sigma}^2 - x^2)\,dx\,dy - \rho^2\,\gamma^2\,\hat{\sigma}^{-5}\,(T-t)^{-1}\,d_1\,d_2^2,\end{aligned}$$

where we now use standard notation. Since

$$\begin{aligned}d_1 &= \frac{M + \hat{\sigma}^2\,(T-t)/2}{\hat{\sigma}\,\sqrt{T-t}}\\ d_2 &= \frac{M - \hat{\sigma}^2\,(T-t)/2}{\hat{\sigma}\,\sqrt{T-t}}\\ d_1\,d_2 &= \frac{M^2 - \hat{\sigma}^4\,(T-t)^2/4}{\hat{\sigma}^2\,(T-t)},\end{aligned}$$

we get

$$\begin{aligned}I = \hat{\sigma} &- \epsilon^{1/2}\,\rho\,\gamma\,\hat{\sigma}^{-3}\,(T-t)^{-1}\left(M - \hat{\sigma}^2\,(T-t)/2\right)\\ &+ \frac{\epsilon}{2\,\hat{\sigma}\,(T-t)}\int_{\sigma}^{\infty}\frac{1}{\beta^2(y)\,p(y)}\int_0^y p(x)\,(\hat{\sigma}^2 - x^2)\,dx\,dy\\ &- \epsilon\,\rho^2\,\gamma^2\,\hat{\sigma}^{-8}\,(T-t)^{-4}\left(M - \hat{\sigma}^2\,(T-t)/2\right)\left(M^2 - \hat{\sigma}^4\,(T-t)^2/4\right) + O(\epsilon^{3/2}),\end{aligned}$$

where

$$M = \log(S\,e^{r\,(T-t)}/K)$$

is the 'forward moneyness.' In terms of M,

$$\begin{aligned}I = \hat{\sigma} &+ \tfrac{1}{2}\,\epsilon^{1/2}\,\rho\,\gamma\,\hat{\sigma}^{-1} - \tfrac{1}{8}\,\epsilon\,\rho^2\,\gamma^2\,\hat{\sigma}^{-2}\,(T-t)^{-1}\\ &+ \frac{\epsilon}{2\,\hat{\sigma}\,(T-t)}\int_{\sigma}^{\infty}\frac{1}{\beta^2(y)\,p(y)}\int_0^y p(x)\,(\hat{\sigma}^2 - x^2)\,dx\,dy\\ &+ \left\{\tfrac{1}{4}\,\epsilon\,\rho^2\,\gamma^2\,\hat{\sigma}^{-4}\,(T-t)^{-2} - \epsilon^{1/2}\,\rho\,\gamma\,\hat{\sigma}^{-3}\,(T-t)^{-1}\right\} M\\ &+ \left\{\tfrac{1}{2}\,\epsilon\,\rho^2\,\gamma^2\,\hat{\sigma}^{-6}\,(T-t)^{-3}\right\} M^2\\ &- \epsilon\,\rho^2\,\gamma^2\,\hat{\sigma}^{-8}\,(T-t)^{-4}\,M^3 + O(\epsilon^{3/2}),\end{aligned} \tag{55.23}$$

which is the main result. Using Equation (55.23), we can easily calibrate a stochastic volatility model to the implied volatilities of traded options. We considered only vanilla options, but a similar analysis can be carried out for any option with a closed-form solution in the Black–Scholes world (like digitals, barriers, ladders and certain compound options). Finally, we note that Equation (55.23) fails for small $T - t$. The reason is that Equation (55.23) only holds if the time to expiry $T - t$ is much larger than the characteristic timescale for volatility fluctuations.

55.9 SUMMARY

Asymptotic analysis is a very, very powerful tool for studying equations when some parameter is large or small. It is used extensively in the hard sciences, and increasingly in quantitative finance.

FURTHER READING

- For other asymptotic work in this area see Fouque, Papanicolaou, & Sircar (1998, 1999, 2000).
- For approximations that can be used close to expiration see Hull & White (1987), Taylor & Xu (1994), Zhu & Avellaneda (1998), Hagan & Woodward (1999) and Lewis (2000).

CHAPTER 56

volatility case study: the cliquet option

In this Chapter. . .

- how the cliquet option depends on volatility
- pricing with volatility ranges

56.1 INTRODUCTION

In Chapter 29 I looked at the pricing of cliquet options in a deterministic volatility setting, including the interpretation of the path dependency in terms of partial differential equations. In this chapter, I will show how the contract value depends on the treatment of volatility. In particular, I shall show results for the uncertain volatility model.

56.2 THE SUBTLE NATURE OF THE CLIQUET OPTION

Traditionally one measures sensitivity to volatility via the vega. This is defined as the derivative of the option value with respect to a (usually constant) volatility. This number is then used to determine how accurate a price might be should volatility change. As part of one's risk management, perhaps one will vega hedge to reduce such sensitivity.

This is entirely reasonable when the contract in question is an exchange-traded vanilla contract and one is measuring sensitivity to the market's (implied) volatility.

However, when it comes to the risk management of exotic options the sensitivity to a constant volatility is at best irrelevant and at worst totally misleading. By now this is common knowledge and I don't need to dwell on the details. It suffices to say that whenever a contract has a gamma which changes sign (as does any 'interesting' exotic) vega may be small at precisely those places where sensitivity to actual volatility is very large.

Confused? As a rule of thumb if you increase volatility when gamma is positive you will increase a contract's value. At points of inflection in the option value (where gamma is zero) the option value may hardly move. But this is sensitivity to a parameter that takes the same value everywhere. What if you increase volatility when gamma is positive and decrease it when gamma is negative? The net effect is an increase in option value even at points of inflection.

Five-year Minimum Coupon Cliquet on ABC Index

Option Buyer	XXXX
Option Seller	YYYY
Notional Amount	EUR 25MM
Start Date	dd/mm/yyyy
Maturity Date	Start Date + Five years
Option Seller Pays at Maturity	$\text{Notional} * \max\left(\sum_{i=1}^{5}\max\left(0,\min\left(\text{Cap},\frac{S_i - S_{i-1}}{S_{i-1}}\right)\right),\text{Floor}\right)$
Index	ABC Index
Cap	8%
Floor	16%
Option Premium	???
Index Levels	S_i = Closing Level of Index on Start Date + i years

This indicative term sheet is neither an offer to buy or sell securities or an OTC derivative product which includes options, swaps, forwards and structured notes having similar features to OTC derivative transactions, nor a solicitation to buy or sell securities or an OTC derivative product. The proposal contained in the foregoing is not a complete description of the terms of a particular transaction and is subject to change without limitation.

Figure 56.1 Example term sheet for a cliquet option.

Skews and smiles can makes matters even worse, unless you are fortunate enough that your skew/smile model forecasts *actual* volatility behavior accurately.

The classical references to this phenomenon are Avellaneda, Levy & Parás (1995) and Lyons (1995).

And the relevance to cliquet options? Take a look at the term sheet in Figure 56.1. To see the importance of volatility you just need to plot the formula

$$\max\left(0, \min\left(\text{Cap}, \frac{S_i - S_{i-1}}{S_{i-1}}\right)\right)$$

against S to see the non-convex nature of the option price; gamma changes sign.

Now comes the subtle part. *The point at which gamma changes sign depends on the relative move in S from one fixing to the next.* The point of inflection is not near any particular value of S. The conclusion has to be that any deterministic, volatility surface model fitted to vanilla prices is not going to be able to model the risk associated with changing volatility. This is true even if you allow the local volatility surface to move up and down *and* to rotate.

For this reason we are going to focus on using the uncertain volatility model described in the above-mentioned references. In this model the actual volatility is chosen to vary with the variables in such a way as to give the option value its worst (or best) possible value. The actual volatility is assumed to lie in the range σ^- to σ^+. The worst option value is when actual volatility is highest for negative gamma and lowest for positive gamma:

$$\sigma(\Gamma) = \begin{cases} \sigma^+ & \text{if } \Gamma < 0 \\ \sigma^- & \text{if } \Gamma > 0. \end{cases}$$

Now let us look at the pricing of the cliquet option.

56.3 PATH DEPENDENCY, CONSTANT VOLATILITY

We will be working in the classical lognormal framework for the underlying

$$dS = \mu S\,dt + \sigma S\,dX.$$

Let's quickly recap what we did in Chapter 29.

In all non-trivial problems we always have the two given dimensions, S and t. In order to be able to keep track, before expiry, of the progress of the possible option payoff we also need the following two new 'state variables'

$$S' \quad \text{and} \quad Q,$$

where

$$S' = \text{the value of } S \text{ at the previous fixing} = S_i$$

and

$$Q = \text{the sum to date of the bit inside the max function} =$$

$$\sum_{j=1}^{i} \max\left(0, \min\left(\text{Cap}, \frac{S_j - S_{j-1}}{S_{j-1}}\right)\right).$$

Here I am using the index i to denote the fixing just prior to the current time, t. This is all made clear in Figure 56.2.

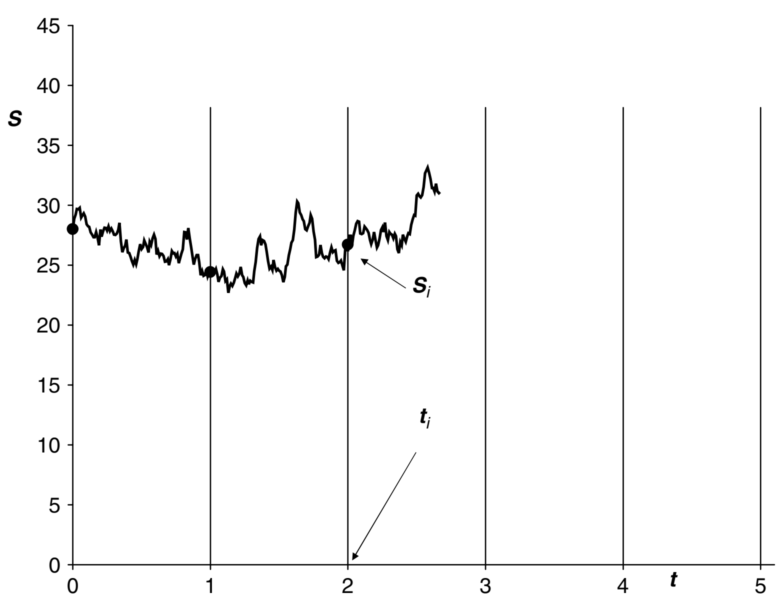


Figure 56.2 Defining the payoff.

Since S' and Q are only updated discretely, at each fixing date, the pricing problem for $V(S, t, S', Q)$ becomes

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

where r is the risk-free interest rate. In other words, the vanilla Black–Scholes equation. The twist is that V is a function of four variables, and must further satisfy the jump condition across the fixing date

$$V(S, t_i^-, S', Q) = V\left(S, t_i^+, S, Q + \min\left(E_1, \frac{S - S'}{S'}\right)\right)$$

and the final condition

$$V(S, T, S', Q) = \max(Q, E_2).$$

Here E_1 is the local cap and E_2 the global floor.

Being a four-dimensional problem it is a toss up as to whether a Monte Carlo or a finite-difference solution is going to be the faster. However, the structure of the payoff, and the assumption of lognormality, mean that a similarity reduction is possible, taking the problem down to only three dimensions and thus comfortably within the domain of usefulness of finite-difference methods. The similarity variable is

$$\xi = \frac{S}{S'}.$$

The option value is now a function of ξ, t and Q. The governing equation for $V(\xi, t, Q)$ (loose notation, but the most clear) is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 \xi^2 \frac{\partial^2 V}{\partial \xi^2} + r\xi\frac{\partial V}{\partial \xi} - rV = 0.$$

The jump condition becomes

$$V(\xi, t_i^-, Q) = V\left(1, t_i^+, Q + \min\left(E_1, \xi - 1\right)\right)$$

and the final condition is

$$V(\xi, T, Q) = \max(Q, E_2).$$

All of the results that I present are based on the finite-difference solution of the partial differential equation. The reason for this is that I want to focus on the volatility dependence, in particular I need to be able to implement the uncertain volatility model and this is not so simple to do in the Monte Carlo framework. (This is because volatility depends on gamma in this model and gamma is not calculated in the standard Monte Carlo implementation.)

56.4 RESULTS

The following results are based on the cliquet option described in the term sheet. In particular, it is a five-year contract with annual fixings, a global floor of 16% and local caps of 8%. The interest rate is 3% and there are no dividends on the underlying.

To understand the following you must remember that the cliquet value is a function of three independent variables, ξ, Q and t. I will be showing plots of value against various variables at certain times before expiry. These will assume a constant volatility. Then we will look at the effect of varying volatility on the prices.

56.4.1 Constant Volatility

In the following five plots, Figures 56.3–56.7, volatility is everywhere 25%. Read the captions to each of these figures very carefully.

56.4.2 Uncertain Volatility

The above shows the evolution of the option value for constant volatility. There is diffusion in the ξ direction and a 'jump condition' to be applied at every fixing. The amount of the diffusion is constant (or rather, is constant on a logarithmic scale.)

To price the contract when volatility is uncertain we must use a volatility that depends on (the sign of) gamma. Some results are shown below.

In Figure 56.8 is plotted the contract value against ξ at five years before expiry with $Q = 0$. Five calculations have been performed.

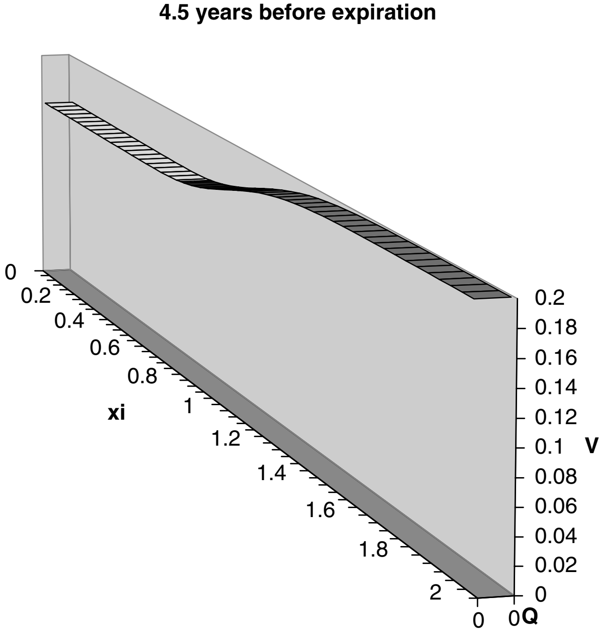


Figure 56.3 This plot shows the cliquet value against Q and ξ at 4.5 years before expiry. The contract has thus been in existence for six months. At this stage there have been no fixings yet and the state variable Q only takes the value 0. The non-convex contract value can be clearly seen.

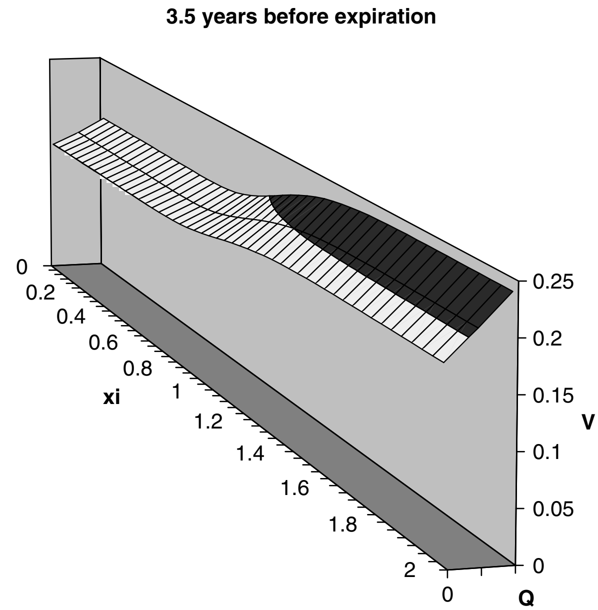


Figure 56.4 The value at 3.5 years before expiry, and therefore 1.5 years into the contract's life, is shown in this plot. The state variable Q now ranges from zero to 8%.

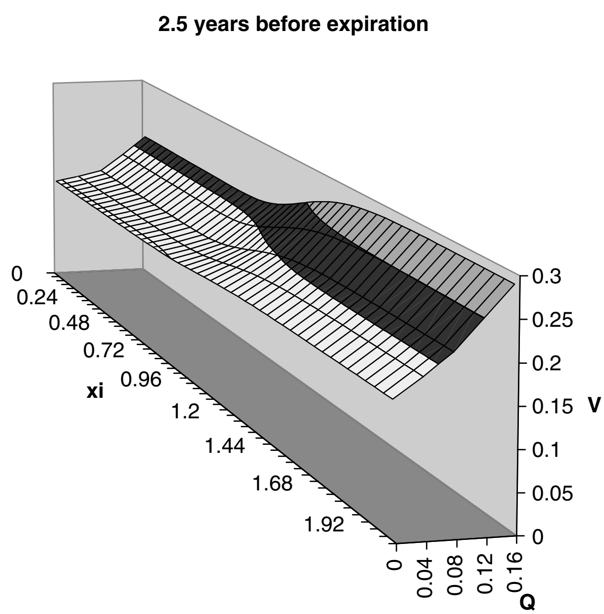


Figure 56.5 One year later, the contract is exactly half way through its life. The state variable Q lies in the range zero to 16%. For small values of Q the option value is very close to being the present value of the 16% floor. This represents the small probability of getting a payoff in excess of the floor at expiry.

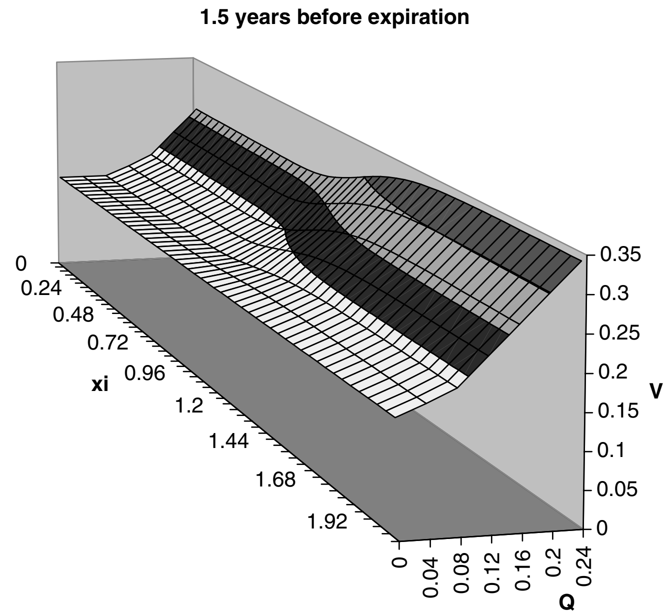


Figure 56.6 After being in existence for 3.5 years, and having only 1.5 years left to run, the cliquet value is as above. Now Q ranges from zero to 24%. When Q is zero there is no chance of the global floor being exceeded and so the contract value there is exactly the present value of 16%.

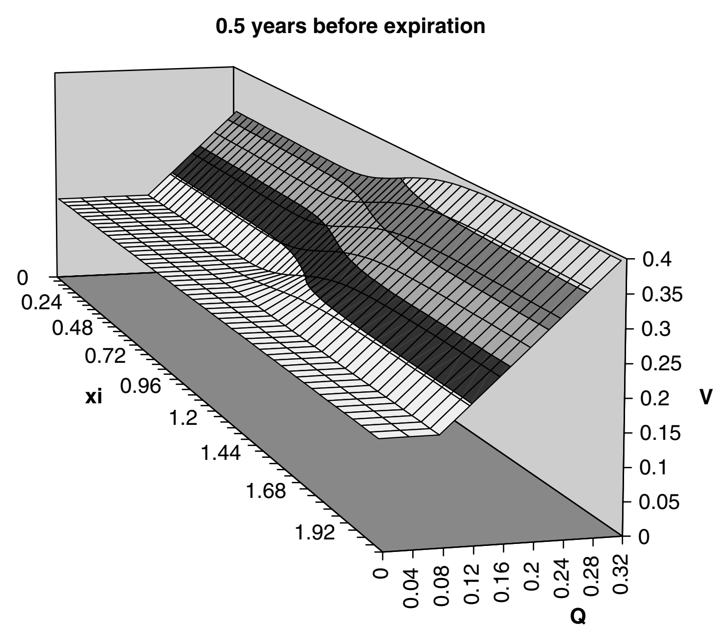


Figure 56.7 Six months before expiry the option value is as shown above. Q ranges from zero to 32% and for any values below 8% the contract is again only worth the present value of 16%.

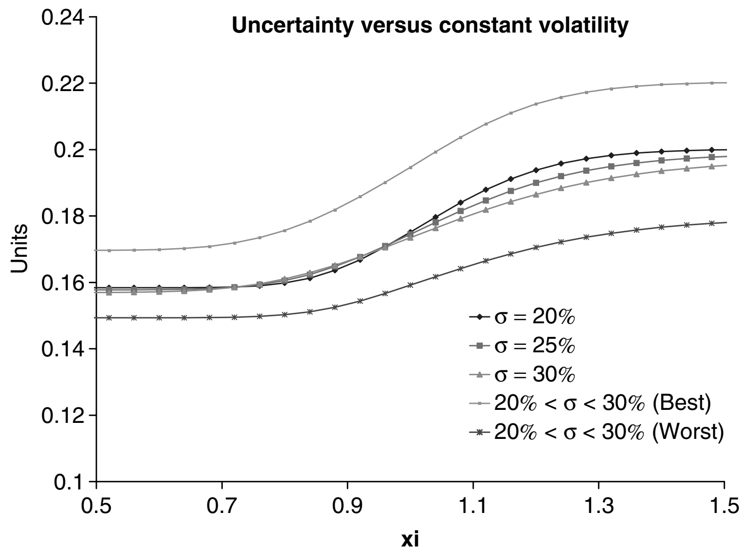


Figure 56.8 The contract value against ξ at five years before expiry with $Q = 0$.

1. The first line to examine is the middle line. This corresponds to a constant volatility of 25%. This is the base case with which we compare other prices.
2. The second line to examine is the one close to the middle line, but slightly above at the right. This is the cliquet value with a constant volatility of just 20%.
3. The third case has a constant volatility of 30%. This is also very close to the middle line, but slightly below at the right.
4. The fourth line represents the cliquet value when the volatility is allowed to range between 20 and 30%, taking a value locally that maximizes the cliquet value overall. This line is much higher than all of the others.
5. The fifth and final curve is the one for which volatility has again been allowed to range from 20 to 30% but now such that it gives the option its lowest possible value. This line is much lower than all of the others.

The first observation to make is how close the constant volatility curves are, i.e. curves 1–3. As stated above, a good rule of thumb is that high volatility and positive gamma give a high option value. Because gamma changes sign in this contract, the result is that there is a ξ value at which the contract value does not appear to be sensitive to the volatility. In this case the value is around 0.95, close to the point of inflection.

Now ask yourself the following question: 'Do I believe that volatility is a constant, and this constant is somewhere between 20% and 30%? Or do I believe that volatility is highly uncertain, but is most likely to stay within the range 20% to 30%?'

If you believe the former, then the calculation we have just done, in curves 1–3, is relevant. If, on the other hand, you think that the latter is more likely (and who wouldn't?) then you must discard the calculations in curves 1–3 and consider the whole spectrum of possible option values by looking at the best and worst cases, curves 4 and 5.

					Vmax						
	0.20	**0.21**	**0.22**	**0.23**	**0.24**	**0.25**	**0.26**	**0.27**	**0.28**	**0.29**	**0.3**
Vmin 0.2	0.1743	0.1720	0.1700	0.1680	0.1662	0.1645	0.1629	0.1615	0.1601	0.1588	0.1576
0.21	0.1763	0.1741	0.1719	0.1699	0.1680	0.1663	0.1646	0.1631	0.1617	0.1603	0.1591
0.22	0.1784	0.1761	0.1739	0.1718	0.1698	0.1680	0.1663	0.1647	0.1632	0.1618	0.1605
0.23	0.1804	0.1780	0.1757	0.1736	0.1716	0.1698	0.1680	0.1663	0.1648	0.1633	0.1620
0.24	0.1824	0.1799	0.1776	0.1754	0.1734	0.1715	0.1696	0.1679	0.1663	0.1648	0.1634
0.25	0.1843	0.1818	0.1794	0.1772	0.1751	0.1731	0.1713	0.1695	0.1679	0.1663	0.1648
0.26	0.1863	0.1837	0.1812	0.1789	0.1768	0.1748	0.1729	0.1711	0.1694	0.1678	0.1662
0.27	0.1881	0.1855	0.1830	0.1807	0.1785	0.1764	0.1744	0.1726	0.1708	0.1692	0.1676
0.28	0.1900	0.1873	0.1847	0.1824	0.1801	0.1780	0.1760	0.1741	0.1723	0.1706	0.1690
0.29	0.1918	0.1890	0.1865	0.1840	0.1817	0.1796	0.1775	0.1756	0.1738	0.1720	0.1704
0.3	0.1935	0.1908	0.1881	0.1857	0.1833	0.1811	0.1790	0.1770	0.1752	0.1734	0.1717

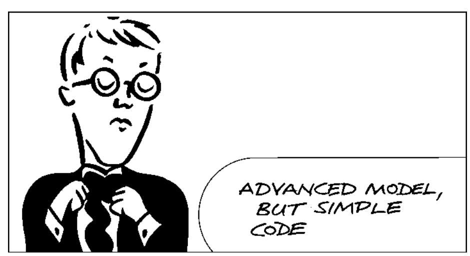
Figure 56.9 How the cliquet value (five years before expiry at $\xi = Q = 0$) varies with the allowed range for volatility.

Such calculations show that the real sensitivity to volatility is much, much larger than a naive vega calculation would suggest.

The next figure, Figure 56.9, shows how the cliquet value (five years before expiry at $\xi = Q = 0$) varies with the allowed range for volatility. The table is to be read as follows. When volatility takes one value only, read along the diagonal to see the contract values. For example, when the volatility is 22% the contract value is 0.1739; and when the volatility is 27% the contract value is 0.1726.

Now consider a range of possible volatilities. Suppose you believe volatility will not stray from the range 22% to 27%. The worst case is in the top right cells, in this case 0.1647. The best case is to be found in the bottom left cells, 0.1830. So, when volatility ranges from 22 to 27% the correct range for the contract value is 0.1647 to 0.1830.

When volatility is a constant, but a constant between 22% and 27%, the contract value range is $0.1739 - 0.1726 = 0.0013$ or 0.75% relative (to mid price) range. When volatility is allowed to vary over the 22–27% range we find that the contract value itself has a value range of $0.1830 - 0.1647 = 0.0183$ or 10.5% relative (to mid price) range. The true sensitivity to volatility is 14 times greater than that estimated by vega.



56.5 CODE: CLIQUET WITH UNCERTAIN VOLATILITY, IN SIMILARITY VARIABLES

Below is some Visual Basic code that can be used for pricing these cliquet options in the uncertain volatility framework.

The range for volatility is `VolMin` to `VolMax`, the dividend yield is `Div`, risk-free interest rate `IntRate`, the local cap is `Strike2` and the global floor `Strike1`. Expiry is `Expiry`. The numerical parameter is `NumAssetSteps`, the number of steps in the S and Q directions.

```
Function cliquet(VolMin, VolMax, Div, IntRate, Strike1, Strike2, _
            NumFixes, Fixing, Expiry, NumAssetSteps)
ReDim xi(-NumAssetSteps To NumAssetSteps)
```

```
Dim Vmax, AssetStep, TStep, QStep, Delta, Gamma, Theta, Tim, Vol, qafter, _
            frac, V1, V2 As Double
Dim i, j, k, M, iafter, kafter, N, NumSoFar, NumQSteps As Integer
ReDim jtest(1 To NumFixes) As Integer

Vmax = Application.Max(VolMin, VolMax)
AssetStep = 1 / NumAssetSteps
TStep = 0.95 * AssetStep ^ 2 / Vmax ^ 2 / 2 ^ 2 ' A) This ensures stability _
            of the explicit method
M = Int(Expiry / TStep) + 1
TStep = Expiry / M
QStep = AssetStep
NumQSteps = Int(Strike2 / QStep) * NumFixes

ReDim Q(0 To NumQSteps)
ReDim VOld(-NumAssetSteps To NumAssetSteps, 0 To NumQSteps) ' First dimension _
            centered on xi = 1
ReDim VNew(-NumAssetSteps To NumAssetSteps, 0 To NumQSteps)

NumSoFar = 1
For j = 1 To NumFixes - 1
jtest(j) = Int(j * Fixing / TStep) ' Used in testing whether fixing date _
                has been passed
Next j

For k = 0 To NumQSteps
Q(k) = k * QStep
For i = -NumAssetSteps To NumAssetSteps
xi(i) = 1 + AssetStep * i ' i = 0 corresponds to xi = 0
VOld(i, k) = Application.Max(Strike1, Q(k) + _
                  Application.Max(0, Application.Min(Strike2, xi(i) - 1))) ' B) _
                  Payoff
Next i
Next k

For j = 1 To M ' C) Time stepping

For k = 0 To NumQSteps
For i = -NumAssetSteps + 1 To NumAssetSteps - 1
Delta = (VOld(i + 1, k) - VOld(i - 1, k)) / 2 / AssetStep ' Central difference
Gamma = (VOld(i + 1, k) - 2 * VOld(i, k) + VOld(i - 1, k)) / AssetStep / AssetStep
Vol = VolMax
If Gamma > 0 Then Vol = VolMin ' Volatility depends on gamma in the uncertain _
                            volatility model
Theta = IntRate * VOld(i, k) - 0.5 * Vol * Vol * xi(i) * xi(i) * Gamma _
                            - (IntRate - Div) * xi(i) * Delta ' _
                            The Black-Scholes equation
VNew(i, k) = VOld(i, k) - TStep * Theta
Next i

VNew(-NumAssetSteps, k) = VOld(-NumAssetSteps, k) * (1 - IntRate * TStep) ' D) _
                            Boundary condition at xi = 0
VNew(NumAssetSteps, k) = VNew(NumAssetSteps - 1, k) ' D) Boundary condition at _
                            xi = infinity. Delta = 0

For i = -NumAssetSteps To NumAssetSteps
```

```
VOld(i, k) = VNew(i, k) ' E) Overwriting with new data
Next i

Next k

If jtest(NumSoFar) = j Then ' F) Test for a fixing date

    For i = -NumAssetSteps To NumAssetSteps
    For k = 0 To NumQSteps
    qafter = Q(k) + Application.Max(0, Application.Min(Strike2, xi(i) - 1)) ' G) _
                  The updating rule
    kafter = Int(qafter / QStep)
    frac = (qafter - QStep * kafter) / QStep

    V1 = 0
    V2 = 0
    If kafter < NumQSteps Then
    V1 = VNew(0, kafter)
    V2 = VNew(0, kafter + 1)
    End If

    VOld(i, k) = (1 - frac) * V1 + frac * V2 ' The jump condition. Linear _
                  interpolation
    Next k
    Next i
    NumSoFar = NumSoFar + 1

End If

Next j

cliquet = VOld ' Output the whole array

End Function
```

This program clearly leaves much to be desired, for example in the discretization, the treatment of the jump condition etc. But it does have the benefit of transparency.

A: The time step is set so that the explicit finite difference method is stable. If the time step is any smaller than this, the method will not converge.

B: Here the payoff is set up; the dependent variable as a function of the independent variables.

C: The time-stepping engine. Delta and gamma are discretized versions of the first- and second-order derivatives with respect to S. This part of the code also treats the uncertain volatility. See how the volatility depends on the sign of gamma.

D: The boundary conditions, for $\xi = 0$ and large ξ.

E: Updating the next step back in the grid.

F: Here the code tests for a fixing date.

G: Across fixing dates the updating rule is applied. This is really the only point in the code that knows we are pricing a cliquet option.

56.6 SUMMARY

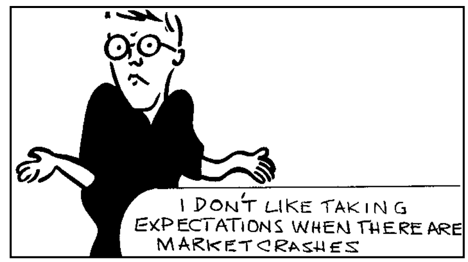
The cliquet can be a subtle instrument, with a dependence on volatility that is far from obvious. In our example there was a high degree of sensitivity to volatility but not to level, nor to current skew. No, that sensitivity was to a rotation of forward volatility skew about a point 4% (half way between the floor and the cap) above the previous year's asset price. And that level just happens to be at approximately where the risk-neutral asset is expected to be. Pure genius, albeit an evil genius!

I hope this chapter has shown you, using a concrete example, just how sensitive contracts can be to volatility, when they might appear to be quite harmless.

FURTHER READING

- This work on cliquet options was first published in Wilmott (2002).

CHAPTER 57
jump diffusion



In this Chapter...

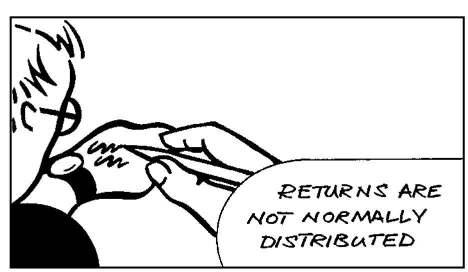
- the Poisson process for modeling jumps
- hedging in the presence of jumps
- how to price derivatives when the path of the underlying can be discontinuous
- jump volatility

57.1 INTRODUCTION

There is plenty of evidence that financial quantities, be they equities, currencies or interest rates, for example, do not follow the lognormal random walk that has been the foundation of almost everything in this book, and almost everything in the financial literature. We look at some of this evidence in a moment. One of the striking features of real financial markets is that every now and then there is a sudden unexpected fall or crash. These sudden movements occur far more frequently than would be expected from a Normally distributed return with a reasonable volatility. On all but the shortest timescales the move looks discontinuous; the prices of assets have jumped. This is important for the theory and practice of derivatives because *it is usually not possible to hedge through the crash*. One certainly cannot delta hedge as the stock market tumbles around one's ankles, and to offload all one's positions will lead to real instead of paper losses, and may even make the fall worse.

In this chapter I explain classical ways of pricing and hedging when the underlying follows a jump-diffusion process.

57.2 EVIDENCE FOR JUMPS



Let's look at some data to see just how far from Normal the returns really are. There are several ways to visualize the difference between two distributions, in our case the difference between the empirical distribution and the Normal distribution. One way is to overlay the two probability distributions. In Figure 57.1 we see the distribution of Xerox returns, from 1986 until 1997, normalized to unit standard deviation. The

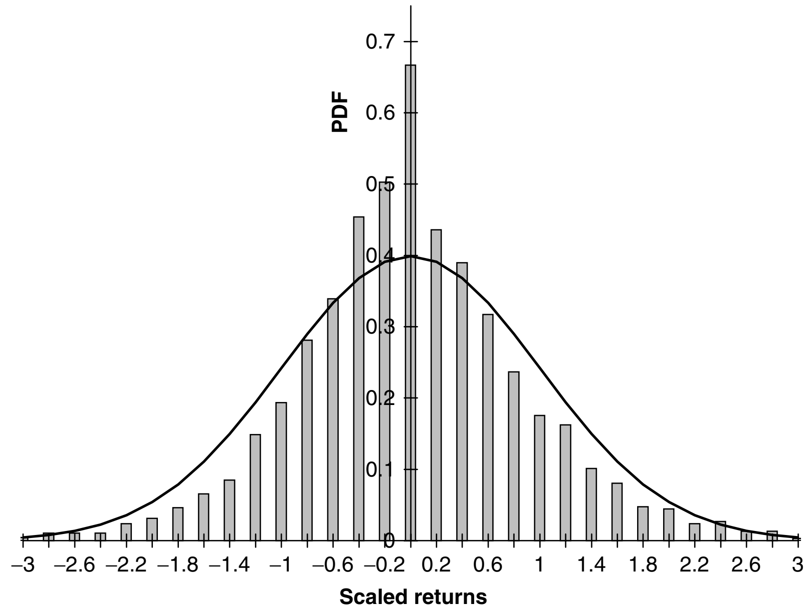


Figure 57.1 The probability density functions for Xerox daily returns, scaled to zero mean and unit standard deviation, and the standardized Normal distribution.

peak of the real distribution is clearly higher than the Normal distribution. Because both of these distributions have the same standard deviation then the higher peak must be balanced by fatter tails, it's just that they would be too small to see on this figure. They may be too small to see here, but they are still very important.

This difference is typical of all markets, even typical of the changes in interest rates. The empirical distribution diverges from the Normal distribution quite markedly. The peak being much higher means that there is a greater likelihood of a small move than we would expect from the lognormal random walk. More importantly, and concerning the subject of this chapter, the tails are much fatter. There is a greater chance of a large rise or fall than the Normal distribution predicts.

In Figure 57.2 is shown the cumulative distribution functions for Xerox daily returns, scaled to zero mean and unit standard deviation, and that for the standardized Normal distribution. And in Figure 57.3 is shown the difference between these two. If you look at the latter figure you will see that there is more weight in the empirical distribution than the Normal distribution from about two standards deviations away from the mean. If you couple this likelihood of an extreme movement with the *importance* of an extreme movement, assuming perhaps that people are hedged against *small* movements, you begin to get a very worrying scenario.

Another useful plot is the logarithm of the pdf for the actual distribution and the Normal distribution. The latter is simply a parabola, but what is the former? Such a plot is shown in Figure 57.4. The log of the pdf of the actual distribution looks more linear than quadratic in the tails.

The final picture that I plot, in Figure 57.5, is called a **Quantile-Quantile** or **Q-Q plot**. This is a common way of visualizing the difference between two distributions when you are particularly interested in the tails of the distribution. This plot is made up as follows.

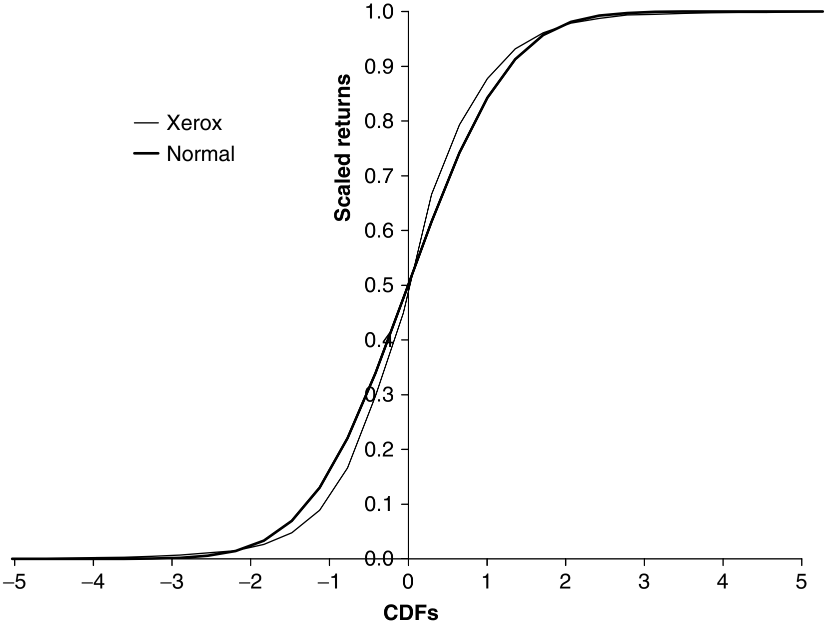


Figure 57.2 The cumulative distribution functions for Xerox daily returns, scaled to zero mean and unit standard deviation, and that for the standardized Normal distribution.

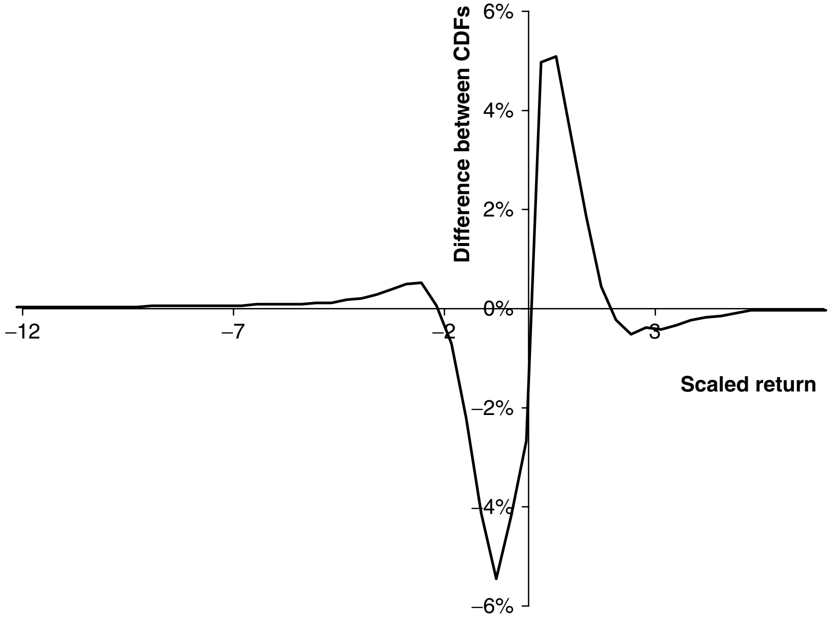


Figure 57.3 The difference between the cumulative distribution functions for Xerox daily returns, scaled to zero mean and unit standard deviation, and the standardized Normal distribution.

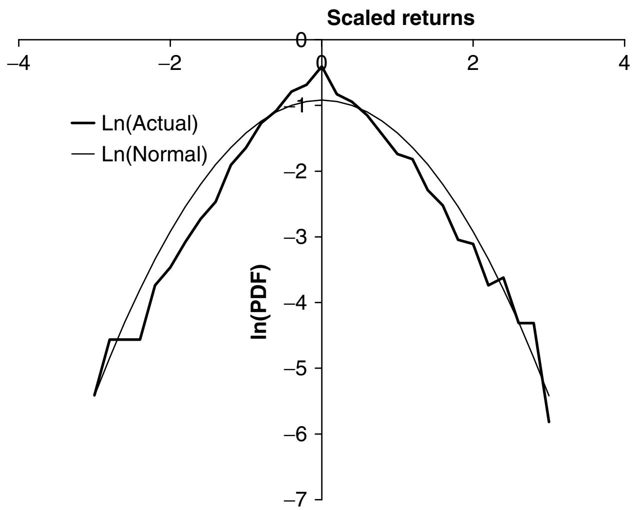


Figure 57.4 Logarithms of the pdfs for Xerox daily returns and the standardized Normal distribution.

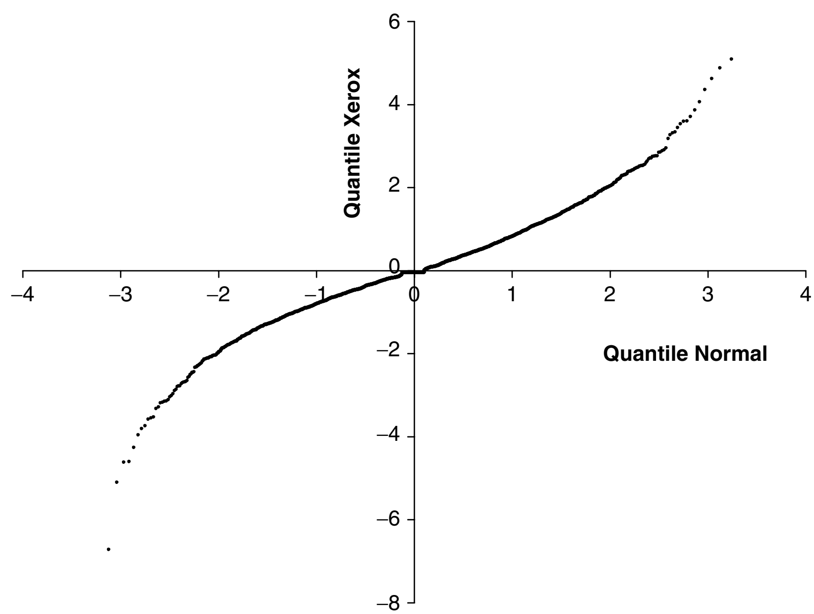


Figure 57.5 Q-Q plot for Xerox daily returns and the standardized Normal distribution.

First, rank the empirical returns in order from smallest to largest, call these y_i with an index i going from 1 to n. Second, for the Normal distribution find the returns x_i such that the cumulative distribution function at x_i has value i/n. Finally, plot each point (x_i, y_i).

The better the fit between the two distributions, the closer the line is to straight. In the present case the line is far from straight, due to the extra weight in the tails.

Several theories have been put forward for the non Normality of the empirical distribution. Three of these are

- Volatility is stochastic
- Returns are drawn from another distribution, a Pareto-Levy distribution, for example
- Assets can jump in value

There is truth in all of these. The first was the subject of Chapter 51. The second is a can of worms; moving away from the Normal distribution means throwing away 99% of current theory and is not done lightly. But some of the issues this raises have to be addressed in jump diffusion (such as the impossibility of hedging). The third is the present subject.

57.3 POISSON PROCESSES

The basic building block for the random walks we have considered so far is continuous Brownian motion based on the Normally distributed increment. We can think of this as adding to the return from one day to the next a Normally distributed random variable with variance proportional to time step. The extra building block we need for the **jump-diffusion model** for an asset price is the **Poisson process**. We saw this process in another context in Chapter 40. A Poisson process dq is defined by

$$dq = \begin{cases} 0 & \text{with probability } 1 - \lambda\, dt \\ 1 & \text{with probability } \lambda\, dt. \end{cases}$$

There is therefore a probability $\lambda\, dt$ of a jump in q in the time step dt. The parameter λ is called the **intensity** of the Poisson process. The scaling of the probability of a jump with the size of the time step is important in making the resulting process 'sensible,' i.e. there being a finite chance of a jump occurring in a finite time, with q not becoming infinite.

This Poisson process can be incorporated into a model for an asset in the following way:

$$dS = \mu S\, dt + \sigma S\, dX + (J-1)S\, dq. \tag{57.1}$$

We assume that there is no correlation between the Brownian motion and the Poisson process. If there is a jump ($dq = 1$) then S immediately goes to the value JS. We can model a sudden 10% fall in the asset price by $J = 0.9$.

We can generalize further by allowing J to also be a random quantity. We assume that it is drawn from a distribution with probability density function $P(J)$, again independent of the Brownian motion and Poisson process.

The random walk in $\log S$ follows from (57.1):

$$d(\log S) = \left(\mu - \tfrac{1}{2}\sigma^2\right) dt + \sigma\, dX + (\log J)\, dq.$$

This is just a jump-diffusion version of Itô.

Figure 57.6 is a spreadsheet showing how to simulate the random walk for S. In this simple example the stock jumps by 20% at random times given by a Poisson process.

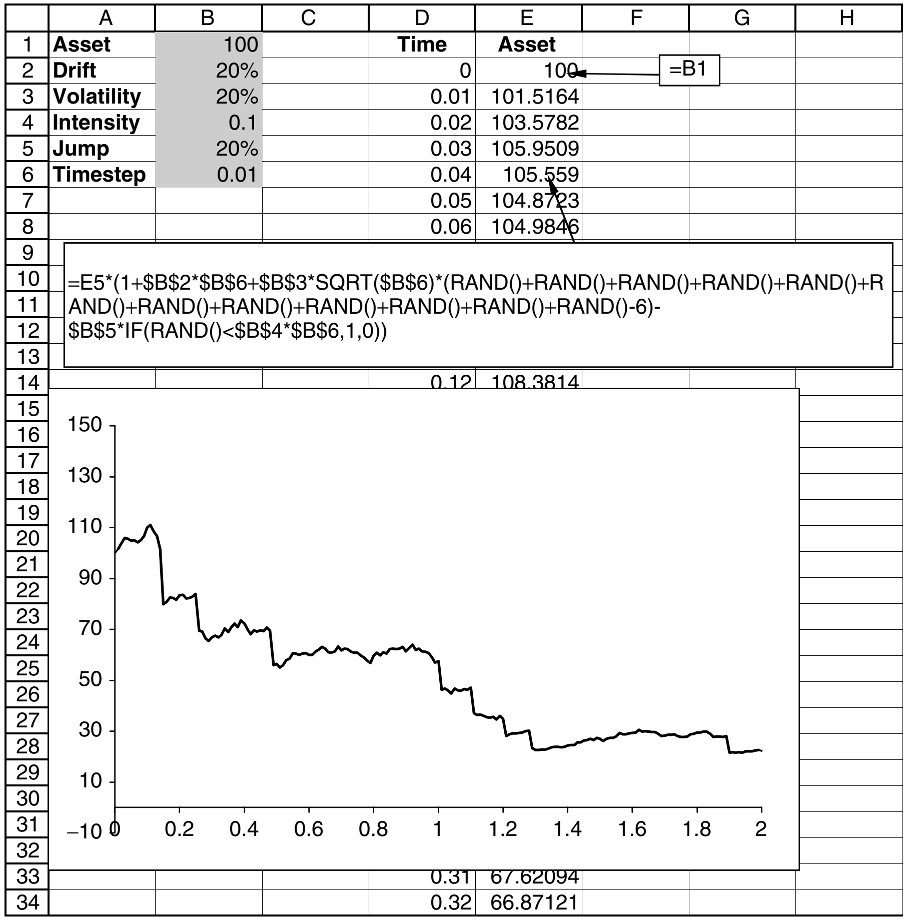


Figure 57.6 Spreadsheet simulation of a jump-diffusion process.

57.4 HEDGING WHEN THERE ARE JUMPS

Now let us build up a theory of derivatives in the presence of jumps. Begin by holding a portfolio of the option and $-\Delta$ of the underlying:

$$\Pi = V(S,t) - \Delta S.$$

The change in the value of this portfolio is

$$\begin{aligned} d\Pi = {} & \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \left(\frac{\partial V}{\partial S} - \Delta\right)(\mu S\, dt + \sigma S\, dX) \\ & + (V(JS,t) - V(S,t) - \Delta(J-1)S)\, dq. \end{aligned} \tag{57.2}$$

Again, this is a jump-diffusion version of Itô.

If there is no jump at time t so that $dq = 0$, then we could have chosen $\Delta = \partial V/\partial S$ to eliminate the risk. If there is a jump and $dq = 1$ then the portfolio changes in value by an $O(1)$ amount, *that cannot be hedged away*. In that case perhaps we should choose Δ to minimize the variance of $d\Pi$. This presents us with a dilemma. We don't know whether to hedge the small(ish) diffusive changes in the underlying which are always present, or the large moves which happen rarely. Let us pursue both of these possibilities.

57.5 HEDGING THE DIFFUSION

If we choose

$$\Delta = \frac{\partial V}{\partial S}$$

we are following a Black–Scholes type of strategy, hedging the diffusive movements. The change in the portfolio value is then

$$d\Pi = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \left(V(JS, t) - V(S, t) - (J-1)S\frac{\partial V}{\partial S}\right) dq.$$

The portfolio now evolves in a deterministic fashion, except that every so often there is a non-deterministic jump in its value. It can be argued (see Merton, 1976) that if the jump component of the asset price process is uncorrelated with the market as a whole, then the risk in the discontinuity should not be priced into the option. Diversifiable risk should not be rewarded. In other words, we can take expectations of this expression and set that value equal to the risk-free return from the portfolio. This is not completely satisfactory, but is a common assumption whenever there is a risk that cannot be fully hedged; default risk is another example of this. If we take such an approach then we arrive at

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \lambda E\left[V(JS, t) - V(S, t)\right] - \lambda \frac{\partial V}{\partial S} SE\left[(J-1)\right] = 0. \tag{57.3}$$

$E[\cdot]$ is the expectation taken over the jump size J, which can also be written

$$E[x] = \int x P(J)\, dJ,$$

where $P(J)$ is the probability density function for the jump size.

This is a pricing equation for an option when there are jumps in the underlying. The important point to note about this equation that makes it different from others we have derived is its non-local nature. That is, the equation links together option values at distant S values, instead of just containing local derivatives. Naturally, the value of an option here and now depends on the prices to which it can instantaneously jump. When $\lambda = 0$ the equation reduces to the Black–Scholes equation.

There is a simple solution of this equation in the special case that the logarithm of J is Normally distributed. If the logarithm of J is Normally distributed with standard deviation σ' and if we write

$$k = E[J-1]$$

then the price of a European non-path-dependent option can be written as

$$\sum_{n=0}^{\infty}\frac{1}{n!}e^{-\lambda'(T-t)}(\lambda'(T-t))^n V_{BS}(S,t;\sigma_n,r_n).$$

In the above

$$\lambda'=\lambda(1+k),\quad \sigma_n^2=\sigma^2+\frac{n\sigma'^2}{T-t}\quad\text{and}\quad r_n=r-\lambda k+\frac{n\log(1+k)}{T-t},$$

and V_{BS} is the Black–Scholes formula for the option value in the absence of jumps. This formula can be interpreted as the sum of individual Black–Scholes values each of which assumes that there have been n jumps, and they are weighted according to the probability that there will have been n jumps before expiry.

If one does not make the assumption that jumps should not be priced in, then one has to play around with concepts such as the market price of risk.

57.6 HEDGING THE JUMPS

In the above we hedged the diffusive element of the random walk for the underlying. Another possibility is to hedge both the diffusion and jumps as much as we can. For example, we could choose Δ to minimize the variance of the hedged portfolio, after all, this is ultimately what hedging is about.

The change in the value of the portfolio with an arbitrary Δ is, to leading order,

$$d\Pi=\left(\frac{\partial V}{\partial S}-\Delta\right)dS+(-\Delta(J-1)S+V(JS,t)-V(S,t))\,dq+\cdots.$$

The variance in this change, which is a measure of the risk in the portfolio, is

$$\operatorname{var}[d\Pi]=\left(\frac{\partial V}{\partial S}-\Delta\right)^2\sigma^2S^2dt+\lambda E\left[(-\Delta(J-1)S+V(JS,t)-V(S,t))^2\right]dt+\cdots. \tag{57.4}$$

This is minimized by the choice

$$\Delta=\frac{\lambda E\left[(J-1)(V(JS,t)-V(S,t))\right]+\sigma^2S\dfrac{\partial V}{\partial S}}{\lambda SE\left[(J-1)^2\right]+\sigma^2S}.$$

(To see this, differentiate (57.4) with respect to Δ and set the resulting expression equal to zero.)

If we value the options as a pure discounted real expectation under this best-hedge strategy then we find that

$$\begin{aligned}&\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial s^2}+S\frac{\partial V}{\partial S}\left(\mu-\frac{\sigma^2}{d}(\mu+\lambda k-r)\right)-rV\\&\quad+\lambda E\left[(V(JS,t)-V(S,t))\left(1-\frac{J-1}{d}(\mu+\lambda k-r)\right)\right]=0,\end{aligned}$$

where

$$d = \lambda E\left[(J-1)^2\right] + \sigma^2.$$

When $\lambda = 0$ this collapses to the Black–Scholes equation. At the other extreme, when there is no diffusion, so that $\sigma = 0$, we have

$$\Delta = \frac{E\left[(J-1)(V(JS,t) - V(S,t))\right]}{SE\left[(J-1)^2\right]}$$

and

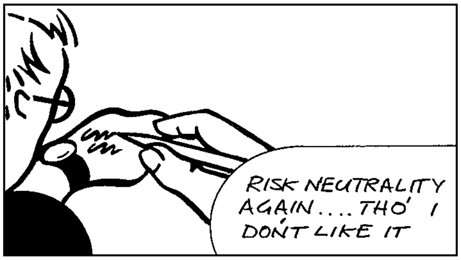
$$\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} - rV + \lambda E\left[(V(JS,t) - V(S,t))\left(1 - \frac{J-1}{d}(\mu + \lambda k - r)\right)\right] = 0.$$

All of the pricing equations we have seen in this chapter are integro-differential equations. (The integral nature is due to the expectation taken over the jump size.) Because of the convolution nature of these equations they are candidates for solution by Fourier transform methods.

57.7 HEDGING THE JUMPS AND RISK NEUTRALITY

The above talks about taking expectations, rather than using the usual hedging arguments, to get to a pricing equation. It is kinda possible to set up a perfectly hedged portfolio, as I'll explain briefly now.

Suppose we know the size of the jump, but not its timing. We can construct a portfolio of *two* options and the underlying that is risk-free. We can hedge the jump risk. But, as always, because we will have one equation (no arbitrage) in two unknowns (the values of the two options) we will inevitably end up with a pricing model that contains a market price of jump risk, an unknown function.



If the jump takes one of two values then we set up a hedged portfolio with three options and the underlying. If there are n jump states then we need $n + 1$ options and the underlying for the perfect hedge. If we have a distribution of jump sizes we need a distribution of options.

I won't go into the details, not being a great fan of anything resulting in market prices of risk.

57.8 THE DOWNSIDE OF JUMP-DIFFUSION MODELS

Jump diffusion as described above is unsatisfactory. Why bother to delta hedge at all when the portfolio will anyway be exposed to extreme movements? Hedging 'on average' is fine, after all that is being done whenever hedging is discrete, but after a crash the portfolio change is so dramatic that it makes hedging appear pointless. The other possibility is to examine the worst-case scenario. What is the worst that could happen, crashwise? Assume that this does happen and price it into the contract. This is discussed in depth in Chapter 58.

57.9 JUMP VOLATILITY

In this section I'm going to return to volatility modeling, now that we know what a Poisson process is. In Chapter 51 we saw Brownian motion models for volatility, but now let's model volatility as a jump process.

Perhaps volatility is constant for a while, then randomly jumps to another value. A bit later it jumps back. Let us model volatility as being in one of two states σ^- or $\sigma^+ > \sigma^-$. The jump from lower to higher value will be modeled by a Poisson process with intensity λ^+ and intensity λ^- going the other way. A realization of this process for σ is shown in Figure 57.7.

If we hedge the random movement in S with the underlying, then take *real* expectations, and set the return on the portfolio equal to the risk-free rate we arrive at

$$\frac{\partial V^+}{\partial t} + \frac{1}{2}\sigma^{+2}S^2\frac{\partial^2 V^+}{\partial S^2} + r\frac{\partial V^+}{\partial S} - rV^+ + \lambda^-(V^- - V^+) = 0,$$

for the value V^+ of the option when the volatility is σ^+. Similarly, we find that

$$\frac{\partial V^-}{\partial t} + \frac{1}{2}\sigma^{-2}S^2\frac{\partial^2 V^-}{\partial S^2} + r\frac{\partial V^-}{\partial S} - rV^- + \lambda^+(V^+ - V^-) = 0$$

for the value V^- when the volatility is σ^-.

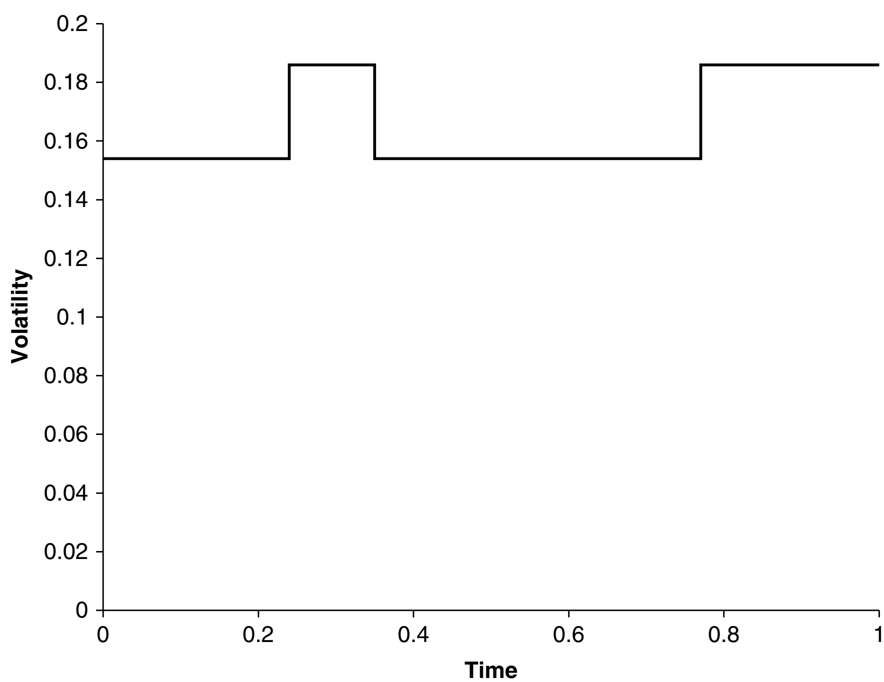


Figure 57.7 Jump volatility.

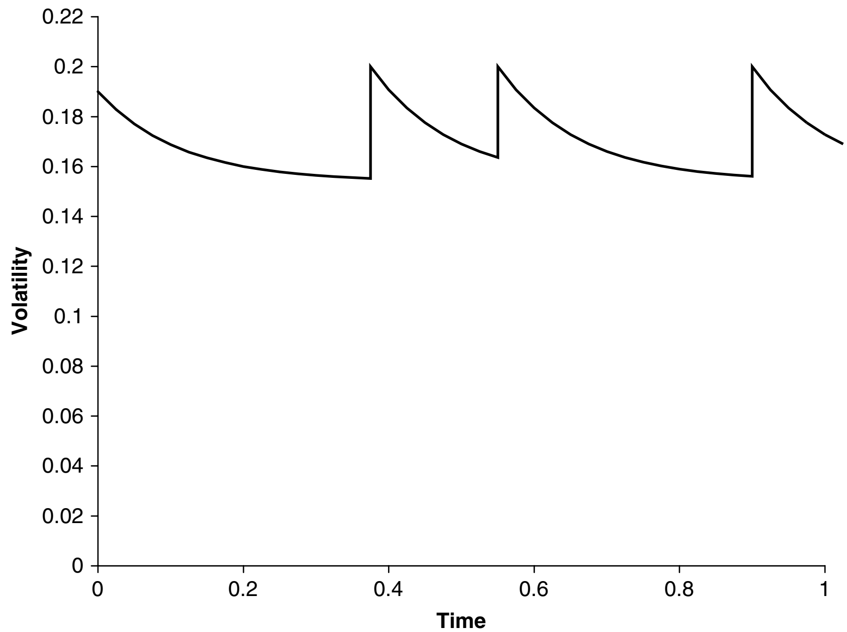


Figure 57.8 Jump volatility with exponential decay.

57.10 JUMP VOLATILITY WITH DETERMINISTIC DECAY

A more sophisticated jump process for the volatility, that resembles the real behavior of volatility, also contains exponential decay of the volatility after the jump. We can write

$$\sigma(\tau) = \sigma^- + (\sigma^+ - \sigma^-)e^{-\nu\tau},$$

where τ is the time since the last sudden jump in the volatility and ν is a decay parameter. It doesn't actually matter what form this function takes as long as it depends only on τ (and S and t if you want). At any time, governed by a Poisson process with intensity λ, the volatility can jump from its present level to σ^+. A realization of this process for σ is shown in Figure 57.8.

The value of an option is given by $V(S, t, \tau)$, the solution of

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial \tau} + \tfrac{1}{2}\sigma(\tau)^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \lambda(V(S, t, 0) - V(S, t, \tau)) = 0.$$

Note that the time since last jump τ is incremented at the same rate as real time t. (In this 'value' means that we have delta hedged with the underlying to eliminate risk due to the movement of the asset but we have taken real expectations with respect to the volatility jump.)

57.11 SUMMARY

Jump diffusion models undoubtedly capture a real phenomenon that is missing from the Black–Scholes model. And they are increasingly being used in practice. However, there are still three main problems with this model: Difficulty in parameter estimation, solution, and impossibility of perfect hedging.

In order to use any pricing model one needs to be able to estimate parameters. In the lognormal model there is just the one parameter to estimate. This is just the right number. More than one parameter is too much work, no parameters to estimate and trading could be done by machine. The jump diffusion model in its simplest form needs an estimate of probability of a jump, measured by λ, and its size, J. This can be made more complicated by having a distribution for J.

The governing equation is no longer a diffusion equation (about the easiest problem to solve numerically), but a partial integro-differential equation. The equation does not just contain local derivatives of the value with respect to its variables, but now links option values at different asset values. The numerical solution of such an equation is certainly not impossible, but is harder than the solution of the basic Black–Scholes equation.

Finally, perfect risk-free hedging is impossible when there are jumps in the underlying. This is because of the non-local nature of the option pricing equation. We have seen two approaches to this hedging, neither of which matches the elegance of the Black–Scholes hedge. The use of a jump-diffusion model acknowledges that one's hedge is less than perfect, which bothers some people. If one sticks to the Black–Scholes model then one can pretend to be hedging perfectly. Of course, the reality of the situation is that there are many reasons why risk-free hedging is impossible, nevertheless the 'eyes wide shut' approach has become market standard.

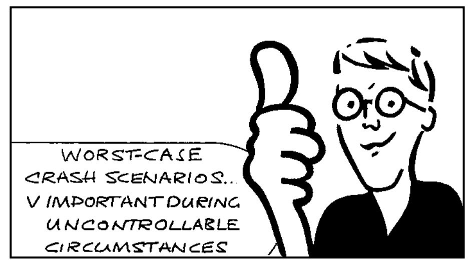
FURTHER READING

- The original jump-diffusion model in finance was due to Merton (1976).
- An impressive tome on the modeling of extremal events is by Embrechts, Klüppelberg & Mikosch (1997).
- Jump volatility is described by Naik (1993) and Lewis (2004a).
- Lewis (2002) gives a great explanation of various jump models, and in Lewis (2003a,b,d) discusses American options with jumps.
- Lewis (2003c, 2004b) analyzes barrier options in the presence of jumps, and Lewis (2004d) analyzes Asian options.
- Sepp & Skachkov (2003) discuss pricing in the presence of all kinds of jumps.
- See d'Halluin, Pooley & Forsyth (2004) for the numerical solution of jump-diffusion problems.
- Penaud (2004) looks at the pricing of portfolios of barrier options in the presence of jumps.

CHAPTER 58
crash modeling

In this Chapter...

- how to price contracts in a worst-case scenario when there are crashes in the prices of underlyings
- how to reduce the effect of these crashes on your portfolio; the Platinum Hedge



58.1 INTRODUCTION

Jump diffusion models have two weaknesses: They don't allow you to hedge and the parameters are very hard to measure. Nobody likes a model that tells you that hedging is impossible (even though that may correspond to common sense) and in the classical jump-diffusion model of Merton (1976) the best that you can do is a kind of average hedging. It may be quite easy to estimate the impact of a rare event such as a crash, but estimating the probability of that rare event is another matter. In this chapter we discuss a model for pricing and hedging a portfolio of derivatives that takes into account the effect of an extreme movement in the underlying but we will make *no assumptions about the timing of this 'crash' or the probability distribution of its size*, except that we put an upper bound on the latter. This effectively gets around the difficulty of estimating the likelihood of the rare event. The pricing follows from the assumption that the worst scenario actually happens, i.e. the size and time of the crash are such as to give the option its worst value. And hedging, delta and static hedging, will continue to play a key role. The optimal static hedge follows from the desire to make the best of this worst value. This, latter, static hedging follows from the desire to optimize a portfolio's value. I also show how to use the model to evaluate the value at risk for a portfolio of options.

58.2 VALUE AT RISK

The true business of a financial institution is to manage risk.

The trader manages 'normal event' risk, where the world operates close to a Black–Scholes one of random walks and dynamic hedging. The institution, however, views its portfolio on a 'big picture' scale and focuses on 'tail events' where liquidity and large jumps are important (Figure 58.1).

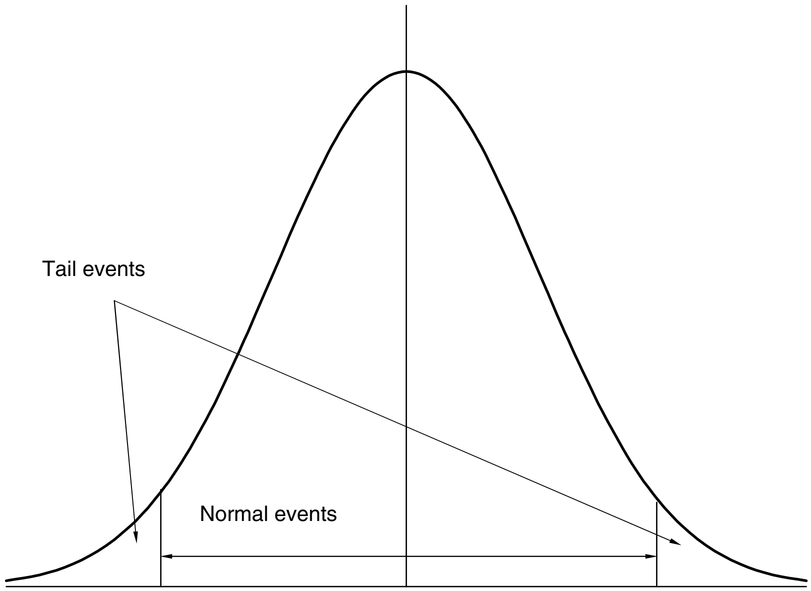


Figure 58.1 'Normal events' and 'tail events.'

Value at Risk (VaR) is a measure of the potential losses due to a movement in underlying markets. It usually has associated with it a time frame and an estimate of the maximum sudden change thought likely in the markets. There is also a 'confidence interval'; for example, the daily VaR is $15 million with a degree of confidence of 99%. The details of VaR and its measurement are discussed in Chapter 19.

A more general and more encompassing definition of VaR will give a useful tool to both book runners and senior management. A true measure of the risk in a portfolio will answer the question 'What is the value of any realistic market movement to my portfolio?'

The approach taken here in finding the value at risk for a portfolio is to model the cost to a portfolio of a crash in the underlying. I show how to value the cost of a crash in a worst-case scenario, and also how to find an optimal static hedge to minimize this cost and so reduce the value at risk.

58.3 A SIMPLE EXAMPLE: THE HEDGED CALL

To motivate the problem and model, consider this simple example. You hold a long call position, delta hedged in the Black–Scholes fashion. What is the worst that can happen, in terms of crashes, for the value of your portfolio? One might naively say that a crash is bad for the portfolio, after all, look at the Black–Scholes value for a call as a function of the underlying, the lower the underlying the lower the call value. Wrong. Remember you hold a *hedged* position; look at Figure 58.2 to see the value of the option, the short asset position and the whole portfolio. The last is the bold line in the figure. Observe that the position is currently delta neutral. Also observe that the portfolio's value is currently at its minimum, at $S = 100$; a sudden fall (or, indeed, a rise) will result in a higher portfolio value: A crash is beneficial. If we are assuming a worst-case scenario, then the worst that could happen is that there is no crash. Changing all the

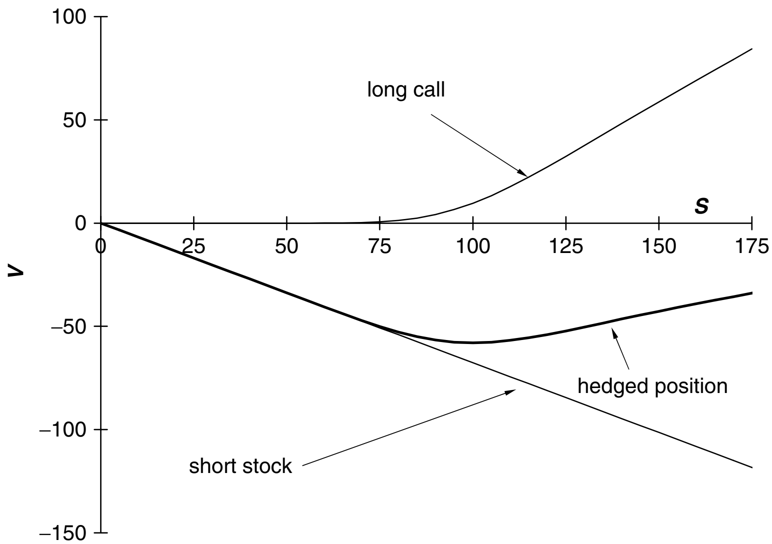


Figure 58.2 The delta-hedged long call portfolio according to Black–Scholes.

signs to consider a short call position we find that a crash is bad, but how do we find the worst case? If there is going to be one crash of 10% when is the worst time for this to happen? This is the motivation for the model below. Note first that, generally speaking, a positive gamma position benefits from a crash, while a negative gamma position loses.

58.4 **A MATHEMATICAL MODEL FOR A CRASH**

The main idea in the following model is simple. We assume that the worst will happen. We value all contracts assuming this, and then, unless we are very unlucky and the worst does happen, we will be pleasantly surprised. In this context, 'pleasantly surprised' means that we make more money than we expected. We can draw an important distinction between this model and the models of Chapter 57, the jump diffusion models. In the latter we make bold statements about the frequency and distribution of jumps and finally take expectations to arrive at a value for a derivative. Here *we make no statements about the distribution of either the jump size or when it will happen*. At most, the number of jumps is limited. Finally, we examine the worst-case scenario so that no expectations are taken.

I will model the underlying asset price behavior as the classical binomial tree, but with the addition of a third state, corresponding to a large movement in the asset. So, really, we have a trinomial walk but with the lowest branch being to a significantly more distant asset value. The up and down diffusive branches are modeled in the usual binomial fashion. For simplicity, assume that the crash, when it happens, is from S to $(1-k)S$ with k given; this assumption will later be dropped to allow k to cover a range of values, or even to allow a dramatic rise in the value of the underlying. Introduce the subscript 1 to denote values of the option before the crash, i.e. with one crash allowed, and 0 to denote values after. Thus V_0 is the value of the option position after the crash. This is a function of S and t and, since I am only permitting one crash, V_0 must be exactly the Black–Scholes option value.

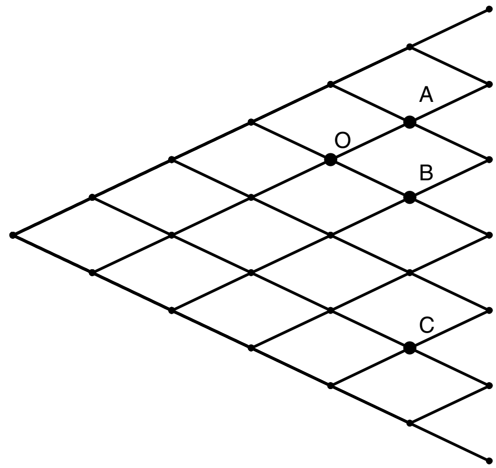


Figure 58.3 The tree structure.

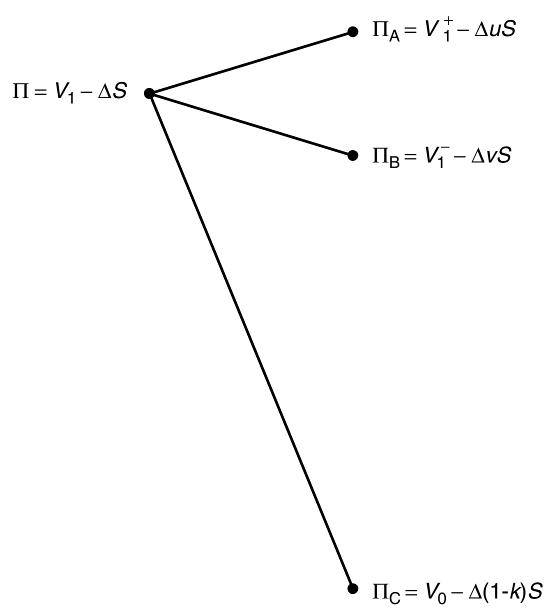


Figure 58.4 The tree and portfolio values.

As shown in Figure 58.3, if the underlying asset starts at value S (point O) it can go to one of three values: uS, if the asset rises; vS, if the asset falls; $(1-k)S$, if there is a crash. These three points are denoted by A, B and C respectively. The values for uS and vS are chosen in the usual manner for the traditional binomial model (see Chapter 15).

Before the asset price moves, we set up a 'hedged' portfolio, consisting of our option position and $-\Delta$ of the underlying asset. At this time our option has value V_1. We must find both an optimal Δ and then V_1. The hedged portfolio is shown in Figure 58.4.

A time δt later the asset value has moved to one of the three states, A, B or C and at the same time the option value becomes either V_1^+ (for state A), V_1^- (for state B) or the Black–Scholes value V_0 (for state C).

The change in the value of the portfolio, between times t and $t + \delta t$ (denoted by $\delta\Pi$) is given by the following expressions for the three possible states:

$$\delta\Pi_A = V_1^+ - \Delta uS + \Delta S - V_1 \text{ (diffusive rise)}$$

$$\delta\Pi_B = V_1^- \Delta vS + \Delta S - V_1 \text{ (diffusive fall)}$$

$$\delta\Pi_C = V_0 + \Delta kS - V_1 \text{ (crash).}$$

These three functions are plotted against Δ in Figures 58.5 and 58.6. I will explain the difference between the two figures very shortly. My aim in what follows is to choose the hedge ratio Δ so as to minimize the pessimistic, worst outcome among the three possible.

There are two cases to consider, shown in Figures 58.5 and 58.6. The former, Case I, is when the worst-case scenario is not the crash but the simple diffusive movement of S. In this case V_0 is sufficiently large for a crash to be beneficial:

$$V_0 \geq V_1^+ + (S - uS - kS)\frac{V_1^+ - V_1^-}{uS - vS}. \tag{58.1}$$

If V_0 is smaller than this, then the worst scenario is a crash; this is Case II.

58.4.1 Case I: Black–Scholes Hedging

Refer to Figure 58.5. In this figure we see the three lines representing $\delta\Pi$ for each of the moves to A, B and C. Pick a value for the hedge ratio Δ (for example, see the dashed vertical line in Figure 58.5), and determine on which of the three lines lies the worst possible value for $\delta\Pi$ (in the example in the figure, the point is marked by a cross and lies on the A line). Change

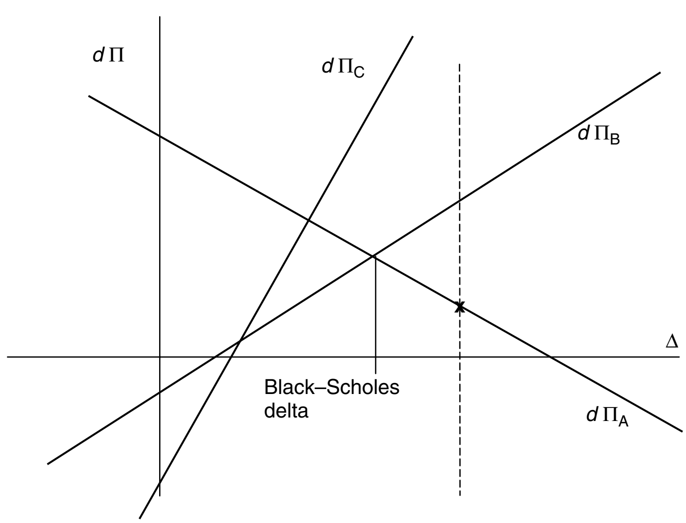


Figure 58.5 Case I: worst case is diffusive motion.

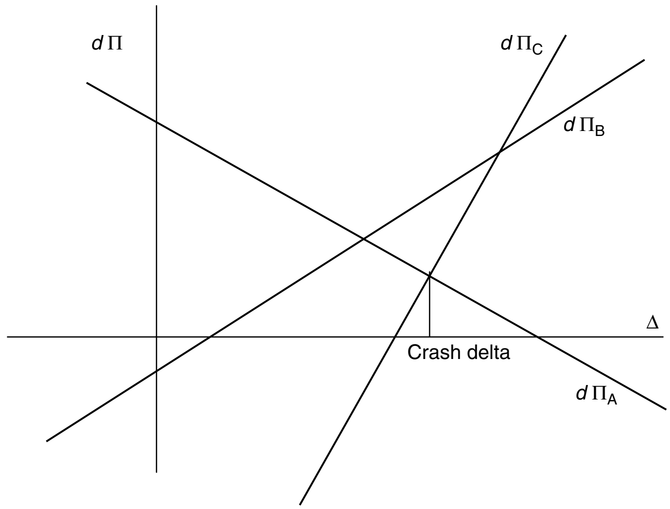


Figure 58.6 Case II: worst case is a crash.

your value of Δ to maximize this worst value. In other words, we want to choose Δ to put us as high up the envelope of the three lines as possible.

In this case the maximal-lowest value for $\delta\Pi$ occurs at the point where

$$\delta\Pi_A = \delta\Pi_B,$$

that is

$$\Delta = \frac{V_1^+ - V_1^-}{uS - vS}. \tag{58.2}$$

This will be recognized as the expression for the hedge ratio in a Black–Scholes world.

Having chosen Δ, we now determine V_1 by setting the return on the portfolio equal to the risk-free interest rate. Thus we set

$$\delta\Pi_A = r\Pi\,\delta t$$

to get

$$V_1 = \frac{1}{1 + r\,\delta t}\left(V_1^+ + (S - uS + rS\,\delta t)\frac{V_1^+ - V_1^-}{uS - vS}\right). \tag{58.3}$$

This is the equation to solve if we are in Case I. Note that it corresponds exactly to the usual binomial version of the Black–Scholes equation; there is no mention of the value of the portfolio at the point C. As δt goes to zero, (58.2) becomes $\partial V/\partial S$ and Equation (58.3) becomes the Black–Scholes partial differential equation.

58.4.2 Case II: Crash Hedging

Refer to Figure 58.6. In this case the value for V_0 is low enough for a crash to give the lowest value for the jump in the portfolio. We therefore choose Δ to maximize this worst case. Again, we want to choose Δ to put us as high up the envelope of the three lines as possible. Thus we choose

$$\delta\Pi_A = \delta\Pi_C,$$

that is,

$$\Delta = \frac{V_0 - V_1^+}{S - uS - kS}. \tag{58.4}$$

Now set

$$\delta\Pi_A = r\Pi\,\delta t$$

to get

$$V_1 = \frac{1}{1 + r\,\delta t}\left(V_0 + S(k + r\,\delta t)\frac{V_0 - V_1^+}{S - uS - kS}\right). \tag{58.5}$$

This is the equation to solve when we are in Case II. Note that this is different from the usual binomial equation, and does not give the Black–Scholes partial differential equation as δt goes to zero (see later in the chapter). Also (58.4) is not the Black–Scholes delta. To appreciate that delta hedging is not necessarily optimal, consider the simple example of the butterfly spread. If the butterfly spread is delta hedged on the right 'wing' of the butterfly, where the delta is negative, a large fall in the underlying will result in a large loss from the hedge, whereas the loss in the butterfly spread will be relatively small. This could result in a negative value for a contract, even though its payoff is everywhere positive.

58.5 AN EXAMPLE

All that remains to be done is to solve equations (58.3) and (58.5) (which one is valid at any asset value and at any point in time depends on whether or not (58.1) is satisfied). This is easily done by working backwards down the tree from expiry in the usual binomial fashion.

As an example, examine the cost of a 15% crash on a portfolio consisting of the call options in Table 58.1.

At the moment the portfolio only contains the first two options. Later I will add some of the third option for static hedging, that is when the bid-ask prices will concern us. The volatility of the underlying is 17.5% and the risk-free interest rate is 6%.

The solution to the problem is shown in Figure 58.7. Observe how the value of the portfolio assuming the worst (21.2 when the spot is 100), is lower than the Black–Scholes value (30.5). This is especially clear where the portfolio's gamma is highly negative. This is because when

Table 58.1 Available contracts.

Strike	Expiry	Bid	Ask	Quantity
100	75 days			–3
80	75 days			2
90	75 days	11.2	12	0

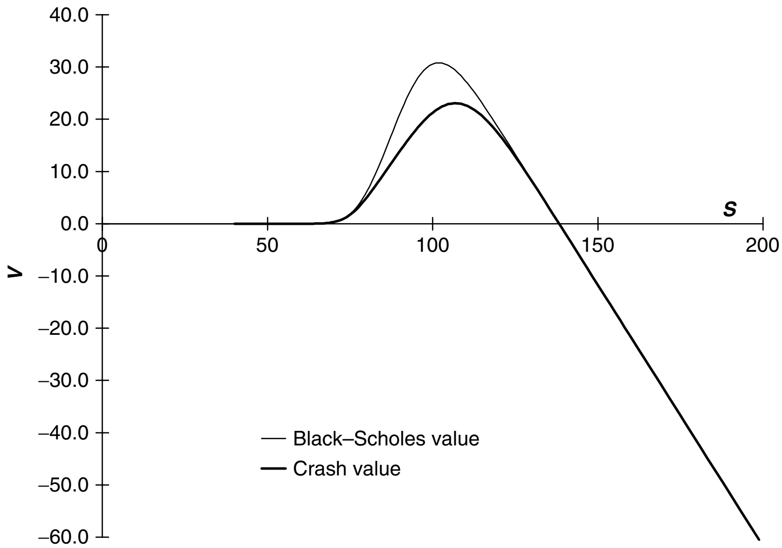
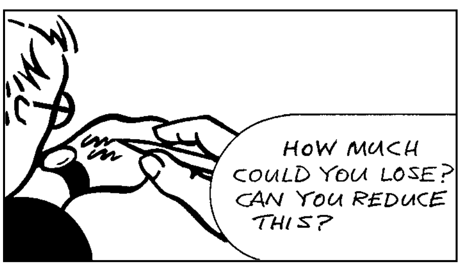


Figure 58.7 Example showing crash value and Black–Scholes value.

the gamma is positive, a crash is beneficial to the portfolio's value. When the gamma is close to zero, the delta hedge is very accurate and the option is insensitive to a crash. If the asset price is currently 100, the difference between the before and after portfolio values is $30.5 - 21.2 = 9.3$. This is the 'Value at Risk' under the worst-case scenario.



58.6 OPTIMAL STATIC HEDGING: VAR REDUCTION

The 9.3 value at risk is due to the negative gamma around the asset price of 100. An obvious hedging strategy that will offset some of this risk is to buy some positive gamma as a 'static' hedge. In other words, we should buy an option or options having a counterbalancing effect on the value at risk. We are willing to pay a premium for such an option. We may even pay more than the Black–Scholes fair value for such a static hedge because of the extra benefit that it gives us in reducing our exposure to a crash. Moreover, if we have a choice of contracts with which to hedge statically we should buy the most 'efficient' one. To see what this means consider the above example in more detail.

Recall that the value of the initial portfolio under the worst-case scenario is 21.2. How many of the 90 calls should we buy (for 12) or sell (for 11.2) to make the best of this scenario? Suppose that we buy λ of these calls. We will now find the optimal value for λ.

The cost of this hedge is

$$\lambda C(\lambda)$$

where $C(\lambda)$ is 12 if λ is positive and 11.2 otherwise (see Table 58.1). Now solve Equations (58.3) and (58.5) with the final total payoffs

$$V_0(S, T) = V_1(S, T) = 2\max(S - 80, 0) - 3\max(S - 100, 0) + \lambda\max(S - 90, 0).$$

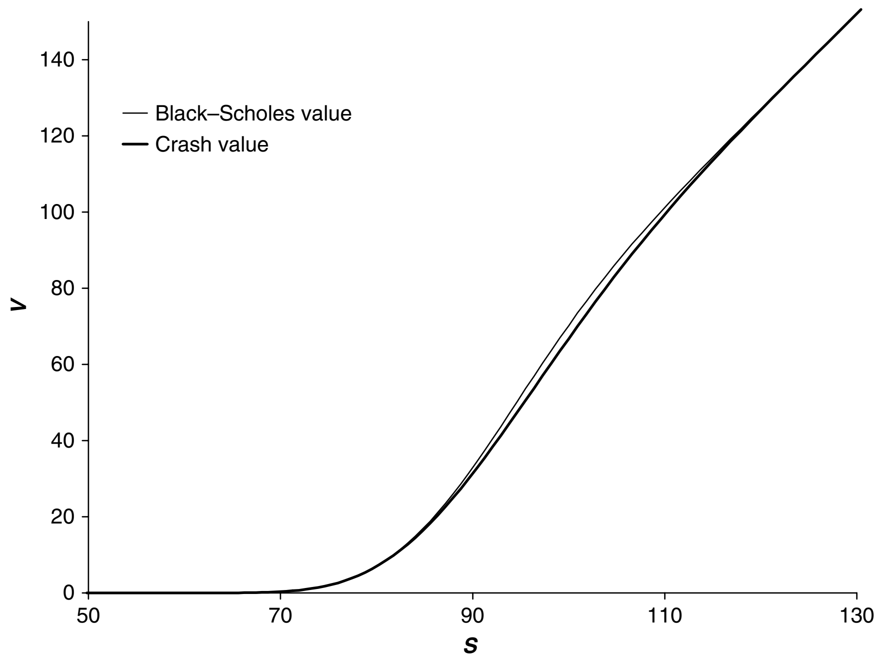


Figure 58.8 Optimally-hedged portfolio, before and after crash.

This is the payoff at time T for the statically hedged portfolio. The *marginal* value of the original portfolio (that is, the portfolio of the 80 and 100 calls) is therefore

$$V_1(100, 0) - \lambda C(\lambda) \tag{58.6}$$

i.e. the worst-case value for the new portfolio less the cost of the static hedge. The arguments of the before-crash option value are 100 and 0 because they are today's asset value and date. The optimality in this hedge arises when we choose the quantity λ to maximize the value, expression (58.6). With the bid-ask spread in the 90 calls being 11.2–12, we find that buying 3.5 of the calls maximizes expression (58.6). The value of the new portfolio is 70.7 in a Black–Scholes world and 65.0 under our worst-case scenario. The value at risk has been reduced from 9.3 to $70.7 - 65 = 5.7$. The optimal portfolio values before and after the crash are shown in Figure 58.8. The optimal static hedge is known as the **Platinum Hedge**.

The issues of static hedging and optimal static hedging are covered in detail in Chapter 60.

58.7 CONTINUOUS-TIME LIMIT

If we let $\delta t \to 0$ in Equations (58.1), (58.2), (58.3), (58.4) and (58.5) we find that the Black–Scholes equation is still satisfied by $V_1(S, t)$ but we also have the constraint

$$\boxed{V_1(S, t) - kS\frac{\partial V_1}{\partial S}(S, t) \leq V_0(S(1 - k), t)}$$

Such a problem is similar in principal to the American option valuation problem, where we also saw a constraint on the derivative's value. Here the constraint is more complicated. To this we must add the condition that the first derivative of V_1 must be continuous for $t < T$.

58.8 A RANGE FOR THE CRASH

In the above model, the crash has been specified as taking a certain value. Only the timing was left to be determined for the worst-case scenario. It is simple to allow the crash to cover a range of values, so that S goes to $(1-k)S$ where

$$k^- \le k \le k^+.$$

A negative k^- corresponds to a rise in the asset.

In the discrete setting the worst-case option value is given by

$$V_1 = \min_{k^- \le k \le k^+} \left(\frac{1}{1 + r\,\delta t} \left(V_0 + S(k + r\,\delta t) \frac{V_0 - V_1^+}{S - uS - kS} \right) \right).$$

This contains the $\min(\cdot)$ because we want the worst-case crash. When a crash is beneficial we still have (58.3).

58.9 MULTIPLE CRASHES

The model described above can be extended in many ways, one of the most important is to consider the effect of multiple crashes. I describe two possibilities below. The first puts a constraint on the total number of crashes in a time period; there can be three crashes within the horizon of one year, say. The second puts a limit on the time between crashes; there cannot be another crash if there was a crash in the last six months, say.

58.9.1 Limiting the Total Number of Crashes

We will allow up to N crashes. We make no statement about the time these occur. We will assume that the crash size is given, allowing a fall of $k\%$. This can easily be extended to a range of sizes, as described above. Introduce the functions $V_i(S, t)$ with $i = 0, 1, \ldots, N$, such that V_i is the value of the option with i more crashes still allowed. Thus, as before, V_0 is the Black–Scholes value.

We must now solve N coupled equations of the following form. If

$$V_{i-1} \ge V_i^+ + (S - uS - kS)\frac{V_i^+ - V_i^-}{uS - vS}$$

then we are in Case I, a crash is beneficial and is assumed not to happen. In this case we have

$$V_i = \frac{1}{1 + r\,\delta t} \left(V_i^+ + (S - uS + rS\,\delta t) \frac{V_i^+ - V_i^-}{uS - vS} \right).$$

Otherwise a crash is bad for the hedged option; this is Case II. We then have

$$V_i = \frac{1}{1 + r\,\delta t} \left(V_{i-1} + S(k + r\,\delta t) \frac{V_{i-1} - V_i^+}{S - uS - kS} \right).$$

In continuous time the equations become

$$\frac{\partial V_i}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_i}{\partial S^2} + rS\frac{\partial V_i}{\partial S} - rV_i = 0$$

for $i = 0, \dots, N$, subject to

$$V_i(S,t) - kS\frac{\partial V_i}{\partial S}(S,t) \le V_{i-1}(S(1-k),t)$$

for $i = 1, \dots, N$. Each of the V_i has the same final condition, representing the payoff at expiry.

58.9.2 Limiting the Frequency of Crashes

Finally, we model a situation where the time between crashes is limited; if there was a crash less than a time ω ago another is not allowed.

This is slightly harder than the N-crash model and we have to introduce a new variable τ measuring the time since the last crash. We now have two functions to consider $V_c(S,t)$ and $V_n(S,t,\tau)$. The former is the worst-case option value when a crash is allowed (and therefore we don't need to know how long it has been since the last crash) and the latter is the worst-case option value when a crash is not yet allowed.

The governing equations, which are derived in the same way as the original crash model, are, in continuous time, simply

$$\frac{\partial V_c}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_c}{\partial S^2} + rS\frac{\partial V_c}{\partial S} - rV_c = 0$$

subject to

$$V_c(S,t) - kS\frac{\partial V_c}{\partial S}(S,t) \le V_n(S(1-k),t,0),$$

and for $V_n(S,t,\tau)$,

$$\frac{\partial V_n}{\partial t} + \frac{\partial V_n}{\partial \tau} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_n}{\partial S^2} + rS\frac{\partial V_n}{\partial S} - rV_n = 0$$

with the condition

$$V_n(S,t,\omega) = V_c(S,t).$$

Observe how the time τ and real time t increase together in the equation when a crash is not allowed.

58.10 **CRASHES IN A MULTI-ASSET WORLD**

When we have a portfolio of options with many underlyings we can still examine the worst-case scenario, but we have two choices. Either (a) we allow a crash to happen in any underlyings completely independently of all other underlyings or (b) we assume some relationship between the assets during a crash. Clearly the latter is not as bad a worst case as the former. It is also easier to write down, so we will look at that model only. Assuming that all assets fall simultaneously by the same percentage k we have

$$V_1(S_1, \dots, S_n, t) - k\sum_{i=1}^{n} S_i \frac{\partial V_1}{\partial S_i}(S_1, \dots, S_n, t) \le V_0((1-k)S_1, \dots, (1-k)S_n, t, t).$$

We examine stock market crashes using CrashMetrics in Chapter 42.

58.11 FIXED AND FLOATING EXCHANGE RATES

Many currencies are linked directly to the currency in another country. Some countries have their currency linked to the US dollar; the Argentine Peso is tied at a rate of one to one to the dollar.

Once an exchange rate is fixed in this way the issue of fluctuating rates becomes a credit risk issue. All being well, the exchange rate will stay constant with all the advantages that stability brings. If economic conditions in the two countries start to diverge then the exchange rate will come under pressure. In Figure 58.9 is a plot of the possible exchange rate, showing a fixed rate for a while, followed by a sudden discontinuous drop and then a random fluctuation. How can we model derivatives of the exchange rate? The models of this chapter are ideally suited to this situation.

I'm going to ignore interest rates in the following. This is because of the complex issues this would otherwise raise. For example, the pressure on the exchange rate and the decoupling of the currencies would be accompanied by changing interest rates. This can be modeled but would distract us from the application of the crash model.

While the exchange rate is fixed, before the 'crash,' the price of an option, $V_1(S, t)$, satisfies

$$\frac{\partial V_1}{\partial t} = 0,$$

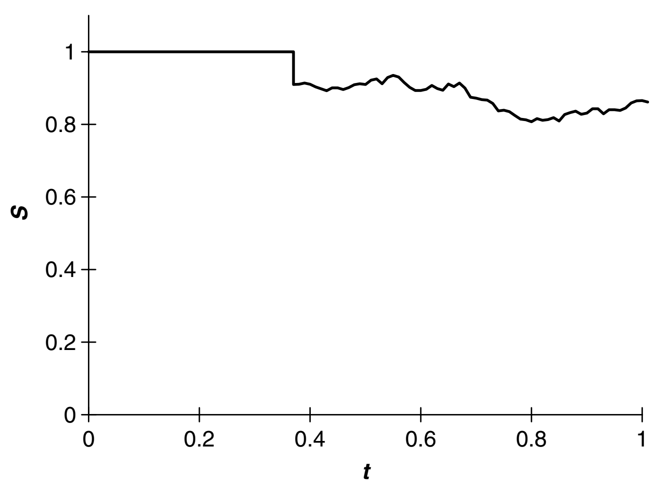


Figure 58.9 Decoupling of an exchange rate.

since I am assuming zero interest rates. Here S is the exchange rate. After the 'crash' we have

$$\frac{\partial V_0}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_0}{\partial S^2} = 0.$$

(But with what volatility?) This is just the Black–Scholes model, with the relevant Black–Scholes value for the particular option payoff.

The worst-case crash model is now almost directly applicable. I will leave the details to the reader.

58.12 **SUMMARY**

I have presented a model for the effect of an extreme market movement on the value of portfolios of derivative products. This is an alternative way of looking at value at risk. I have shown how to employ static hedging to minimize this risk. In conclusion, note that the above is not a jump-diffusion model since I have deliberately not specified any probability distribution for the size or the timing of the jump: we model the worst-case scenario.

We have examined several possible models of crashes, of increasing complexity. One further thought is that we have not allowed for the rise in volatility that accompanies crashes. This can be done with ease. There is no reason why the after-crash model (V_0 in the simplest case above) cannot have a different volatility from the before-crash model. Of particular interest is the final model where there is a minimum time between crashes; we could easily have the volatility post crash being a decaying function of the time since the crash occurred, τ. This would involve no extra computational effort.

FURTHER READING

- For further details about crash modeling see Hua & Wilmott (1997) and Hua (1997).
- Derman & Zou (1997) describe and model the behavior of implied volatility after a large move in an index.

CHAPTER 59
speculating with options

In this Chapter...

- how to find the present value of the *real* expected payoff (and why you should want to know this)
- several ways to model asset price drift
- how to close your option position optimally
- when and when not to hedge

59.1 INTRODUCTION

Almost everything I have shown you so far is about finding pricing equations for derivative contracts using hedging arguments. There is a very powerful reason why such arguments should give option prices in the market regardless of whether or not investors actually hedge the option. The reason is simple: If the option is not priced at the Black–Scholes fair value and hedging is possible then either

- the option price is too low, in which case someone will buy it, hedge away the risk and make a riskless profit or
- the option price is too high, in which case someone will artificially replicate the payoff and charge less for this contract.

Both of these are examples of market inefficiencies which would disappear quickly in practice (to a greater or lesser extent).

However, a few minutes spent on a typical trading floor will convince you that option contracts are often bought for speculation and that the value of an option to a speculator and the choice of expiry date and strike price depend strongly on his **market view**, something that is irrelevant in the Black–Scholes world. Moreover, many OTC and other contracts are used to offset risks outside the market, and if every derivative is hedged by both writer and purchaser then the whole thing is pointless, or at best a series of hedging and modeling competitions. In this chapter I show possible ways in which a speculator can choose an option contract so as to profit from his view if it turns out to be correct.

Throughout this chapter I use the word 'value' to mean the *worth of an option to a speculator* and this will not necessarily be the Black–Scholes 'fair' value. To get around the obvious criticism that what follows is nonsense in a complete market, in which delta hedging is possible, I'm going to take the point of view that the investor can value a contract at other than the Black–Scholes value because of market incompleteness, for whatever reason you like. The cause of incompleteness may be transaction costs, restrictions on sales or purchases, uncertain parameters, general model errors etc. Thus any 'arbitrage' that you may see because values are different from Black–Scholes cannot be exploited by our investor. He is going to buy a contract as an investment, a risky investment.

One of the motivations for this chapter is the observation that in the classical Black–Scholes theory the drift of the asset plays no part. Two people will agree on the price of an option even if they differ wildly in their estimates for μ—indeed, they may not even have estimates for μ—as long as they agree on the other parameters such as the volatility.

59.2 A SIMPLE MODEL FOR THE VALUE OF AN OPTION TO A SPECULATOR

Imagine that you hold very strong views about the future behavior of a particular stock: You have an estimate of the volatility and you believe that for the next twelve months the stock will have an upwards drift of 15%, say, and that this is much greater than the risk-free rate. How can you benefit from this view if it is correct? One simple way is to follow the principles of Modern Portfolio Theory, Chapter 18, and buy a call, just out of the money, maybe, perhaps with an expiry of about one year because this may have an appealing risk/return profile (also see later in this chapter). You might choose an out of-the-money option because it is cheap, but not so far out of the money that it is likely to expire worthless. You might choose a twelve-month contract to benefit from as much of the asset price rise as possible. Now imagine it is a couple of months later, and the asset has fallen 5% instead of rising. What do you do? Maybe, the stock *did* have a drift of 15% but, since we only see one realization of the random walk, the volatility caused the drop. Alternatively, maybe you were wrong about the 15%.

This example illustrates several points.

- How do you determine the future parameters for the stock: The drift and volatility?
- How do you subsequently know whether you were right?
- And if not, what should you do about it?
- If you have a good model for the future random behavior of a stock, how can you use it to measure an option's value *to you*?
- Which option do you buy?

A thorough answer to the first of these questions is outside the scope of this book; but I have made suggestions in that direction. The other four questions fall more into the area of modeling and we discuss them here.

59.2.1 The Present Value of Expected Payoff

Speculation is the opposite side of the coin to hedging. The speculator is taking risks and hopes to profit from this risk by an amount greater than the risk-free rate. When gambling like this, it is natural to ask what is your expected profit. In option terms, this means we would ask what is the present value of the expected payoff. I say 'present value' because the payoff, if it comes, will be at expiry (assuming that we do not exercise early if the option is American, or close the position).

If the asset price random walk is

$$dS = \mu S\,dt + \sigma S\,dX$$

and the option has a payoff of

$$\max(S - E, 0)$$

at time T then the *present value of the expected payoff*[1] satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} - rV = 0$$

with

$$V(S, T) = \max(S - E, 0).$$

This is related to the ideas in Chapter 10 for transition density functions.

This equation differs from the Black–Scholes equation in having a μ instead of an r in front of the delta term. Seemingly trivial, this difference is of fundamental importance. The replacement of μ by r is the basis of 'risk-neutral valuation.' The absence of μ from the Black–Scholes option pricing equations means that the only asset price parameters needed are the volatility and the dividend structure. For the above equation, however, we need to know the drift rate μ in order to calculate the expected payoff.

Let us return to the earlier example, and fill in some details. Suppose we buy a call, struck at 100, with an expiry of one year, the underlying has a volatility of 20% and, we believe, a drift of 15%. The interest rate is 5%. In Figure 59.1 is shown the Black–Scholes value of the option (the lower curve) together with the present value of the *real* expected payoff (the upper curve) both plotted against S. These two curves would be identical if μ and r were equal. The lower curve is a risk-neutral expectation and the upper curve a real expectation.

It is easy to show, and financially obvious, that if $\mu > r$ then the present value of the expected payoff for a call is greater than the Black–Scholes fair value. In other words, if we expect the stock to drift upwards with a rate higher than the interest rate then we expect to make more money than if we had invested in bonds. Here the word 'expect' is of paramount importance: There is nothing certain about this investment. The greater expected return is compensation for the greater risk of speculation, absent in a hedged portfolio. Nevertheless, for the speculator the expected payoff is an important factor in deciding which option to buy. The other important factor is a measure of the risk; when we know the risk and the return for an option we are in a position to apply ideas from Modern Portfolio Theory, Chapter 18, for example.

59.2.2 Standard Deviation

As I have said, the return from an unhedged option position is uncertain. That uncertainty can be measured by the standard deviation of the return about its mean level.[2] This standard deviation can be determined as follows. Suppose that today's date is $t = 0$, and introduce the

[1] This is not dissimilar to how options used to be priced in the 1960s before Black and Scholes.

[2] This is not entirely satisfactory for many products, especially those with a non-unimodal payoff distribution.

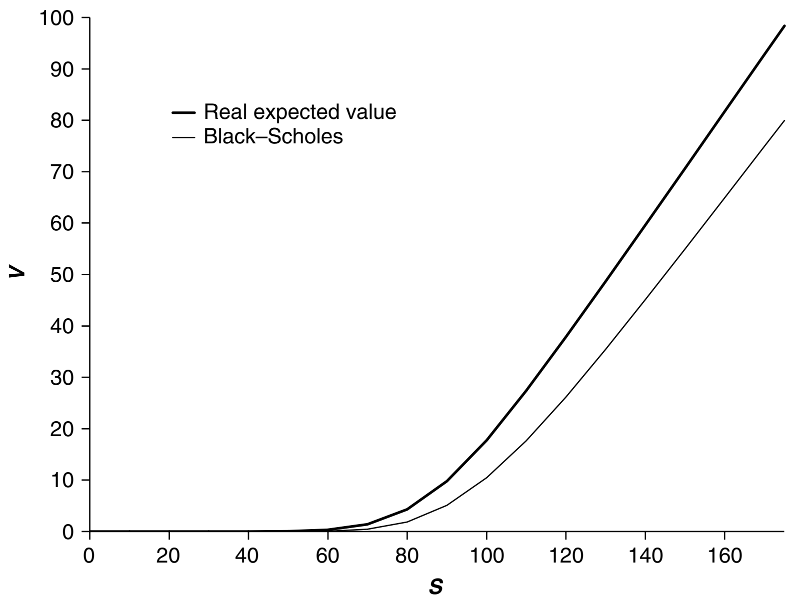


Figure 59.1 The Black–Scholes value of a call option (the lower curve) and the present value of the expected payoff (the upper curve). See text for data.

function $G(S,t)$ as the expected value of the square of the present value of the payoff. This function, being an expectation satisfies the backward Kolmogorov equation (see Chapter 10):

$$\frac{\partial G}{\partial t}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 G}{\partial S^2}+\mu S\frac{\partial G}{\partial S}=0 \tag{59.1}$$

(which, of course, is simply the Black–Scholes equation with no discounting and with μ instead of r).[3] The final condition is

$$G(S,T)=\left(e^{-rT}\max(S-E,0)\right)^2.$$

Then from its definition, the standard deviation today is given by

$$\sqrt{G(S,0)-(V(S,0))^2}.$$

In Figure 59.2 is shown this standard deviation plotted against S, using the same data as in Figure 59.1.

A natural way to view these results is in a risk/reward plot. In this model the reward is the logarithm of the ratio of the present value of the expected payoff to the cost of the option, divided by the time to expiry of the option. But how much will the option cost? A not unreasonable assumption is that the rest of the market is using Black–Scholes to value the option so we shall assume that the cost of the option is simply the Black–Scholes value V_{BS}. The risk is

[3] The astute reader will notice that there is no $-rG$ in this 'present value' calculation; we have absorbed this into the final condition. Read on.

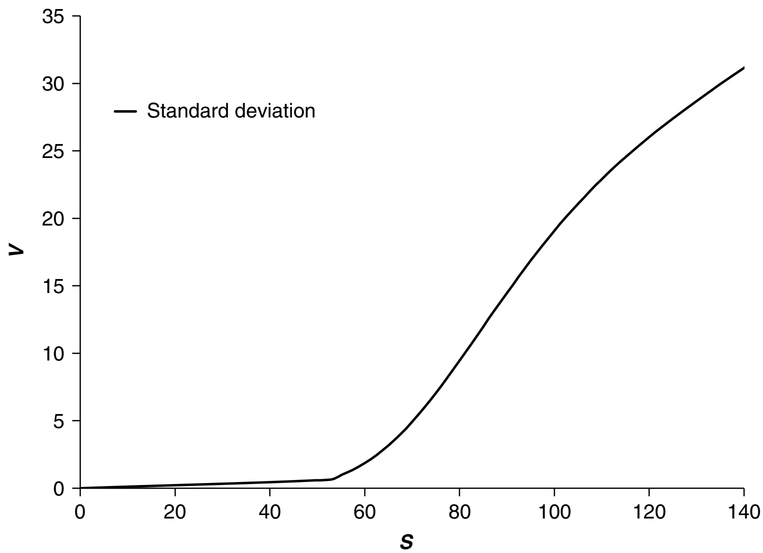


Figure 59.2 The standard deviation of the present value of the expected payoff for a call option. See text for data.

the standard deviation, again scaled with the Black–Scholes value of the option and with the square root of the time to expiry. Thus we define

$$\text{Reward} = \frac{1}{T}\log\left(\frac{V(S,0)}{V_{BS}(S,0)}\right)$$

and

$$\text{Risk} = \frac{\sqrt{G(S,0)-(V(S,0))^2}}{V_{BS}(S,0)\sqrt{T}}.$$

These definitions, in particular the scalings with time, have been chosen to tie in with the traditional measures of reward and risk.

In Figure 59.3 we show a plot of this reward versus the risk for a call option and the same data as the previous two figures. In Figure 59.4 we show the same plot but for a put option and with the same data. Observe how, with our data assumptions, the put option has decreasing return for increasing risk. Obviously, in this case, call options are more attractive.

(In the above we have considered only the standard deviation as a measure of risk. This measure attaches as much weight to a better outcome than average as it does to a worse outcome. A better choice as a measure of risk may be a one-sided estimate of only the downside risk; I leave the formulation of this problem as an exercise for the reader.)

59.3 MORE SOPHISTICATED MODELS FOR THE RETURN ON AN ASSET

Despite modeling the 'value' of an option to a speculator, we know from experience that as time progresses we are likely to change our view of the market and have to make a decision

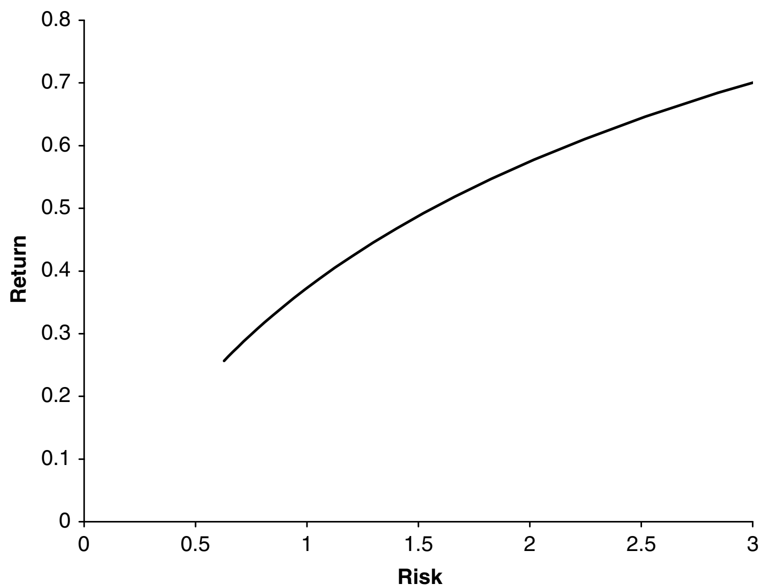


Figure 59.3 The scaled present value of expected payoff versus the scaled standard deviation for a call option. See text for details and data.

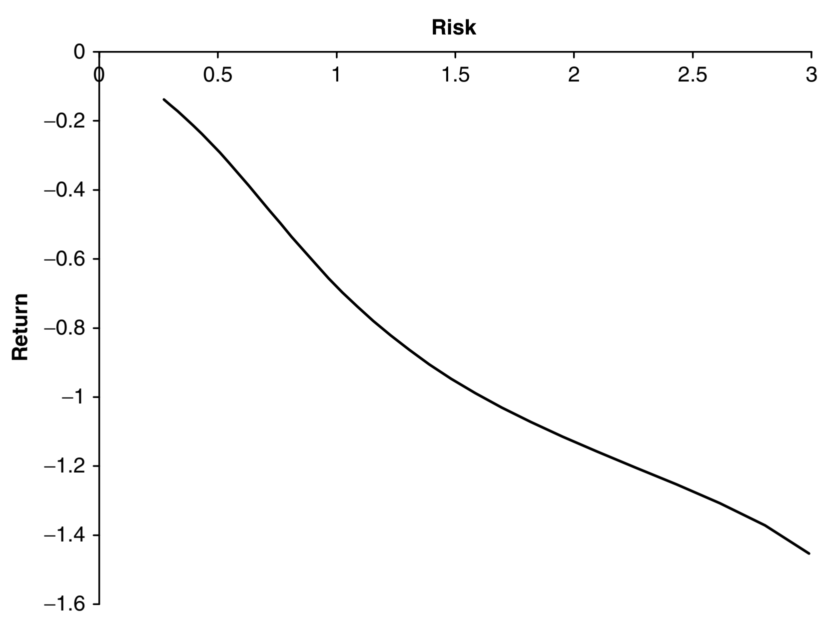


Figure 59.4 The scaled present value of expected payoff versus the scaled standard deviation for a put option. See text for details and data.

regarding the success, or otherwise, of our position: Perhaps six months into the life of our call option we decide that we were wrong about the direction of the market and we may therefore close our position. Is there any way in which we can build into our model of the market, *a priori*, our experience that market conditions change?

I am now going to present two models for a randomly varying drift rate. The first assumes that the drift rate follows some stochastic differential equation. The second model assumes that the drift can be in one of two states: High drift or low drift. The drift jumps randomly between these two states.

59.3.1 Diffusive Drift

The asset price still satisfies

$$dS = \mu S\,dt + \sigma S\,dX_1.$$

Now assuming that the drift rate μ satisfies

$$d\mu = \eta(S,\mu,t)\,dt + \nu(S,\mu,t)\,dX_2$$

then the present value of the expected return $V(S,\mu,t)$ satisfies the two-factor partial differential equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 V}{\partial S\,\partial\mu} + \tfrac{1}{2}\nu^2 \frac{\partial^2 V}{\partial \mu^2} + \mu S \frac{\partial V}{\partial S} + \eta \frac{\partial V}{\partial \mu} - rV = 0$$

where ρ is the correlation coefficient between the two random walks.[4]

In Figure 59.5 I show the value of a call option for the model

$$d\mu = (a - b\mu)\,dt + \beta\,dX_2.$$

In this figure I have used $a = 0.3$, $b = 3$ and $\beta = 0.1$ with $\rho = 0$. The option is a call struck at 100 with an expiry of one year and a volatility of 20%. The interest rate is 5%. When the asset is 100 and the drift is zero the option value is 11.64.

The standard deviation of these payoffs may be found in a similar manner to that above for constant drift. The relevant equation is simply the two-factor version of (59.1).

59.3.2 Jump Drift

The above model for diffusive drift is perhaps unnecessarily complicated for describing the market view of a typical trader. A simpler model, which I describe now, allows the drift to be in one of two states, either high or low. The asset random walk is allowed to jump from one state to the other. For example, we believe that in the short term the asset in question will have a drift of 15%. However, this may only last for six months. If the drift does change then it

[4] At this point had we been using a Black–Scholes hedging argument we would have arrived at this equation but with r as the coefficient of the $\partial V/\partial S$ term. With final conditions independent of μ we would then find that $\partial V/\partial\mu = 0$ everywhere and thus that V satisfies the Black–Scholes equation. Thus even with stochastic drift the Black–Scholes option value is independent of μ. There are some technical details, concerning the 'support' of the real and risk-neutral random walks, that must be satisfied. Simply put, the range of values that can be attained by the two random walks must be the same.

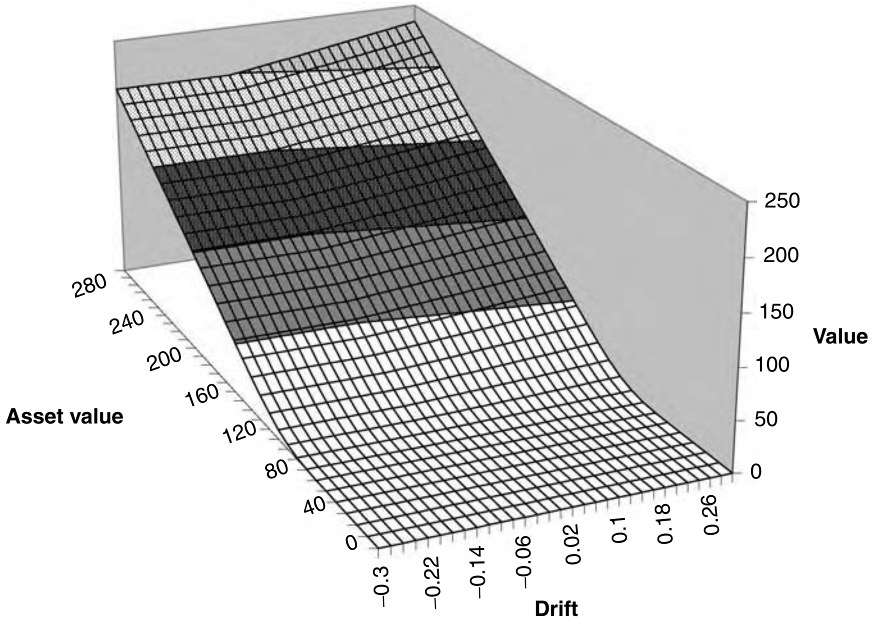


Figure 59.5 The value of a call option when the drift is diffusive. See text for details and data.

will drop to a level of 0%. In this example, call options look appealing in the short term, but we may not want to buy longer-dated calls since we will probably suffer when the downturn arrives. We assume that the two states have constant drift and that the jump from one state to another is governed by a Poisson process. In other words, the probability of changing from a high drift to a low drift in a period dt is $\lambda^- dt$, and the probability of changing from a low drift to a high drift is $\lambda^+ dt$.

We can value options to a speculator by again looking at the present value of the expected payoff. First, we must introduce the notation. Let $V^+(S,t)$ be the option 'value' when the asset has a drift of μ^+. Similarly, the option has value $V^-(S,t)$ when the asset has a drift of $\mu^- < \mu^+$. We can easily examine expected payoffs by considering how the option value changes at each time step.

At time t the asset value is S and we assume without loss of generality that it is in the higher drift state, with the drift of μ^+. At the later time $t + dt$ the asset price changes to $S + dS$, and at the same time the drift rate may jump to μ^- but only with a probability of $\lambda^- dt$.

With the real expected return equal to the risk-free rate we arrive at the following equation:

$$\frac{\partial V^+}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V^+}{\partial S^2} + \mu^+ S \frac{\partial V^+}{\partial S} - rV^+ + \lambda^-(V^- - V^+) = 0.$$

Similarly, we find that

$$\frac{\partial V^-}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V^-}{\partial S^2} + \mu^- S \frac{\partial V^-}{\partial S} - rV^- + \lambda^+(V^+ - V^-) = 0$$

when the drift starts off in the lower state. The final conditions for these equations are, for example,

$$V^+(S,T) = V^-(S,T) = \max(S-E,0)$$

for a call option.

The standard deviation also takes two forms, depending on which state the drift is in. We must solve

$$\frac{\partial G^+}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 G^+}{\partial S^2} + \mu^+ S \frac{\partial G^+}{\partial S} + \lambda^-(G^- - G^+) = 0$$

and

$$\frac{\partial G^-}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 G^-}{\partial S^2} + \mu^- S \frac{\partial G^-}{\partial S} + \lambda^+(G^+ - G^-) = 0$$

with

$$G^+(S,T) = G^-(S,T) = \left(e^{-rT}\max(S-E,0)\right)^2.$$

Then the standard deviation at time $t=0$ is given by

$$\sqrt{G^{\pm}(S,0) - (V^{\pm}(S,0))^2},$$

depending on whether the drift is in the high or low drift state at that time.[5]

In Figure 59.6 we show the value of a call option against asset price. The top curve shows the option value when the asset is in the high drift state, and the bottom curve is when the asset is in the low drift state. Between the two, the bold curve, is the Black–Scholes fair value. The

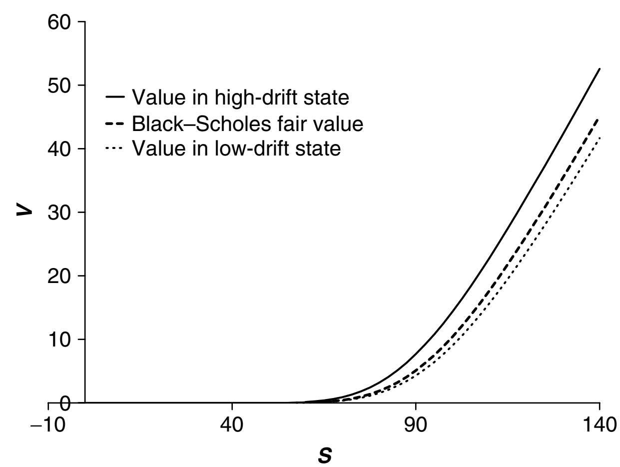


Figure 59.6 The value of a call option when the drift jumps between two levels. See text for details and data.

[5] How do we know what state we are in? That's a statistical question. Alternatively, we could say that if the underlying has fallen by a certain amount this means a change of state; this is also easy to model.

option is a call struck at 100 with one year to expiry. The underlying has a volatility of 20% and the interest rate is 5%. The high drift is 15% and the low drift is zero. The intensity of the Poisson process taking the asset from high to low is 1 and 0.5 going from low to high. In this example it is easier for the asset to sink to the low drift state than to recover, and this is why the expected value in the low drift case is below Black–Scholes.

59.4 EARLY CLOSING

So far we have modeled the value to a speculator of an option on which he has a strong view on the drift rate and which he will hold until expiry. The diffusive drift and the jump drift models allow for the possibility that the drift leaves its advantageous state and changes to a disadvantageous state. In our example of Section 59.2, the asset drift was initially high, but has fallen to below the risk-free rate, and our call option no longer seems like such a good bet. What can we do? Although our models price this into the option's value, they do not allow for the obvious trader reaction: sell the option and close the position.

Within the present models it is natural for the trader to close his position if the value to him falls below the market value. Why hold onto a position in which you expect to make a loss? He may want to sell earlier than this but we use this very simple constraint as an example. We also assume that the market value of the option is simply the Black–Scholes value, with the same risk-free rate of interest and volatility; again, this could be generalized.

Having seen the mathematical analysis for American options in Chapter 9, it is obvious that our early-closing problem is identical in spirit: Take either the diffusion model or the jump model for the drift and impose the constraint that the option value must always be greater than the Black–Scholes value and that the first derivatives must be continuous. In the diffusion model this amounts to

$$V(S,\mu,t) \geq V_{BS}(S,t) \quad \text{and} \quad \frac{\partial V}{\partial S} \quad \text{continuous.}$$

For the jump model we have

$$V^{+}(S,t) \geq V_{BS}(S,t), \quad V^{-}(S,t) \geq V_{BS}(S,t) \quad \text{and} \quad \frac{\partial V^{\pm}}{\partial S} \quad \text{continuous.}$$

In Figure 59.7 is shown the value of the option against the underlying and the drift for the diffusive drift model. The parameter values are as earlier. The option value at an asset value of 100 and a drift of zero is now 11.84. The extra value comes from the ability to close the position. In Figure 59.8 is plotted the difference between the option values with and without the ability to close the position.

In Figure 59.9 is shown the value of the option against the underlying in the two-state drift model. This two-drift state model is particularly interesting because the option value in the low-drift state is the same as Black–Scholes. The interpretation of this is that you should sell the option as soon as you believe that the drift is in the low state. The extra value in being able to sell the position is plotted in Figure 59.10.

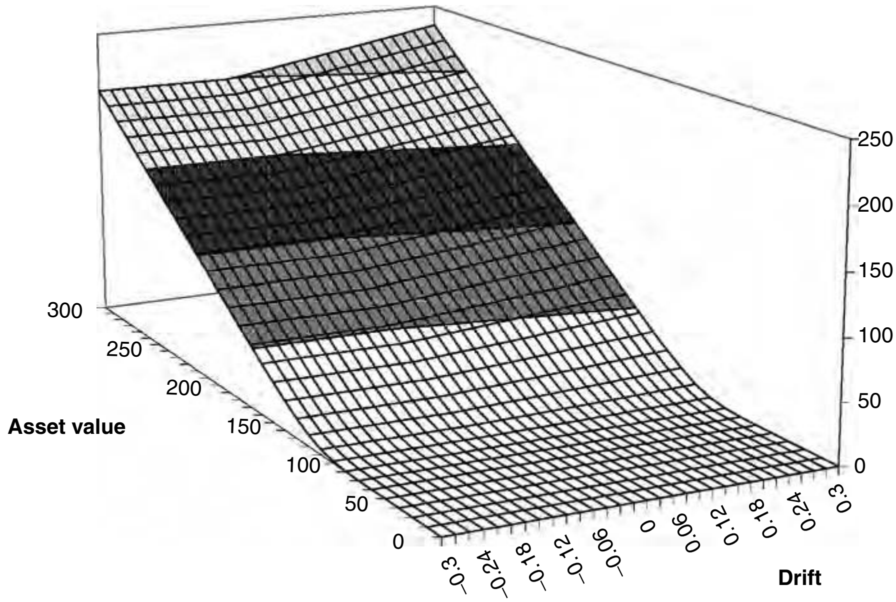


Figure 59.7 The option value to a speculator with closure of the position to ensure that it never falls below the market value: Diffusive drift model.

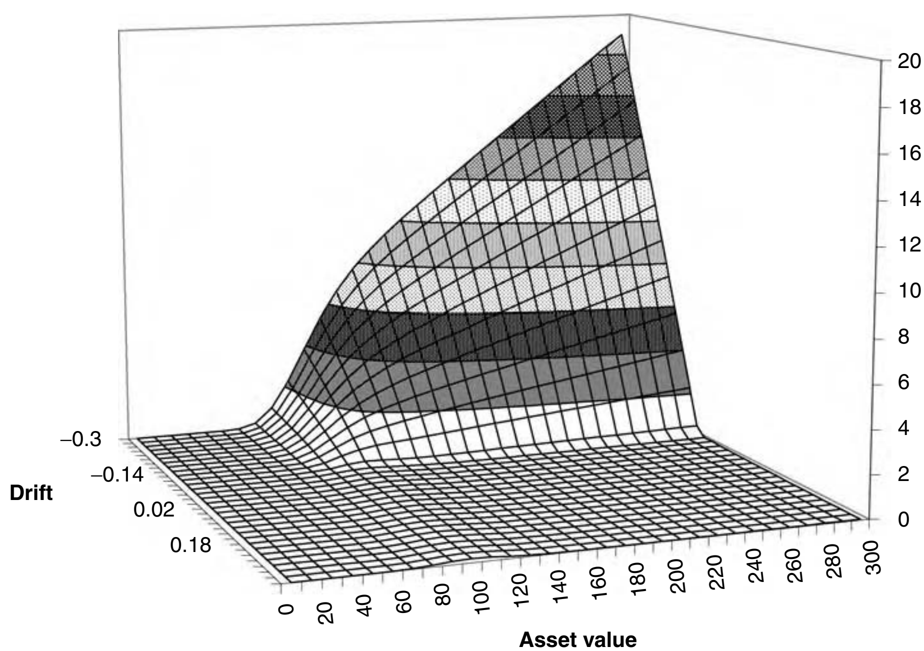


Figure 59.8 The added value of being able to close the position: Diffusive-drift model.

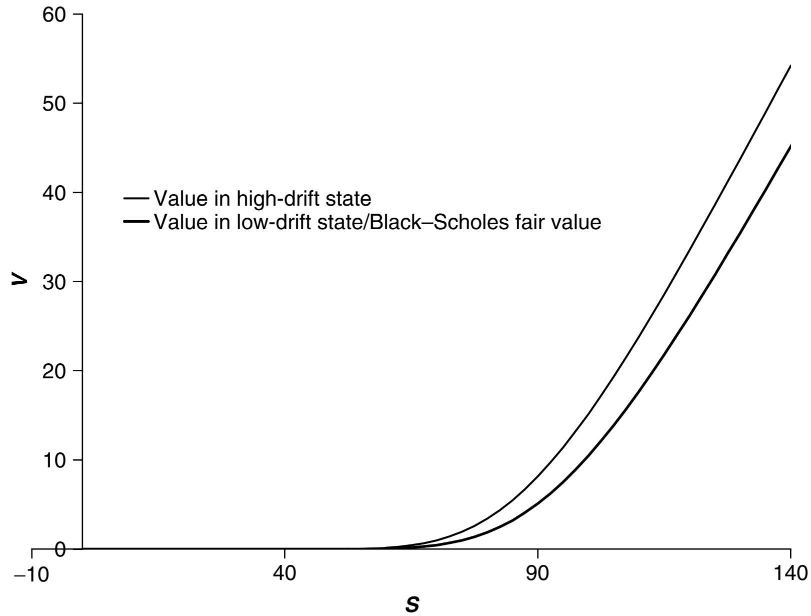


Figure 59.9 The option value to a speculator with closure of the position to ensure that it never falls below the market value: Jump drift model.

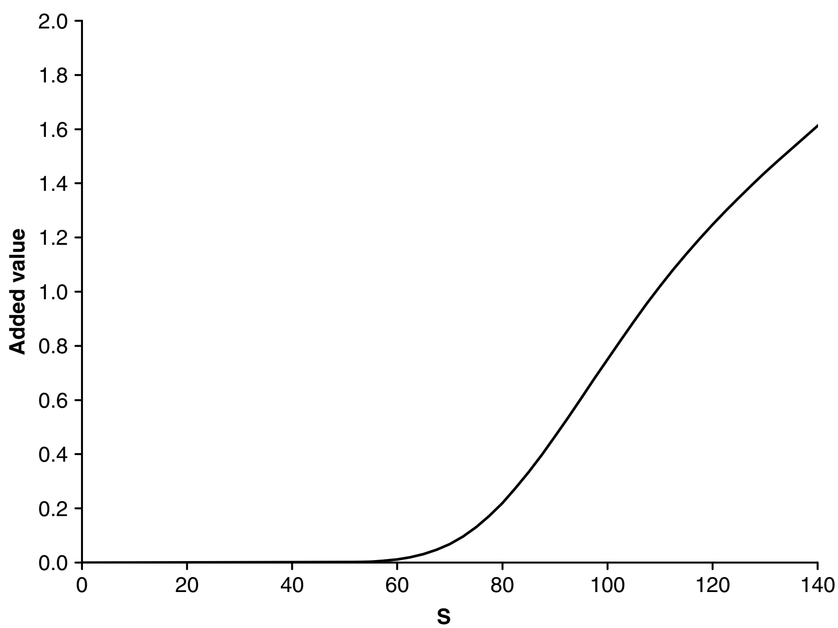


Figure 59.10 The added value of being able to close the position: Two-state model.

The problem for the standard deviation is slightly more complicated when there is early closing because of the appearance of the early-closing boundary. For example, with the two-state drift model we have

$$\frac{\partial G^+}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 G^+}{\partial S^2} + \mu^+ S \frac{\partial G^+}{\partial S} + \lambda^-(G^- - G^+) = 0$$

$$G^+(S,T) = \left(e^{-rT}\max(S-E,0)\right)^2$$

and

$$G^+(S_e(t),t) = \left(e^{-rt}V^+(S_e(t),t)\right)^2.$$

Here $S_e(t)$ is the position of the early-closing boundary given by the solution of the free boundary problem for V^+ and V^-. There is a similar equation for G^-.

Again, the standard deviations are given by

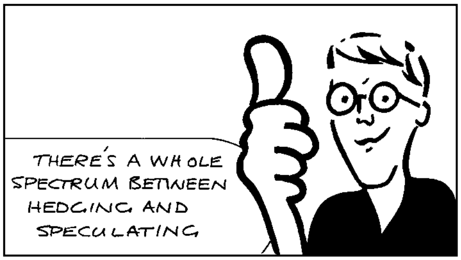
$$\sqrt{G^\pm(S,0) - V^\pm(S,0)^2}.$$

The above illustrates the importance of not only having a model for the underlying but also, when relevant, having a model for the pricing mechanism *of the market*. This is relevant if one expects or hopes to sell the contract before its expiry, since one must sell it for its market value and not one's theoretical value. One of the benefits inherent in this is that having a model for the market value enables you to make even more money than simply having a model for the underlying. This gives you the option of either holding the contract to expiry or, if the market value rises higher than your theoretical value, closing out the position. The value of the contract to you, under this scenario, cannot be less than your theoretical value without the option to close the position. This principle can easily be extended. For example, suppose you have a model for the market's perception of volatility, you can use this to give a lower bound to the contract in question, regardless of the validity of the market's view.

In the above framework for valuing contracts it is obvious that some contracts are more appealing than others. For an asset with a drift above the risk-free rate calls would be preferred over puts for example. There are timescales associated with changes in the drift rate and these may have some bearing on the choice of expiry dates.

In practice you don't have to liquidate the entire portfolio, you may decide to sell off only part of the portfolio. Having this extra freedom means that you can squeeze even more value from your position. But how can this be done optimally? You should be able to do this, at least theoretically, by the end of this part of the book.

59.5 TO HEDGE OR NOT TO HEDGE?



So far in this chapter we have looked at 'valuing' options in the complete absence of hedging. As I have said, speculation is the opposite side of the coin to hedging. Perhaps it is possible to bring these two sides closer together.

In this section we look at the expected values of options, assuming a quite sophisticated hedging strategy, one that is intuitively highly appealing. This strategy is to speculate when we expect to make money and hedge when we would otherwise expect to lose it.

We will assume that the drift rate of the asset, μ, is constant and greater than the risk-free interest rate. This is the simplest problem that can be used to introduce the ideas. Extensions to other cases, $\mu < r$, non-constant μ etc. are obvious. With this assumption, our strategy can be modeled mathematically by a portfolio of value Π where

$$\Pi = V(S,t) - \overline{\Delta} S, \tag{59.2}$$

and V is the 'value' of our option. We choose the quantity $\overline{\Delta}$ so as to maximize the expected growth in the value of our portfolio. We also impose the constraints on the value of $\overline{\Delta}$

$$\alpha \frac{\partial V}{\partial S} \leq \overline{\Delta} \leq \beta \frac{\partial V}{\partial S}.$$

Such a constraint could be used to bound the risk in the portfolio.

Since the expected drift of the portfolio in excess of the risk-free rate is

$$\left(\mu S \frac{\partial V}{\partial S} - rV - (\mu - r) S \overline{\Delta} \right) dt,$$

it is simple to show that the choice

$$\overline{\Delta} = \begin{cases} \alpha \dfrac{\partial V}{\partial S} & \text{if} \quad \dfrac{\partial V}{\partial S} \geq 0 \\ \beta \dfrac{\partial V}{\partial S} & \text{if} \quad \dfrac{\partial V}{\partial S} \leq 0 \end{cases}$$

maximizes the expected drift of the portfolio. (These inequalities would swap for the case $\mu < r$.) The reasoning behind this choice is simple; if our option has a positive $\Delta = \partial V / \partial S$ then we *expect* to make more money by speculating than hedging since we expect the asset price to rise faster than the risk-free rate and hence we expect the portfolio value to rise faster. When Δ is negative we expect to lose out on a speculating position in comparison to a hedged position. If we now calculate the present value of the expected payoff we find that it satisfies

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + S \frac{\partial V}{\partial S} \left(\mu - (\mu - r) F(\Delta) \right) - rV = 0,$$

with

$$F(\Delta) = \begin{cases} \alpha \Delta & \text{if} \quad \Delta \geq 0 \\ \beta \Delta & \text{if} \quad \Delta \leq 0 \end{cases}$$

This is another non-linear diffusion equation.

In Figure 59.11 we see the solution of this equation for a position of two long calls at 90 and three short 100 calls. A model such as this is an obvious improvement over pure speculation.

The hedger-speculator model can be combined with the early closing constraint.

With the parameters $\alpha = \beta = 1$ we have Black–Scholes and with $\alpha = \beta = 0$ we return to the hedging-speculating model (59.5).[6]

[6] There are many other obvious ways to bound risk. For example, by choosing α and β suitably we can overhedge if we think that this is advantageous. We could also allow α and β to depend on V or Π; we may want to take less risk with a larger portfolio.

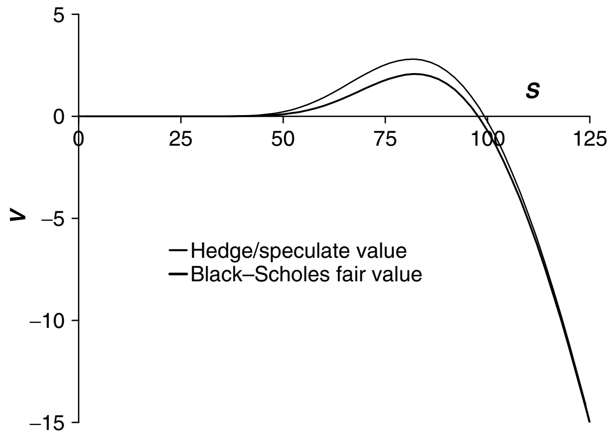


Figure 59.11 The option value to a hedger-speculator and the Black–Scholes value; see text for details.

59.6 OTHER ISSUES

An accurate model of the drift of the underlying is especially important when speculating with barrier options or binary/digital options. Since binary options have a discontinuous payoff it is not possible in practice to follow the Black–Scholes hedging strategy no matter how frequently one hedges. Thus there is always some exposure to the drift; to some extent, one is always speculating with binary options.

We have seen how the drift rate of the underlying comes into the equation for the real expected payoff. Another simple but important calculation in a similar vein is the calculation of the real expected value of an option part way into its life. For example, we hold a five-year warrant and want to know how much this will be worth in two years' time. We may want to know this so as to have some idea of the value of our portfolio at that time. We calculate the *real expected* value of the warrant in two years as follows.

First, note that the value of the option in the Black–Scholes world satisfies the Black–Scholes equation and does not depend on μ. Let us assume that this is how the market will value the contract throughout its life. However, we want to know the real expected value in two years. We can do this by simply solving the equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \nu S \frac{\partial V}{\partial S} - rV = 0,$$

where

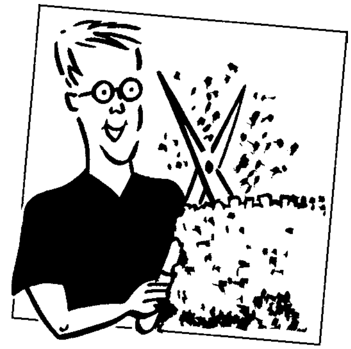
$$\nu = \begin{cases} \mu & \text{for} \quad 0 \le t \le 2 \\ r & \text{for} \quad 2 < t \le 5. \end{cases}$$

59.7 SUMMARY

One mustn't get too hung up on delta hedging, no arbitrage and risk neutrality. The concepts are extremely important but there are times when their relevance is questionable, and one of those times is when you are speculating. Sure, the price for which you buy a contract has something to do with these concepts via Black–Scholes, but from then on, if you are not hedging, there are more important things to worry about.

FURTHER READING

- See Korn & Wilmott (1996, 1998) for details about option pricing with a view.
- The best choice to make from a selection of risky investments is just asset allocation, see for example Markowitz (1959), but this is discussed in Chapter 18.



CHAPTER 60
static hedging

In this Chapter...

- matching a contract's value at a set of points using standard contracts
- vega hedging
- how non-linear equations make a mockery of 'parameter fitting'
- statically hedging with traded options to improve your prices
- how to optimally statically hedge to make the most out of your contract

60.1 INTRODUCTION

Delta hedging is a wonderful concept. It leads to preference-free pricing (risk neutrality) and a risk-elimination strategy that can be used in practice. There are quite a few problems, though, on both the practical and the theoretical side. In practice, hedging must be done at discrete times and is costly. These issues were discussed in Chapters 47 and 48. Sometimes one has to buy or sell a prohibitively large number of the underlying in order to follow the theory. This is a problem with barrier options and options with discontinuous payoff.

On the theoretical side, we have to accept that the model for the underlying is not perfect; at the very least we do not know parameter values accurately. Delta hedging alone leaves us very exposed to the model; this is model risk.

Many of these problems can be reduced or eliminated if we follow a strategy of static hedging as well as delta hedging: Buy or sell more liquid contracts to reduce the cashflows in the original contract. The static hedge is put into place now, and left until expiry.[1]

60.2 STATIC REPLICATING PORTFOLIO

The value of a complex derivative is usually very model dependent. Often these contracts are difficult to delta hedge, either because of transaction costs in the underlying or because of discontinuities in the payoff or the delta. Problems with delta hedging can be minimized by

[1] In practice, if conditions become favorable, one can reassess the static hedge. It is not set in stone.

static hedging or **static replication**, a procedure which can reduce transaction costs, benefit from economies of scale and, importantly, reduce model risk.

There are three forms of static hedging that I want to discuss here:

- Payoff matching
- Vega hedging
- Non-linear

The first form of static hedging is about constructing a portfolio of traded options whose value matches the 'target' contract's value as closely as possible at a set of dates and asset values. The second form tries to reduce sensitivity of the price of an exotic to volatility. The third form depends crucially on the governing equation being non linear.

60.3 MATCHING A 'TARGET' CONTRACT

The first idea is simple, and the implementation not much harder. Let's think of a concrete example, the up-and-out barrier option. Referring to Figure 60.1, the aim is to construct a portfolio of vanilla puts and calls, say, to have the same value as the target contract, the barrier option, at expiry and at the points marked in the figure. In this example the barrier option has zero value along the points marked in the figure.

Suppose that the target contract has a call payoff at expiry T, but expires worthless on the upper boundary marked in the figure. This is a up-and-out call option. Thus on the upper boundary, our target contract has value zero. We have sold this contract and we want to statically hedge it by buying vanilla calls and puts. That is, we want to match the call payoff at T and have zero value on the boundary. This is easy.

To match the payoff at expiry we buy a single call option with the same strike as the knock-out call. This will immediately give us zero payoff at expiration, but will mess up all the zeros along the top boundary. We now have to buy and sell other vanillas to try and eliminate all the non-zero values along the top boundary.

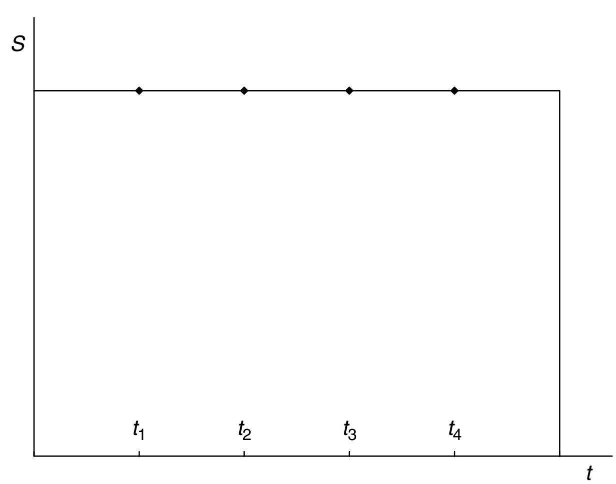


Figure 60.1 The target contract is to be replicated at the marked points.

Ideally, we would like to have zero value on the entire boundary, but this would involve us buying an infinite number of vanilla options. It is natural, therefore, since there are only a finite number of traded contracts, to match the portfolio value at the expiry dates of the traded options. In the figure, this means matching values at the four points on the upper boundary. Typically this will require us to buy four more vanilla options, in quantities to be determined.

We have matched the payoff at expiry with a single call option. Now, working back from expiry, value until time t_4 the contract consisting of the knock-out call held short and the single long vanilla call. In this 'valuation' we use whatever model we like. We use the model to give us the theoretical value of the two-contract portfolio at time t_4 at the point marked in Figure 60.1. We want to buy or sell further vanilla contracts to make the net portfolio value zero at this point. The obvious contracts to choose for this hedging are calls with expiry at T and strikes such that they are out-of-the-money for all values of S on the target-option boundary at time T. We need only one contract to achieve this, an out-of-the-money call, say.

Now value this new portfolio (it now contains our original target contract and two vanillas) back until time t_3. We can find its value at the marked point at this time. To hedge this value buy or sell vanilla contracts *that expire at time* t_4, that are out-of-the-money for all values of S within the target-option boundary at time t_4, and such that between them they make the portfolio value zero on the upper boundary at time t_3. We now hold three options in our static hedge (plus the target contract) and are 'hedged' from t_3 until T. In this way we work backwards until the present.

We can obviously include as many expiry dates as exist in traded contracts, and the finer their resolution the better will be our static hedge. As long as we only use a finite number of options in our static hedge, then there will be some residual risk that is not statically hedged. We can either delta hedge this, or if it is small and with bounded downside, we could leave it unhedged (assuming we have allowed for this when we sold the target contract).

I have illustrated the idea with a barrier option hedged with calls. Obviously the idea can be extended in any number of ways. One possibility is to use one-touch options; these are no longer thought of as being exotic and are actively traded.

Static hedging is obviously a useful technique. It is not perfect since there is still some exposure to the model. We had to value the portfolio using some 'model,' and the accuracy of the hedge is reflected to some degree in the accuracy of the model. However, the resulting portfolio is not as model dependent as a contract that is only delta hedged.

60.4 VEGA MATCHING

Why do we like to statically hedge? Because sometimes delta hedging is just not good enough. Why not? Because we don't necessarily know the correct delta and this is because we don't know volatility. It all boils down to volatility.

Suppose you have used the deterministic volatility surface model of Chapter 50. Obviously your price, your delta and your profit depend on the accuracy of your model. As I mentioned in that chapter, this model is scientifically dodgy. So what people do to make it less shaky is to statically hedge with vanillas.

Suppose you have calibrated your deterministic volatility model at time t^* to the implied volatilities σ_{imp}. Now use the resulting local volatility surface to price your barrier option, say. The price etc. will be very sensitive to the implied volatilities. Now construct a portfolio consisting of the barrier option and some of the vanillas to which you have just calibrated. To determine how many of the vanillas to hedge with do the following experiment. Increase the value of one of the vanillas by \$1. By so doing you will change its implied volatility, and thus the local volatility surface. Once the local volatility surface changes then so will the price of the barrier option. Suppose the barrier option increases in value by \$1.5. This tells us how to hedge. Simply sell 1.5 of the vanilla option. You now have a portfolio consisting of the original barrier option and short 1.5 of this vanilla. Your portfolio is now insensitive to that one implied volatility. Then, one at a time, do the same for every traded vanilla option to which you have calibrated, ending up with the barrier option and a whole basket of vanillas hedging it.

In some, loose, sense, your barrier option is now vega hedged, and is insensitive to volatility.

If only it were that simple. Unfortunately there are many inconsistencies in this approach, mainly concerning the confusion of implied and actual volatilities, not to mention the foundation being the deterministic volatility model.

Both of the above two static hedging methods are less than satisfactory because of inherent inconsistencies between the assumptions of the model and the so-called 'hedging' technique.

In the rest of this chapter we discuss static hedging when the governing equation is non linear. Some of these non-linear models get a lot closer to model-independent pricing; at least they are 'parameter-insensitive' models. Note that the idea of 'optimal' static hedging that we will be seeing can also be applied to the above replicating portfolio: It may be cheaper to statically hedge with one contract than with another.

60.5 STATIC HEDGING: NON-LINEAR GOVERNING EQUATION

Many of the models described in this part of the book are non linear. To summarize, these models are the following.

- **Transaction costs:** Purchase and sale of the underlying for delta hedging when there are costs leads to non-linear equations for the option value. There are many models, all of them non linear. See Chapter 48.
- **Uncertain parameters:** When parameters, such as volatility, dividend rate or interest rate, are permitted to lie in a range, options can be valued in a worst-case scenario, for example the model of Avellaneda, Levy & Parás and Lions. The governing equation for uncertain volatility is mathematically identical to the Hoggard–Whalley–Wilmott transaction cost model. See Chapter 52.
- **Crash modeling:** Allowing the underlying to jump and valuing contracts in a worst-case scenario, with no probabilistic assumptions about the crash, leads to a non-linear equation. This contrasts with the linear classical jump-diffusion model. See Chapter 58.
- **Speculating with options:** Some of the strategies for the 'valuation' of contracts when speculating lead to non-linear equations. For example, optimal closure of an option position has a linear governing equation but the problem is non linear because it has a free boundary. Choosing to hedge or speculate also leads to a non-linear governing equation. See Chapter 59.

Later, in Chapter 68, we will see a non-probabilistic interest rate model that is non linear; all of the ideas of this chapter can be applied to that model.

60.6 NON-LINEAR EQUATIONS

Many of the new partial differential equations that we have derived are non linear. This non-linearity has many important consequences. Most obviously, we must be careful when we refer to an option position, making it clear whether it is held long or short. For example, for a long call we have the final condition

$$V(S,T) = \max(S-E, 0)$$

and for a short call

$$V(S,T) = -\max(S-E, 0).$$

Because of the nonlinearity, the value of a portfolio of options is not necessarily the same as the sum of the values of the individual components. This is a very important point to understand: The value of a contract depends on what else is in the portfolio.

These two points are key to the importance of non-linear pricing equations: They give us a bid-offer spread on option prices, and they allow *optimal* static hedging.

For the rest of this chapter we discuss the pricing and hedging of options when the governing equation is non linear. The ideas are applicable to any of the non-linear models mentioned above. We use the notation $V_{NL}(S,t)$ to mean the solution of the model in question, whichever model it may be. So that the explanation of the issues does not get too confusing I will always refer to the concrete example of the uncertain volatility/transaction cost model, but remember *the ideas apply equally well to the other models*.

60.7 PRICING WITH A NON-LINEAR EQUATION

One of the interesting points about non-linear models is the prediction of a spread between long and short prices. If the model gives different values for long and short then this is in effect a spread on option prices. This can be seen as either a good or a bad point. It is good because it is realistic; spreads exist in practice. It only becomes bad when this spread is too large to make the model useful.

60.7.1 Example: Non-linear Model, Unhedged

Let us consider a realistic, uncertain volatility/transaction cost model example. The reader is reminded that this model is

$$\frac{\partial V_{NL}}{\partial t} + \tfrac{1}{2}\sigma(\Gamma)^2 S^2 \frac{\partial^2 V_{NL}}{\partial S^2} + rS\frac{\partial V_{NL}}{\partial S} - rV_{NL} = 0 \tag{60.1}$$

where

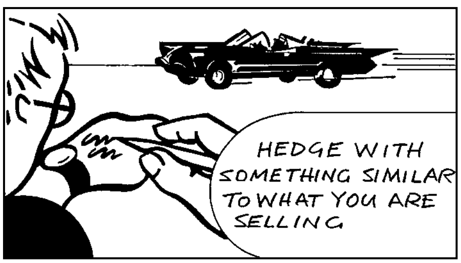
$$\Gamma = \frac{\partial^2 V_{NL}}{\partial S^2}$$

and

$$\sigma(\Gamma) = \begin{cases} \sigma^+ & \text{if } \Gamma < 0 \\ \sigma^- & \text{if } \Gamma > 0. \end{cases}$$

(But we are using it as a proxy for any of the non-linear equations.)

This model can result in very wide spreads on options. For example, suppose that we have a European call, strike price \$100, today's asset price is \$100, there are six months to expiry, no dividends on the underlying and a spot interest rate of 5%. We will assume that the volatility lies between 20% and 30%. We are lucky with this example. Because the gamma is single-signed for a vanilla call, we can calculate the values for long and short calls, assuming this range for volatility, directly from the Black–Scholes formulae. A long call position is worth \$6.89 (the Black–Scholes value using a volatility of 20%) and a short call is worth \$9.63 (the Black–Scholes value using a volatility of 30%). This spread is much larger than that in the market, and is due to the uncertain volatility. The market prices may, for example, be based on a volatility between 24% and 26%. The uncertain parameter model is useless unless it can produce tighter spreads. In the next section I show how the simple idea of static hedging can be used to reduce these spreads significantly. The idea was originally due to Avellaneda & Parás.



60.7.2 Static Hedging: A First Attempt

Suppose that we want to sell an option with some payoff that does not exist as a traded contract. (For the moment think in terms of a European contract with no path-dependent features; we will generalize later.) We want to determine how low a price can we sell it for (or how high a price we can buy it for), with the constraint that we guarantee that we will not lose money as long as our range for volatility is not breached. There are two related reasons for wanting to solve this problem. If we can sell it for more than this minimum then we are guaranteed to make money, and if our spread is tight then we will have more customers.

I motivate the idea of static hedging with a simple example. Suppose that options on a particular stock are traded with strikes of \$90 and \$110 and with six months to expiry. The stock price is currently \$100. However, we have been asked to quote prices for long and short positions in \$100 calls, again with six months before expiry. This call is *not* traded.

As above, assume that volatility lies between 20% and 30%. Remember that our pessimistic prices for the 100 call were \$9.63 to sell and \$6.85 to buy, calculated using the Black–Scholes formulae with volatilities 30% and 20% respectively. This spread is so large that we will not get the business. However, we are missing one vital point: The 90 and 110 calls are trading, so can't we take advantage of them in some way?

Suppose that the market prices the 90 and 110 calls with an implied volatility of 25% (forget bid-offer spreads for the moment). The market prices, i.e. the Black–Scholes prices, are therefore 14.42 and 4.22 respectively. These numbers are shown in Table 60.1. The question marks are to emphasize that we can buy or sell as many of these contracts as we want, but in a fashion which will be made clear shortly. Shouldn't our quoted prices for the 100 call reflect the availability of contracts with which we can hedge?

Table 60.1 Available contracts.

Strike	Expiry	Bid	Ask	Quantity
90	180 days	14.42	14.42	?
110	180 days	4.22	4.22	?

Consider first the price at which we would sell the 100 call. If we sell the 100, and 'statically hedge' it by buying 0.5 of the 90 and 0.5 of the 110, then we have a residual payoff as shown in Figure 60.2. We call this a static hedge because we put it in place now and do not expect to change it. This contrasts with the delta hedge, for which we expect to hedge frequently (and would like to hedge continuously).

The statically-hedged portfolio has a much smaller payoff than the original unhedged call, as shown in Figure 60.2. It is this *new portfolio* that we value by solving the non-linear Equation (60.1), and that we must delta hedge; I emphasize this last point, *the residual payoff must be delta hedged*. To value the residual payoff in our uncertain parameter framework we solve Equation (60.1) with final condition

$$V_{NL}(S,T) = -\max(S-100,0) + \tfrac{1}{2}\left(\max(S-90,0) + \max(S-110,0)\right).$$

Let us see what effect this has on the price at which we would *sell* the 100 call.

First of all, observe that we have paid

$$0.5 \times \$14.42 + 0.5 \times \$4.22 = \$9.32$$

for the static hedge made up of 0.5 of each of the 90 and 110 calls.

Now, solve Equation (60.1) using the payoff shown in Figure 60.2 as the final condition. The solution gives a value for the residual contract today of \$0.58.

The net value of the call is therefore

$$\$9.32 - \$0.58 = \$8.74.$$

To determine how much we should pay to *buy* the 100 call we take as our starting point the sale of 0.5 of each of the 90 and 110 calls. This nets us \$9.32. But now we solve Equation (60.1)

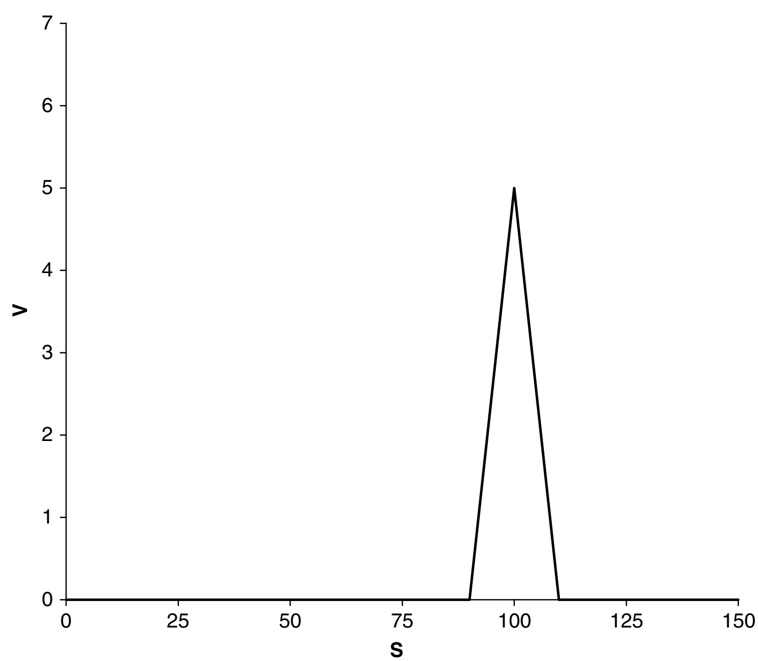


Figure 60.2 The residual risk after hedging a 100 call with 0.5 each of the 90 and 110.

using the *negative* of the payoff shown in Figure 60.2 as the final condition; the effect of the nonlinearity is different from the previous case because $\sigma(\Gamma)$ takes different values in different places in S, t-space. We get a value of \$1.96 for the hedged position. Thus we find that we would pay

$$\$9.32 - \$1.96 = \$7.36.$$

Note how the use of a very simple static hedge has reduced the spread from \$6.85–9.6 to \$7.36–8.74. This is a substantial improvement, as it represents a volatility range of 22–27%; while our initial estimate for the volatility range was 20–30%. The reason for the smaller spread is that the residual portfolio has a smaller absolute spread, only \$0.58–1.96, and this is because it has a much smaller payoff than the unhedged 100 call.

In the trivial case where the option on which we are quoting is also traded then we would find that our quoted price was the same as the market price. This is because we would hedge one for one and the residual payoff, which we would normally delta hedge, would be identically zero. This means that we always match market prices; 'fitting' as described in Chapter 50 is redundant.

60.7.3 Static Hedging: The Best Hedge

In the above example we decided to hedge the 100 call using 0.5 of each of the 90 and 110 calls. What prompted this choice? In this case we chose the numbers 0.5 for each option purely on the grounds of symmetry: It made the residual payoff look simple (a typical mathematician's choice).

There was no reason why we should not choose other numbers. Suppose, for example, we decided to hedge with 0.4 of the 90 call and 0.7 of the 110. Would this have an effect on the value of our 100 call? Since our problem is non linear *the value of our OTC option depends on the combination of options with which we hedge*. So, generally speaking, we expect a different OTC option value if we choose a different static hedge portfolio. Of course, we now ask 'If we get different values for an option depending on what other contracts we hedge it with, then is there a *best* static hedge?'

Continuing with our earlier example, what are the optimal static hedges for long and short positions? Buying 0.5 each of the two calls to hedge the short 100 call we find a marginal value of \$8.74 for the 100 call. This is actually a very good hedge. Slightly better, and optimal, is to buy 0.51 of each call giving a marginal value of \$8.73. The optimal hedge for a long position is vastly different. We should sell 1.07 of the 90 call and sell 0.34 of the 110 call. The marginal value of the 100 call is then \$7.47.

60.8 OPTIMAL STATIC HEDGING: THE THEORY

Suppose that we want to find the lowest price for which we can sell a particular OTC or 'exotic' option with payoff $\Lambda(S)$. Suppose that we can hedge our exotic with a variety of traded options. These options will have payoffs (at the same date as our exotic to keep things simple for the moment) which we call

$$\Lambda_i(S).$$

At this point we can introduce bid and offer prices for the traded options: C_i^+ is the ask price of the ith option and C_i^- the bid, with $C_i^- < C_i^+$.

Now we set up our statically hedged portfolio: we will have λ_i of each option in our hedged portfolio. The cost of setting up this static hedge is

$$\sum_i \lambda_i C_i(\lambda_i),$$

where

$$C_i(\lambda_i) = \begin{cases} C_i^+ & \text{if } \lambda_i > 0 \\ C_i^- & \text{if } \lambda_i < 0 \end{cases}.$$

If $\lambda_i > 0$ then we have a positive quantity of option i at the offer price in the market, C_i^+; if $\lambda_i < 0$ then we have a negative quantity of option i at the bid price in the market, C_i^-.

We let $V^-(S, t)$ be the pessimistic value of our *hedged* position. The residual payoff for our statically hedged option is

$$V^-(S, T) = \Lambda(S) + \sum_i \lambda_i \Lambda_i(S).$$

Now we solve (60.1) with this as final data, to find the *net* value of our position (today, at time $t = 0$, say) as

$$V^-(S(0), 0) - \sum_i \lambda_i C_i(\lambda_i) = F(\lambda_1, \dots, \lambda_n).$$

This is a mathematical representation of the type of problem we solved in our first hedging example; $V^-(S, T)$ was the residual payoff shown in Figure 60.2.

Our goal now is to choose the λ_i to minimize $F(\cdots)$ if we are selling the exotic, and maximize if buying. (Thus the best hedge in the two cases will usually be different.) This is what we mean by 'optimal' static hedging.

I like to write the definition of an exotic option's value in this model in the following, pseudo-mathematical, way.

$$\text{Exotic option value} = \max_{\lambda \mathrm{S}} \left(V_{NL} \left(\text{Exotic} + \sum_i \lambda_i \ \text{Vanilla}_i \right) - \sum_i \lambda_i \ \text{CostofVanilla}_i \right). \tag{60.2}$$

V_{NL} just means the solution of one of our non-linear equations. The above is far removed from just solving some partial differential equations, involving as it does pricing portfolios and then optimizing.

Sometimes the results are not always immediately obvious. This is because traditionally, when we think of hedging, we tend to choose hedges so as to minimize the risky payoff. In the above algorithm the cost of the hedge plays as important a role as its effect on the residual payoff. In an extreme case, if we had an option with an implied volatility outside our best-guess range, then the algorithm would tell us to buy or sell an infinite amount of that option: The optimization algorithm finds arbitrage opportunities.

There are two other aspects to this static hedging. The first is that we are not necessarily restricted to hedging with options of the same maturity as the contract in question; they can have shorter or longer maturities. The optimization procedure will still find the best choice. The second point is that the static hedge need not be fixed. We could return at any time (or even continuously in theory) to examine the choice of optimal hedge portfolio. If the optimal hedge changes then, of course, we can buy and/or sell options to improve the worst-case value. This may not happen very often if the bid-offer spread on the traded options is large, but if it is negligible then the static hedge may often be rebalanced. However, this depends on the evolution of the implied volatilities, something which we are deliberately not modeling or predicting. If we do decide to change the static hedge then this is because we can improve the worst-case option value and thus we 'lock in' more guaranteed profit.

In this section we have concentrated on hedging unusual payoffs, or non-exchange traded contracts, with vanillas having the same expiry dates. The ideas are easily extended to hedging with options having other expiry dates, either before or after the expiry of the main contract. This is simply done via jump conditions across each expiry date. As the non-linear diffusion equation is solved backwards in time we must make the portfolio of options jump by the relevant payoff.

60.9 CALIBRATION?

Inevitably people ask how well do these non-linear models price traded contracts. Is it necessary to calibrate them? Take a look at Equation (60.2) and imagine what would happen if you tried to price, not an exotic contract, but a vanilla option, vanilla option 13, say.

What would happen is that the optimizer would decide to sell one of the 13th vanilla for hedging, leaving you with a portfolio containing nothing, and according to (60.2) the theoretical price of the vanilla would be $\text{CostofVanilla}_{13}$, i.e. the market price. End result, perfect calibration. Better than that, just change the signs and you will see that this method calibrates to bid *and* offer, whereas with other calibration techniques you have to pick one or the other or the mid-point to match. Furthermore, this technique calibrates to liquidity, because the market price of 1000 options may be different from 1000 times the market price of a single option. This technique knows all about that.

Really calibration is not the right word for what this technique does. All theoretical prices match the market by definition.

As a final thought in this section, observe how I never said that we 'believed' the price of market traded contracts, only that they *may possibly turn out to be correct*. This is contrast to the less satisfying philosophy of Chapter 50 where I showed how to 'fit' the volatility structure to match traded prices.

It is a slightly more complicated matter to extend the ideas to path-dependent options. In the following sections I show how this is done for American and barrier options.

60.10 HEDGING PATH-DEPENDENT OPTIONS WITH VANILLA OPTIONS, NON-LINEAR MODELS

Valuation with a non-linear model and optimal static hedging are more complicated for American and barrier options (and some other path-dependent contracts) because at the onset of the contract we do not know whether it will still exist by the final expiry date: We may start out with an excellent static hedge for a contract but after a few months we may find ourselves using the same hedge for a nonexistent option.

60.10.1 American Options

The trick to valuing American options in the non-linear framework is to realize that there are two possible states for the American option: Exercised and not exercised. The reason why this is important is to do with static hedging. If we put a static hedge in place it is supposed to be optimal. But what is optimal while the option remains unexercised will clearly not be optimal afterwards. Let's go straight to looking at a portfolio containing one American option and other Europeans.

Let $V_{NL}(S, t, 0)$ and $V_{NL}(S, t, 1)$ be the values of the portfolio of options before and after the American option has been exercised.

Both of $V_{NL}(S, t, 0)$ and $V_{NL}(S, t, 1)$ will satisfy the non-linear governing equation. The final conditions and jump conditions will be different for $V_{NL}(S, t, 0)$ and $V_{NL}(S, t, 1)$. The former will include the American option conditions whereas the latter won't. Optimality of exercise will be guaranteed by having

$$V_{NL}(S, t, 0) \geq V_{NL}(S, t, 1) + \max(S - E, 0),$$

for all times prior to the expiration date of the American option (assuming that the American option is a call option). Apart from the nonlinearity this is the same kind of problem as for American options generally.

To figure out the optimal static hedge you must add into the portfolio (into both $V_{NL}(S, t, 0)$ and $V_{NL}(S, t, 1)$) an arbitrary portfolio of vanillas and then optimize over the quantities of exotics so as to maximize today's value of $V_{NL}(S, t, 0)$.

60.10.2 Barrier Options

The trick to valuing barrier options is, just as with American options, to realize that there are two possible states for the option: Untriggered and triggered. We can go from the former to the latter but not vice versa. Whatever static hedge we choose must take this into account. I introduce the terms 'active' and 'retired' to describe options that still exist and have been triggered, respectively.

Let $V_{NL}(S, t, 0)$ and $V_{NL}(S, t, 1)$ be the values of the portfolio of options before and after the barrier has been triggered. Both of these functions satisfy Equation (60.1), with final conditions at time T depending on the details of the option contracts. At the barrier the values of the two portfolios will be the same.

As a first example suppose that we hold a down-and-out barrier call only, i.e. there is no static hedge, and after the option is retired we hold an empty portfolio. The problem for $V_{NL}(S, t, 1)$ is simply Equation (60.1) with zero final conditions so that

$$V_{NL}(S, t, 1) = 0.$$

Of course, the solution for $V_{NL}(S, t, 1)$ is zero, since it corresponds to an empty portfolio. The problem for $V_{NL}(S, t, 0)$ is Equation (60.1) with

$$V_{NL}(S, T, 0) = \max(S - E_1, 0).$$

Finally, on the barrier

$$V_{NL}(X, t, 0) = V_{NL}(X, t, 1) = 0.$$

The problem for $V_{NL}(S, t, 1)$ only holds for $S > X$.

Now consider the slightly more complicated problem of hedging the barrier call with a vanilla call; we will hold short λ of a vanilla call, both contracts having the same expiry (although this can easily be generalized). The strike price is E_2. Thus we have hedged a barrier call with a vanilla call. The problem for $V_{NL}(S, t, 1)$ is simply Equation (60.1) with

$$V_{NL}(S, T, 1) = -\lambda \max(S - E_2, 0).$$

The problem for $V_{NL}(S, t, 0)$ is Equation (60.1) with

$$V_{NL}(S, T, 0) = \max(S - E_1, 0) - \lambda \max(S - E_2, 0).$$

Finally, on the barrier

$$V_{NL}(X, t, 0) = V_{NL}(X, t, 1).$$

The problem for $V_{NL}(S, t, 0)$ is to be solved for all S, but the problem for $V_{NL}(S, t, 1)$ only holds for $S > X$.

The ideas of optimal static hedging carry over; for example, what is the optimal choice for λ in the above?

Example

The stock price is \$100, volatility is assumed to lie between 20% and 30%, the spot interest rate is 5%. What is the optimal static hedge, using vanilla calls with strike \$110 and costing \$4.22, for a long position in a down-and-out barrier call with strike \$90? Both options expire in six months. In other words, what is the optimal number of vanilla calls to buy or sell to get as much value as possible out of the barrier option?

With no static hedge in place we simply solve the above problem with $V^{-}(S, t, 1) = 0$, since if the option knocks out that is the end of the story. In this case we find that the barrier option is worth \$6.32.

With a hedge of -0.47 of the vanillas the barrier is now worth \$6.6. This is the optimal static hedge.

60.10.3 Pricing and Optimally Hedging a Portfolio of Barrier Options

The pricing of a *portfolio* of barriers is a particularly interesting problem in the non-linear equation framework. Again, this is best illustrated by an example. What is the worst value for a portfolio of one down-and-out call struck at \$100 with a barrier at \$90 and an up-and-in put struck at \$110 with a barrier at \$120? The volatility lies between 20 and 30%, the spot interest rate is 5%, there are no dividends, there are six months to expiry and the underlying asset has value \$102.

The first step is to realize that instead of there being two states for the portfolio (as in the above where the single barrier is either active or retired), now there are *four* states: Each of the two barrier options can be either active or retired. In general, for n barrier options we have 2^n states; the vanilla component of the portfolio is, of course, always active. As long as the barriers are not intermittent then there is a hierarchy of barrier options that can be exploited to reduce the computational time dramatically. This hierarchy exists because the triggering of one barrier means that barriers closer in are also triggered. This issue is discussed by Avellaneda & Buff (1997).

Returning to the two-option example, the four states are:

1. Both options are active; this is the initial state.
2. If the asset rises to 120 without first falling to 90 then the up barrier option is retired.
3. If the asset falls to 90 before rising to 120 then the down barrier option is retired.
4. If both barriers are triggered then both options are retired. The only active options will be any vanillas in the static hedge.

To solve this problem we must introduce the function $V_{NL}(S, t; i, j)$. The portfolio value with both options active is $V_{NL}(S, t; 0, 0)$, with the down-and-out option retired but the up-and-in still active $V_{NL}(S, t; 1, 0)$, with the down option active but the up retired $V_{NL}(S, t; 0, 1)$ and with both options retired $V_{NL}(S, t; 1, 1)$. In the absence of any static hedge, the problem to solve is Equation (60.1) for each V_{NL} with the following final conditions:

- $V_{NL}(S, T; 0, 0) = \max(S - 90, 0)$ (since at expiry, if both options are active only the down-and-out call pays off);
- $V_{NL}(S, T; 1, 0) = 0$ (if the down-and-out has been triggered it has no pay off);
- $V_{NL}(S, T; 0, 1) = \max(S - 90, 0) + \max(120 - S, 0)$ (if the up knocks-in it has the put payoff);
- $V_{NL}(S, T; 1, 1) = \max(120 - S, 0)$ (if both options are triggered the payoff is that for the 'in' option).

The retirement of an option is expressed through the following boundary conditions:

- $V_{NL}(90, t; 0, 0) = V_{NL}(90, t; 1, 0)$ and $V_{NL}(90, t; 0, 1) = V_{NL}(90, t; 1, 1)$ (the option value just before knock-out is equal to that just after, regardless of the status of the in option);
- $V_{NL}(120, t; 0, 0) = V_{NL}(120, t; 0, 1)$ and $V_{NL}(120, t; 1, 0) = V_{NL}(120, t; 1, 1)$ (the option value just before knock-in is equal to that just after regardless of the status of the out option).

60.11 THE MATHEMATICS OF OPTIMIZATION

We have encountered several non-linear problems, and will meet some more later on. I have explained how the nonlinearity of these models allows static hedging to be achieved in an optimal fashion. Mathematically this amounts to finding the values of parameters such that some function is maximized. This is an example of an **optimization problem**. In this section I will briefly describe some of the issues involved in optimization problems and suggest a couple of methods for the solution.

Typically we are concerned with a maximization or a minimization. And since the minimization of a function is the same as the maximization of its negative, I will only talk about maximization problems. We will have some function to maximize, over a set of variables. Let's call the variables $\lambda_1, \ldots, \lambda_N$ and the function $f(\lambda_1, \ldots, \lambda_N)$. Note that there are N variables

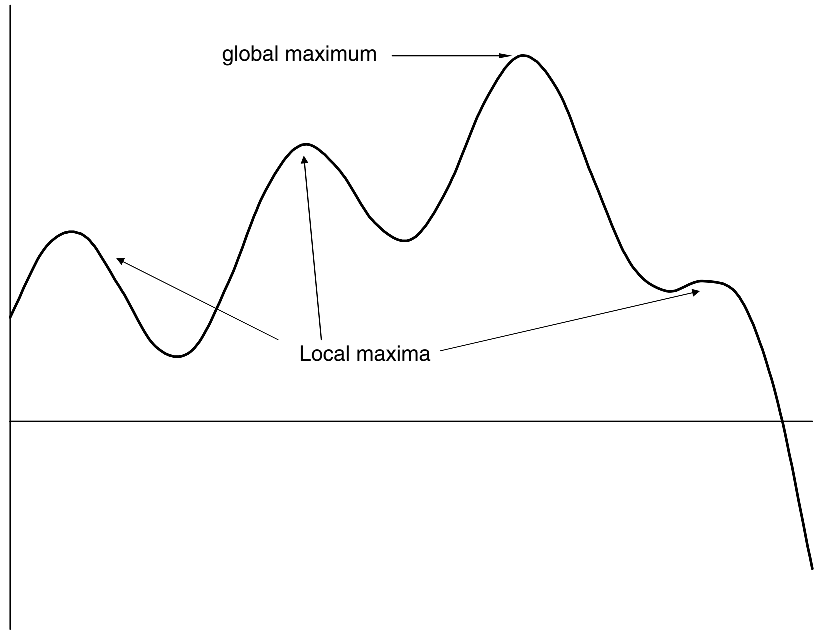


Figure 60.3 Local and global maxima.

so that we are trying to find the maximum over an N-dimensional space. In our financial problems we have, for example, a set of λs representing the quantities of hedging contracts and the function f is the value of our target contract.

There are two kinds of maximum, **local** and **global**. There may be many local maxima; each is the highest point in its neighborhood, but it may not be the highest point over all of the N-dimensional space. What we really want to find is the global maximum, the point having the highest function value over the whole of the space of the variables. In Figure 60.3 is shown a one-dimensional function together with a few local maxima but only one global maximum. Whatever technique we use for finding the maximum must be able to find the global maximum.

60.11.1 Downhill Simplex Method

The first method I want to describe is the **downhill simplex method**. This method is not particularly fast, but it is very robust and easy to program. If there is more than one local maxima then it may not find the global maximum; instead it might get stuck at a local high spot. I will describe a way around this weakness later. Traditionally, one talks about function minimization, and so the word 'downhill' makes sense. Because our problems are typically ones of maximization perhaps we should use the word 'uphill.'

A **simplex** in N dimensions is made up of $N + 1$ vertices. In two dimensions this gives a triangle, in three a tetrahedron. The downhill simplex method starts with an initial simplex, think of it as a ballpark guess for the position of the maximum. Suppose that the $N + 1$ vertices of this simplex are at the points given by the vectors $\boldsymbol{\lambda}_i$ for $i = 1, \dots, N + 1$. Calculate the values of the function at each of the vertices; these are $f(\boldsymbol{\lambda}_i)$. One of these values will be lower than the other N values. If this value is particularly bad, i.e., low, perhaps we should look far

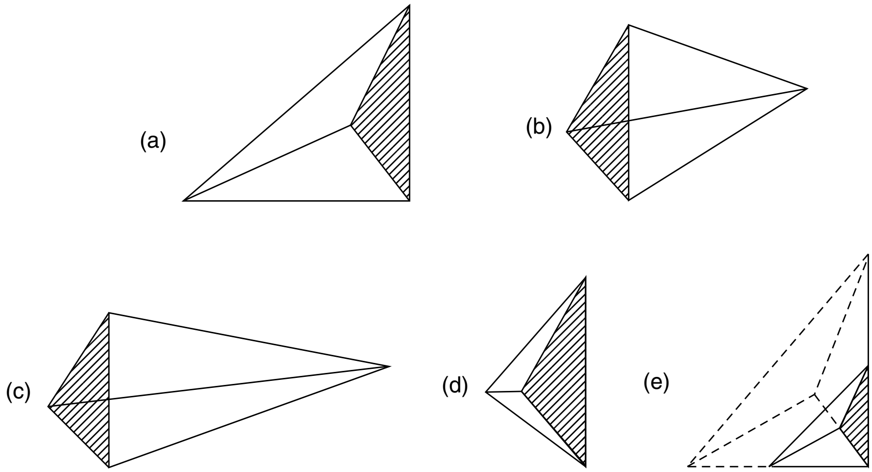


Figure 60.4 (a) Initial simplex, (b) reflection, (c) reflection and expansion, (d) contraction in one dimension and (e) multiple contractions to the highest point.

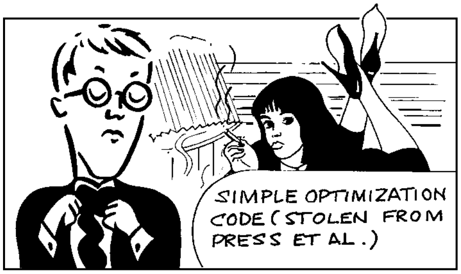
away from this vertex for another guess at the maximum. One systematic way of doing this is to reflect the worst point in the 'plane' made up by the remaining N points. This is easy to see for the tetrahedral simplex shown in Figure 60.4. In this figure (a) is the initial state of the simplex. The worst/lowest point is reflected in the 'base' of the tetrahedron resulting in (b). If the function evaluated at the new vertex is even better/higher than all of the values of the function at the other (old) vertices, then expand the tetrahedron a bit to exploit this. This gives (c). On the other hand, if the new vertex has a function value that is worse than the second lowest then the tetrahedron should be contracted as shown in (d), otherwise the new simplex would be worse than before. Finally, if the one-dimensional contraction doesn't make things better then the simplex is contracted about the best vertex in all dimensions simultaneously, giving (e). This algorithm sees the simplex tumbling over itself, rising to the peak of the function where it begins to contract. The process is repeated until the simplex has shrunk enough to give a sufficiently precise estimate of the maximum.

The following code finds a local maximum of the function `func` in `N` dimensions. The code is a translation of the C code, in Press *et al.* (1992), into VB. The number of dimensions, the maximum number of iterations `MaxIts` and the tolerance `tol` must be input, as must the initial simplex. The initial guess goes in the $N+1$ by N array `FirstGuess`; each row is a different vertex.

The subroutine `Ranking` finds the value of the function at, and the index of, the two lowest and the highest vertices. The subroutine `FindCenter` finds the center of the simplex, given by

$$\frac{1}{N+1}\sum_{i=1}^{N+1}\boldsymbol{\lambda}_i.$$

The subroutine `Deform` reflects the lowest point in the opposite face of the simplex by an amount `d`. If `d` $=1$ there is no reflection, if `d` $=1/(N+1)$ then vertex goes to the center of the simplex and if `d` $=0$ then the vertex goes into the opposite plane. Here we only use `d` $=-1$ for case (b), `d` $=2$ for case (c) and `d` $=\frac{1}{2}$ for case (d).




```
Function maximize(FirstGuess As Object, N As _
 Integer, tol As Double, MaxIts As Integer)
ReDim outputarray(1 To N + 1, 1 To N + 1) As Double
ReDim y(1 To N + 1) As Double
ReDim p(1 To N + 1, 1 To N) As Double
ReDim ptemp(1 To N) As Double
Dim ranktest(1 To 3, 1 To 2) As Double

its = 0
rtol = 1

For i = 1 To N + 1
For j = 1 To N
p(i, j) = FirstGuess(i, j)
ptemp(j) = p(i, j)
Next j
y(i) = func(ptemp)
Next i

While (rtol > tol) And (its < MaxIts)

Call Ranking(y, N, ilo, ylo, inlo, ynlo, ihi, yhi)

Call Deform(y, p, N, ilo, ylo, -1)
ytest = ylo

If ytest >= yhi Then
    Call Deform(y, p, N, ilo, ylo, 2)
    ytest = ylo
Else
    If ytest <= ynlo Then
    ysave = y(ilo)
    Call Deform(y, p, N, ilo, ylo, 0.5)
    ytest = ylo
        If ytest <= ysave Then
            For i = 1 To N + 1
            If i <> ihi Then
                For j = 1 To N
                p(i, j) = 0.5 * (p(i, j) + p(ihi, j))
                ptemp(j) = p(i, j)
                Next j
            End If
            y(i) = func(ptemp)
            Next i
        End If
   End If
End If

    its = its + 1
    rtol = 2 * Abs(yhi - ylo) / (Abs(yhi) + Abs(ylo))
Wend

For i = 1 To N + 1
For j = 1 To N
outputarray(i, j) = p(i, j)
```

```
Next j
outputarray(i, N + 1) = y(i)
Next i

maximize = outputarray

End Function

Sub FindCenter(p, N As Integer, center)
For j = 1 To N
center(j) = 0
For i = 1 To N + 1
center(j) = center(j) + p(i, j) / (N + 1)
Next i
Next j
End Sub

Sub Ranking(y, N As Integer, ilo, ylo, inlo, ynlo, ihi, yhi)
ilo = 1
inlo = 1
ihi = 1
ylo = y(ilo)
ynlo = y(inlo)
yhi = y(ihi)
For i = 2 To N + 1
    If y(i) < ynlo Then
        If y(i) < ylo Then
            inlo = ilo
            ynlo = ylo
            ilo = i
            ylo = y(i)
        Else
            inlo = i
            ynlo = y(i)
        End If
    End If
    If y(i) > yhi Then
        ihi = i
        yhi = y(i)
    End If
Next i
End Sub

Sub Deform(y, p, N As Integer, ilo, ylo, d)
ReDim ptest(1 To N) As Double
ReDim center(1 To N + 1) As Double
Call FindCenter(p, N, center)
For j = 1 To N
    ptest(j) = d * p(ilo, j) + (1 - d) * ((N + 1) / N * center(j) - p(ilo, j) / N)
Next j
ytest = func(ptest)
If ytest > ylo Then
    y(ilo) = ytest
    ylo = ytest
End If
```

```
For j = 1 To N
p(ilo, j) = ptest(j)
Next j
End Sub
```

60.11.2 Simulated Annealing

The main downside of the above maximization algorithm is that it will not necessarily find the global maximum. This is a problem with many optimization techniques. Currently, one of the most popular methods for overcoming this is **simulated annealing**. This gets its name from the form of crystallization that occurs as molten metals are cooled. In the optimization algorithm there is a degree of randomness in the search for a maximum, and as the number of iterations increases, so the randomness gets smaller and smaller. The analogy is with the random motion of atoms in the liquid state of the metal; as the metal cools so the motion decreases. The hope is that the random motion will find the neighborhood of the global maximum, and as the 'temperature'/random motion decreases, the search will home in on the maximum.

One of the simplest simulated annealing techniques, related to the **Metropolis algorithm**, uses the addition or subtraction of a random number to the function value at each vertex, but is otherwise the same as the downhill simplex method. Suppose that we have a simplex and know the function values at each vertex. We now subtract a positive random number from the function value at each vertex. When we come to test a new replacement point we add a positive random number. In this way we still always accept a move to a better vertex, but occasionally we accept a move to a worse vertex. However, this worse vertex may actually be closer to the global maximum. How often we accept such a move depends on the size of the random variables. We must choose a distribution for the random variable, and importantly its scale. As the number of iterations increases, so we decrease this scale so that it tends to zero as the number of iterations increase.

The scale of the random moves must decrease slowly enough for the simplex to have a good look around before converging on the global maximum.

60.12 SUMMARY

In this chapter I showed how to statically hedge a complex product with simpler contracts. We began with the classical static hedge, by setting up a portfolio having exactly the same theoretical value as our target contract at a specified set of asset/time points. The resulting hedge then reduces the risk from delta hedging. But it is not perfect; the hedge is still sensitive to the pricing model and its parameters.

Then I described the idea of static hedging for some non-linear models. Because some of these models are less sensitive to parameter values (such as volatility) the resulting hedged contract is also insensitive. We also saw how to optimize this static hedge by maximizing the *marginal value* of the target contract, after allowing for the real cost of setting up the hedge.

FURTHER READING

- See Bowie & Carr (1994), Carr & Chou (1997), Carr, Ellis, & Gupta (1998) and Derman, Ergener & Kani (1997) for more details of static replicating portfolios.

- See Avellaneda & Parás (1996) for optimal static hedging as applied to uncertain volatility, and Oztukel (1996) for uncertain correlation.
- See Epstein & Wilmott (1997, 1998) for optimal static hedging in a fixed-income world.
- For a general description of optimization, with many references, algorithms and code, see the excellent Press *et al.* (1992).
- Avellaneda & Buff (1997) explain how to value large portfolios of barrier options in the uncertain volatility framework.
- Joshi (2003) is very good on the principles and practice of static hedging.

CHAPTER 61

the feedback effect of hedging in illiquid markets

In this Chapter...

- how delta hedging an option position or replication can influence and move the market in the underlying

61.1 INTRODUCTION

I have referred constantly to the 'underlying' asset. The implication of this was that the asset price leads and the option price follows; the option value is contingent upon the value and probabilistic properties of the asset, and not the other way around. In practice in many markets, however, the trade in the options on an asset can have a nominal value that exceeds the trade in the asset itself. So, can we still think of the asset price as leading the option?

In the traditional derivation of the Black–Scholes option pricing equation it is assumed that the replication trading strategy has no influence on the price of the underlying asset itself, that the asset price moves in a random way. This is a bit like saying that options trading is a zero-sum game; for every winner there is an equal and opposite loser. Sometimes the justification for this is the action of **noise traders** or the random flow of information concerning the asset or the economy. Nevertheless, a significant number of trades are for hedging or replication purposes. And, crucially, these trades are for predictable amounts. What is their impact on the market?

Usually it is assumed that the effect of individual trading on the asset price is too small to be of any importance and is neglected when the strategy is derived. This seems justifiable if the market in question has many participants and a high degree of liquidity, which is usually true for modern financial markets. On the other hand, the portfolio insurance trading strategies are very often implemented on a large scale, and the liquidity of the financial markets is sometimes very limited. In the case of the October 1987 stock market crash some empirical studies and even the official report of the investigations carried out by the Brady commission suggest that portfolio insurance trading contributed to aggravate the effects of the crash.

In this chapter we are going to address the problem of the influence of these trading strategies on the price of the underlying asset and thus, in a feedback loop, onto themselves.

61.2 THE TRADING STRATEGY FOR OPTION REPLICATION

In theory, any simple option can be replicated by following the appropriate trading strategy. These trading strategies have enjoyed tremendous popularity among portfolio managers who use them to insure themselves against large movements in the share price. This strategy is called **portfolio insurance**. One of the most popular portfolio insurance strategies is the replication of a European put option. Any simple option having a value $V(S,t)$ can be *replicated* by holding

$$\Delta(S,t) = \frac{\partial V}{\partial S}(S,t) \tag{61.1}$$

shares at time t if the share price is S. (Contrast this with *hedging* an option for which we must hold this number short.) In particular, Black and Scholes found an explicit formula for the value $V(S,t)$ of the put option. As derived in Chapter 5, this formula is

$$V(S,t) = Ee^{-r(T-t)}N(-d_2) - S\,N(-d_1)$$

where

$$d_1 = \frac{\log(S/E) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

For the European put Equation (61.1) results in

$$\Delta(S,t) = N(d_1) - 1.$$

We can now think of $\Delta(S,t)$ as corresponding to a trading *strategy*. Put replication is one of the trading strategies, corresponding to one example of portfolio insurance, that we are going to

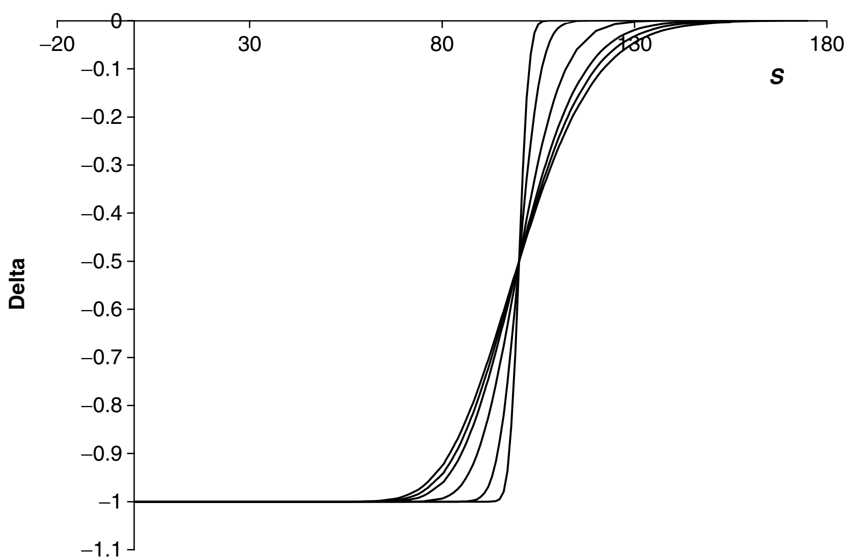


Figure 61.1 The delta for a European put at several times before expiry.

analyze in detail. Figure 61.1 shows the delta of a European put at various times before expiry; the steeper the curve, the closer the option is to expiry.

In the next section I propose a form for the reaction of the market to a trading strategy Δ and will derive the modification of the price process that takes account of the effect of trading.

61.3 THE EXCESS DEMAND FUNCTION

In order to quantify the effects of replication on the movement of the asset we need a model describing the relationship between supply of and demand for an asset and its price. Consider the difference between demand and supply in the market. This is the **excess demand** which we assume takes the form $\mathcal{X}(S, t, x)$ i.e. it is a function of price S, time t and, importantly, a random influence x. The random influence will ensure that our model does not stray too far from the classical lognormal; this will be a limiting case of our more complex model. We will model x as a random variable shortly. Its influence can be regarded as the effect of randomly arriving new information on the value of the underlying asset or the action of noise traders. Generally $(\partial/\partial S)\mathcal{X}$ is negative; rising price leads to falling demand.[1]

At any given time the equilibrium price $S_{Eq}(t)$ is the price for which demand is equal to supply, or excess demand is zero;

$$\mathcal{X}(S_{Eq}(t), t, x) = 0. \tag{61.2}$$

Typically, there is a strong tendency in any market to return to the equilibrium price if it has been disturbed. Surplus demand pushes the price up, towards the equilibrium price; excess supply makes the price drop, again towards the equilibrium price. The market equilibrium is stable. Note that this argument supposes that the excess demand function is negatively sloped; for a positive slope the equilibrium would still exist but it would be unstable. We will encounter locally positively sloped excess demand functions later on.

Disequilibrium is obviously possible but given the speed of the flow of information in these markets and the large number of professionals on the stock markets a full equilibrium in stocks and flows in modern financial markets is a good approximation. This does not mean that these markets are static: In our model both demand and supply can change in time because of the stochastic parameter x.

61.4 INCORPORATING THE TRADING STRATEGY

In this section we add to the original demand the extra demand resulting from the hedging of the put option. Not only will we have the random demand due to the noise traders but also a completely deterministic demand, due to the trading strategy Δ. With this additional demand of the form $\Delta(S, t)$ the equilibrium condition (61.2) becomes

$$\mathcal{X}(S, t, x) + \Delta(S, t) = 0 \tag{61.3}$$

[1] This is necessarily a simplification of the real story, but nevertheless a simplification often made in economic theory and one which holds the essence of truth.

(dropping the subscript on the equilibrium price) and the same must hold for the changes in $\mathcal{X}$ and Δ

$$d\mathcal{X} + d\Delta = 0.$$

We could consider arbitrary excess demand function $\mathcal{X}$, but for simplicity we assume from now on that

$$\mathcal{X}(S, t, x) = \frac{1}{\epsilon}(x - S).$$

Here ϵ is a positive real number and x follows the stochastic process

$$dx = \mu_x \, dt + \sigma_x \, dX;$$

μ_x and σ_x can be functions of x and t. We can now think of x as being the 'intrinsic value' of the stock.

The parameter ϵ shows how strongly the excess demand function reacts to changes in the price. If the price changes by dS then the excess demand changes by $-dS/\epsilon$. It gives an indication of the liquidity of the market. Liquid markets react very strongly to changes in the price; for such markets ϵ is small:

- A *liquid market* is a market in which ϵ is *small*.
- An *illiquid market* is a market in which ϵ is *large*.

As mentioned before most—but not all—financial markets are liquid markets.

If the parameter ϵ is assumed to be zero this reduces the price process S to the random walk x. We will not take this step but merely assume ϵ to be small. We will see later that in certain cases no matter how small ϵ is the individual effect on the excess demand cannot be neglected.

Now (61.3) becomes

$$(x - S) + \epsilon \Delta(S, t) = 0. \tag{61.4}$$

In the undisturbed equilibrium, with appropriate choice of scalings, ϵ^{-1} is also equal to the **price elasticity of demand**.

Applying Itô's lemma to (61.4) we find that the stochastic process followed by S is

$$dS = \mu_S(S, t) \, dt + \sigma_S(S, t) \, dX \tag{61.5}$$

with μ_S and σ_S given by

$$\mu_S = \frac{\sigma_S}{\sigma_x} \left(\mu_x + \epsilon \left(\frac{\partial \Delta}{\partial t} + \tfrac{1}{2} {\sigma_S}^2 \frac{\partial^2 \Delta}{\partial S^2} \right) \right)$$

and

$$\sigma_S = \frac{\sigma_x}{1 - \epsilon \dfrac{\partial \Delta}{\partial S}}.$$

The details are left to the reader; the application of Itô's lemma is straightforward.

So far all analysis has been done for an arbitrary stochastic process for x. Usually the price process of share price is assumed to be lognormal, i.e.

$$dS = \mu S\,dt + \sigma S\,dX \tag{61.6}$$

with μ and σ constant. Our analysis of the model on the other hand yielded the modified price process (61.5). As the equilibrium price S is a known function of x we can choose

$$\mu_x = \mu S \quad \text{and} \quad \sigma_x = \sigma S$$

to achieve consistency between (61.5) and (61.6) in the sense that (61.5) reduces to (61.6) when $\epsilon = 0$.

61.5 THE INFLUENCE OF REPLICATION

One of the most important portfolio insurance strategies Δ is put replication as explained in Section 61.2. Look again at Figure 61.1 to see the delta of the European put. As the expiry date of the option is approached Δ goes towards a step function with step from -1 to 0 at the exercise price.

Recalling Equation (61.4) the equilibrium condition can be written as

$$-x = -S + \epsilon\Delta. \tag{61.7}$$

The effect of the trading strategy is a small (i.e. of order ϵ) perturbation added to the original demand function. Far from expiry the right-hand side of (61.7) is simply $-S$ since $\epsilon\Delta$ is

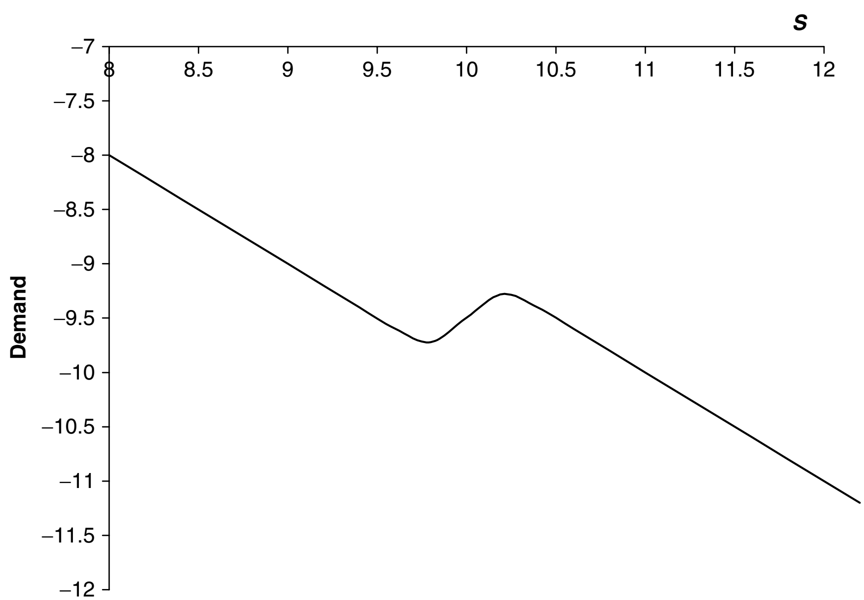


Figure 61.2 The sum of the original linear demand function and the extra demand due to put replication.

small. Close to expiry the Δ term becomes important and the shape of the demand curve alters dramatically, becoming as shown in Figure 61.2. Figure 61.2 shows the right-hand side of (61.7) for the put replication strategy; we are sufficiently close to expiry for the curve to no longer be monotonic.

As expiry approaches the sequence of events is as follows. Far from expiry there is hardly any deviation from the normal linear function. In an interval around the expiry price the demand curve's slope becomes less and less negative until it is positively sloped. Thus there are times very close to expiry and close to the strike for which there is an unstable equilibrium.

Geometrically, Equation (61.7) can be interpreted as follows. For a given t, the market equilibria are defined as the points where the randomly moving horizontal line $-x$ intersects the function $-S+\epsilon\Delta$. An example of this is shown in Figure 61.3. As long as $-S+\epsilon\Delta$ is monotonically decreasing the equilibrium is unique for any value of x. But because Δ approaches a step function as expiry is approached, $-S+\epsilon\Delta$ must become multi-valued at some point. From this time on the situation becomes more complicated.

Refer to Figure 61.3. In this figure the horizontal line denotes the left-hand side in the equilibrium condition (61.7) and the curve is the corresponding right-hand side for t close to expiry T. There are various possible situations depending on the value of x at that point in time:

1. x is sufficiently large or small as to be outside the critical region where there are several equilibria. The equilibrium asset price is unique, corresponding to the asset value where the horizontal line crosses the curve.

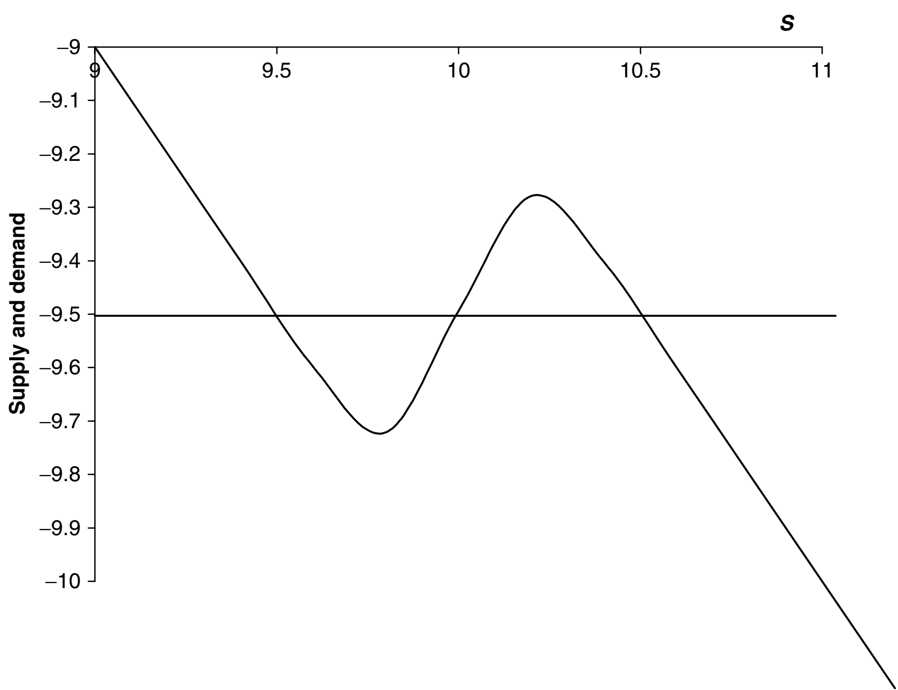


Figure 61.3 The intersection of the supply and demand curves. There are four cases depending on the relative position of the intersection and the maxima and minima.

2. x is inside the critical region where there are three possible equilibria. This is the case shown in the figure. The middle equilibrium value for the asset is unstable, corresponding to a positively sloped demand function. The other two are stable.
3. The limiting case between cases 1 and 2. If the horizontal line just touches the curve, then we have two equilibria. One of these is stable, the other, corresponding to the asset value where the line just *touches* the curve, is stable only for movements in the random variable x of one sign. At this point either the equilibrium asset value will move continuously or will jump to the next equilibrium point: if a jump occurs there is a discontinuity in the price.

The third case is of particular interest. What happens to the equilibrium asset value is sketched in Figure 61.4. The points A, B, C and D are defined by

$$(A(t), -A(t) + \epsilon\Delta(A(t), t)),$$

etc. as follows. $S = B(t)$ and $S = C(t)$ are the local extrema of the demand function, B is the local minimum and C is the local maximum. The definitions of these points are

$$-1 + \epsilon\frac{\partial\Delta}{\partial S}(B, t) = 0 \tag{61.8}$$

$$-1 + \epsilon\frac{\partial\Delta}{\partial S}(C, t) = 0 \tag{61.9}$$

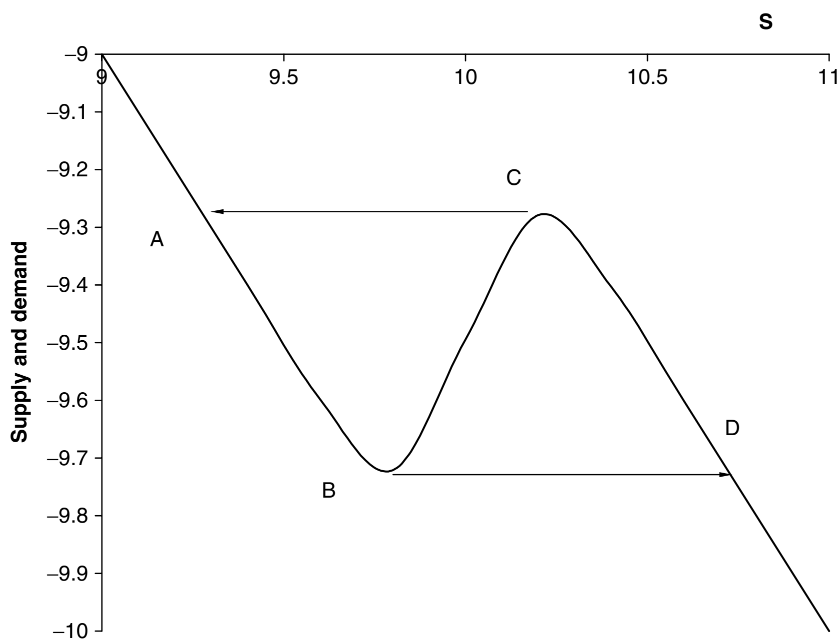


Figure 61.4 The positions of the four points A, B, C and D.

and

$$\frac{\partial^2}{\partial S^2}\Delta(B,t) \geq 0 \geq \frac{\partial^2}{\partial S^2}\Delta(C,t). \tag{61.10}$$

$S = A(t)$ and $S = D(t)$ are defined as:

$$A \leq B \leq C \leq D \tag{61.11}$$

$$[-S + \epsilon\Delta]\,(A,t) = [-S + \epsilon\Delta](C,t) \tag{61.12}$$

$$[-S + \epsilon\Delta]\,(D,t) = [-S + \epsilon\Delta](B,t). \tag{61.13}$$

Thus there are four different points of interest: The two extrema *from* which the jumps come and the two points *to* which the price jumps.

The arrows in Figure 61.4 mean that the asset can jump from one point of the curve to another, that is from B to D and from C to A. But that only happens if the increments of x are of the right sign when the asset is at B or C. If the increment is of the opposite sign then the asset price at B could fall by an infinitesimal amount, or at C rise by an infinitesimal amount. I emphasize the point that the jump is only one way.

The same result can be found by an examination of the drift μ_S and variance σ_S of the modified price process. Both have a term of the form

$$1 - \epsilon\frac{\partial\Delta}{\partial S}$$

in the denominator. This is the negative of the derivative of the total demand function. When this becomes equal to zero the demand function has zero slope, as case 3 above. Here μ_S and σ_S approach infinity.

But even when $\partial\Delta/\partial S < \epsilon^{-1}$, a positive $\partial\Delta/\partial S$ has the effect of increasing both σ_S and the absolute value of μ_S: The market becomes more volatile. If conversely $\partial\Delta/\partial S$ is negative, its effect is to decrease the volatility of the market. If Δ is regarded as the trading strategy to replicate a derivative security V, the relation $\Delta = \partial V/\partial S$ yields that:

- Replication of a derivative security V with positive gamma $\Gamma = \partial^2 V/\partial S^2$, i.e. with concave payoff profile, destabilizes the market of the underlying.

Long positions in put and call options have positive gamma.

61.6 THE FORWARD EQUATION

We can analyze the new stochastic process (61.5) by examining the probability density function $p(S,t)$ which gives the probability density of the share price being at S at time t subject to an initial distribution. Here S and t are the forward variables, usually written as S' and t'.

The probability density function satisfies the Kolmogorov forward equation

$$\frac{\partial p}{\partial t}(S,t) = \tfrac{1}{2}\frac{\partial^2}{\partial S^2}(p(S,t)\,{\sigma_S}^2(S,t)) - \frac{\partial}{\partial S}(p(S,t)\,\mu_S(S,t)).$$

We will use a delta function initial condition, meaning that we know exactly where the asset starts out.

61.6.1 The Boundaries

As pointed out in Section 61.5 there are two pairs of jump boundaries in the area close to expiry (see Figure 61.4). Jumps[2] can occur from point B to point D and from point C to point A.

The positions of these points change in time. A typical graph for a put-replicating trading strategy is shown in Figure 61.5. For any continuous time-dependent trading strategy all four boundary points have to arise at the same point. This point is characterized as a point of inflexion of the full demand function $x - S + \epsilon\Delta$, which obviously satisfies the conditions (61.8) to (61.13). In Figure 61.5 this point is marked as I. Subsequently the points will constantly maintain the order $A < B < C < D$ and fan out as shown.

The put-replicating strategy shown in Figure 61.4 approaches a step function as expiry approaches. B and C disappear together at the exercise price and expiry date (point II). Points A and D approach $E - \epsilon$ and $E + \epsilon$ respectively resulting in a curve looking like a tulip lying on its side. Note that all the boundary curves are confined to an area of order $(\epsilon \times \epsilon^2)$ around (E, T).

Points B and C act locally as absorbing boundaries. The appropriate boundary conditions are therefore

$$p(B, t) = 0 = p(C, t).$$

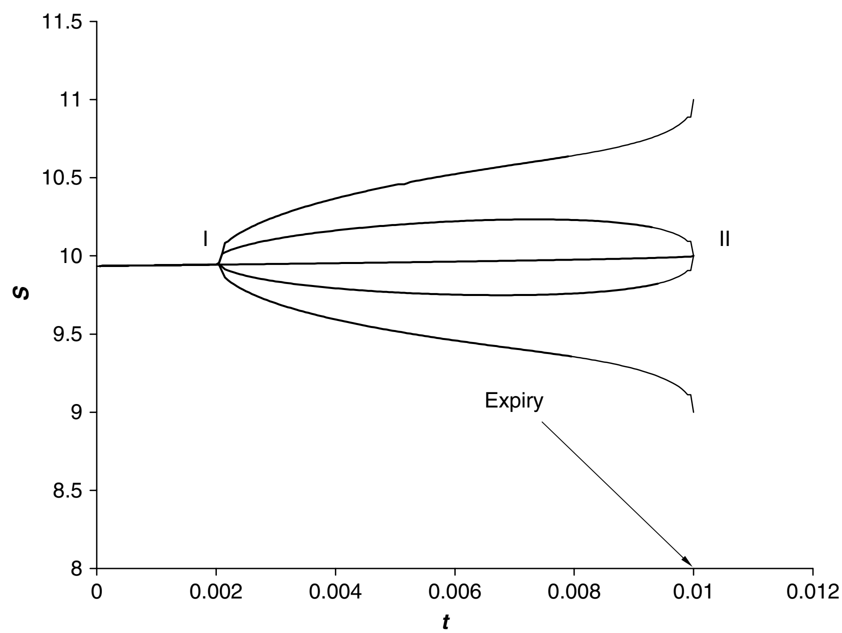


Figure 61.5 The time-dependent positions of the four points A, B, C and D for the put-replicating strategy.

[2] Deterministic jumps are not allowed in classical models of assets since they lead to arbitrage opportunities. We discuss this point later.

The boundary conditions for points A and D are not so obvious because μ_S and σ_S have singularities as the jump points B and C are approached. They can be derived from a conservation of probability argument, see Schönbucher (1993) for further details.

The boundary conditions at $S = 0$ and $S \to \infty$ are

$$p(0, t) = 0 \quad \text{and} \quad \lim_{S\to\infty} p(S, t) = 0.$$

61.7 NUMERICAL RESULTS

In this section I present numerical results for the solution of the forward equation in three cases. The parameters are as follows (unless otherwise stated): $E = 10$, $r = 0.1$, $\mu_x = 0.2$, $\sigma_x = 0.2$, and $\epsilon = 1$.

The first example is the evolution of the probability density function in the absence of any feedback. Thus $\epsilon = 0$. In Figure 61.6 is shown a three-dimensional plot of the probability density function against asset price and time. The starting condition is a delta function at $S = 10$. As time increases the curve flattens out as a lognormal density function. In Figure 61.7 is shown a contour map of this same function.

The first non-trivial case is a time-invariant trading strategy that is strong enough to give rise to the four new boundaries A, B, C and D. The time-independent addition to the demand, Δ, was taken to be that from a put-replicating trading strategy with time fixed at $t = T - 0.05$, shortly before expiry. The initial price is assumed to be known $S = 11$, thus the probability density function contains a delta function at $S = 11$. This case helps to visualize the effects of the boundaries. This may perhaps model the effect of a large number of investors all delta

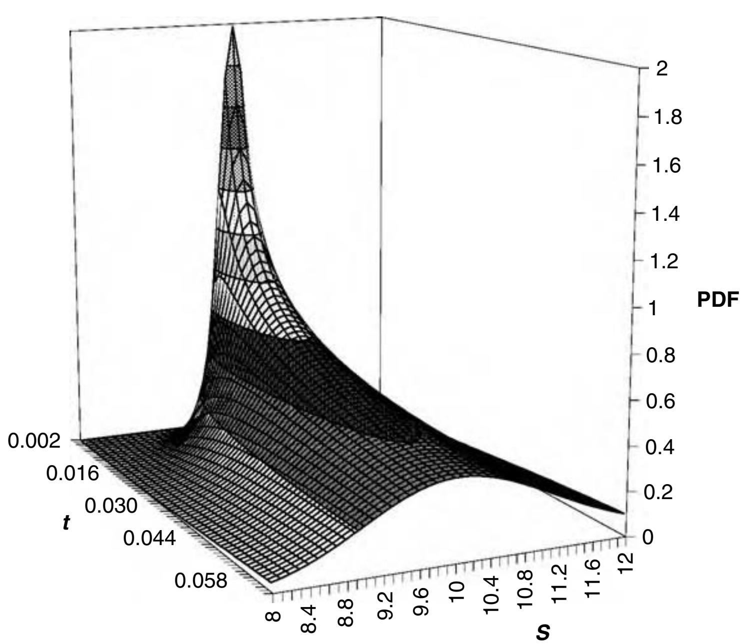


Figure 61.6 The probability density function for an asset with no feedback. The initial condition is a delta function.

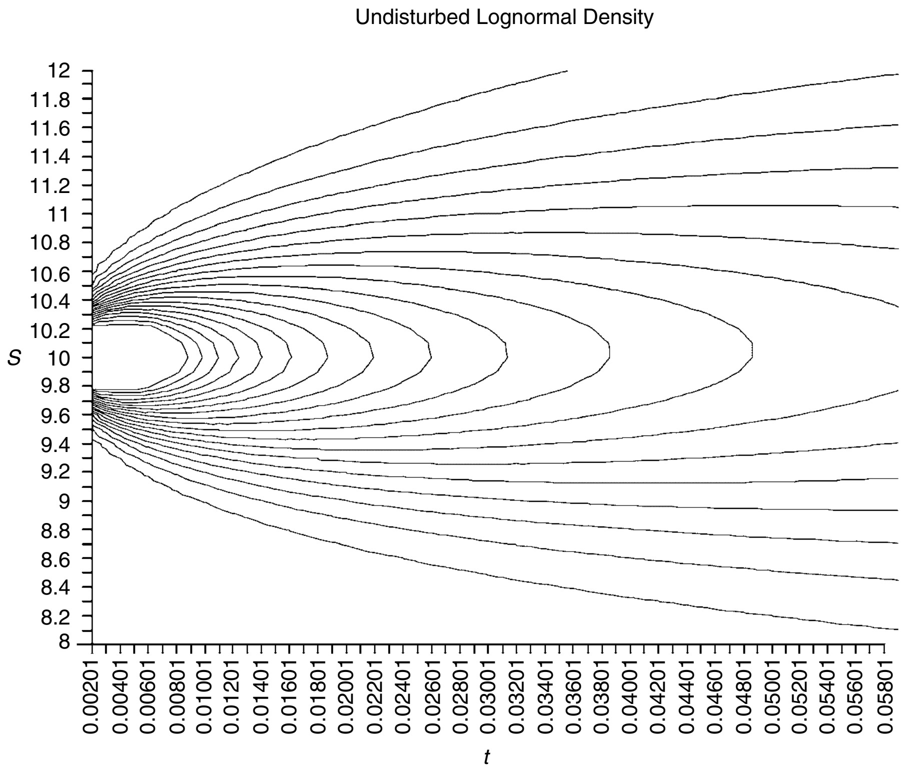


Figure 61.7 The contour plot of Figure 61.6.

hedging using the delta of an option with the same strike but each with a different expiry date such that the overall effect of these hedgers is to add a time-independent function to the demand. This is the simplest case to consider since the coefficients in the forward equation are time-independent.

The final case is simple put replication with only one investor moving the market. The trading strategy is the replication strategy for a European put option. This is one of the most popular trading strategies used by portfolio managers and of direct relevance for option pricing. The addition to the demand due to this one large investor is now time-dependent.

Let's take a look at the last two examples in detail.

61.7.1 Time-independent Trading Strategy

Time-invariant hedging schemes are mainly used to maintain the general performance of the portfolio without the need to satisfy any precise conditions at fixed points in time (such as the potential liabilities from writing an option). Here we will use the delta of a European put at time 0.05 before expiry.

Figure 61.8 shows the full development of the probability density function according to the feedback model as time proceeds. At $t = 0$ the probability density function is a delta function at $S = 11$, meaning that the price at $t = 0$ is known to be 11. Later the probability density function spreads out. Figure 61.9 shows the contour plot for the probability density function.

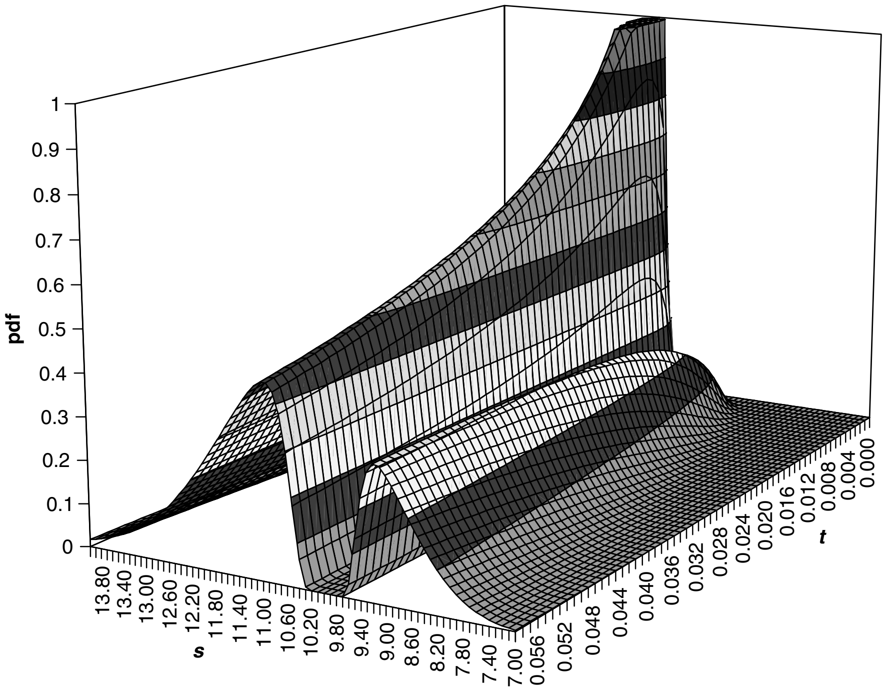


Figure 61.8 The probability density function for a time-independent trading strategy. The initial condition is a delta function.

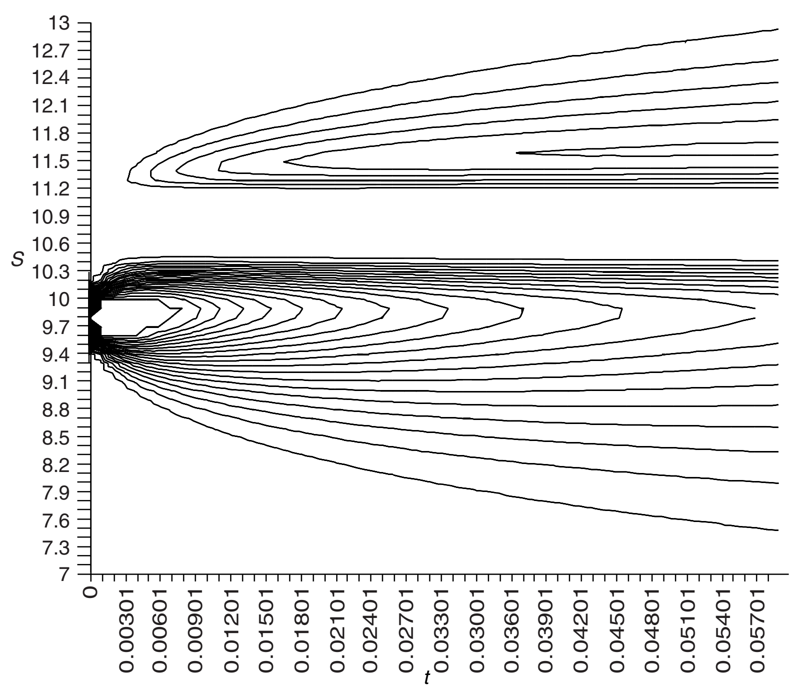


Figure 61.9 The contour plot of Figure 61.8.

The contour plot shows very clearly the unattainable region, the 'corridor,' between asset values of approximately 9.6 and 10.3 which the asset price can never reach. This region is that between the two points previously labeled B and C. Since the replication strategy is time-independent this corridor does not change shape. Even though the starting value for the asset ($S = 11$) is above this region the asset can still reach values less than 9.6 by reaching the barrier C and jumping across to the point A. For more realistic values of ϵ this corridor is very narrow and, away from the corridor, is effectively only a small perturbation to the usual lognormal probability density function as shown in Figures 61.6 and 61.7.

61.7.2 Put Replication Trading

The more interesting case is the time-dependent put-replicating trading and the development in the area around the 'Tulip curve.'

The probability density function is shown in Figure 61.10, the corresponding contour plot is shown in Figure 61.11. Since this replication strategy is genuinely time dependent the corridor that we saw in the above example is now the tulip shape shown in Figure 61.5. The unattainable barren region (the center of the 'Tulip curve') is most easily seen on the contour plot.

In this example, the asset price starts off at $S = 10$ and then evolves. The effect of the replication strategy is felt immediately but the tulip curve itself does not appear until about $t = 0.03$. At this time there appears the barren region around the exercise price (10) which the asset price avoids. This zone is shown most clearly in the contour plot of Figure 61.11. After expiry of the replicated option, $t = 0.04$, the barren zone disappears and all values of S are attainable.

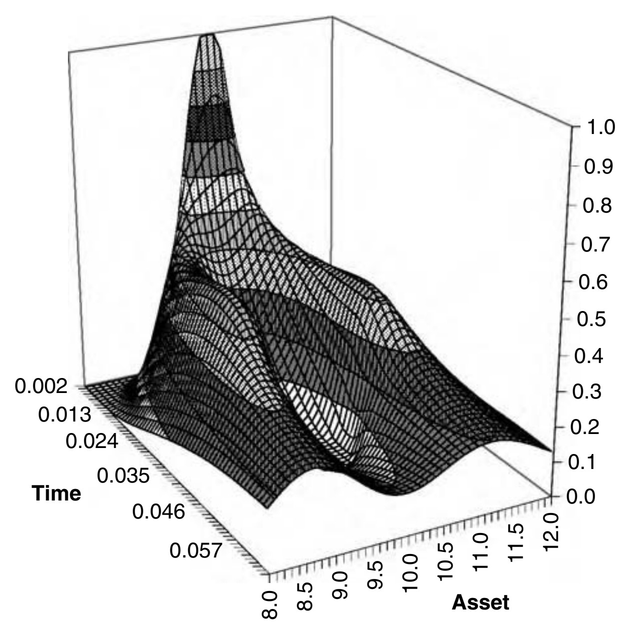


Figure 61.10 The probability density function for put replication; the initial condition is a delta function.

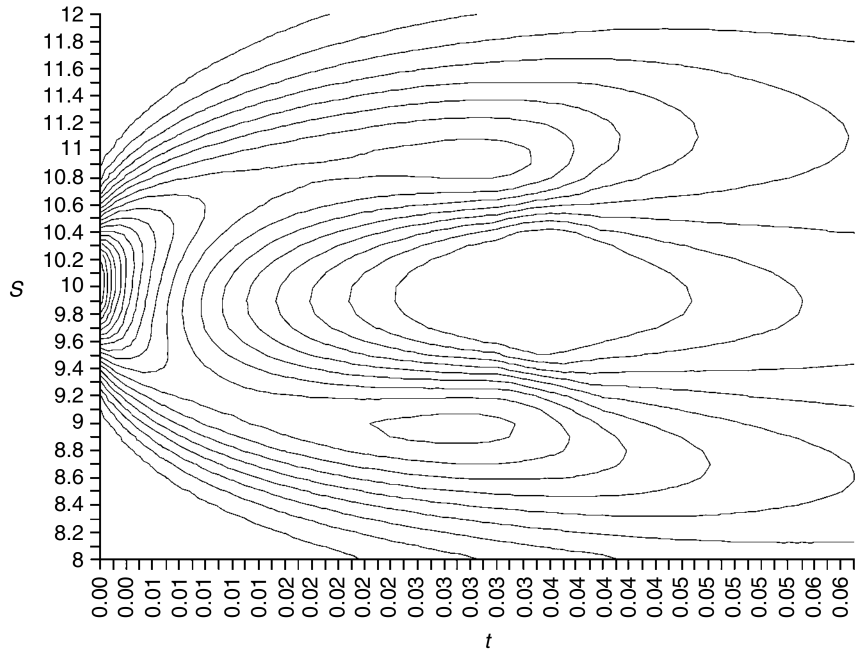


Figure 61.11 The contour plot of Figure 61.10.

61.8 ATTRACTION AND REPULSION

The above analysis focuses on the hedging of short gamma positions. The end result is the zone of repulsion around the strike close to expiration.

If people are, on balance, hedging long gamma positions then we get the opposite effect, a zone of attraction. The mathematics is very similar to much of the above, just with some obvious sign changes. Now the stock price is drawn towards the strike price as expiration approaches. This is not uncommonly seen in practice. A related phenomenon happens with convertible bonds. Upon the issuance of new convertible bonds the volatility of the underlying instrument can be dampened for a while. This is caused by the action of convertible bond arbitragers delta hedging their new bond positions.

61.9 SUMMARY

The influence of trading strategies has been the subject of much discussion but—apart from empirical studies—little theoretical research. In this chapter we have seen a way to incorporate trading strategies formally into the stochastic process followed by the underlying asset.

Many trading strategies that are used today are derived from replicating strategies for derivative securities. This class of trading strategies is also central to the theory of option pricing. We found that the securities whose replication destabilizes the market and their effects on the market can be described by their payoff profile and its gamma.

The effects are especially strong in markets with low liquidity and can even induce discontinuities in the price process. Such price discontinuities are not allowed in classical models of asset prices since they lead to arbitrage. The present model does allow such arbitrage and this can be justified in several ways. First, such effects as I describe do occur in practice and traders with a knowledge of the positions of other market players and their hedging requirements can take advantage of this knowledge. Second, this is just a simple first attempt at modeling feedback. It may be possible to remove certain arbitrage opportunities by incorporating elasticity in the response of the market price to large trades. The barren zone that we have found above may still be unstable, but no longer unattainable.

FURTHER READING

- See Schönbucher (1993), Schönbucher & Wilmott (1995a,b), Frey & Stremme (1995) and Sircar & Papanicolaou (1996) for further examples and analysis of feedback models.
- The book by O'Hara (1995) discusses many of the issues in market microstructure.
- Taleb (1997) discusses the reality of options markets, and feedback, market manipulation and discontinuous price paths are part of that reality. A particularly interesting account of the behavior of asset prices close to the barriers of knock-out options is given there.
- Porter & Smith (1994) describe some laboratory experiments in which simulated stock market trading led to interesting bubbles and crashes.

CHAPTER 62
utility theory

In this Chapter...

- utility definitions and axioms
- risk aversion
- certainty equivalent wealth
- maximization of expected utility
- ordinal and cardinal utility
- von Neumann–Morgenstern utility functions

62.1 INTRODUCTION

Most of derivatives theory is about hedging and the elimination of risk. For a variety of reasons, some of which we have discussed, perfect hedging is not always possible. In that case we must decide how to 'value' the unhedged residual risk. One way is to simply ignore it, and just concern oneself with real expectations of outcomes. Another is to examine both the average outcome and the standard deviation. We've seen both of these approaches in previous chapters.

Another approach is to consider the **utility** of each outcome, that is we associate a number measuring the 'happiness' (in a loose sense) that each outcome gives us. Although \$10 billion is twice as much as \$5 billion, we wouldn't necessarily be twice as happy with the former; we'd be happier, probably, but since both numbers are so staggeringly large and unimaginable, the difference in our happiness would not be that great. We'll see how this fits into quantitative finance theory in later chapters. In this chapter we set the framework.

62.2 RANKING EVENTS

As a general rule people have (at least roughly) a ranking of preferences for various 'commodities.' An E-type Jaguar is preferable to any BMW, for example. Sometimes it is impossible to decide between two commodities, peanut butter cheesecake versus (my own) pumpkin pie. The commodities don't have to be physical quantities, they could be events, or outcomes.

Definition

Let's suppose we have commodities or events A, B, C, ... We write

$$\mathrm{A} \succ \mathrm{B}$$

if A is strictly preferred to B. Note that this is a personal choice.[1] If we can't decide between A and B then we write

$$\mathrm{A} \sim \mathrm{B}.$$

We can also write

$$\mathrm{A} \succeq \mathrm{B}$$

if A is weakly preferred to, or is at least as good as, B.

We must have some axioms, to ensure that the relationship between events makes sense. The usual axioms are as follows.

Axiom 1: Completeness

Given two events A and B, one of the following three must hold.

$$\text{Either} \quad \mathrm{A} \sim \mathrm{B} \quad \text{or} \quad \mathrm{A} \succ \mathrm{B} \quad \text{or} \quad \mathrm{B} \succ \mathrm{A}.$$

Axiom 2: Reflexivity

$$\mathrm{A} \sim \mathrm{A}.$$

Axiom 3: Transitivity

$$\text{If} \quad \mathrm{A} \succeq \mathrm{B} \quad \text{and} \quad \mathrm{B} \succeq \mathrm{C} \quad \text{then} \quad \mathrm{A} \succeq \mathrm{C}.$$

These axioms are sufficient to give us a ranking between various events. But they do not ensure the existence of a utility function. For this we need another definition and a final axiom.

Definition

Let A and B be two events and let $0 \leq \phi \leq 1$. By

$$\{\phi\mathrm{A} + (1 - \phi)\mathrm{B}\}$$

we mean the lottery having the two possible outcomes, A with probability ϕ and B with probability $1 - \phi$.

Axiom 4: Continuity

Let A, B and C be events such that

$$\mathrm{A} \succ \mathrm{C} \succ \mathrm{B}$$

[1] Some people like BMWs.

then there exists some $0 \le \phi \le 1$ such that

$$\{\phi \mathrm{A} + (1-\phi)\mathrm{B}\} \sim \mathrm{C}.$$

62.3 THE UTILITY FUNCTION

A **utility function** represents the 'worth' or 'happiness' associated with events or outcomes or, from now on, levels of wealth. The last is measured in units of currency, dollars, say. Although rarely used in practice they are often found in the financial and economic literature.

I will denote the utility function by $U(W)$ where W is the wealth, measured at some time horizon. Generally speaking, you would expect utility functions to have the following properties. Here $'$ denotes differentiation with respect to W.

- The function $U(W)$ can vary among investors, each will have a different attitude to risk for example.
- $U'(W) \ge 0$: More is preferred to less. If it is a strict inequality then satiation is not possible, the investor will always like more than he has. This slope measures the marginal improvement in utility with changes in wealth.
- Usually $U''(W) < 0$: The utility function is strictly concave. Since this is the rate of change of the marginal 'happiness,' it gets harder and harder to increase happiness as wealth increases. An investor with a concave utility function is said to be **risk averse**. This property is often referred to as the law of diminishing returns.

62.4 RISK AVERSION

The **absolute risk aversion function** is defined as

$$A(W) = -\frac{U''(W)}{U'(W)}.$$

The **relative risk aversion function** is defined as

$$R(W) = -\frac{WU''(W)}{U'(W)} = WA(W).$$

62.5 SPECIAL UTILITY FUNCTIONS

When it comes to choosing particular utility functions there are some popular choices.

Constant Absolute Risk Aversion (CARA)

The choice

$$U(W) = -\frac{1}{\eta}e^{-\eta W} \quad \text{with } \eta > 0$$

is a **Constant Absolute Risk Averse** (**CARA**) utility function. The absolute risk aversion function is just

$$A(W) = \eta, \quad \text{a constant.}$$

Constant Relative Risk Aversion (CRRA)

The choice

$$U(W) = \frac{W^\gamma - 1}{\gamma} \quad \text{with } \gamma < 1 \quad \text{and} \quad \gamma \neq 0$$

is a **Constant Relative Risk Averse** (**CRRA**) utility function. The function $U(W) = \log W$ is also a member of this family; it's the limit of the above as $\gamma \to 0$.

The relative risk aversion function is just

$$R(W) = 1 - \gamma, \quad \text{a constant.}$$

Hyperbolic Absolute Risk Aversion (HARA)

The choice

$$U(W) = \frac{1-\gamma}{\gamma}\left(\beta W/(1-\gamma) + \eta\right)^\gamma$$

is a **Hyperbolic Absolute Risk Averse** (**HARA**) utility function. It is only valid for $\beta W/(1-\gamma) + \eta > 0$ but is otherwise a very broad family.

62.6 CERTAINTY EQUIVALENT WEALTH

When the end of period wealth is uncertain, and all outcomes can be assigned a probability, one can ask what amount of certain wealth has the same utility as the expected utility of the unknown outcomes. In other words, solve

$$U(W_c) = E[U(W)].$$

The quantity of wealth W_c that solves this equation is called the **certainty equivalent wealth**. One is therefore indifferent between the average of the utilities of the random outcomes and the guaranteed amount W_c.[2]

Investor 1 is more risk averse than Investor 2 if for every portfolio

$$W_c|_1 < W_c|_2.$$

In words, the certainty equivalent for 1 is less than the certainty equivalent of 2 for every portfolio.

The concavity of the utility function ensures that

$$E[U(W)] < U(E[W]).$$

It follows that

$$U(W_c) < U(E[W])$$

and since $U'(W) > 0$

$$W_c < E[W];$$

i.e., if possible always accept the average outcome in place of a random one.

[2] $U(\text{A bird in the hand}) = \sum_{i=1}^{2} \text{Prob(Catching bird } i) \; U(\text{Bird } i \in \text{bush})$.

Example

Faced with winning or losing one dollar on the toss of a coin what should you, a risk-averse person, do?

Assuming that you start off without any money at all, the expected utility after the coin toss would be

$$\tfrac{1}{2}U(1) + \tfrac{1}{2}U(-1).$$

This is less than the utility of the expected wealth $U\left(\frac{1}{2} - \frac{1}{2}\right)$ which in this case also corresponds to not taking part in the contest in the first place. If you want to know how much the bet is (certainly) equivalent to, you must solve

$$U(W_c) = \tfrac{1}{2}U(1) + \tfrac{1}{2}U(-1).$$

The answer will be less than the average outcome of zero.

Let's do that calculation again but starting with a wealth of W and with a very small bet of ϵ (so that we can do a Taylor series approximation).

The expected utility after the coin toss is

$$\tfrac{1}{2}U(W+\epsilon) + \tfrac{1}{2}U(W-\epsilon) \sim U(W) + \tfrac{1}{2}\epsilon^2 U''(W) + \cdots.$$

(The two first derivative terms cancel each other.) This is less than the utility of the average which is just $U(W)$.

The certainty equivalent is given by

$$U(W_c) = U(W) + \tfrac{1}{2}\epsilon^2 U''(W) + \cdots.$$

It follows that[3]

$$W_c \sim W + \tfrac{1}{2}\epsilon^2 \frac{U''(W)}{U'(W)} + \cdots$$

Of course, this analysis can easily be extended to arbitrary distribution of outcomes as long as the Taylor series is still valid...

The expected utility after a 'bet/investment' having a probability density function of $p(W)$ is

$$\int p(w)U(w)\,dw \approx \int p(w)\left(U(W) + (w-W)U'(W) + \tfrac{1}{2}(w-W)^2U''(W) + \cdots\right)dw$$

$$= U(W) + \left(\int wp(w)\,dw - W\right)U'(W) + \tfrac{1}{2}U''(W)\int (w-W)^2 p(w)\,dw + \cdots. \quad (62.1)$$

The utility of the average is just

$$U\left(\int wp(w)\,dw\right) = U\left(\int (w-W)p(w)\,dw + W\right) \approx U(W) + \left(\int wp(w)\,dw - W\right)U'(W)$$

$$+ \tfrac{1}{2}U''(W)\left(\int wp(w)\,dw - W\right)^2 + \cdots. \quad (62.2)$$

(Under what conditions are these expansions valid?)

[3] So that's where the absolute risk aversion function comes from.

Comparing (62.1) and (62.2) the first two terms of each are the same, while the third term makes the former expression smaller. And the certainty equivalent?

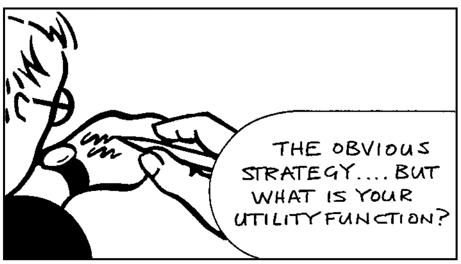
The certainty equivalent is given by

$$U(W_c) = U(W) + \left(\int wp(w)\,dw - W\right) U'(W) + \tfrac{1}{2}U''(W) \int (w - W)^2 p(w)\,dw + \cdots.$$

It follows that

$$W_c \approx \int wp(w)\,dw + \frac{1}{2}\frac{U''(W)}{U'(W)} \int (w - W)^2 p(w)\,dw + \cdots.$$

The first term in this is just the average future wealth and the second term is the correction due to the risk aversion of the investor.



62.7 MAXIMIZATION OF EXPECTED UTILITY

One of the main uses of utility theory is in choosing the optimal investment from a choice of investments. Typically these investments will be risky, having unknown outcomes at the end of the time horizon.

Example

If there are N assets, each having a random return R_i, then one way to optimize our portfolio is to choose the weights w_i of each asset such that the expected utility

$$E\left[U\left(W_0 \sum_{i=1}^{N} w_i R_i\right)\right]$$

is maximized. Here W_0 is the wealth initially invested. This must be maximized subject to the budget constraint

$$\sum_{i=1}^{N} w_i = 1.$$

If we add an $N + 1$th risk-free asset earning interest of r over the horizon then the optimization problem becomes

$$\max_{w_1,\ldots,w_n} E\left[U\left(W_0\left(r + \sum_{i=1}^{N} w_i(R_i - r)\right)\right)\right].$$

This incorporates the budget constraint.

62.7.1 Ordinal and Cardinal Utility

If we are only concerned with the ranking of events then we could equally use as our utility function $\Phi(U(W))$ where Φ is any strictly monotonically increasing function. Sometimes, the actual value of the utility function is important.

Two investors having different utility functions may rank events in the same order (this is **ordinal utility**) but when faced with uncertain outcomes/investments they may decide differently from each other (they have different **cardinal utility**).

We've just seen how to maximize expected utility to solve problems such as finding an optimal portfolio. If we are interested in finding this maximum then the new function

$$aU(W) + b \quad \text{with } a > 0$$

will do just as well since it has a maximum at the same value of W as the original $U(W)$. For an individual, the utility function is arbitrary up to a positive affine transformation (i.e. multiply by a positive constant and translate by a constant). For this reason specifying either the function $A(W)$ or $R(W)$ is sufficient information to rank an investor's choice of investments in terms of their cardinal utility.

Example

Let's see how the ordinal/cardinal distinction works in practice. An investor has the utility function $U(W) = W$. (We should have $U'' < 0$ but I want to keep things as simple as possible.) He has a choice between entering a lottery with outcomes of 0 or 9, with equal probabilities, or receiving 4. The expected utility of the first choice is

$$\tfrac{1}{2} \times 0 + \tfrac{1}{2} \times 9 = 4.5.$$

This beats the utility of the second choice which is 4; he takes the gamble.

Now consider a different utility function, which is in a sense just a monotonic transformation of the first function, such as $U(W) = W^{1/2}$. The outcomes 0, 4, 9 are still ranked in that order, but what happens to the utility when considering expectations? The first choice has expected utility

$$\tfrac{1}{2} \times 0 + \tfrac{1}{2} \times 3 = 1.5.$$

This is now not as good as the second choice which has utility of 2. He takes the certain outcome. The cardinal utility function or **von Neumann–Morgenstern utility function** attaches special meaning to the numerical value of the utility, and is more than just a mechanism for ranking certain outcomes.

We will be seeing more of expected utility maximization in later chapters.

62.8 SUMMARY

Utility theory is a useful framework for valuing risk. It's not popular outside of academic, and in particular, economic, circles but it is a subject than I am warming to. It gets a lot more complicated, but what I've shown you here will serve our purposes.

FURTHER READING

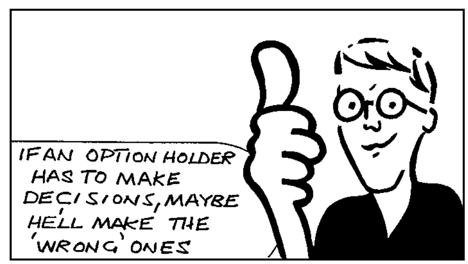
- For a more general discussion of utility theory see Ingersoll (1987) (from whom some of the above examples have been borrowed) and Owen (1995).

CHAPTER 63

more about American options and related matters

In this Chapter...

- why some people might exercise American options at non-optimal times
- the effect of non-optimal exercise on the profit of the option writer



63.1 INTRODUCTION

In 1998 Dr Hyungsok Ahn, then my postdoctoral researcher at Oxford University, and I wrote a paper on the exercise of American options.

In this chapter I reproduce the paper with just a few minor changes. However, to start with, here's what the nice *Derivatives Week* published by us on the subject of early exercise.

63.2 WHAT DERIVATIVES WEEK PUBLISHED © 1999 Institutional Investor Inc. (article reproduced with kind permission)

8 DERIVATIVES WEEK
JANUARY 11, 1999
www.iinews.com

LEARNING CURVE® EXERCISING AMERICAN OPTIONS

INTRODUCTION

In this article we explain the ideas behind the valuation of options with early exercise features, so-called American options. We also aim to clarify some popular misconceptions about when an American option should be exercised.

HOW TO PRICE AMERICAN OPTIONS

If an option can only be exercised at expiry it is called European. If it can be exercised at any time prior to expiry, it is called American. Bermudan options have prespecified exercise dates which may be particular days or whole periods of time. Because they give the holder more rights, the American option is at least as valuable as an equivalent Bermudan option which in turn is at least as valuable as an equivalent European contract.

The idea behind valuing options with early exercise is to decide when the option should be exercised. Is there, in some sense, a best or optimal time for exercise? To correctly price American options we must place ourselves in the shoes of the option-writer. We must be clear about the principles behind his strategy. From the modeling point of view we assume that the writer of the option also is hedging his option position by trading in the underlying asset. The hedging strategy is dynamic and referred to as delta hedging. The position in the underlying asset is maintained delta neutral so as to be insensitive to movement of the asset. By maintaining such a hedge, the writer does not care about the direction in which the underlying moves, he eliminates all asset price risk. However, he does remain exposed to the exercise strategy of the option holder. If the writer makes an assumption about when the holder will exercise his option and this assumption turns out to be incorrect, this will have an impact on the writer's profit. Since the writer cannot possibly know what the holder's strategy will be, how can the writer reduce his exposure to this strategy?

The answer is simple. The writer assumes that the holder exercises at the worst possible time for the writer. He assumes that the option is exercised at the moment that gives the writer the least profit. This is often referred to as the optimal stopping time, although as far as the writer is concerned it is the last thing he wants to happen. So, out of all the possible exercise strategies we must find the one that gives the option the least value to the writer or the highest value to the holder. This sounds very complicated but anyone who has implemented the binomial method knows that it is just a matter of adding one line of code to the program. That line of code simply tests at each node in the tree whether the theoretical option value is greater than the payoff, if it is not then the payoff is used instead, and this corresponds to a time at which the option should be exercised.

To summarize this section, the assumptions are that the option writer is delta hedging and prices the option at the highest possible value over all exercise strategies.

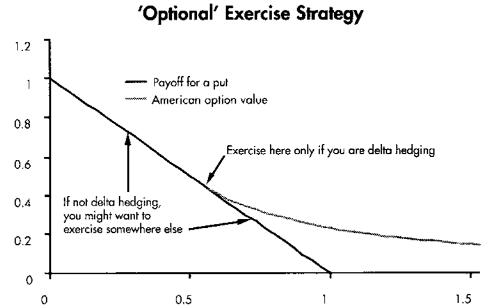
WHEN SHOULD THE HOLDER EXERCISE?

The holder of the option rarely delta hedges. Perhaps he has bought the option as a static hedge for the rest of his portfolio, or perhaps as a speculative investment. Either way it is unlikely that he is insensitive to the direction of the underlying asset. The initial assumption concerning the writer of the option does not apply to the holder. Should the holder therefore act in the optimal way that follows from the two assumptions summarized at the end of the previous section?

Consider the simplest scenario. You buy a call because you believe the underlying asset is going to rise significantly. If you are correct you will make a substantial return. If there are no dividends on the underlying then it is 'theoretically' never optimal to exercise before expiry. We put the word theoretically between inverted commas because the 'theory' is only relevant to someone who is delta hedging. Now suppose that the stock does indeed rise, but the economic situation makes you believe that a sudden fall is imminent. What should you do? The obvious solution is to sell the option and lock in your profit. But this may not be possible, for example if the option is over-the-counter. The only way of locking in the profit may be to exercise the option early. The theory says don't exercise, but if the stock does fall then you lose the profit. At this stage it is important to remember that the theory is not relevant to you.

The writer and the holder of the option have different priorities, what is optimal to one is not necessarily optimal

'Optional' Exercise Strategy



Derivatives Week is now accepting submissions from industry professionals for Learning Curve®, the tutorial for new or potential users of derivatives. For details and guidelines on writing a Learning Curve®, please call Sophie Belcher in London at 44-171-303-1733, Paula Froelich in New York at (212) 224-3253 or Sarah Yavorsky in Hong Kong on 852-2912-8097.

DERIVATIVES WEEK 9
JANUARY 11, 1999
www.iinews.com

to the other. The holder of the option simply may have a gut feeling about the stock and decide to exercise. That is perfectly valid. Or he may have a stop-loss strategy in place. He may even have a complex utility maximization strategy. Anyway, it is highly unlikely his exercise time will correspond to that calculated by the writer of the option.

The figure shows the payoff and theoretical value for an American put option some time before expiry. If you are delta hedging there is an optimal asset value at which to exercise.

HOW DOES THE WRITER FEEL ABOUT EXERCISE?

The writer has received a sum of money in exchange for the option. That sum of money was calculated assuming that the option holder exercises at a certain optimal time. This optimal exercise strategy gives the option its highest theoretical value. The writer receives this maximum amount even though the holder may exercise at any time.

It is clear that the writer can never lose. The worst that can happen to him is that the option is exercised at this theoretical optimal time. But this has already been priced into the premium he received. On the other hand, if the holder exercises at some other time he can only benefit. How does the writer feel about exercise? At worst he has no feelings, at best he is very happy!

CONCLUSION

Since there is a clever but complex theory behind the exercise of American options there has arisen some misunderstanding about the optimal time to exercise. It is commonly believed that there is only one optimal time. This is far from being true. Finally, we would like to make the observation that if you want to sell someone an option and early exercise doesn't add too much to the theoretical value then you should always make it American. This gives you the possibility of a surprise windfall profit if the holder exercises at an unexpected time.

*This week's Learning Curve was written by **Hyungsok Ahn**, a visiting academic at **Oxford University** and a Director of **Woniing Financial Consultants** based in Korea, and **Paul Wilmott**, a Royal Society University Research Fellow at **Oxford University** and **Imperial College**, London.*

63.3 HOLD THESE THOUGHTS

I don't want you to get the wrong idea while reading this so I want you to keep the following in mind.

- The writer is in the Black–Scholes world, including no transaction costs.
- The option holder can't delta hedge (why would he buy the option if he could?).
- The option cannot be sold, think of it as OTC.

This means that the holder cannot close his position other than by exercising.

63.4 CHANGE OF NOTATION

Since this is the original paper I have kept the original notation. The main differences are

- For S_t and s read S
- For $v(t, s)$ read $V(S, t)$
- For v_t read $\dfrac{\partial V}{\partial t}$
- For v_s read $\dfrac{\partial V}{\partial S}$
- For v_{ss} read $\dfrac{\partial^2 V}{\partial S^2}$
- For K read E

63.5 AND FINALLY, THE PAPER ...

On Trading American Options

Hyungsok Ahn and **Paul Wilmott**
OCIAM, Oxford University

Abstract. In this paper we consider the effect that early exercise has on the profit of the writer of an American option. The American option is correctly priced in the Black–Scholes framework assuming that the holder exercises at the worst possible time for the writer. This early exercise time is unique. But why should the holder exercise at this time? If the holder of the American option wants to both eliminate all risk by delta hedging and maximize his wealth then certainly he should exercise here. But if he is following this strategy why does he bother to buy the option? Complete markets make purchasing the option unnecessary. So let's assume that he is following some other strategy, as he is free to do. Now, clearly he would always be better off selling the option than exercising it, but what if the contract is OTC, and he can only close his position by exercise? If the option holder follows some strategy that results in early exercise at a time not given by the classical optimal free boundary then the writer makes more profit than might be expected. As examples, we assume that the holder exercises according to the maximization of his own utility function. We illustrate our results by applying them to several families of utility functions, namely the CARA, the HARA, and the expected return. While the option holder maximizes his utility, the issuer gains from the difference between the price-maximizing exercise boundary and the exercise boundary of the option holder. We provide numerical results that describe the effect of the physical drift and the risk aversion on the issuer's expected profit.

63.6 INTRODUCTION

The American option is an option contract that allows the option holder to exercise before maturity if he is better off doing so. Because of the flexibility of choosing the exercise time, the price of the option is calculated as the value of the option in the worst case for the issuer among all feasible exercise strategies that the option holder may perform. Typically the price-maximizing exercise time, and hence the least favorable exercise time for the issuer, is described as an optimal stopping time, and the resulting pricing equation becomes a free-boundary partial differential equation (PDE). Although it may appear that it requires rather sophisticated mathematics to price an American option, the fundamental concept of the absence of arbitrage is still an integral part of determining the price. The issuer can construct a hedging portfolio involving the trading of the underlying assets in such a way that the value of the replicating portfolio (i.e., the upfront premium for the option plus the result of the trading) is not less than his liability even when his customer exercises the option at the least favorable time. Throughout this paper, we will assume that the American option is sold at this classical price and that the market is complete.

If he so desired, the option holder could delta hedge his position as well, constructing his portfolio exactly opposite to that of the issuer's and exercising his option at the price-maximizing exercise time. In this case, the balances of both the issuer and the option holder would be equal to zero. If this was what they intended, they could have chosen not to trade the option at the beginning and saved their effort in maintaining their hedge positions. In fact, option traders may choose not to construct a risk-free portfolio, especially when they use options as a means of investment. This has nothing to do with market completeness: Those who visit Las Vegas play slot machines even when they can sustain a risk-free wealth by not playing. Thus it is reasonable to assume that the option holder engages in some strategies. For example, he may adopt a stop-loss strategy: Buy-and-hold the option until he decides to exercise it. Unlike the previous case, the option holder may gain or lose depending upon the behavior of the underlying asset price while the potential loss is not more than the premium he paid. The issuer gains unless the option holder exercises at the first time that the asset price reaches the price-maximizing exercise boundary. One of the questions we address in this paper is ‘Should the option *holder* exercise at the price-maximizing exercise time?’ Clearly, the option holder would always be better off selling the option than exercising it, but if the contract is OTC this possibility may not exist. We could, for example, assume that exercise is the only way available for him to close his position.

Trading an option is not a two-person zero-sum game because both the issuer and the holder can trade the underlying asset with other investors. For example, if the issuer maintains a risk-free portfolio by trading the underlying asset while the holder leaves the option unhedged, the issuer’s balance will eventually become zero while the holder’s depends upon where the underlying asset price lands at the exercise time. Hence, the worst case for the issuer is not necessarily the best case for the holder. It is not always true that the option holder is better off exercising the option at the price-maximizing exercise time. First, we consider the physical drift of the underlying asset. The price of an option depends upon the risk-neutral drift, not the physical drift of the underlying asset. The reason is that the presence of an option immediately allows one to construct locally riskless portfolios, and hence the risk-free rate is the only one that governs its price. As a result, the price-maximizing exercise boundary is also independent of the physical drift. Can we assert that the optimal exercise boundary for the option holder is not affected by the market direction? Almost certainly not. For example, it is well known that the price-maximizing exercise time for the American call is the maturity of the contract, provided that the underlying stock pays no dividend. If there is any evidence that the price of the underlying asset is expected to fall, however, a wise investor would exercise his call earlier before it expires worthless. Second, each investor has his own risk preference. Two different rational decision makers may exercise differently, even when they agree on the probability distribution of outcomes. It is nonsense to argue that a single exercise strategy is the optimal strategy for every investor. As an aside, there is evidence (Overdahl and Martin (1994), for example) that a substantial proportion of all exercised American call options are exercised before their classical optimal stopping time. This is probably due to friction factors, errors or differences of opinion about parameter values. Our paper also applies in these situations.

This paper establishes the optimal exercise boundary provided that the option holder is a utility-maximizing investor. The optimal exercise boundary, or the utility-maximizing exercise boundary, depends upon the risk aversion and the physical drift. we confirm the following: (i) If the option holder is sufficiently risk averse, early exercise is optimal even for a call; (ii) The optimal exercise time is a non-decreasing function of the physical drift, if the option is a call; (iii) If the option is a put, the optimal exercise time is a non-increasing function of the physical drift. We illustrate these results with several families of utility functions: The constant absolute risk averse (CARA), the hyperbolic absolute risk averse (HARA), and the linear utility (i.e. the expected return). Some of the highlights are:

(a) If the option holder's utility is of the CARA type, early exercise prevails for both call and put regardless of the absolute risk aversion parameter.

(b) Certain HARA utilities may yield two separate exercise boundaries.

(c) Upon the expected return criterion, a call option is exercised early only when the physical drift is surpassed by the risk-free rate.

Another result of this paper is the equation for the expected profit selling American options. As we stated earlier, the issuer gains from the difference between the price-maximizing exercise time and the exercise time chosen by his customer. The profit grows as the occupation time of the asset price in the region between the exercise boundaries of the price maximization and the utility maximization. The difference between the value of the option and the exercise value is the final piece of the profit. We provide numerical results on how the physical drift and the risk aversion affect the issuer's profit.

The paper is structured as follows. In the next section we review the classical results of pricing and hedging American options. In Section 63.8, we find the optimal exercise time for the utility maximizing investor. In Section 63.9, we analyze the effect of the option holder's optimal exercise strategy on the issuer's profit. Section 63.10 contains concluding remarks.

63.7 PRELIMINARY: PRICING AND HEDGING

The early-exercise feature makes the valuation of the American option more intriguing than that of the European counterpart. The main concepts are the optimal stopping and the corresponding parabolic variational inequalities. Myneni (1992) surveyed the literature on the subject and summarized key results. Here we state the assumptions for the rest of the paper and review the variational inequalities.

The classical theory of option pricing is predicated on many assumptions for market completeness. We assume that the market is frictionless, that short-selling is allowed without restriction, that one can trade assets as frequently as one wishes, that all risk-free assets grow at the common rate r which is known *a priori*, and that there is a unique risk-neutral equivalent martingale measure. The last assumption becomes less abstract when we assume that the price of the underlying asset follows a geometric Brownian motion and that market participants are not capable of foreseeing the future. Thus, in what follows the price of the underlying asset evolves as

$$dS_t = \sigma S_t \, dW_t + \mu S_t \, dt \tag{63.1}$$

where W is a standard Brownian motion. In addition the filtration is natural, meaning that the stream of information consists of the observations of the asset price only.

As shown in Harrison and Pliska (1981), the complete market assumption allows a trader to replicate the payoff of an arbitrary contingent claim by trading the underlying assets. We start by assuming that the issuer of the option maintains Δ shares of the underlying asset to hedge his position. In other words, the value of the issuer's portfolio is given by $\Delta S - v$ where v is the value of the option. This portfolio must grow at least at the risk-free rate r:

$$\Delta \, dS_t - dv \ \geq \ r(\Delta S_t - v) \, dt. \tag{63.2}$$

Because of the Markovian nature of the underlying asset price (63.1), the value of the option v at time t is a function of t and the asset price S_t. For the time being, we assume that v is continuously differentiable with respect to t and twice continuously differentiable with respect to s. Then we have

$$dv(t, S_t) = v_t(t, S_t)\,dt + v_s(t, S_t)\,dS_t + \tfrac{1}{2}\sigma^2 S_t^2 v_{ss}(t, S_t)\,dt \tag{63.3}$$

which follows from Itô's formula. It is required for the issuer to pick $\Delta = v_s$ in order to fulfill (63.2) because the random growth dS_t is of the order $\sqrt{dt}$ and is much bigger than the dt terms. Rearranging (63.2) after replacing Δ by v_s yields

$$\mathcal{L}v = v_t + \tfrac{1}{2}\sigma^2 s^2 v_{ss} + r(sv_s - v) \le 0 \tag{63.4}$$

for each s. The value of the option will never fall below an immediate exercise value. Otherwise the issuer loses. This yields the second condition for v:

$$v \ge \phi \tag{63.5}$$

where ϕ is the payoff of the option: $\phi(s) = \max(s - K, 0)$ for a call with strike K and $\phi(s) = \max(K - s, 0)$ for a put. At each time t, the option holder may or may not exercise his option. If $v > \phi$ at this moment, then exercising the option is not the least favorable outcome for the issuer because he can claim a non-zero profit $v - \phi$ instantly. In this case, $\mathcal{L}v = 0$ because the issuer has an arbitrage opportunity if $\mathcal{L}v$ were strictly less than zero. Therefore we obtain the third condition:

$$(\mathcal{L}v) \cdot (v - \phi) = 0. \tag{63.6}$$

The inequalities (63.4), (63.5), and (63.6) subject to $v(T, s) = \phi(s)$ form a parabolic obstacle problem. We refer to Friedman (1988) for the existence and the uniqueness of the solution to such problems. Jaillet, Lamberton, and Lapeyre (1990) showed that the solution of the parabolic variational inequalities (63.4), (63.5), and (63.6) has a continuous gradient at the free boundary (i.e. a smooth fit), and Van Moerbeke (1976) showed that the optimal stopping boundary is continuously differentiable. Thus Itô's formula (63.3) is valid at least in a weak sense: See San Martin and Protter (1993) for details.

In the theory of optimal stopping, the space-time domain defined by $v > \phi$ is called the continuation region as the stopping is premature in this region and the graph of its boundary is called the optimal stopping boundary. In this paper we call this the price-maximizing exercise boundary, distinguishing it from the optimal stopping boundary from the utility maximization problem in the next section. We stress that the contents of this paper is in no way conflicting with the Black–Scholes–Merton theory of option pricing and hedging, that we are in a complete market and that there is thus only one fair value for an American option and this is given by the classical approach.

63.8 UTILITY-MAXIMIZING EXERCISE TIME

We assume that the option holder possesses a utility function $U : \Re \to \Re$ that is strictly increasing and twice continuously differentiable. The investor, who purchases an American option at time 0, will select his exercise time by maximizing the expected utility of his discounted wealth. The class of feasible exercise times consists of stopping times that are less than or equal to

T, the maturity of the option. This includes exercise times that are strategically selected based upon the price of the asset up to that date as well as pre-scheduled times (i.e., non-random). A feasible exercise time will be denoted by τ. If the option holder never exercises the option we set $\tau = T$. As before ϕ is designated for the payoff. Then, at time t, the option holder faces the following optimal stopping problem:

$$u(t,s) = \sup_{t \le \tau \le T} E^{t,s}\Big[U(e^{-r\tau}\phi(S_\tau))\Big] \tag{63.7}$$

where $E^{t,s}$ is the conditional expectation given that $S_t = s$, τ is the option holder's exercise time, and ϕ is the payoff. The essential supremum is taken over all the feasible exercise times. Finally the expectation is governed by the physical measure not the risk-neutral equivalent martingale measure. We consider only situations when (63.7) is well defined. A sufficient condition is that $U \circ \phi$ is bounded by a polynomial.

We could have defined u as the expected utility of $e^{-r\tau}\phi(S_\tau) - v(0, S_0)$, the discounted payoff minus the option price. In our definition, the option price is a part of the utility function U, as we treat the option price as a constant.

As in the case of the price maximization, the optimal stopping problem (63.7) is equivalent to a parabolic obstacle problem. Thus u satisfies a set of variational inequalities. We will describe the variational inequalities financially, omitting technical details. For notational convenience, we define $g(t,s) = U(e^{-rt}\phi(s))$. First we check that

$$u \ge g. \tag{63.8}$$

This is because the maximum expected utility is not smaller than the utility of immediate exercise which is a special case of feasible stopping times. Next we will explain the following inequality for $t < T$:

$$u(t,s) \ge E\Big[u(t+\delta, S_{t+\delta})\Big] \tag{63.9}$$

for each δ that makes $t + \delta$ a feasible exercise time. Note that the right side of (63.9) coincides with the expected utility when the option holder pursues the optimal stopping only after δ elapses. In other words, the option holder is dormant until time $t + \delta$ and he tries to find an optimal exercise time from then on. Thus the value of this expected utility cannot exceed the maximum expected utility which is on the left side of (63.9). The implication of (63.9) is the following inequality

$$\mathcal{L}_\mu u = u_t + \mu s u_s + \tfrac{1}{2}\sigma^2 s^2 u_{ss} \le 0 \tag{63.10}$$

which is obtained by applying Itô's formula to u. If $\mathcal{L}_\mu u < 0$, then the maximum expected utility is expected to fall in an infinitesimal time, and hence the optimal strategy is to exercise the option immediately. That is, $u = g$. Therefore u must satisfy

$$(\mathcal{L}_\mu u) \cdot (u - g) = 0. \tag{63.11}$$

The set of variational inequalities (63.8), (63.10), and (63.11) with terminal data $g(T,s)$ characterizes the maximum expected utility u. The optimal exercise time is the first time that the asset price S_t hits the free boundary of the inequalities.

Next we consider $h = e^{rt}U^{-1} \circ u$, the maximum expected certainty equivalence. U^{-1}, the inverse of U, is well defined as U is an increasing function of wealth. The merit of using this change of variable is that it facilitates us comparing the utility maximization to the price maximization. We confirm that h must satisfy the following variational inequalities:

$$h \geq \phi$$
$$\mathcal{D}h = h_t + \tfrac{1}{2}\sigma^2 s^2 \Big(h_{ss} + \frac{U''}{U'}(e^{-rt}h)e^{-rt}(h_s)^2 \Big) + \mu s h_s - rh \leq 0 \tag{63.12}$$
$$(\mathcal{D}h) \cdot (h - \phi) = 0$$

subject to $h(T, s) = \phi(s)$. Therefore the utility-maximizing exercise boundary depends upon the physical drift and Pratt's measure of absolute risk aversion $-U''/U'$, and is different from the price-maximizing exercise boundary. The distortion in discount is caused by the nonlinearity of the utility function.

The utility-maximizing exercise time for an American option has the following properties:

(i) If the absolute risk aversion is sufficiently large, then there is a positive probability of early exercise for both call and put.

(ii) The exercise time is non decreasing in μ, when the option is a call.

(iii) The exercise time is non increasing in μ, when the option is a put.

Proof. Note that the exercise region coincides with the space-time domain of $\mathcal{D}h < 0$. If the absolute risk aversion $-U''/U'$ tends to infinity uniformly in its argument, then $\{(t, s) : \mathcal{D}h < 0,\ 0 \leq t < T,\ s > 0\}$ is a set of a positive measure. Since the support of a non-degenerate geometric Brownian motion (i.e. $\sigma^2 > 0$) occupies the entire positive plane, the utility-maximizing exercise time can be less than the maturity with a positive probability. This proves (i). When the option is a call, h_s is positive. Thus $\mathcal{D}h$ becomes more negative when μ becomes smaller. If the option is a put, h_s is negative, and hence $\mathcal{D}h$ becomes more negative when μ becomes larger. Therefore we have (ii) and (iii).

Our next task is to locate the boundary when the time to maturity is arbitrarily close to zero. Note that the certainty equivalence h tends to ϕ as $t \to T$ and the utility-maximizing exercise boundary (as a function of time) is continuously differentiable. Thus when t is near T, the utility-maximizing exercise boundary is close to the boundary of $\mathcal{D}\phi < 0$. This is *the boundary at maturity*. If $\phi(s) = \max(s - K, 0)$ (i.e., a call option), then the boundary is above the strike K for each $t \in [0, T)$ and hence the boundary at maturity is

$$\partial \Big[s > K : \tfrac{1}{2}\sigma^2 s^2 \frac{U''}{U'} \Big(e^{-rT}(s - K) \Big) e^{-rT} + (\mu - r)s + rK < 0 \Big]. \tag{63.13}$$

Here, the symbol ∂ is used for indicating the boundary of a set. Similarly, if $\phi(s) = \max(K - s, 0)$ (i.e. a put option), the boundary at maturity is

$$\partial \Big[s < K : \tfrac{1}{2}\sigma^2 s^2 \frac{U''}{U'} \Big(e^{-rT}(K - s) \Big) e^{-rT} - (\mu - r)s - rK < 0 \Big]. \tag{63.14}$$

Sometimes (63.13) and (63.14) may contain more than one element. In such a case, we have more than one free boundary. In the remaining of this section, we provide an explicit expression for the boundary at maturity when the option holder's utility belongs to one of the following categories: The CARA, the HARA, and the expected return (i.e. the linear utility).

63.8.1 Constant Absolute Risk Aversion

This is the case when the absolute risk aversion is a constant regardless of the wealth of the investor. That is, $-U''/U' \equiv \lambda$ for a positive constant λ. Up to a constant, the utility is of the form $U(\omega) = -\alpha e^{-\lambda\omega}$ for a positive constant α.

First we consider a call option. We confirm that the boundary at maturity (63.13) reduces to

$$\max\left(K, \frac{1}{\lambda\sigma^2}\left(\mu - r + \sqrt{(\mu-r)^2 + 2\lambda\sigma^2 K r e^{-rT}}\right)e^{rT}\right). \tag{63.15}$$

Note that (63.15) tends to infinity as λ tends to zero. Hence as the risk aversion of the option holder vanishes, the utility-maximizing exercise time tends to the maturity which coincides with the price-maximizing exercise time. Next we consider a put option. The inequality in (63.14) is

$$-\tfrac{1}{2}\sigma^2\lambda e^{-rT}s^2 - (\mu - r)s - rK < 0. \tag{63.16}$$

If the physical drift is at least the risk-free rate ($\mu \geq r$), (63.16) is true for all positive s. Thus the boundary at maturity is K. Suppose that $\mu < r$. The quadratic inequality (63.16) is always satisfied if

$$d = (m - r)^2 - 2\lambda\sigma^2 K r e^{-rT} < 0.$$

In this case the boundary at maturity is also K. Now suppose that $d \geq 0$ as well as $\mu < r$. Solving the quadratic inequality (63.16), we obtain the boundary at maturity:

$$\min\left(K, \frac{1}{\lambda\sigma^2}\left(r - \mu + \sqrt{(r-\mu)^2 - 2\lambda\sigma^2 K r e^{-rT}}\right)e^{rT}\right).$$

63.8.2 Hyperbolic Absolute Risk Aversion

Merton (1990) provides a complete description of this family of utility functions. The hyperbolic absolute risk aversion means $-U''/U'(\omega) = \lambda/(\omega + \alpha)$ for a positive constant λ. This utility applies in the case when the wealth of the investor is bounded below $\omega + \alpha > 0$. Thus the richer the investor is, the less he is risk averse. Up to a constant shift,

$$U(\omega) = \begin{cases} \dfrac{1}{\beta^\lambda}\dfrac{(\omega+\alpha)^{1-\lambda}}{1-\lambda}, & \text{if } \lambda \neq 1 \\[2ex] \dfrac{1}{\beta}\log(\omega+\alpha), & \text{otherwise} \end{cases}$$

where $\beta > 0$. The parameter α is assumed positive as the option payoff could be zero.

Simple algebra reduces the inequalities in (63.13) and (63.14) to quadratic inequalities. For example, (63.13) is equivalent to

$$\partial\left[s > K : As^2 + Bs + C < 0\right] \tag{63.17}$$

where $A = (\mu - r - \frac{1}{2}\sigma^2\lambda)e^{-rT}$, $B = (\mu - r)(\alpha - e^{-rT}K) + re^{-rT}K$, and $C = rK$ $(\alpha - e^{-rT}K)$. The continuation region and the exercise boundary depend upon the choice of parameters. An unusual case is when the parameters satisfy the following:

$$r + \tfrac{1}{2}\sigma^2\lambda < \mu < \tfrac{1}{2}\sigma^2\lambda\frac{e^{-rT}K}{\alpha}.$$

In this case, the continuation region near the maturity is separated by the exercise region:

$$\left[s > K : As^2 + Bs + C < 0\right] = \left[s : K < s < \frac{-B + \sqrt{B^2 - 4AC}}{2A}\right].$$

If the physical drift is sufficiently large, the option is very valuable to the holder when the option is very in-the-money. If not, the curvature reduces the holder's utility. Also note that there is no exercise boundary if

$$\mu > r + \tfrac{1}{2}\sigma^2\lambda \qquad \text{and} \qquad \alpha > e^{-rT}K.$$

This is the case when the physical drift is large while the risk aversion is not.

63.8.3 The Expected Return

This is a special case of $U(\omega) = \alpha\omega + \beta$ for a positive constant α. As U'' vanishes in this case, our analysis on the boundary at maturity becomes straightforward.

When the option is a call, the inequality in (63.13) becomes $(\mu - r)s + rK < 0$. This is never satisfied if $\mu \geq r$. Thus the utility-maximizing exercise time is the maturity when the physical drift is at least the risk-free rate. If $\mu < r$, on the other hand, the boundary at maturity is

$$\max\left(K, \frac{r}{r-\mu}K\right).$$

Next we consider a put option. If $\mu \geq r$, then the inequality in (63.14) is always satisfied. Thus the boundary at maturity is K in this case. If $\mu < r$, the boundary at maturity becomes

$$\min\left(K, \frac{r}{r-\mu}K\right).$$

63.9 PROFIT FROM SELLING AMERICAN OPTIONS

In the previous section, we observed that the option holder's exercise time could differ from the price-maximizing exercise time, when he optimizes his utility. When this happens, the issuer gains from the difference. In this section we examine the profit from selling American options to utility-maximizing investors.

The issuer charges $v(0, S_0)$ at time 0 as he sells an American option. He will hedge his short position as described in Section 63.7 until his customer exercises the option or the option expires. The discounted potential liability of the issuer is $e^{-r\tau}\phi(S_\tau)$ where τ is the actual time that his customer exercises. When the option holder never exercises, $\tau = T$ by convention. Thus the present value of the issuer's profit becomes:

$$P = v(0, S_0) + \int_0^\tau e^{-rt}\Delta\Big(dS_t - rS_t\,dt\Big) - e^{-r\tau}\phi(S_\tau) \tag{63.18}$$

The second term in the right side of (63.18) is the result of delta hedging with the cost of carry. First we add and subtract $e^{-r\tau}v(\tau, S_\tau)$ from the profit P. Applying Itô's formula to v yields:

$$v(0, S_0) + \int_0^\tau e^{-rt}\Delta\Big(dS_t - rS_t\,dt\Big) - e^{-r\tau}v(\tau, S_\tau) = -\int_0^\tau dt\; e^{-rt}\mathcal{L}v$$

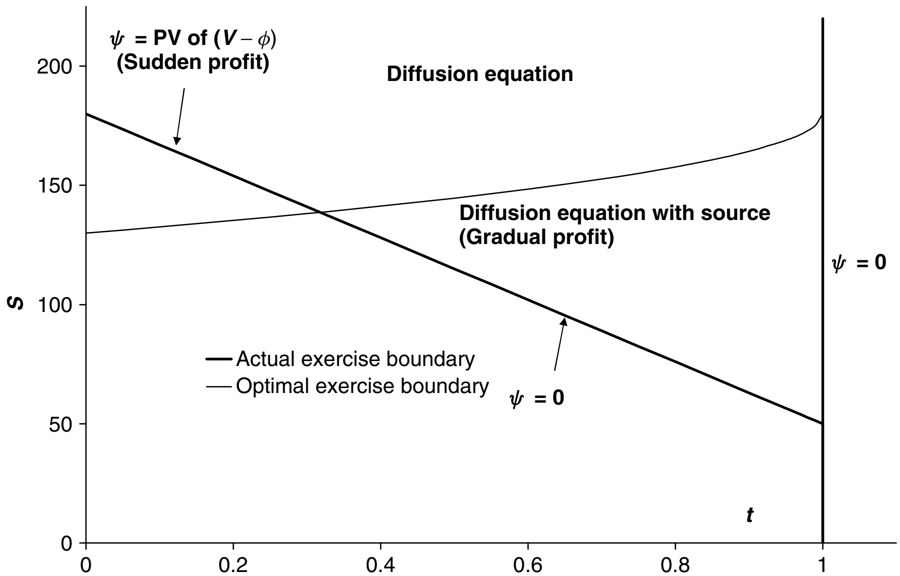


Figure 63.1 Overlapping exercise boundaries

where $\mathcal{L}$ is the Black–Scholes differential operator defined in (63.4). Thus we may rewrite the profit (63.18) as

$$P = -\int_0^{\tau} dt\ e^{-rt}\mathcal{L}v + e^{-r\tau}\Big(v(\tau, S_\tau) - \phi(S_\tau)\Big). \tag{63.19}$$

We define the expected profit at time t as

$$\psi(t, s) = E^{t,s}\Big[-\int_t^{\tau} dt\ e^{-rt}\mathcal{L}v + e^{-r\tau}\Big(v(\tau, S_\tau) - \phi(S_\tau)\Big)\Big] \tag{63.20}$$

We will show that ψ satisfies a diffusion equation with a moving boundary which is known *a priori*. Recall that h is the maximum expected certainty equivalence of the option holder and its free boundary gives the optimal exercise time τ. Let $\mathcal{H}$ and $\mathcal{V}$ be the domains defined by $h > \phi$ and $v > \phi$, respectively. These are the regions of continuation for the utility maximization and the price maximization. We also define $\mathcal{G} = \mathcal{H}\backslash\mathcal{V}$, see Figure 63.1. Since $\mathcal{L}v$ vanishes on $\mathcal{V}$, the expected profit ψ satisfies

$$\psi_t + \mu s\psi_s + \tfrac{1}{2}\sigma^2 s^2\psi_{ss} - e^{-rt}\mathcal{L}v\ \mathcal{I}_{\mathcal{G}} = 0 \tag{63.21}$$

subject to $\psi(T, s) = 0$ and $\psi = e^{-rt}(v - \phi)$ on $\partial\mathcal{H}$, the utility-maximizing exercise boundary. The indicator $\mathcal{I}_{\mathcal{G}}$ is one if (t, s) belongs to $\mathcal{G}$ and zero otherwise. If the option is a call, the left side of (63.21) vanishes because $\mathcal{G}$ is empty. If the option is a put, then $v = \phi$ on the complement of $\mathcal{V}$, and therefore

$$e^{-rt}\mathcal{L}v\ \mathcal{I}_{\mathcal{G}} = e^{-rt}\mathcal{L}\phi\ \mathcal{I}_{\mathcal{G}} = -re^{-rt}K\mathcal{I}_{\mathcal{G}}$$

where K is the strike price. Here we have used the fact that the price-maximizing exercise boundary is not above the strike when the option is a put.

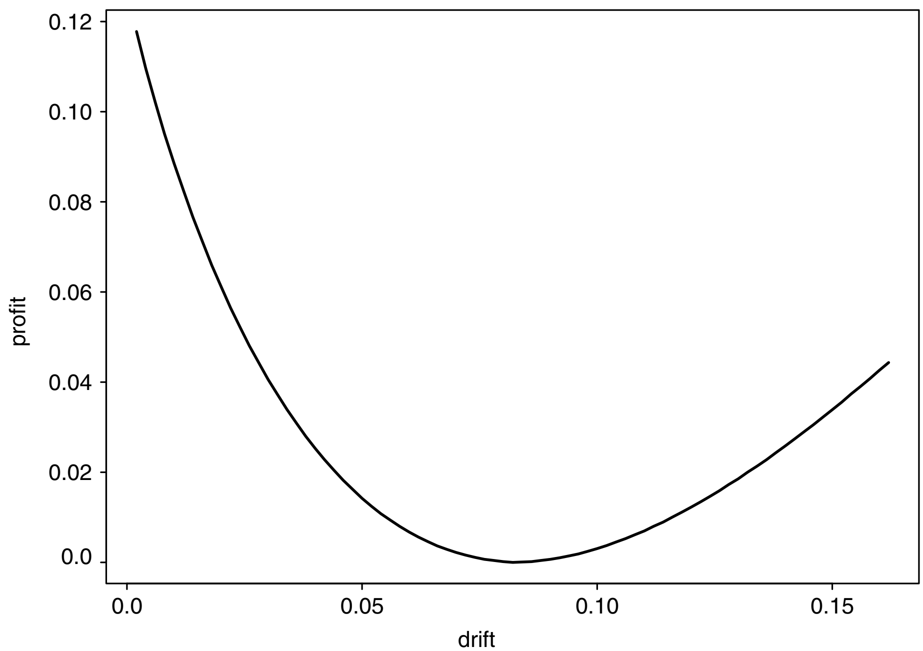


Figure 63.2 The effect of the physical drift.

Figure 63.2 shows the expected profit from selling an at-the-money American put to an investor who maximizes his expected return. In this case, the option holder's criterion in choosing the exercise time is free from risk aversion, and hence the outcome can be considered as the marginal effect of the physical drift to the issuer's expected gain. The initial asset price is 50, the asset volatility is 20% per annum, the maturity of the option is 6 months, and the risk-free rate is 8% per annum. When the physical drift coincides with the risk free rate, the exercise boundary that maximizes the expected return coincides with the price-maximizing exercise boundary, and hence there is no profit for the issuer. When the physical drift surpasses the risk-free rate, the holder's exercise boundary is inside the price-maximizing exercise boundary. In this case, $\mathcal{G}$ is empty and the only source of the issuer's profit is the difference between the value of the option and the exercise value (i.e., the value of ψ on the moving boundary $\partial\mathcal{H}$). If the physical drift is less than the risk-free rate, then the holder's exercise boundary is outside of the price-maximizing exercise boundary, and hence the issuer's profit grows with the occupation time of the asset price in between the two boundaries. This explains the asymmetry in the picture.

Figure 63.3 is the issuer's expected profit as a function of the absolute risk aversion. The option holder's exercise time maximizes the expected CARA utility, while the physical drift coincides with the risk-free rate 8% per annum. Thus, the outcome is the marginal effect of the absolute risk aversion to the issuer's expected profit. Again, the option is an at-the-money American put, the initial asset price is 50, the asset volatility is 20% per annum, and the maturity is 6 months. If the absolute risk aversion vanishes and the physical drift and the risk-free rate coincide, then the utility-maximizing exercise boundary coincides with the price maximizing one, and hence the expected profit vanishes.

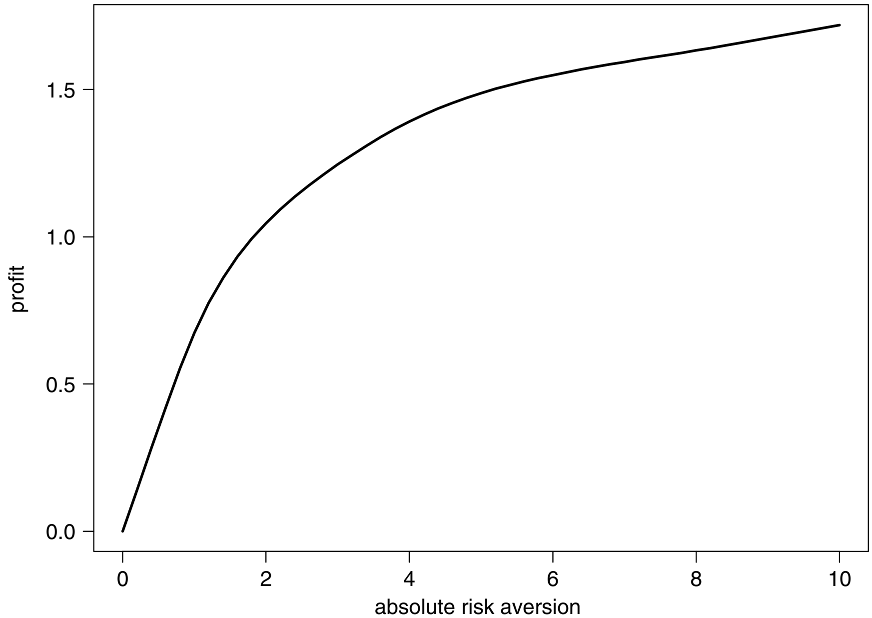


Figure 63.3 The effect of the absolute risk aversion.

63.10 CONCLUDING REMARKS

The theory of optimal stopping has been applied to the valuation of the American option. People are prone to using the terminologies of the theory of optimal stopping when they talk about American options. For example, the price-maximizing exercise boundary has been referred to the optimal exercise boundary, while it is optimal to neither the issuer nor the option holder. This causes confusion to students, practitioners, and even academic researchers in the field. There are two obvious sources of confusion. First, financial software packages almost invariably value contracts, and find exercise strategies, from the option writer-hedger point of view. Rarely do they have anything to say about what is optimal for the contract holder. Some research papers even suggest that we price an American option by estimating the price-maximizing exercise boundary from the empirical data of exercise boundaries. In this paper, we have explained why these ideas are wrong. The holder of an American option should pursue his own profit maximization and choose the right exercise time for himself regardless of the price-maximizing exercise boundary, unless he is determined that the issuer should not gain. As the exercise times of market participants are affected by the market direction, their risk aversion factors, and their financial structures, an estimate of the price-maximizing exercise boundary from empirical data is also invalid.

Acknowledgments

We would like to thank Professor ____ _____, the editor of the _______ *of* _______, for persevering with us after reading our original rather incomprehensible first draft. Paul Wilmott would like to also thank the Royal Society for their support. Hyungsok Ahn would like to thank David Epstein for his service in improving the readability of the text.

REFERENCES

Friedman, A 1988 *Variational Principles and Free Boundary Problems*, Robert E. Kreiger Publishing, New York.

Harrison, JM & Pliska, SR 1981 Martingales and stochastic integrals in the theory of continuous trading, *Stoch. Proc. and Their Appl.*, **11**, 215–260.

Jaillet, P, Lamberton, D & Lapeyre, B 1990 Variational inequalities and the pricing of American options, *Acta Appl. Math.*, **21**, 263–289.

Merton, RC 1990 *Continuous-Time Finance*, Blackwell, Oxford.

Myneni, R 1992 The pricing of the American option, *Annals Appl. Probab.*, **2**(1), 1–23.

Overdahl, JA & Martin, PG 1994 The exercise of equity options: theory and empirical tests, *J. of Derivatives*, Fall, 38–51.

San Martin, J & Protter, P 1993 General change of variables formulas for semimartingales in one and finite dimensions, *Probab. Theory Rel. Fields*, **97**, 363–381.

Van Moerbeke, P 1976 On optimal stopping and free boundary problems, *Arch. Rational Mech. Anal.*, **60**, 101–148.

Wilmott, P, Dewynne, JN & Howison, SD 1993 *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, Oxford.

63.11 WHO WINS AND WHO LOSES?

Trading an option is not a two-person zero-sum game, because both the issuer and the holder can trade the underlying asset with other investors. If the issuer maintains a risk-free portfolio by trading the underlying asset while the holder leaves his position naked, the issuer's balance will be non negative in the end while the holder's depends upon whether or not he guessed market direction correctly. If the holder exercises his option at his own optimal time, there is also a possibility that both issuer and holder can make a profit at the same time. Does this violate the conservation law? Not really, even if we consider an extreme case that everyone in the option market makes a profit. There would then have been an influx of capital from the spot market due to delta hedging.

Let us consider the potential cashflows in some detail. Please refer to Figures 63.4 and 63.5.

In Figure 63.4 we see the profit made by shareholders in a stock that has risen in value over some period, say a year. At the moment there are no options on this stock.

Now let us introduce an American option (or any contract that incorporates some choice/decisions for the holder) into this scenario, the expiry of the option coinciding with the horizon in the previous case, that is, one year. The stock will evolve dynamically exactly as in the non-option case.

What are the cashflows now?

1. If the writer sells the option for 'fair value', according to the delta-hedging-and-price-maximizing method, *and* the holder exercises at the price-maximizing boundary then the writer makes no profit. The premium paid for the option is paid by the holder and is added to the profit made by the shareholders. The mechanism for the transfer of premium from holder to shareholders is simply the process of delta hedging.
2. On exercise or expiry the holder of the option may get a payoff. If the contract is a simple call and ends up in the money then the holder gets the difference between the share price and the strike price. This again does not have any impact on the writer of the option, but via delta hedging, it is taken from the profit of the shareholders.

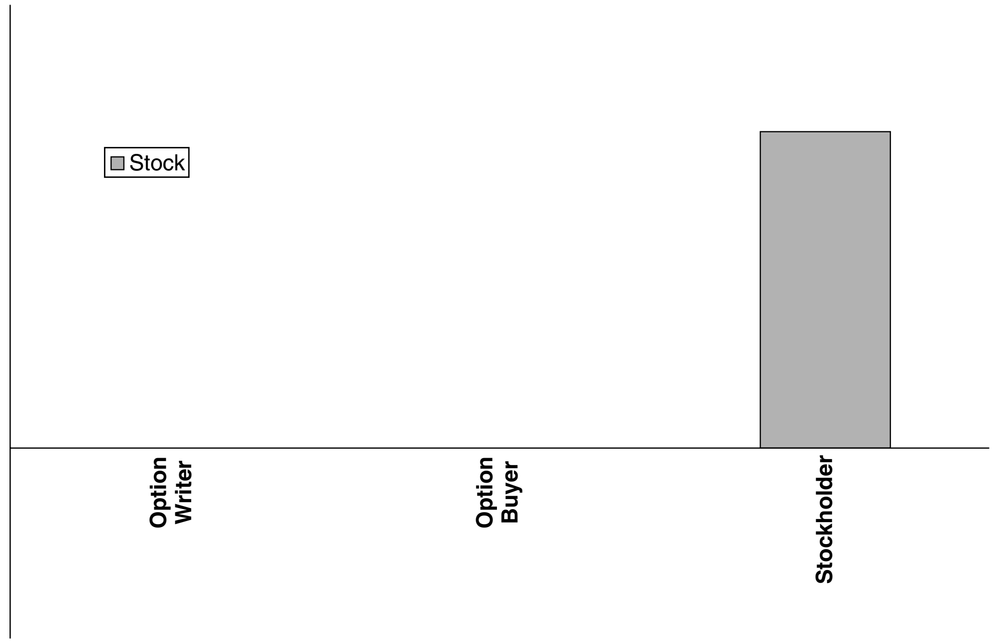


Figure 63.4 Profit made by shareholders in a rising market: No options traded

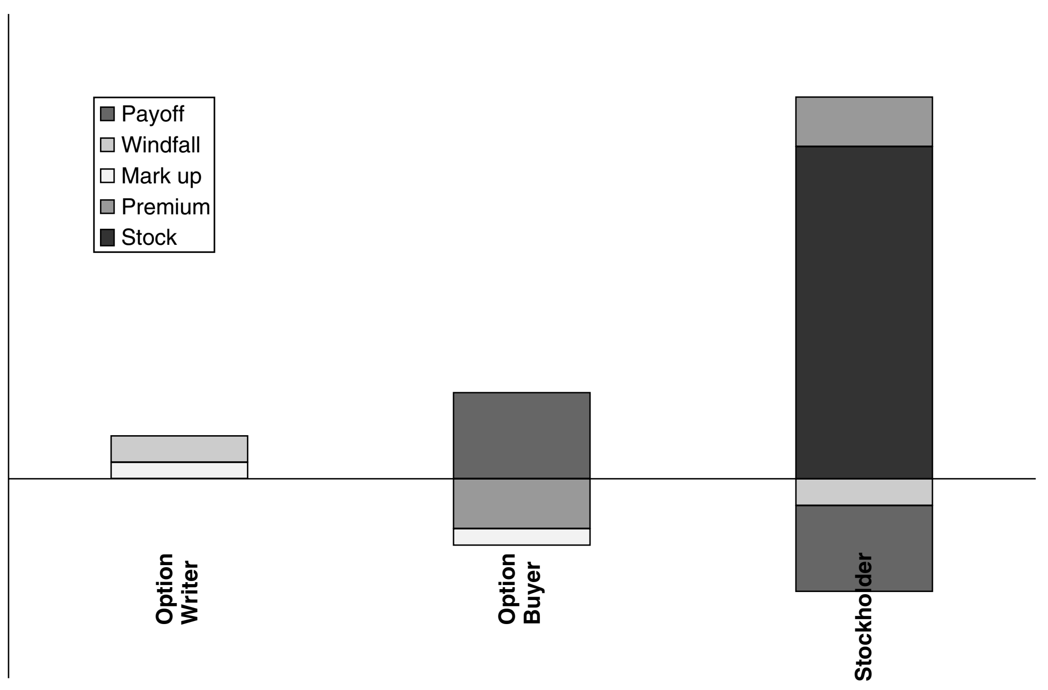
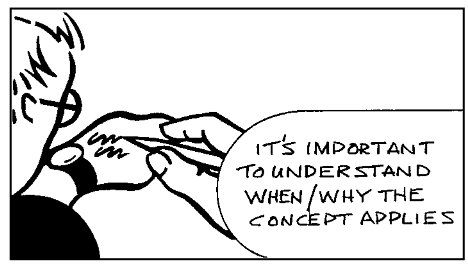


Figure 63.5 Profit made by shareholders in a rising market: American options traded

3. Of course, the writer of the option is in practice going to add a mark-up, his profit margin. This is paid for by the option holder.
4. The final cashflow in this picture is concerned entirely with the non-optimal, in the price-maximizing sense, exercise of the option. There is a windfall profit made by the option writer whenever the holder exercises 'non optimally.' (This profit could be thought of as part of the payoff, but in the diagram we have subtracted it from the shareholders.)

Figure 63.5 shows all of these cashflows. In the perfect Black–Scholes world the option writer profits only from mark-up and from the windfall. The option holder will gain or lose according to the balance between the premium they pay and the payoff they receive. The market value of all the positions at expiry/exercise adds up to the value of the shareholders' profit in a world without options.

63.12 **FAQ**



Q How can there be another value for American options in a Black–Scholes world?
A We are definitely not saying that there is another value. In a world where perfect delta hedging is possible and everyone agrees on the volatility of the underlying which follows a lognormal random walk there should be only one value for an American option. Concentrate, man.

Q Wouldn't the owner always be better off selling the option than exercising it?
A Yes, he would. But this is not always possible. Perhaps there just isn't the necessary liquidity in the market. Perhaps the contract is OTC and the original writer will not let the option holder off the hook without paying a stiff penalty.

Q Well, wouldn't the holder simply delta hedge until maturity?
A Again, yes, he would if he could. But if he can delta hedge why is he buying the option in the first place. Not everyone has access to the underlying asset sufficiently free of transaction costs to allow them to delta hedge.

Q Do people exercise options at 'non-optimal' times in practice?
A Yes, definitely with OTC contracts and any contracts in which the owner has to make complex decisions. Even vanilla options are sometimes exercised at odd times, and they can definitely be sold instead of exercised.

63.13 **ANOTHER SITUATION WHERE THE SAME IDEA APPLIES: PASSPORT OPTIONS**

Remember the passport option? This is a contract that pays off the positive part of a trader's trading account. If he has lost money at the expiry of the contract then he is not liable. If he has made money he keeps it. We saw in Chapter 27 how this turned into a stochastic control problem with the trader having to make decisions about how to invest. From the writer's point of view we have to assume that the trader invests in such a way as to give the option its highest value. In practice he will do whatever he thinks best to make the most money. This is very, very similar to the US option situation above. Before we go on with the analysis it might be best if you read through Chapter 27 again briefly to familiarize yourself with the basic problem and the mathematics involved.

63.13.1 Recap

To value the passport option we introduce a new variable π that is the value of the trading account. This quantity satisfies the following

$$d\pi = r(\pi - qS)\,dt + q\,dS,$$

where q is the amount of stock held at time t. I will restrict the size of the position in the stock by insisting that $|q| \leq 1$.

The contract pays off

$$\max(\pi, 0)$$

at time T. This will be the final condition for our option value $V(S, \pi, t)$. Note that the option value is a function of three variables.

Now let us hedge this option:

$$\Pi = V - \Delta S.$$

... usual Itô stuff here ...

Since $d\pi$ contains a dS term the correct hedge ratio is

$$\Delta = \frac{\partial V}{\partial S} + q\frac{\partial V}{\partial \pi}.$$

From the no-arbitrage principle follows the pricing equation

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + q\sigma^2 S^2 \frac{\partial^2 V}{\partial S\,\partial \pi} + \tfrac{1}{2}q^2\sigma^2 S^2 \frac{\partial^2 V}{\partial \pi^2} + rS\frac{\partial V}{\partial S} + r\pi\frac{\partial V}{\partial \pi} - rV = 0.$$

If we are selling this contract then we should assume that the holder acts optimally, making the contract's value as high as possible. The highest value for the contract occurs when q is chosen to maximize the q terms in the above:

$$\max_{|q|\leq 1}\left(q\sigma^2 S^2 \frac{\partial^2 V}{\partial S\,\partial \pi} + \tfrac{1}{2}q^2\sigma^2 S^2 \frac{\partial^2 V}{\partial \pi^2}\right).$$

Call this strategy q^*.

The sentence in italics is crucial. In practice the option holder *will not* act in this fashion. Let's suppose that he acts to maximize his expected utility; this is a nice framework to analyze this problem, but not necessarily realistic.

63.13.2 Utility Maximization in the Passport Option

In this section we examine how the option holder trades the underlying and how much the issuer gains by selling the option. The investor who owns a passport option may construct his trading strategy to maximize his utility, trying to predict the market movement. When the physical trend of the market differs from the risk-neutral drift significantly, the option holder will benefit as long as he has a correct view on the market. At the same time, the issuer will gain from the difference between the price-maximizing trading strategy and the trading strategy performed by his customer.

First we assume that the option holder finds his strategy by solving the value of the maximum expected utility of the payoff:

$$u(S,\pi,t) = \max_{|q|\le 1} E^{S,\pi,t}\left[e^{-r(T-t)}U\Big(\max(\pi(T),0)\Big)\right] \tag{63.22}$$

where E is the expectation under the real drift and U is the option holder's utility function which is increasing in its argument. u satisfies the following Bellman equation:

$$-\frac{\partial u}{\partial t} = r\pi\frac{\partial u}{\partial \pi} + \mu S\frac{\partial u}{\partial S} - ru + \max_{|q|\le 1}\left(qS(\mu-r)\frac{\partial u}{\partial \pi} + \tfrac{1}{2}\sigma^2 S^2\left(q^2\frac{\partial^2 u}{\partial \pi^2} + 2q\frac{\partial^2 u}{\partial \pi\,\partial S} + \frac{\partial^2 u}{\partial S^2}\right)\right), \tag{63.23}$$

$$u(S,\pi,T) = U\Big(\max(a,0)\Big)$$

where μ is the physical drift of the underlying asset.

If $U(x) = x$, $\partial^2 u/\partial\pi^2$ stays positive and hence the utility maximizing strategy has its value at either ± 1. The interpretation of the linear utility is that the option holder maximizes real expected return. A motivation for studying such utility is that the investor's portfolio is already insured by the passport option he owns and that it is tractable. In this case u has a similarity solution of the form $u(S,\pi,t) = Sh(\pi/S,t)$. Furthermore $h(y,t)$ satisfies the following equation:

$$-\frac{\partial h}{\partial t} = (\mu - r)\left(h - y\frac{\partial h}{\partial y}\right) + \max_{|q|\le 1}\left(\tfrac{1}{2}\sigma^2(y-q)^2\frac{\partial^2 h}{\partial y^2} + q(\mu-r)\frac{\partial h}{\partial y}\right) \tag{63.24}$$

with the terminal data $\max(y, 0)$. From this, we obtain the option holder's trading strategy:

$$q = \operatorname{sign}\left(\frac{\mu - r}{\sigma^2}\cdot\frac{\dfrac{\partial h}{\partial y}}{\dfrac{\partial^2 h}{\partial y^2}} - y\right). \tag{63.25}$$

When μ coincides with the risk-free rate r, (63.24) agrees with the price-maximizing value function for the option and q in (63.25) coincides with the price-maximizing strategy q^*. If μ differs from r, then the option holder's choice will be different from the price-maximizing strategy.

Next we discuss the issuer's hedging strategy, Δ. In Chapter 27, I explained that the hedging strategy must be in tune with the *actual* trading strategy performed by the option holder, and that it is given by

$$\Delta = \frac{\partial V}{\partial S} + q\frac{\partial V}{\partial \pi}.$$

Then the profit of the issuer becomes:

$$P = V(S_0, 0, 0) + \int_0^T e^{-rt}\Delta\Big(dS - rS\,dt\Big) - e^{-rT}V(S(T),\pi(T),T). \tag{63.26}$$

The first term is the price of the option he collects in cash, the second is the result of the delta hedging, and the third is the present value of the potential liability.

Applying Itô's formula to V yields:

$$P = -\int_0^T e^{-rt}\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2\left(\frac{\partial^2 V}{\partial S^2} + 2q\frac{\partial^2 V}{\partial S\partial\pi} + q^2\frac{\partial^2 V}{\partial \pi^2}\right)\right)dt$$

$$= \tfrac{1}{2}\sigma^2\int_0^T dt\, e^{-rt}S^2\cdot\left((q^{*2} - q^2)\frac{\partial^2 V}{\partial\pi^2} + 2(q^* - q)\frac{\partial^2 V}{\partial S\partial\pi}\right) \qquad (63.27)$$

where q^* is the price-maximizing strategy and q is the strategy performed by the option holder.

Recall that $V(S,\pi,t)$ has a similarity solution $SH(\pi/S,t)$. In particular, we have

$$\frac{\partial^2 V}{\partial S\partial\pi} = -\frac{\pi}{S^2}\frac{\partial^2 H}{\partial\xi^2} \qquad \text{and} \qquad \frac{\partial^2 V}{\partial\pi^2} = \frac{1}{S}\frac{\partial^2 H}{\partial\xi^2},$$

where $\xi = \pi/S$. Also recall that $q^* = -\operatorname{sign}(\xi)$. Thus we have a further reduction in the integrand of (63.27):

$$S^2\cdot\left((q^{*2} - q^2)\frac{\partial^2 V}{\partial\pi^2} + 2(q_\star - q)\frac{\partial^2 V}{\partial S\partial\pi}\right) = \frac{\partial^2 H}{\partial\xi^2}\cdot\Big(2|\pi| + 2q\pi + (1-q)^2 S\Big) \qquad (63.28)$$

Now suppose that the option holder finds his strategy by maximizing the expected return. Then, as we computed earlier in (63.25), the option holder's strategy q depends on π and S only through the ratio $\pi/S = \xi$ and has its value either ± 1. Hence the last term in (63.28) drops out, and the profit of the issuer becomes:

$$P = \sigma^2\int_0^T e^{-rt}S\Big(|\xi| + q(\xi,t)\xi\Big)\frac{\partial^2 H}{\partial\xi^2}\,dt. \qquad (63.29)$$

To obtain the expected profit $E[P]$ of the issuer, we define

$$g(S,\pi,t) = \sigma^2 E^{S,\pi,t}\left[\int_t^T e^{-r\tau}S\,(|\xi| + q\xi)\,\frac{\partial^2 H}{\partial\xi^2}\,d\tau\right].$$

Again we observe that g has a similarity solution of the form $g(S,\pi,t) = S\psi(\pi/S,t)$ and that $\psi(\xi,t)$ satisfies the following:

$$-\frac{\partial\psi}{\partial t} = (\mu - r)(q-\xi)\frac{\partial\psi}{\partial\xi} + \mu\psi + \tfrac{1}{2}\sigma^2(q-\xi)^2\frac{\partial^2\psi}{\partial\xi^2} + \sigma^2 e^{-rt}\Big(|\xi| + \xi q\Big)\frac{\partial^2 H}{\partial\xi^2}.$$

subject to $\psi(\xi,T) = 0$. To solve this equation, we need to obtain H from (27.3) and q from (63.24).

Figure 63.6 shows the expected gain by the issuer as a function of μ, the physical drift, that is $\psi(0,0)$ against μ. The asset volatility is 20% per annum and the maturity of the option is six months. We calculate the profit $100\cdot\psi$ at different values of the physical drift from zero to 16%. When the drift coincides with the risk-free rate of 8% per annum, the gain vanishes. As explained earlier, the issuer gains more as the gap between the drift and the risk-free rate become larger.

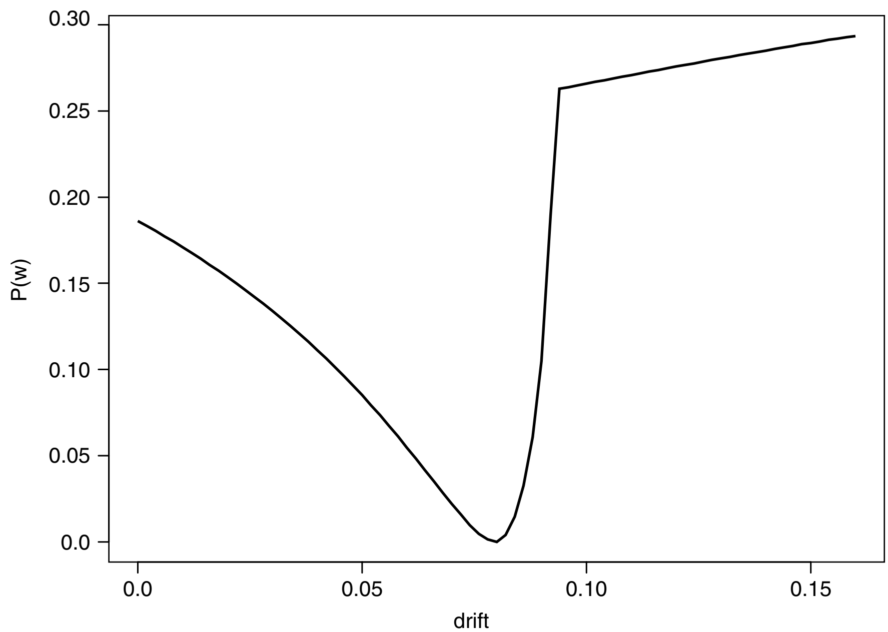


Figure 63.6 Issuer's expected gain versus the drift of the underlying asset.

63.14 SUMMARY

I'd like to hear what you have to think on this subject and I'll keep a poll of people's views. Email me on `paul@wilmott.com`.

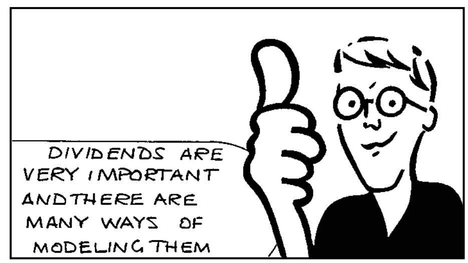
FURTHER READING

- More details about the passport option are contained in Ahn, Penaud & Wilmott (1998).
- The American option paper was finally published in *Wilmott* magazine Ahn & Wilmott (2003).

CHAPTER 64
advanced dividend modeling

In this Chapter...

- the importance of dividend modeling
- the stochastic dividend model
- uncertainty in dividend amount and timing

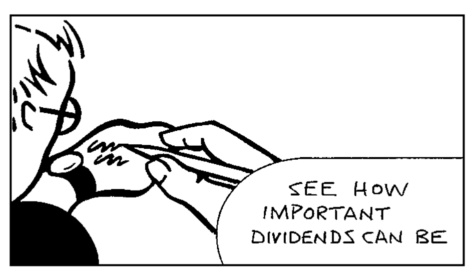


64.1 INTRODUCTION

We've already seen in Chapter 8 how to incorporate dividends into option models. But that supposes that we have a decent knowledge of the timing and the amount of the dividend. I relaxed the requirement for perfect knowledge of the amount of the dividend in Chapter 52 when we looked at best and worst cases. Here I present some more sophisticated models.

64.2 WHY DO WE NEED DIVIDEND MODELS?

The primary reason why the financial literature extensively analyzes dividends on equity is the perfect market hypothesis, i.e. that stocks are valued as their discounted dividend stream in perpetuity. To a large extent, equity analysts use this model and its variations to estimate a stock's fair value and forecast its future performance. Derivatives analysts, however, in order to value and hedge a portfolio of options on a stock or an index, normally require as standard parameters the current spot price, exercise price, risk-free rate, expiry time, volatility and, often, the stock's dividend yield over the life of the option. For many practitioners the dividend yield is considered to be relatively unimportant compared with the other parameters.



The amount of literature available on dividend models applied to derivatives is tiny. But if we compare the sensitivity ratios for different types of options, it becomes apparent that this might not describe the real world accurately. Compare, for example, option price sensitivity to the dividend yield with the sensitivity to the volatility, the vega. Figure 64.1 shows the ratio of these two sensitivities against volatility and dividend yield. The option is a European call with one year to expiry, struck at 98, the risk-free rate is 5%. The underlying asset is

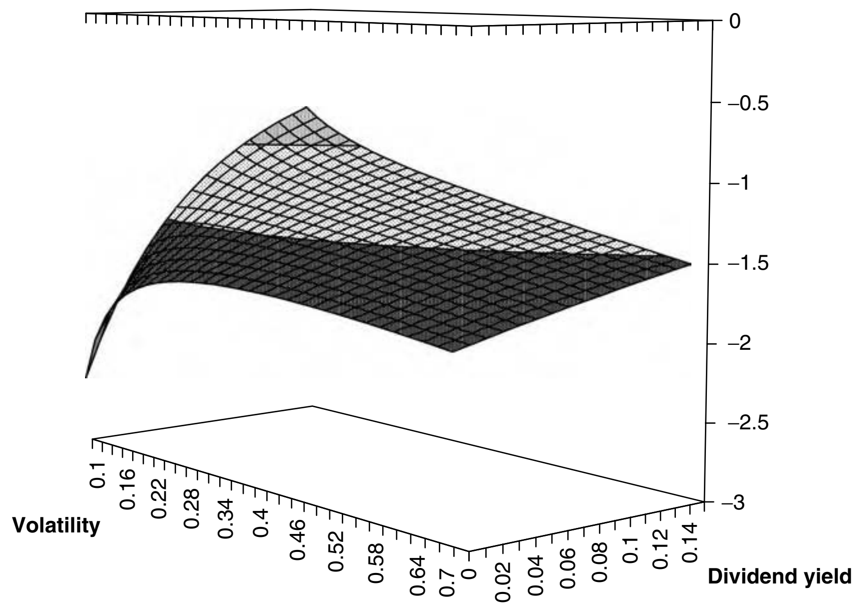


Figure 64.1 Ratio of dividend sensitivity to vega, call option.

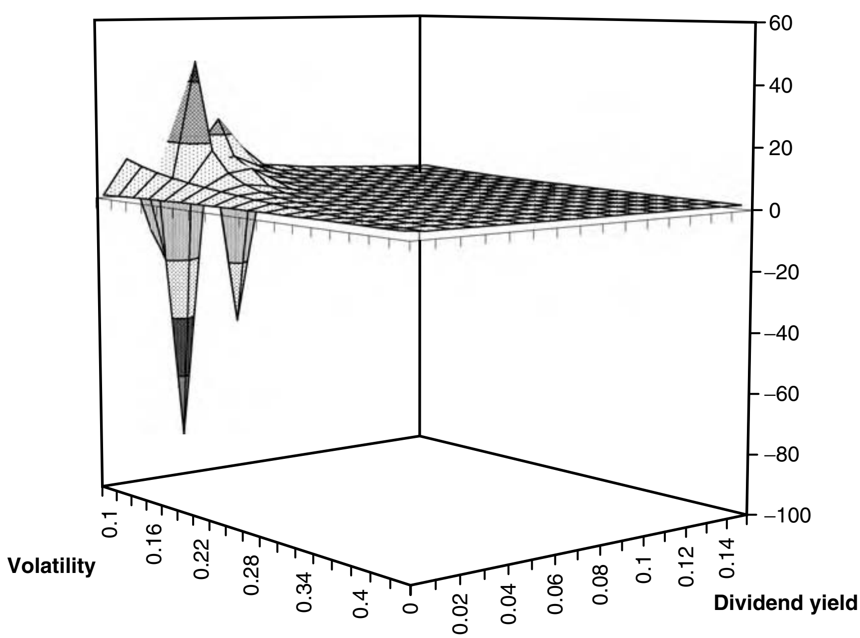


Figure 64.2 Ratio of dividend sensitivity to vega, binary option.

100. We can see that for most parameter values this ratio is greater than one. In other words, the sensitivity to the dividend is more important than the sensitivity to the volatility. This is even more pronounced for a binary call option (see Figure 64.2).

In this chapter I attempt to give a concise summary of the relation between dividends on equity and the value of derivatives. You'll see various approaches to modeling the dividend process of a security or an index, elaboration of existing models and some new ideas. Choices normally have to be made as to whether to use deterministic dividends or stochastic dividends and discrete dates or a continuous flow. We'll see all of these here.

64.3 EFFECTS OF DIVIDENDS ON ASSET PRICES

In most financial theory the standard assumption is that the spot price S is lognormally distributed, satisfying a stochastic differential equation. The paths of the random walk are continuous almost everywhere, but smooth almost nowhere. However, such a model is only valid for non-dividend paying stocks. If the underlying stock is paying discrete dividends there are two points in time where the stock and/or the underlying asset might jump discontinuously. The two dates to examine closely are the date that the dividend is announced and the date when the asset goes ex-dividend.

Figure 64.3 shows the paths of a stock and option price, with dividend announcement date and ex-dividend date marked.

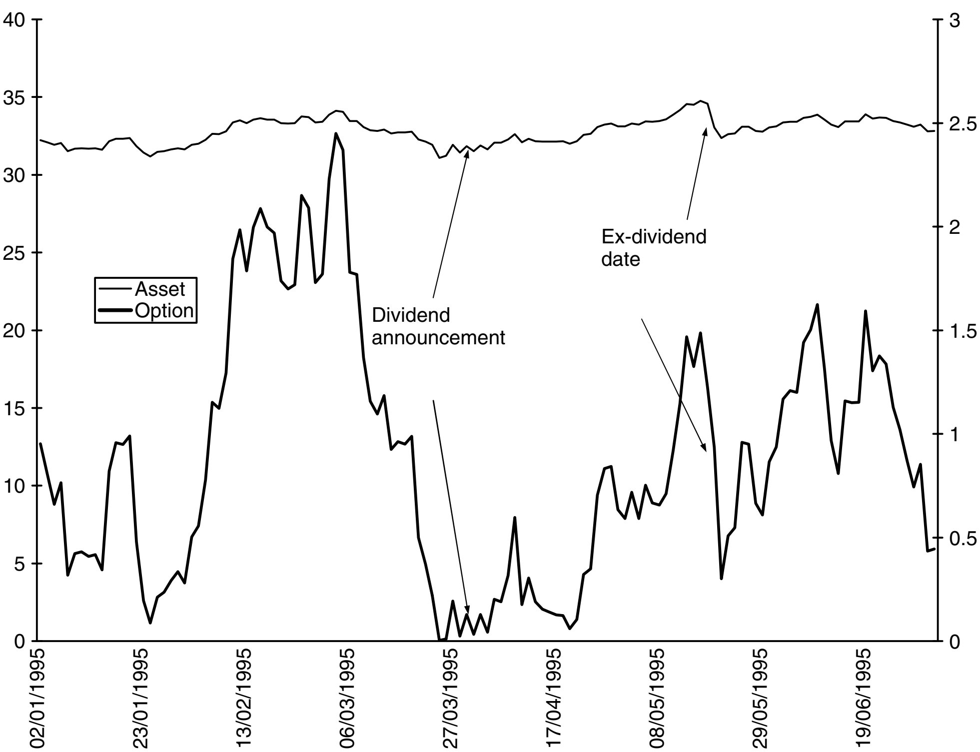


Figure 64.3 Path of a stock and an option on that stock.

Consider first the ex-dividend date t_i. If markets are efficient and frictionless, the stock price drops discontinuously by exactly the, possibly time- and spot-dependent, dividend D;

$$S(t_i^+) = S(t_i^-) - D.$$

However any option written on the underlying stock stays unchanged in its price due to no-arbitrage arguments, thus

$$V(S, t_i^-) = V(S - D, t_i +)$$

We saw this in Chapter 8.

The other key date t_a, which itself may be random, is the time when the company announces the dividend size and ex-dividend date on its stock. Firstly, the value of the derivative may jump discontinuously because of a discrepancy between implied and actual dividend and/or the exact ex-dividend date. The implied values, in practice, are inferred from prices of traded futures. We'll come back to implied dividends shortly. Across the dividend announcement date the price of the derivative can change. The path of the derivative will stay continuous only if the size of the dividend and its ex-dividend date are as expected by the market.

On the other hand, in practice, the stock price itself may change in light of this new information, but it is not always clear in which direction. A growing dividend may signal good earnings prospects and a dividend decline, financial distress; but not necessarily. However on a more theoretical level, the famous theorem by Modigliani & Miller states that stock returns should not change with the dividend policy, hence there should be no discontinuous jump in S across t_a.

64.3.1 Market Frictions

In reality the jump in stock-price across the ex-dividend date is not exactly the amount of the dividend due to frictions in the market. The error term is mostly attributed to different taxation of dividends and capital gains.

Some empirical tests conclude that due to short-term trading the tax effects are removed. That means, because of low transaction costs, market makers who unlike other investors have nearly equal tax rates, buy stocks with high dividend yield. Indeed, the trading volume in stocks around t_i is exceptionally high (see Figure 64.4).

A second market friction is introduced if the dividend payout date is not the same as the ex-dividend date. Typically, German companies pay their dividend on the ex-dividend date, UK companies five weeks after the ex-dividend date and other European companies up to six months after the ex-dividend date. Hence the stock price should jump by an amount that discounts the dividend accordingly.

We've seen the theory behind deterministic dividends in Chapter 5. We'll take a quick look at the term structure of dividends and then concentrate on random or uncertain dividend models.

64.3.2 Term Structure of Dividends

For individual stocks and short-dated index instruments practitioners rely on fundamental research when estimating the dividend correction to an option value. That means they try and estimate every single dividend amount, timing and the anticipated percentage drop. However for long-dated derivatives or for large indices it may be more efficient to rely on implied

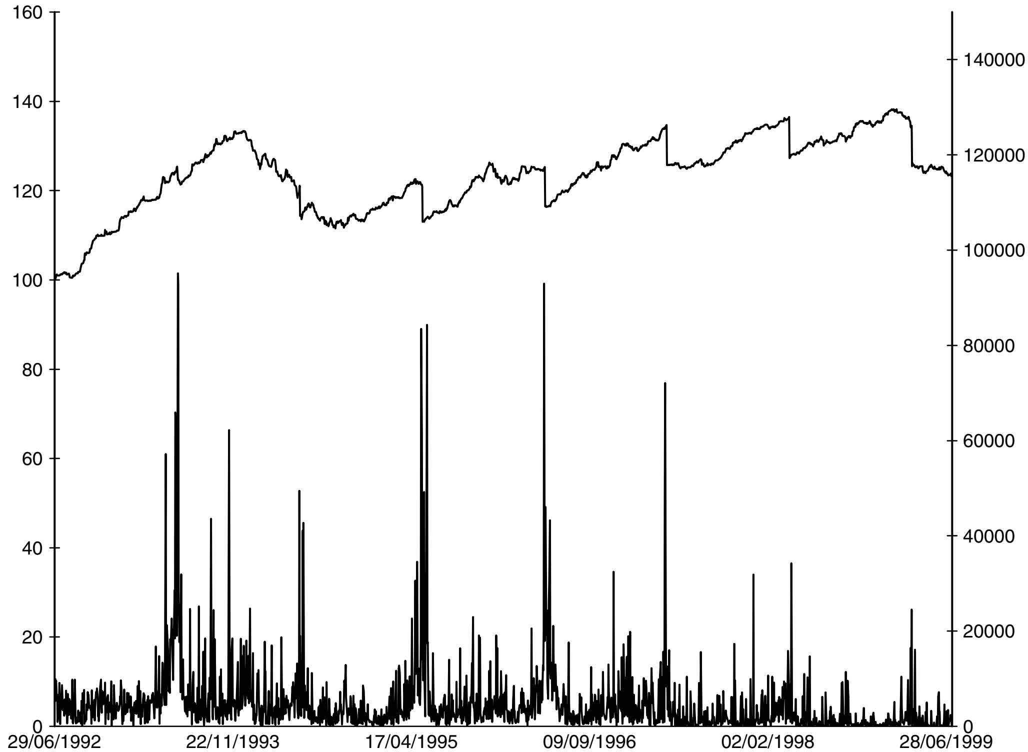


Figure 64.4 Price and volume.

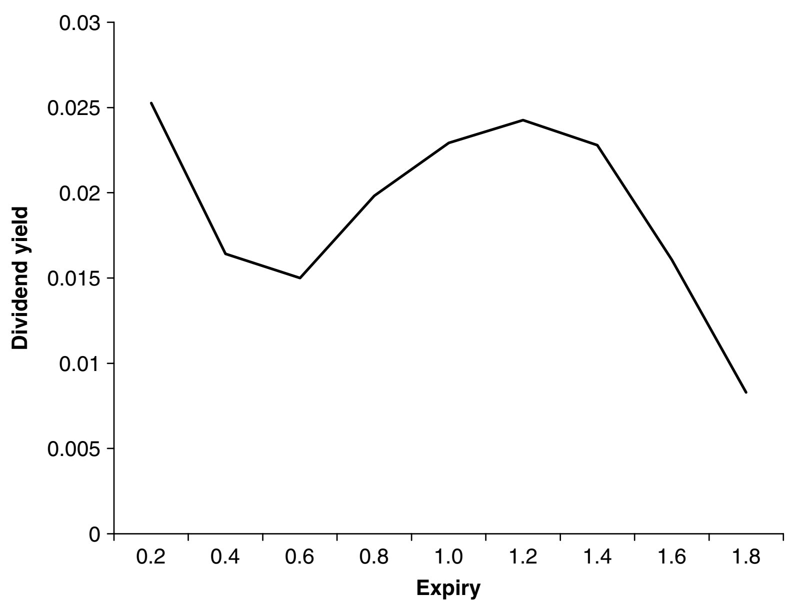


Figure 64.5 Term structure of dividend yield

dividend yields. The term structure of dividends can be inferred from traded futures and the interest rate yield curve. The relationship between the forward prices F and the spot price S is

$$F = Se^{(r-D(t;T))(T-t)}.$$

Here r is the interest rate, which I've assumed to be constant for all time, and $D(t; T)$ the dividend yield over the period t to T.

Typically, there are only a limited number of liquid futures and therefore the implied dividend term structure will consist of only a discrete number of actual points. If the dividend yield of a non-traded interim expiry is required it can be interpolated (see Figure 64.5).

64.4 STOCHASTIC DIVIDENDS

Let's now look at a more traditional stochastic model, this time for dividend yield.

$$dD = p(S, D, t)\,dt + q(S, D, t)\,dX_2$$

with a correlation of ρ with stock returns.

If we want to obtain the value of a derivative on the dividend-yielding stock we will have to deal with an additional source of randomness that has to be hedged away by another instrument. To derive an equation we again set up a hedged portfolio, hedging one option with the underlying and another option.

Naturally, we end up with

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \sigma q S \frac{\partial^2 V}{\partial S \partial D} + \tfrac{1}{2}q^2 \frac{\partial^2 V}{\partial D^2} + (r - D)S\frac{\partial V}{\partial S} + (p - \lambda q)\frac{\partial V}{\partial D} - rV = 0$$

and the territory is familiar.

64.5 POISSON JUMPS

One phenomenon that sometimes happens during the life of an option is the announcement of special, i.e. non-recurring, dividends. These tend to be paid out when the firm is restructuring its financing or after mergers. These dividends lead to jumps in the option prices since they are not usually anticipated. These events can be incorporated as a Poisson jump process when evaluating the price of the option. This is exactly the same as modeling a stock market crash, which could be done in the Merton, Chapter 57, or worst-case, Chapter 58, sense.

64.6 UNCERTAINTY IN DIVIDEND AMOUNT AND TIMING

Instead of approximating the averages of a set of parameters or making assumptions about the respective stochastic processes that they follow, it may be more robust to come up with estimates of intervals that cannot be breached, then the option is valued in a worst-case scenario. We've seen this idea in Chapters 52 and 58 and will see it applied to interest rates later on. Indeed, in Chapter 52 we even saw how to apply the idea to uncertain dividend amounts or

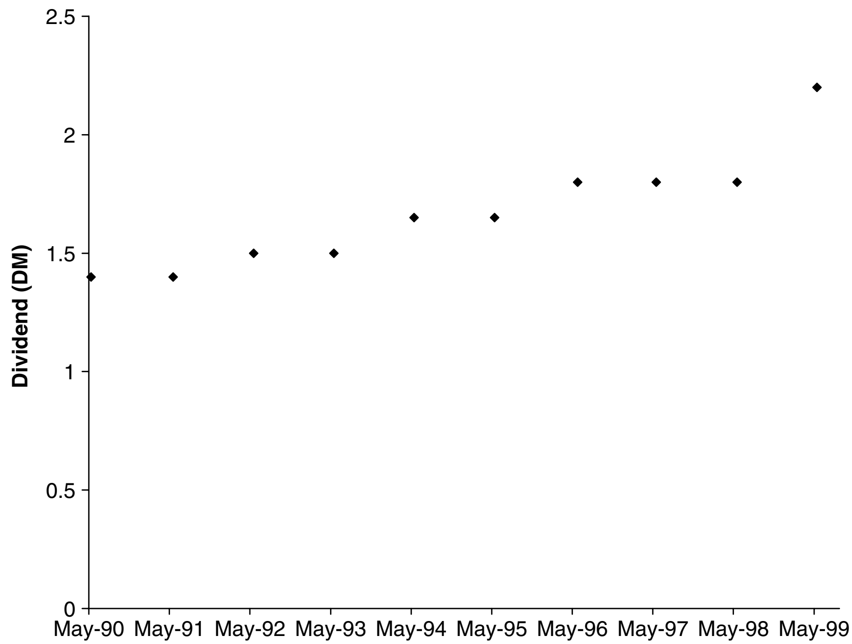


Figure 64.6 Deutsche Bank dividends.

yields. Applied to dividends the simplest model is to assume that the dividend or the dividend yield lies within a range. This can be applied to a continuously or discretely paid dividend. See Figure 64.6 for a short time series of Deutsche Bank dividends.

Figure 64.7 shows the ex-dividend months for components of the Eurostoxx50, 1996–1998. We can see from this that companies are not necessarily consistent in the timing of their dividends, never mind the amounts.

A more realistic scenario is to assume that both dividends and ex-dividend dates are uncertain, both lying in ranges. Let's write the range for dividends as

$$D^- \leq D \leq D^+$$

and the time range

$$t_i^- \leq t_i \leq t_i^+.$$

We are now very much in the discretely paid dividend world.[1]

Assuming that ranges for one ex-dividend date do not overlap ranges for another, we must introduce the functions $V_0(S, t)$ and $V_1(S, t)$ representing option values with dividend having been paid and not yet paid, respectively. These two functions exist during the time range containing the ex-dividend date. We also have the function $V(S, t)$ as the option value outside of these ranges. These functions all satisfy the basic Black–Scholes

[1] These definition for $t_i^{+/-}$ are different from our usual definitions, encountered when there are jump conditions.

1996

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
					elf						
			schneider	veba	telitalia						
			portel	unicredito	stgobain						
			nokia	rdutch	lvmh						
		vivendi	lufthansa	repsol	loreal		unilever				
		telefonica	endesa	petrofina	fortis		repsol				
		rhone	elsevier	electrabel	eni		ing	rdutch			
	siemens	philips	carrefour	deubank	endesa	generali	elsevier	kpn			
telefonica	paribas	ing	Bayer	ahold	alcatel	fiat	allianz	ahold	kpn	unilever	rwe
mannes	allirish	airliquide	akzonobel	aegon	abnamro	akzonobel	abnamro	aegon	BBV	lvmh	allirish

1997

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
				veba	elf						
				unicredito	telitalia						
				telefonica	stgobain	repsol					
				rdutch	lvmh	kpn	vivendi				
				petrofina	lufthansa	ing	nokia				
telefonica	siemens			paribas	francetel	generali	metro				
lvmh	schneider		portel	electrabel	eni	fiat	elsevier				
loreal	mannes	rhone	elsevier	deubank	deutel	endesa	ahold	unilever		unilever	
kpn	Bayer	philips	carrefour	airliquide	allirish	BBV	aegon	rdutch	BBV	fortis	
ing	allirish	endesa	akzonobel	aegon	alcatel	ahold	abnamro	abnamro	akzonobel	allianz	rwe

1998

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
				veba							
				unicredito							
				telefonica	elf						
				portel	telitalia						
				petrofina	stgobain						
				paribas	mannes	unilever					
				electrabel	lvmh	repsol				unilever	
		rhone		deutel	lufthansa	metro	ing			rwe	
	siemens	philips	kpn	deubank	francetel	generali	elsevier			rdutch	
vivendi	schneider	nokia	carrefour	airliquide	eni	fiat	allirish	rdutch		ing	
telefonica	endesa	kpn	Bayer	ahold	endesa	ahold	aegon	repsol	BBV	fortis	
lvmh	allirish	elsevier	akzonobel	aegon	alcatel	abnamro	abnamro	allianz	akzonobel	BBV	loreal

Figure 64.7 Ex-dividend months for components of the Eurostoxx50, 1996–1998.

equation. They also satisfy the conditions

$$V_1(S,t) \le \min_{D^- \le D \le D^+} \left(V_0(S-D,t)\right).$$

We must then have

$$V(S,t_i^-) = \min\left(V_1(S,t_i^-), \min_{D^- \le D \le D^+} \left(V_0(S-D,t_i^-)\right)\right)$$

and

$$V_0(S,t_i^+) = V_1(S,t_i^+) = V(S,t_i^+).$$

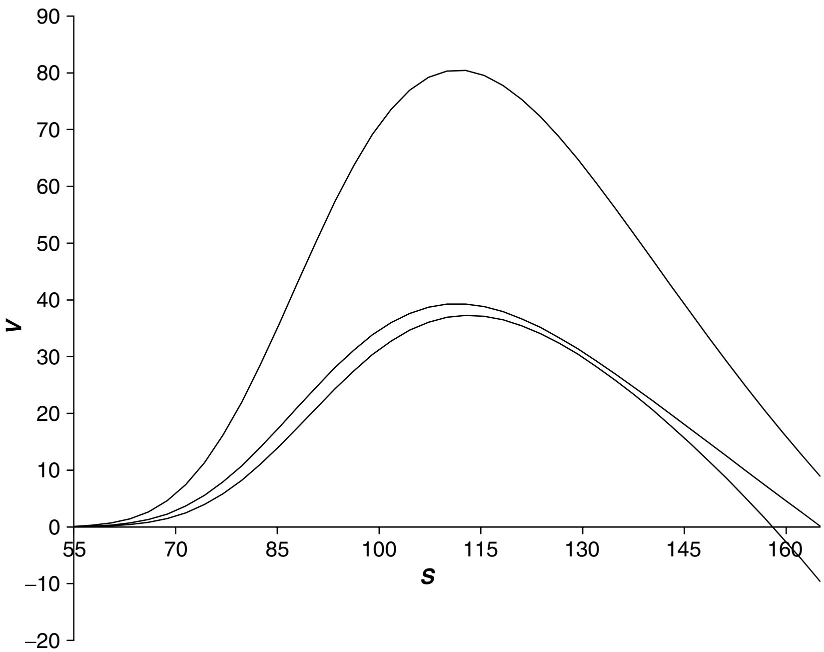


Figure 64.8 Best and worst cases when amount and timing of dividend are uncertain.

In Figure 64.8 are shown the best and worst cases for a portfolio of long two calls struck at 80 and short three struck at 110. There are six months to expiry. This is the same example as in Chapter 52. There we looked at a single dividend paid just before expiry lying in the range zero to five. Here we increase the uncertainty by making the dividend date lie between 0.49 and 0.51 years. Since the expiry of the portfolio is in 0.5 years, the ex-dividend date could be before or after expiry. The best-worst range can be so big that static hedging is necessary to reduce the spread.

64.7 SUMMARY

Dividends can have a major impact on the price of derivatives. Knowledge about the size and timing of dividends are often more important than having accurate volatility knowledge. Here we've seen some of the latest approaches to the subject of modeling dividends. Much of this is quite familiar by now, such as stochastic dividends. However, the uncertain timing of dividends is a new concept to us.

FURTHER READING

- There is a good description of dividends in Ingersoll (1987).
- For more on the term structure of dividends see Fabozzi & Kipnis (1984).
- The idea of stochastic dividends is due to Geske (1978). See also Hilliard & Reis (1998).

- Barone–Adesi & Whaley (1986) fail to reject the hypothesis that there is a tax-factor for dividends on US stocks. On the other hand, however, Kaplanis (1986) finds significant difference for implied dividends in option prices on UK stocks.
- For more details of the models described above see Bakstein & Wilmott (1999).
- See Lewis, Haug & Haug (2003) for work on discrete dividends.

CHAPTER 65

serial autocorrelation in returns

In this Chapter...

- measuring serial autocorrelation
- modeling serial autocorrelation
- the telegraph equation

65.1 INTRODUCTION

One of the many assumptions underlying the random walks used in finance, at least for equities, is that of independent returns. The return on one day is independent of whatever happened in the past. The fact that a stock price has risen ten days in a row is irrelevant, the next return is a completely fresh start.

Now this may be true for a coin toss but should it be true for financial series? And if it's not, how can we incorporate this into a model?

65.2 EVIDENCE

Calculating serial autocorrelation is simple. It's like calculating correlation but you only use one time series.

Suppose you have a series of returns from stock price data, and you've subtracted off the mean and divided by the standard deviation to make a series, R_i, with zero mean and unit standard deviation.

The serial autocorrelation with a lag of one day is then calculated as

$$s_1 = \sum_{i=1}^{N-1} R_i R_{i+1}.$$

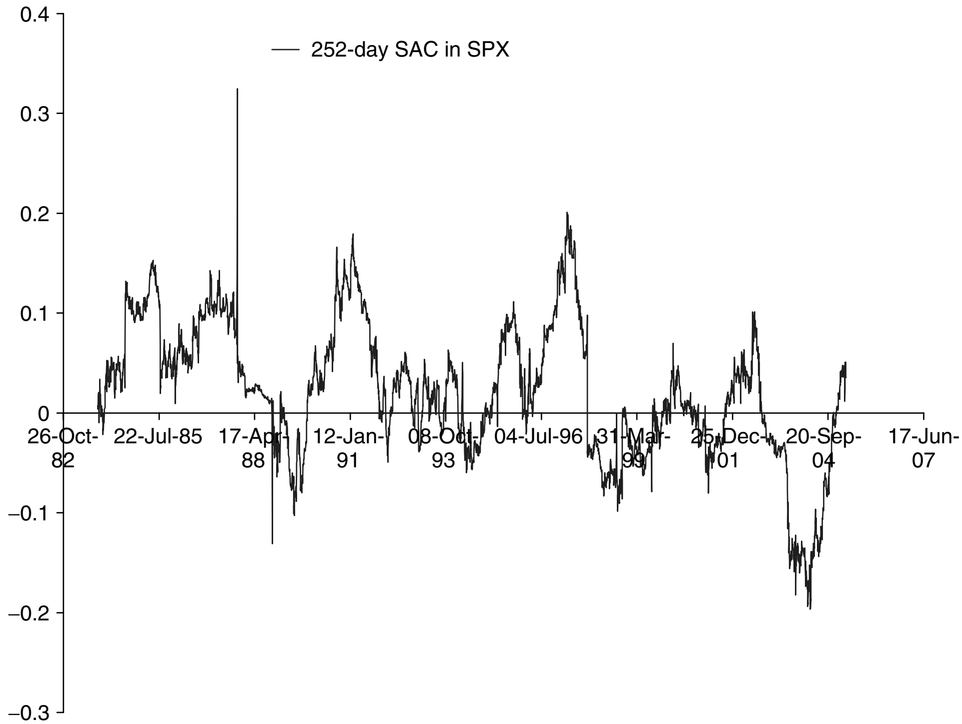


Figure 65.1 Moving 252-day, one-day lag, serial autocorrelation in SPX.

You can similarly calculate the autocorrelation with a two-day lag etc.:

$$s_2 = \sum_{i=1}^{N-2} R_i R_{i+2}.$$

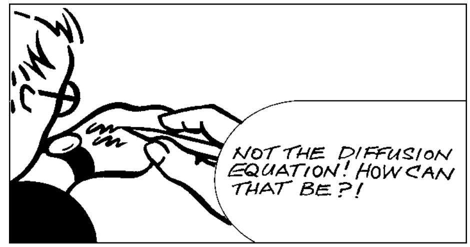
In Excel this is particularly simple to do. Just use the worksheet function `CORREL`. This takes in two arrays as its arguments. The first array will be the returns series, not including the first one. The second array will be the returns series from the first until the penultimate. In other words, calculate the correlation between a series and the series lagged by one day. To calculate the two-day lagged autocorrelation is similar, just shift the series by two days instead of one in the `CORREL` function.

Figure 65.1 shows the moving 252-day, one-day lagged, serial autocorrelation in SPX. The average level of this correlation is about 0.02. However, prior to the year 2000 the average level had been above 4%. Is this significant, is there something structural going on causing serial autocorrelation to venture into negative territory of late? Or is this all just sampling error, and really there isn't any correlation at all?

You can interpret positive serial autocorrelation as being trendlike,[1] whereas a negative serial autocorrelation might be associated with profit taking. Are the ubiquitous hedge funds affecting the traditional behavior of the markets?

[1] 'Trend' is technically not the correct word, of course, since trending is really an $O(1)$ timescale phenomenon, caused by the drift coefficient, whereas serial autocorrelation is an infinitesimal timescale phenomenon.

65.3 THE TELEGRAPH EQUATION



Before we look at pricing derivatives in the presence of serial autocorrelation, let's look at the equation for the probability density function for a random variable that follows a serially autocorrelated random walk.

I'm not going to do the analysis in its full generality, just for a quite specific case, so that you can see how the structure of the governing equation is different from that for the diffusions we have seen so far. My probabilities will be constants and I won't include a drift. So I'll end up with a relatively simple equation for the probability density function.

Start a random walk for a variable y at the origin at time zero. This variable will increase or decrease by δy in a time step δt. Crucially, the probability of the variable increasing or decreasing will depend on what it did the previous time step.

Toss a coin to get started. So it's 50-50 whether y goes up or down. Thereafter the probability will depend on the previous direction. The probability of y going to $y + \delta y$ will be $0.5 + \phi$ if the previous move was up, and $0.5 - \phi$ if it was down, ϕ is a constant between plus and minus a half. The serial autocorrelation is 2ϕ for this random walk.

Figures 65.2 and 65.3 show simulations of such random walks. The trending nature of the walk with positive serial autocorrelation can be seen clearly in Figure 65.2, as can the oscillatory nature of the walk with negative serial autocorrelation.

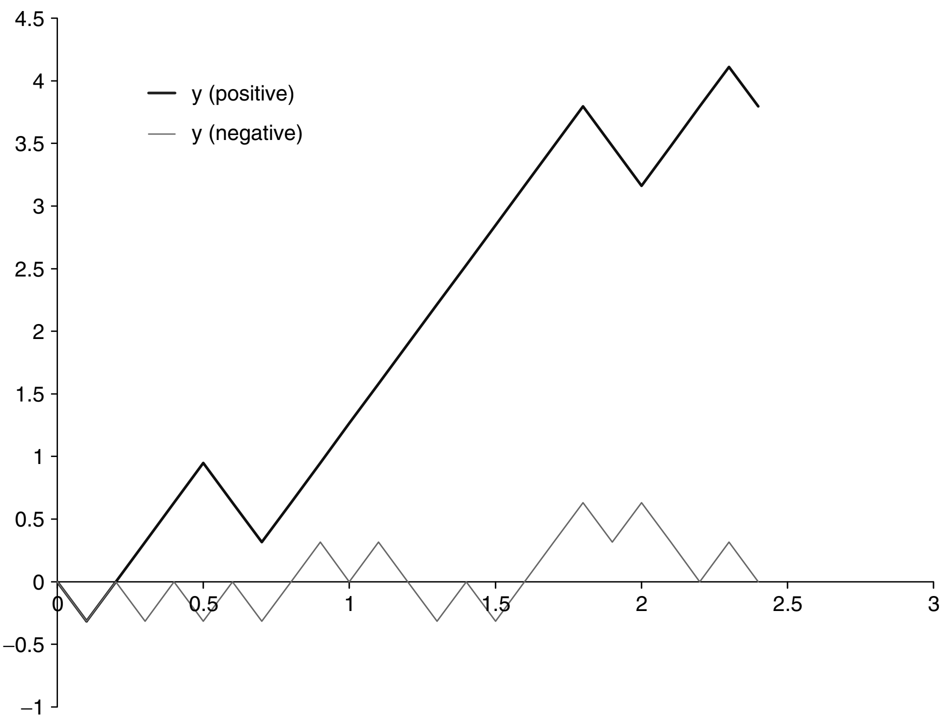


Figure 65.2 Simulation of random walks with serial autocorrelation, close up. $\phi = 0.3$.

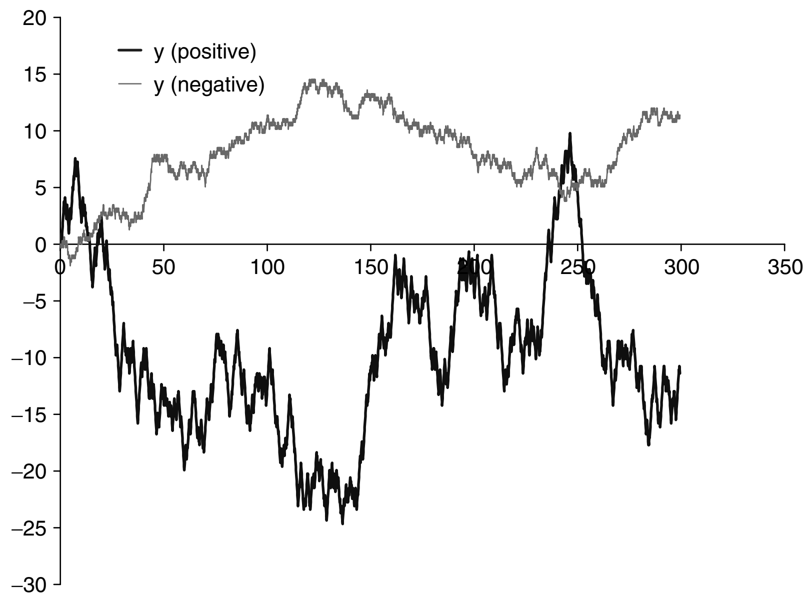


Figure 65.3 Simulation of random walks with serial autocorrelation, longer time scale. $\phi = 0.3$.

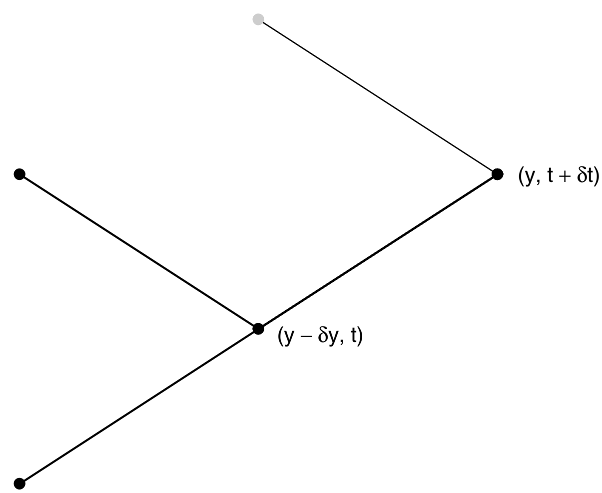


Figure 65.4 Representation of the random walk.

Introduce $p^{+}(y, t)$ as the probability that the random variable is at y at time t having risen. $p^{-}(y, t)$ is the probability that it is at y having fallen. Referring to Figure 65.4 we can see that

$$p^{+}(y, t+\delta t) = \left(\tfrac{1}{2} + \phi\right) p^{+}(y-\delta y, t) + \left(\tfrac{1}{2} - \phi\right) p^{-}(y-\delta y, t).$$

Similarly,

$$p^{-}(y, t+\delta t) = \left(\tfrac{1}{2} + \phi\right) p^{-}(y+\delta y, t) + \left(\tfrac{1}{2} - \phi\right) p^{+}(y+\delta y, t).$$

Expanding in Taylor series for small δt and δy we get

$$\delta t\ \frac{\partial p}{\partial t} = -2\phi\ \delta y\ \frac{\partial v}{\partial y}$$

and

$$(1-2\phi)v + \delta t\ \frac{\partial v}{\partial t} = -\delta y\ \frac{\partial p}{\partial y}.$$

where

$$p = p^{+} + p^{-} \quad \text{and} \quad v = p^{+} - p^{-}.$$

Eliminating v between these two equations results in

$$\frac{\partial^2 p}{\partial y^2} - a\ \frac{\partial p}{\partial t} = c^2\ \frac{\partial^2 p}{\partial t^2}. \tag{65.1}$$

where a and c are the constants

$$a = \frac{(1-2\phi)\ \delta t}{2\phi\ \delta y^2} \quad \text{and} \quad c^2 = \frac{\delta t^2}{2\phi\ \delta y^2}.$$

Equation (65.1) is the **Telegraph Equation**.

This equation differs from all other equations we have derived in this book in one crucial aspect. It contains a second derivative with respect to time. Such a term is associated with momentum.[2] When ϕ is positive this equation is hyperbolic, a wave equation, and not diffusive. The word 'momentum' itself conveys the role of this new term.[3]

65.4 PRICING AND HEDGING DERIVATIVES

So, if the equation for the probability density function is no longer diffusive, what happens to the equation for pricing derivatives? Does this also become a wave equation? No, it doesn't. The pricing equation is still Black–Scholes as you go to the continuous-time limit. The reason for this is that the action of hedging the option with the underlying eliminates any dependence on the probability of a stock rising or falling and so the fact that this probability depends on the history of the stock is also irrelevant. We are right back to risk-neutral pricing and the diffusion equation.[4]

If you are interested in hedging only discretely in time, as you must be in practice, then you can easily modify the delta to allow for serial autocorrelation along the lines explained in Chapter 47. Finally, if you find yourself in a market with significant negative autocorrelation, and also such that implied volatilities rise when stocks fall, you can time the sale of put options to benefit from drops in the market. At such times the implied volatility will rise, so that you bring in more premium, but then you will benefit from the probable bounce back due to negative autocorrelation.

---

[2] Think Newton, force equals rate of change of momentum.

[3] The astute reader will observe that as $\delta t \to 0$ there are some severe restrictions on the sizes of ϕ and δy.

[4] Having said that, there is the technical requirement that the domain for the stock in the real and risk-neutral worlds are the same. This is not the case for a diffusive equation having infinite speed of propagation and a wave equation with only a finite speed.

65.5 SUMMARY

The governing equation for the distribution of a serially autocorrelated stock may be quite interesting, but it seems to have minimal impact on the prices of derivatives.

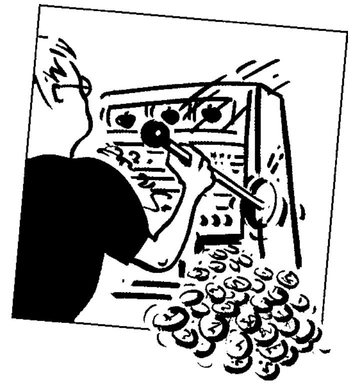
Finally, here's an exercise for you. How would you value and hedge an exotic derivative that pays off the realized serial autocorrelation in a stock price? Or serial autocovariance? You would still price these in a risk-neutral framework, by Monte Carlo simulations, perhaps. These are interesting contracts whose theoretical value would always be zero[5] but would nevertheless be potentially valuable to those with views on serial autocorrelation.

FURTHER READING

- There is little in the literature on pricing derivatives in the presence of serial autocorrelation. But see Bouchaud, Potters & Cornalba (2002) for a detailed analysis of several such models, having correlation in Gaussian and non-Gaussian processes.

[5] Under some assumptions about interest rates, volatility, frequency of sampling and exact formula for serial autocorrelation.

CHAPTER 66
asset allocation in continuous time



In this Chapter...

- the wealth process
- optimizing over asset holdings and consumption
- the Bellman equation
- the effect of transaction costs

66.1 INTRODUCTION

In Chapter 18 we looked at ways in which to invest optimally assuming that we cannot change our minds later; this is the one-period investment. In practice, we can change our minds as often as we want according to the way that the future unveils itself. In this chapter we examine models for continuous-time investment and portfolio rebalancing. Most of this is taken from Merton's work. One of the restrictions on rebalancing our portfolio is the cost of transacting. At the end of the chapter we take a brief peek at transaction costs.

66.2 ONE RISK-FREE AND ONE RISKY ASSET

To get the ball rolling, let's first look at investing in just two assets, one risk-free earning a sure rate of r and the other risky following the random walk

$$dS = \mu S\,dt + \sigma S\,dX. \tag{66.1}$$

We are going to invest a fraction $w(t)$ of our wealth in the risky asset and the remaining fraction $1 - w$ in the risk-free. We will also choose to consume or spend, spend, spend, at a rate of $C(t)$. Our aim is to decide how to split our money between the two assets and how much to consume; that is we must decide on the best $w(t)$ and $C(t)$. Obviously, we will have to optimize some problem as we'll shortly see.

66.2.1 The Wealth Process

How does our wealth evolve as the asset prices change and we consume? Can we find a stochastic differential equation for the wealth W?

From one time step to the next our wealth changes for three reasons: The asset S changes randomly; We consume; We earn interest on our cash. The fraction w only changes at the end of the time step, not during (it's rather like the Δ in the Black–Scholes analysis).

Let's build up the stochastic differential equation for W in stages.

The easiest bit is the consumption. Wealth changes by the amount we consume:

$$dW = -C(t)\,dt + \cdots.$$

The $\cdots$ represent all the terms to come.

We get interest on the cash. The cash we hold is $(1-w)W$ and so the interest received is $r(1-w)W\,dt$:

$$dW = -C(t)\,dt + r(1-w)W\,dt + \cdots.$$

The final part of the stochastic differential equation is due to the change in the risky asset. Since the asset price evolves according to (66.1) and we hold a proportion w of our wealth in this asset, there is a change in our wealth of $w\,(\mu W\,dt + \sigma W\,dX)$. Note that the asset price itself does not explicitly appear in this expression. This is because the wealth held in this asset is directly proportional to the asset value. The change in this contribution to the wealth is therefore proportional to dS, and since the asset follows a lognormal random walk where drift and randomness scale with S, the drift and randomness of our wealth (held in the risky asset) must also scale with the wealth.[1] The conclusion is

$$dW = -C(t)\,dt + r(1-w)W\,dt + w\,(\mu W\,dt + \sigma W\,dX)\,.$$

On rearranging, this becomes

$$dW = ((w(\mu - r) + r)\,W - C)\,dt + w\sigma W\,dX. \tag{66.2}$$

This is our wealth process.

66.2.2 Maximizing Expected Utility

Now we have to decide how much to spend, $C(t)$, and how to allocate our wealth between the two assets, $w(t)$. One way of doing this (and there are others, of course) is to maximize our expected utility. I'll write down the equation and explain the various terms.

$$\max_{C,w} E|_0\left[\int_0^T e^{-\rho\tau}U(C(\tau))\,d\tau + B(W)\right].$$

- $\max_{C,w}$: Maximize something by choosing $C(t)$ and $w(t)$ optimally.
- $E|_0[\cdots]$: We are going to maximize an expectation, using information available to us at time $t=0$.

[1] This wouldn't be the case if we didn't have a scalable, lognormal random walk for S.

- $\int_0^T e^{-\rho\tau} U(C(\tau))\,d\tau$: The expectation has two parts. The first is the utility of our consumption $U(C(t))$ over our time horizon, from $t = 0$ to T. Consumption in the present is more valuable to us than consumption in the future, so we discount the utility at a rate ρ, which is the investor's choice and measures his impatience.[2]
- $B(W)$: The second part of the expectation is the **bequest function**. If we have wealth left at the end of the time horizon, what is it worth? If we die at time T, how much is it worth to us to pass on what's left of our wealth to our children?

66.2.3 Stochastic Control and the Bellman Equation

The neatest way to set up the problem for solution is to introduce a new function

$$J(W,t) = \max_{C,w} E|_t \left[\int_t^T e^{-\rho\tau} U(C(\tau))\,d\tau + B(W) \right], \tag{66.3}$$

called the derived utility of wealth. So we want to maximize

$$J(W_0, 0),$$

where W_0 is the starting wealth.

Expression (66.3) can be written as

$$\begin{aligned} J(W,t) &= \max_{C,w} E|_t \left[\int_t^{t'} e^{-\rho\tau} U(C(\tau))\,d\tau + \int_{t'}^T e^{-\rho\tau} U(C(\tau))\,d\tau + B(W) \right] \\ &= \max_{C,w} E|_t \left[\int_t^{t'} e^{-\rho\tau} U(C(\tau))\,d\tau + J(W', t') \right]. \end{aligned} \tag{66.4}$$

Here W' is our (unknown) wealth at time $t' > t$.

We are now in a position to apply some stochastic calculus, by letting t' be a small increment greater than t, i.e. $t' = t + dt$. Equation (66.4) becomes

$$\begin{aligned} J(W,t) &= \max_{C,w} E|_t \left[\int_t^{t+dt} e^{-\rho\tau} U(C(\tau))\,d\tau + J(W + dW, t + dt) \right] \\ &= \max_{C,w} E|_t \left[e^{-\rho t} U(C(t))\,dt + J(W,t) + \frac{\partial J}{\partial t}\,dt + \frac{\partial J}{\partial W}\,dW + \tfrac{1}{2} w^2 \sigma^2 W^2 \frac{\partial^2 J}{\partial W^2}\,dt \right]. \end{aligned}$$

After subtracting $J(W,t)$ from both sides, taking expectations and dividing through by dt we get

$$0 = \max_{C,w} \left(e^{-\rho t} U(C) + \frac{\partial J}{\partial t} + ((w(\mu - r) + r)\,W - C)\,\frac{\partial J}{\partial W} + \tfrac{1}{2} w^2 \sigma^2 W^2 \frac{\partial^2 J}{\partial W^2} \right). \tag{66.5}$$

Nearly there. All we need do now is the maximization. Let's maximize with respect to C first.

---

[2] Of course, we could have had a term of the form $U(C, t)$, i.e. a time-dependent utility function.

There are just two terms inside the large parentheses which contain C. The maximization yields the condition

$$e^{-\rho t}U'(C) = \frac{\partial J}{\partial W}. \tag{66.6}$$

Call the solution of this C^*.

Similarly, the maximization over w gives

$$w = -\frac{(\mu - r)\dfrac{\partial J}{\partial W}}{\sigma^2 W \dfrac{\partial^2 J}{\partial W^2}} = w^*. \tag{66.7}$$

(This is only true when $(\partial^2 J/\partial W^2) < 0$ otherwise there won't be a maximum.)

Between them, (66.5), (66.6) and (66.7), are three equations for three unknowns, J, C and w. One equation (66.5) is a partial differential equation for J, the other two are algebraic equations. In general these would have to be solved numerically. However, there are some special cases with explicit solutions. These are particularly interesting because of the insight they give.

To complete the formulation of the problem we must have the final condition

$$J(W, T) = B(W).$$

66.2.4 Constant Relative Risk Aversion

When the utility function is of the form

$$U(C) = \frac{C^\gamma - 1}{\gamma},$$

with $\gamma < 1$, the relative risk aversion function $-U''C/U'$ is constant. Equations (66.5)–(66.7) become

$$\frac{\partial J}{\partial t} - \frac{(\mu - r)^2}{2\sigma^2}\frac{\left(\dfrac{\partial J}{\partial W}\right)^2}{\dfrac{\partial^2 J}{\partial W^2}} + rW\frac{\partial J}{\partial W} + \frac{1-\gamma}{\gamma}e^{-\frac{\rho t}{1-\gamma}}\left(\frac{\partial J}{\partial W}\right)^{(\gamma/1-\gamma)} = 0$$

with

$$w^* = -\frac{(\mu - r)\dfrac{\partial J}{\partial W}}{\sigma^2 W \dfrac{\partial^2 J}{\partial W^2}}$$

and

$$C* = \left(e^{\rho t}\frac{\partial J}{\partial W}\right)^{(1/\gamma - 1)}.$$

The first equation must be solved subject to the final condition being the bequest function. Let's consider the simplest case where the bequest is identically zero.[3]

[3] The 'you can't take it with you' condition.

The equation for J is a highly non-linear partial differential equation and as such it would be a miracle if we could find an explicit solution. However, as seems to happen a lot in finance, there is a miracle and lo and behold ...

$$J(W,t) = \frac{e^{-\rho t}}{\gamma}\left(\frac{1-e^{-\nu(T-t)}}{\nu}\right)^{1-\gamma} W^{\gamma}$$

where

$$\nu = \frac{\alpha}{1-\gamma} \quad \text{and} \quad \alpha = \rho - \gamma\left(\frac{\mu - r}{2\sigma^2(1-\gamma)} + r\right).$$

It follows that

$$C^* = \frac{\nu}{1-e^{-\nu(T-t)}} W$$

and

$$w^* = \frac{\mu - r}{\sigma^2(1-\gamma)}.$$

The financial interpretation of the last two equations are that consumption is proportional to wealth (with a weak time dependence) and that a constant fraction of one's wealth should remain in the risky asset.

66.2.5 Constant Absolute Risk Aversion

The same problem can also be solved when the utility function is of CARA type,

$$U(C) = -\frac{1}{\eta} e^{-\eta C}.$$

The equations become

$$\frac{\partial J}{\partial t} - \frac{(\mu - r)^2}{2\sigma^2} \frac{\left(\frac{\partial J}{\partial W}\right)^2}{\frac{\partial^2 J}{\partial W^2}} + \frac{\partial J}{\partial W}\left(rW - \frac{1}{\eta} + \frac{1}{\eta}\log\left(\frac{\partial J}{\partial W}\right)\right) - \rho J = 0$$

with again

$$w^* = -\frac{(\mu - r)\frac{\partial J}{\partial W}}{\sigma^2 W \frac{\partial^2 J}{\partial W^2}}$$

but now

$$C^* = -\frac{1}{\eta}\log\left(\frac{\partial J}{\partial W}\right).$$

Look for a solution of the form

$$J(W,t) = f(t)e^{\alpha W}$$

and you will find that

$$\alpha = -r\eta$$

and

$$\dot{f} + \left(-\frac{(\mu - r)^2}{2\sigma^2} + r - r\log(-r\eta f) - \rho\right) f = 0,$$

where $\dot{}$ denotes differentiation with respect to t.

This can be written as

$$\dot{f} + (\beta - r\log(-f))\, f = 0,$$

where

$$\beta = -\frac{(\mu - r)^2}{2\sigma^2} + r - r\log(r\eta) - \rho.$$

This doesn't have a non-zero solution for zero bequest; however, the steady-state solution is

$$f = -e^{\beta/r} = f_\infty.$$

Consumption is then

$$C^* = -\frac{1}{\eta}\log(-r\eta f_\infty) + rW$$

and

$$w^* = \frac{\mu - r}{\eta r\sigma^2 W}.$$

So consumption contains a part that is independent of wealth. The fraction of wealth held in the risky asset decreases as wealth increases.

66.3 MANY ASSETS

Suppose that we have N risky assets governed by

$$dS_i = \mu_i S_i\, dt + \sigma_i S_i\, dX_i$$

with correlation of ρ_{ij} between assets i and j. We will hold a proportion w_i in each of the N assets and the remainder, $1 - \sum_{i=1}^{N}$, in the risk-free asset.

The wealth process is surprisingly simple considering the apparent high dimensionality of the problem:

$$dW = \left(\left(r + \sum_{i=1}^{N}(\mu_i - r)w_i\right) W - C\right) dt + W\sum_{i=1}^{N} w_i\sigma_i\, dX_i.$$

The random component of this can be written as

$$\sigma\, dX$$

where

$$\sigma^2 = \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}$$

and where $\boldsymbol{\Sigma}$ is the covariance matrix and $\boldsymbol{w}$ is the column vector with ith entry w_i.

The corresponding Bellman equation is now

$$\max_{C,w}\left(e^{-\rho t}U(C) + \frac{\partial J}{\partial t} + \left(\left(r + \sum_{i=1}^{N}(\mu_i - r)w_i\right)W - C\right)\frac{\partial J}{\partial W} + \tfrac{1}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}\,W^2\frac{\partial^2 J}{\partial W^2}\right) = 0.$$

Optimizing over consumption yields

$$e^{-\rho t}U'(C) = \frac{\partial J}{\partial W}.$$

Optimizing over the w_is yields

$$\sum_{j=1}^{N} w_j\sigma_{ij} = -\frac{\dfrac{\partial J}{\partial W}}{W\dfrac{\partial^2 J}{\partial W^2}}(\mu_i - r)$$

where σ_{ij} are the entries in $\boldsymbol{\Sigma}$. These are N linear simultaneous equations for the w_i. The solution is

$$\boldsymbol{w} = -\frac{\dfrac{\partial J}{\partial W}}{W\dfrac{\partial^2 J}{\partial W^2}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1}),$$

where $\boldsymbol{\Sigma}^{-1}$ is the inverse of $\boldsymbol{\Sigma}$, and $\boldsymbol{\mu}$ and $\mathbf{1}$ are column vectors, the latter consisting of one in each entry.

For special utility functions these can be solved explicitly.

66.4 MAXIMIZING LONG-TERM GROWTH

There is a very special portfolio, one in which there is no consumption so that all growth is reinvested back in the portfolio. Clearly, in the above there is no consumption term, $C = 0$, so the utility of consumption is irrelevant. But what about the bequest function? So far this has been something that we haven't paid much attention to. This plays a key role when we want to *maximize the long-term growth*. Remember the Blackjack example from Chapter 18 while reading the following.

The average growth rate of the portfolio from time t until T is

$$\frac{1}{T-t}\log\left(\frac{W(T)}{W(t)}\right).$$

The **growth-optimum portfolio** is thus one for which we maximize

$$E|_t\left[\log W(T)\right].$$

This can be achieved in the above framework by solving the zero-consumption equation

$$\max_w \left(\frac{\partial J}{\partial t} + ((w(\mu - r) + r)\, W) \frac{\partial J}{\partial W} + \tfrac{1}{2} w^2 \sigma^2 W^2 \frac{\partial^2 J}{\partial W^2} \right) = 0.$$

subject to the final condition

$$J(W) = \log W.$$

This is equivalent to solving a logarithmic bequest function problem.

Again we have the optimal weight

$$w^* = -\frac{(\mu - r)\dfrac{\partial J}{\partial W}}{\sigma^2 W \dfrac{\partial^2 J}{\partial W^2}}.$$

The equation for J is then

$$\frac{\partial J}{\partial t} + rW\frac{\partial J}{\partial W} - \frac{1}{2}\frac{(\mu - r)^2 \left(\dfrac{\partial J}{\partial W}\right)^2}{\sigma^2 \dfrac{\partial^2 J}{\partial W^2}} = 0.$$

The solution of this problem is

$$J(W, t) = \log W + \left(r + \frac{(\mu - r)^2}{2\sigma^2} \right)(T - t).$$

The optimal weight is therefore

$$w^* = \frac{(\mu - r)}{\sigma^2}$$

which is independent of current wealth and time. It's the same as the Kelly criterion we saw with reference to Blackjack, except that in Blackjack we didn't worry about the risk-free rate of interest that our money could have earned while we were sitting at the tables.

The extension to N-risky assets is straightforward yielding

$$w^* = \mathbf{\Sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1}).$$

66.5 A BRIEF LOOK AT TRANSACTION COSTS

Incorporating transaction costs into any financial model is usually far from simple. This is especially true of asset allocation problems where the dimensionality is potentially very large. There is one rather elegant exception to this rule and this is the case of long-term growth maximization where each rebalancing trade results in a cost, in this case proportional to the value of the entire portfolio. This is a rather unrealistic transaction cost structure, since normally the cost would depend only on the size of the trade, but it allows significant progress to be made and insight to be gained. If the portfolio has value W before the rebalancing, it will have value $(1 - \epsilon)W$ after. The cost term ϵ will be small, and we shall exploit this in the analysis.

Recall from the previous section that the ideal portfolio is weighted such that

$$\boldsymbol{w}^{\mathrm{M}} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1}).$$

In other words, *the proportion of wealth in each of the $n+1$ assets should be kept constant.* The amount held in each asset depends on the covariance matrix, the expected returns and the risk-free rate. This is the classical Merton solution. Think of this as a point in $w_1, \ldots, w_N$ space.

Clearly, as the asset prices change randomly, so the proportion $\boldsymbol{w}$ quickly moves away from the optimal one $\boldsymbol{w}^{\mathrm{M}}$. Ideally, the portfolio would be continually rebalanced, but in the presence of costs this would be suicidal. What now is the optimal rebalancing strategy?

The math is rather complicated to go into here, so references are given below. One must solve an N-dimensional free-boundary problem. This would be almost impossible to do either analytically or numerically, but we are saved by the smallness of the parameter ϵ. We can perform an asymptotic analysis as $\epsilon \to 0$.

I'm going to cut right to the chase and focus on the action in the neighborhood of the optimal Merton point, $\boldsymbol{w}^{\mathrm{M}}$. It turns out that the optimal strategy in the presence of costs is (approximately) to rebalance *back to the Merton point* as soon as the current portfolio weighting moves to the edge of a very special N-dimensional ellipsoid. This ellipsoid, which is centered on the Merton point, is defined by

$$\boldsymbol{w} = \boldsymbol{w}^{\mathrm{M}} + \epsilon^{\frac{1}{4}}\overline{\boldsymbol{w}}$$

where

$$\overline{\boldsymbol{w}}^T \mathbf{M} \overline{\boldsymbol{w}} = 1$$

for some positive definite symmetric matrix $\mathbf{M}$. This matrix is given by the solution of

$$8\mathbf{MAM} + 4\mathrm{Tr}(\mathbf{AM})\mathbf{M} = \boldsymbol{\Sigma}. \tag{66.8}$$

The matrix $\mathbf{A}$ has entries

$$\mathrm{A}_{ij} = w_i^{\mathrm{M}} w_j^{\mathrm{M}} (\mathbf{e}_i^T - \boldsymbol{w}^T)\boldsymbol{\Sigma}(\mathbf{e}_i - \boldsymbol{w})$$

where $\mathbf{e}_i$ is the column vector with a one in the ith row and zero everywhere else.

Neat result, huh?

The matrix $\mathbf{M}$ can be easily found from (66.8) by an iterative procedure.

In Figure 66.1 is shown the results of a two-dimensional problem. In this figure the axes are the two weights w_1 and w_2. Data are $r = 0.07$, $\mu_1 = 0.13$, $\mu_2 = 0.15$, the covariance matrix has diagonal elements 0.1 and 0.17 with off-diagonal elements both 0.07. The cost parameter is $\epsilon = 0.001$. The oval shape is the outline of the exact no-trade region (calculated numerically by Morton & Pliska, 1995) and the two straight lines are the axes of the asymptotically derived ellipse. The center of the ellipse is the Merton point. Note that in more than two dimensions the approximate ellipse is still easy to find whereas the numerical solution of the full problem becomes prohibitively time consuming.

Since we know the boundary of the no-trade region we can ask questions about the expected first exit time, the expected time between trades. Again, I won't go into the details, but with realistic parameters the time between rebalances turns out to be of the order of a year or two.

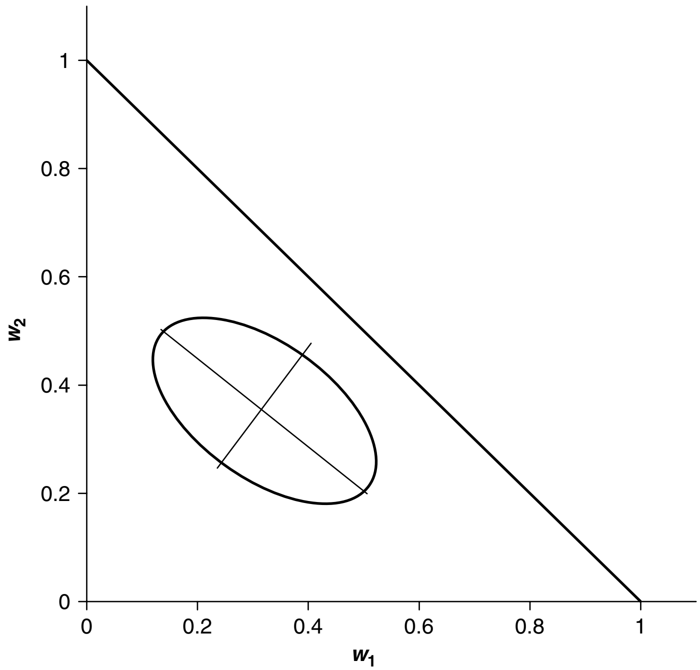


Figure 66.1 A two-dimensional example. See text for details.

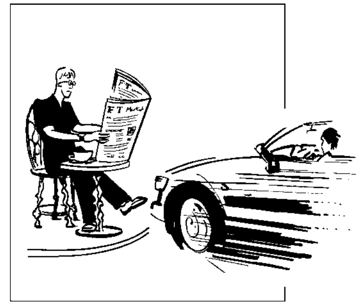
This is much longer than happens in practice. More realistically, rebalances are more likely because of perceived changes in parameter values, such as drift and volatility, than due to restrictions imposed by frictions such as costs.

66.6 SUMMARY

This chapter addressed the issue of how to invest optimally in assets in continuous time. The theory has been around for a long time but is generally not as well known, or used, as the one-period investment models, MPT and CAPM.

FURTHER READING

- Merton's (1992) book is the classic reference for this subject.
- Transaction costs are dealt with in detail by Morton & Pliska (1995), Atkinson & Wilmott (1995) and Atkinson, Pliska & Wilmott (1997).



CHAPTER 67

asset allocation under threat of a crash

In this Chapter...

- how to allocate your money to an investment when you are worried about a market crash

67.1 INTRODUCTION

The following by Ralf Korn[1] and myself was originally published in the *International Journal of Theoretical and Applied Finance* **5** 171–188 (2002).

We consider the determination of optimal portfolios under the threat of a crash. Our main assumption is that upper bounds for both the crash size and the number of crashes occurring before the time horizon are given. We make no probabilistic assumption on the crash size or the crash time distribution. The optimal strategies in the presence of a crash possibility are characterized by a balance problem between insurance against the crash and good performance in the crash-free situation. Explicit solutions for the log-utility case are given. Our main finding is that constant portfolios are no longer optimal ones.

It is well known that the classical lognormal stock/Black–Scholes model is not able to explain large jumps in stock prices appearing in real-world security markets. In particular, it does not contain the possibility of a crash of the stock prices. Although these crashes are rare events, they do occur in real life; the October 1987 crash being the largest in recent memory. Therefore, there is a long tradition of modeling jumps in stock prices. A seemingly obvious candidate of a class of suitable stochastic processes is that of jump-diffusion processes (see Merton (1976) for an early reference). However, jump-diffusion approaches only lead to strategies that hedge a crash situation in the mean which is no real protection against the consequences of a jump at all. In particular, an investor following such a strategy will suffer large losses during a crash. As a contrast, by implementing the strategy that we will propose an investor need not be in fear of a crash as our strategy gives him full protection in a sense that will be made more precise in the following sections.

An alternative for modeling crashes is given in Hua and Wilmott (1997). There, the underlying assumptions are that both the maximal number of crashes in a given time interval and the biggest

[1] Ralf Korn, Fachbereich Mathematik, Universität Kaiserslautern, Germany korn@mathematik.uni-kl.de

possible size of the crashes are known in advance. Otherwise, we do not assume we know the exact time of the crashes (or even if they occur at all) and their exact sizes. In this model, a crash is characterized by the fact that all random securities traded on the market become highly correlated at the crash time and all fall together at the same time. We will base our considerations on this model.

In the current work, the modeling of crashes is not our main concern. As the title indicates, we will look at optimal portfolios when the threat of a crash before the time horizon is given. Our main aim is to show that suitable investment in stocks still can be more profitable than playing safe and investing all the funds in the riskless bond. We will arrive at a typical balance problem between obtaining good worst-case bounds for the case of a crash and also a reasonable performance should no crash occur at all. Obviously, the pure bond strategy yields excellent lower bounds but a poor performance if there is no crash. On the other hand, the portfolio processes which are optimal in the classical lognormal stock model are far too risky under the threat of a crash.

In Section 67.2 we will start by presenting the most basic case, that of a single risky security when at most one crash can occur.

Section 67.3 is devoted to the multi-asset case in the log-utility setting while Section 67.4 presents some refinements and extensions.

67.2 OPTIMAL PORTFOLIOS UNDER THE THREAT OF A CRASH: THE SINGLE STOCK CASE

We consider a security market consisting of a riskless bond and a single risky security with prices given by

$$dP_0(t) = P_0(t)r\,dt, \quad P_0(0) = 1$$
$$dP_1(t) = P_1(t)\,(\mu\,dt + \sigma\,dW(t)) \quad P_1(0) = p_1$$

where the market coefficients are constants satisfying $\mu > r$ and $\sigma \neq 0$. The ‘crash feature’ of the stock price is now given by the possibility of a sudden fall of the stock price before the time horizon T. We (temporarily) only assume that this could happen at most once between now ($t = 0$) and the time horizon T. Our knowledge of the size of the crash is subsumed in the fact that the relative, i.e. percentage, fall in the stock price lies in the interval $[0, k^*]$ where the constant $0 < k^* < 1$ (‘the worst possible crash’) is given. Otherwise we do not make any assumption on a probabilistic distribution of the exact crash time and the exact crash size. So, before and after the crash, we are in the classical lognormal stock world. (However, we will see later that our work is not limited to this situation.)

Our main aim is to find optimal portfolio processes in the following sense (for the rigorous definition of a portfolio process and the notion of a utility function as below we refer to Korn (1997)):

Let $U(x)$ be a utility function, $\pi(t)$ a portfolio process (i.e. the process of the fraction of the total wealth invested in the stock), and let $X^{\pi}(t)$ be the wealth process corresponding to the portfolio process $\pi(t)$ and the initial wealth of $x > 0$. We then look for the maximal worst-case bound for the expected utility of final wealth,

$$\sup_{\pi(.)\in A(x)} E\left(\inf_{0\le t\le T, 0\le k\le k^*} U\left(X^{\pi}(T)\right)\right)$$

where the final wealth $X^\pi(T)$ in the case of a crash of size k at time t is given by

$$X^\pi(T) = (1 - \pi(t)k)\,\tilde{X}^\pi(T)$$

with $\tilde{X}^\pi(t)$ the usual wealth process corresponding to π if there is no crash. More precisely $\tilde{X}^\pi(t)$ is given as the unique solution to the stochastic differential equation

$$d\tilde{X}^\pi(t) = \tilde{X}^\pi(t)((r(t) + \pi(t)(\mu(t) - r(t)))\,dt + \pi(t)\sigma(t)\,dW(t))$$
$$\tilde{X}^\pi(0) = x.$$

The above representation of $X^\pi(T)$ clearly shows that there are two different competing effects. Of course, for obtaining a high utility from the final wealth (in the case with or without crash) it is necessary to follow a sufficiently high portfolio process (always assumed that the mean rate of stock return μ exceeds the riskless rate of r). On the other hand, a high portfolio process at the time of the crash leads to a significant decrease of the total wealth. In particular, a portfolio process exceeding $1/k^*$ bears the risk of bankruptcy if a crash occurs.

Before solving the above worst-case problem, we will highlight its main features and in particular support the above remarks by looking at the following two extreme strategies in the case of the logarithmic utility function.

Two extreme strategies

i) $\pi(t) \equiv 0$ before the crash: 'Playing safe'. For this strategy (the pure bond investment) the worst-case scenario is that no crash occurs at all. Why is this so? Of course, a crash would do this strategy no harm, but it would give it the possibility to switch to the optimal portfolio process in the log-utility case, $\pi(t) \equiv \pi* := (\mu - r)/\sigma^2$ afterwards. In the no-crash scenario the pure bond strategy would lead to the following worst-case bound of

$$WCB_0 = E\left(\log\left(X^0(T)\right)\right) = \log(x) + rT.$$

ii) $\pi(t) \equiv \pi* := b - r/\sigma^2$ before the crash: 'Optimal investment in the crash-free world'. Of course, a crash would lead to losses in this case as we have a big stock investment. The worst-case scenario is here given by a crash of maximum size k^* (independent of time), leading to the following worst-case bound of

$$WCB_{\pi^*} = E\left(\log\left(X^{\pi^*}(T)\right)\right) = \log(x) + rT + \tfrac{1}{2}\left(\frac{b-r}{\sigma}\right)^2 T + \log\left(1 - \pi^* k^*\right).$$

Insights

- Which of the above strategies yields the better worst-case bound (mainly) depends on time to maturity.
- As a consequence of the form of the above worst-case bounds one can easily infer that a constant portfolio process cannot be the optimal one (in contrast to the crash-free setting).
- Strategy i) takes too few risks to be good if no crash occurs while strategy ii) is too risky to perform well if a crash occurs, thus, an optimal strategy should balance this out.

To make the above insights more precise on a technical level, let $v_1(t, x)$ be the value function of the problem (before $\pi(t)$ has been chosen at time t) if we know that on $[t, T]$

at most one crash can occur. Further, let $v_0(t, x)$ be the value function corresponding to the optimization problem in the usual, crash-free Black–Scholes setting. It can also be interpreted (and we will do so) as the value function of the above problem after the crash has already happened.

Important remarks

1. If we compare two different investment strategies with respect to their worst-case bound then we do not compare them pathwise ('scenario-wise'). We look separately at the worst-case for both strategies which then yields the worst-case bound. So typically two different strategies have two different worst-case scenarios (even if they might have the same worst-case bound).
2. As we have assumed $\mu > r$, we do not have to consider portfolio processes $\pi(t)$ that can attain negative values as long as the utility function is increasing in x (which we will always assume when not stated otherwise). The reason for this is that the corresponding portfolio process $\pi(t)^+$ would yield a higher expected final utility if no crash occurs at all and that the worst-case bound given a crash occurs would only be better than that of $\pi(t)^+$ if $\pi(t)$ would be strictly negative for all t. But then the worst-case scenario would be the absence of a crash. And of course, then the pure bond strategy would yield a better worst-case bound than $\pi(t)$.

Proposition 1

a) An optimal portfolio process $\pi_{\text{opt}}(t)$ for the worst-case problem has to satisfy

$$v_0\left(t, x\left(1 - \pi_{\text{opt}}(t)k^*\right)\right) \geq U\left(xe^{r(T-t)}\right),$$
$$\pi_{\text{opt}}(T) = 0.$$

b) We have

$$v_1(t, x) \geq U\left(xe^{r(T-t)}\right).$$

c) In the case of log-utility $U(x) = \log(x)$ we must have

$$E\left(\int_t^T \left(\pi_{\text{opt}}(s)(\mu - r) - \tfrac{1}{2}\pi_{\text{opt}}(s)^2\sigma^2\right) ds\right) \geq 0.$$

Proof

a) Both assertions of a) follow from comparison of the optimal portfolio process with the pure bond investment. The first inequality says that the optimal portfolio (i.e. the one delivering the best worst-case bound) should yield a final expected utility at least as big as the one obtained by pure bond investment if an immediate crash (of highest size) happens. The final condition $\pi_{\text{opt}}(T) = 0$ is implied by the fact that a crash at the time horizon should have no impact. This requirement also follows from the comparison of the optimal strategy with the pure bond one.

b) is a direct consequence of a) and the fact that the best bound should always be at least as big as the pure bond bound.

c) To see assertion c), consider the effect of no crash in the log-utility case. Then the expected final utility of the optimal strategy should be no worse than that of the pure bond investment, i.e. the assertion is implied by the inequality

$$E\left(\log(x)+r\,(T-t)+\int_t^T\left(\pi(s)\,(\mu-r)-\tfrac{1}{2}\pi(s)^2\sigma^2\right)ds\right)\geq E\left(\log(x)+r\,(T-t)\right)$$

As the above assertions were all necessary but not sufficient conditions for the existence of an optimal strategy, we will at least in the log-utility case show that there exist explicit examples of strategies that perform better than the pure bond investment even under the threat of a crash. The key for constructing such a strategy lies in (the proof) of the above assertions a) and b).

Corollary 2

Assume $U(x)=\log(x)$. Then we have:

a) There exist strategies $\pi(.)$ with a strictly higher worst-case bound than the pure bond strategy.

b) There exists a strategy $\hat{\pi}(.)$ such that the corresponding expected log-utility after an immediate crash equals the expected log-utility given no crash occurs if there exists a solution $\hat{\pi}(.)$ to the differential equation

$$\dot{\pi}(t)=\frac{1}{k^*}\left(1-\pi(t)k^*\right)\left(\pi(t)\,(\mu-r)-\tfrac{1}{2}\left(\pi(t)^2\sigma^2+\left(\frac{\mu-r}{\sigma}\right)^2\right)\right)$$

with

$$\pi(T)=0.$$

$$0\leq\hat{\pi}(.)<\frac{1}{k^*}$$

c) If there exists an optimal portfolio process for the worst-case problem then it is a non-constant one (in contrast to the problem without the threat of a crash).

Proof

a) Let

$$a(t):=\frac{1}{k^*}\left[1-e^{-\frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2(T-t)}\right]\wedge\frac{\mu-r}{\sigma^2}.$$

We then choose $\pi(t):=\dfrac{a(t)}{2}$ and thus obtain

$$E\left(\int_t^T\left(\pi(s)\,(\mu-r)-\tfrac{1}{2}\pi(s)^2\sigma^2\right)ds\right)>0$$

(due to $\mu>r$) while the explicit form of $v_0(t,x)$,

$$v_0(t,x)=\log(x)+\left(r+\tfrac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2\right)(T-t)\,,$$

yields

$$v_0\left(t, x\left(1-\pi(t)k^*\right)\right) > U\left(xe^{r(T-t)}\right).$$

Hence, if there is no crash at all $\pi(.)$ yields a higher final log-utility than the pure bond strategy and in the case of an immediate crash (at all time instants) the worst-case bound $\pi(.)$ still exceeds the one of the pure bond strategy.

b) The above requirement on $\hat{\pi}(.)$ translates to the integral equation

$$\log\left(1-\pi(t)k^*\right) = E\left(\int_t^T \left(\pi(s)(\mu-r) - \tfrac{1}{2}\pi(s)^2\sigma^2\right)ds\right) - \tfrac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2(T-t)$$

If now there exists a solution $\hat{\pi}(.)$ to the differential equation

$$\dot{\pi}(t) = \frac{1}{k^*}\left(1-\pi(t)k^*\right)\left(\pi(t)(\mu-r) - \tfrac{1}{2}\left(\pi(t)^2\sigma^2 + \left(\frac{\mu-r}{\sigma}\right)^2\right)\right)$$

$$\pi(T) = 0$$

(also satisfying the additional side constraint in b)), then the deterministic strategy $\hat{\pi}(.)$ obviously solves the above integral equation.

c) is a direct consequence of Proposition 1a) and part a) of the corollary proved above.

Remark

If we draw the right conclusions out of the above proposition and corollary then the solution of our problem is nearly obvious. Again, look at the two extreme strategies that we considered at the beginning of this section. The pure bond strategy is too safe. As part a) above shows one can in fact increase the risk by investing in the stock and still be better off than with the pure bond investment. So in this case the risky position can be increased so long until the crash scenario and the no crash scenario both lead to the same worst-case bound. The situation is similar for the optimal constant portfolio of the crash-free model. Here, the risk of a high stock position should be reduced until again the crash scenario and the no crash scenario both lead to the same worst-case bound. In this sense there is a balance problem between total hedging against immediate crashes and taking full risk for obtaining a high expected final log-utility (and hoping that no crash will occur) which should be taken into account to solve the worst-case problem. As a consequence of these considerations we now look at the strategy for which the worst-case bound is attained for both an immediate crash and by the final expected log-utility if no crash occurs at all.

Theorem 3 'Best portfolio under the threat of a crash'

In the log-utility case, the portfolio process $\hat{\pi}(.)$ such that the corresponding expected log-utility after an immediate crash equals the expected log-utility given no crash occurs, which is given as the solution $\hat{\pi}(.)$ of the differential equation

$$\dot{\pi}(t) = \frac{1}{k^*}\left(1 - \pi(t)k^*\right)\left(\pi(t)(\mu - r) - \tfrac{1}{2}\left(\pi(t)^2\sigma^2 + \left(\frac{\mu - r}{\sigma}\right)^2\right)\right)$$

$$\pi(T) = 0$$

and satisfies

$$0 \le \hat{\pi}(.) < \frac{1}{k^*},$$

is an optimal portfolio process for the worst-case problem.

Proof

i) By the explicit form of

$$E\left(\log\left(\tilde{X}^{\pi}(T)\right)\right) = \log(x) + rT + E\left(\int_0^T \left(\pi(t)(\mu - r) - \tfrac{1}{2}\pi(t)^2\sigma^2\right) dt\right),$$

a portfolio process $\pi(.)$ with a higher final expected log-utility than $\hat{\pi}(.)$ if no crash occurs at all has to satisfy

$$E(\pi(t)) > \hat{\pi}(t) \tag{67.1}$$

for some t. Let $\hat{v}(t, x)$ denote the expected log-utility of terminal wealth from following the portfolio process $\hat{\pi}(.)$. Then, due to

$$v_0\left(t, \hat{X}(t)\left(1 - \hat{\pi}(t)k^*\right)\right) = \hat{v}\left(t, \hat{X}(t)\right) = \hat{v}(0, x) + \int_0^t \hat{v}_x\left(s, \hat{X}(s)\right)\hat{\pi}(s)\sigma\, dW(s)$$

$$+ \int_0^t \left(\hat{v}_t\left(s, \hat{X}(s)\right) + \hat{v}_x\left(s, \hat{X}(s)\right)(r + \hat{\pi}(s)(\mu - r)) + \tfrac{1}{2}\hat{v}_{xx}\left(s, \hat{X}(s)\right)\hat{\pi}(s)^2\sigma^2\right) ds$$

we have exactly the same worst-case bounds for all possible future times of the crash,

$$E\left(v_0\left(t, \hat{X}(t)\left(1 - \hat{\pi}(t)k^*\right)\right)\right) = E\left(\hat{v}\left(t, \hat{X}(t)\right)\right) = \hat{v}(0, x) = E\left(\hat{v}\left(T, \hat{X}(T)\right)\right)$$

for the portfolio process $\hat{\pi}(.)$. If we would now have $\pi(0) > \hat{\pi}(0)$ then due to the construction of $\hat{\pi}(.)$ the strategy $\pi(.)$ would have a strictly lower worst-case bound. To see this, note that an immediate crash would lead to a strictly smaller wealth than that corresponding to $\hat{\pi}(.)$. Of course if the two portfolio processes coincide at the initial time then again due to the construction of $\hat{\pi}(.)$ the worst-case bound of $\pi(.)$ cannot exceed the one for $\hat{\pi}(.)$. Thus, we may assume $\pi(0) < \hat{\pi}(0)$. But due to this assumption and to (67.1) there exists a first time $t \in [0, T]$ with

$$E(\pi(t)) > \hat{\pi}(t) \quad \text{and} \quad E\left(\log\left(X^{\pi}(t)\right)\right) \le E\left(\log\left(\hat{X}(t)\right)\right). \tag{67.2}$$

However, at exactly that time we have

$$E\left(\log\left(1-\pi(t)k^*\right)\right) \le \log\left(1-k^* E\left(\pi(t)\right)\right) < \log\left(1-\hat{\pi}(t)k^*\right) \tag{67.3}$$

which together with (67.2) and the explicit form of $v_0(t,x)$ lead to

$$E\left(v_0\left(t, X^{\pi}(t)\left(1-\pi(t)k^*\right)\right)\right) < E\left(v_0\left(t, \hat{X}(t)\left(1-\hat{\pi}(t)k^*\right)\right)\right).$$

So, again due to the construction of $\hat{\pi}(.)$ the worst-case bound of $\pi(.)$ cannot exceed the one for $\hat{\pi}(.)$. Thus, a strategy $\pi(.)$ with a higher final expected log-utility as above cannot be an optimal one for the worst-case problem.

ii) By i) a portfolio process $\pi(.)$ with a higher final expected log-utility than $\hat{\pi}(.)$ cannot have a higher worst-case bound than $\hat{\pi}(.)$. On the other hand, due to the construction of $\hat{\pi}(.)$, a portfolio process $\pi(.)$ leading to a smaller final expected log-utility than $\hat{\pi}(.)$ automatically has a smaller worst-case bound than $\hat{\pi}(.)$. Putting i) and ii) together yields the assertion of the above proposition.

Remark: 'Uniqueness of the optimal strategy'

Due to the above theorem there only exists one (deterministic) equilibrium strategy. The main reason for this is the fact that we have only one risky stock in our market model. It is exactly the one-dimensionality that allows us to conclude the relation (67.1) in the above proof. To obtain the analogous result in the multi-stock setting we have to put in more information as there we typically have more than one equilibrium strategy. So we have the additional problem of figuring out the best such one (see Section 67.3).

Example 4 log-utility ('maximizing growth rate')

We will from now on specify to the use of the log-utility function which can also be expressed as maximizing the growth rate of the wealth process. As implied by Theorem 3 above we obtain the optimal portfolio process in this setting via solving the corresponding differential equation. By separation of variables, we arrive at the following non-linear equation for $\hat{\pi}(t)$:

$$C - \frac{1}{2}\frac{\sigma^2}{k^*}t = \alpha \log\left(\frac{|\hat{\pi}(t)-\pi^*|}{|1-\hat{\pi}(t)k^*|}\right) + \frac{\beta\hat{\pi}(t)}{\hat{\pi}(t)-\pi^*}$$

with

$$C := \alpha\log\left(\left|\pi^*\right|\right) + \frac{1}{2}\frac{\sigma^2 T}{k^*}, \quad \alpha := \frac{k^*}{(1-\pi^*k^*)^2}, \quad \beta := -\frac{1}{(1-\pi^*k^*)\,\pi^*}, \quad \pi^* := \frac{\mu - r}{\sigma^2}.$$

Uniqueness and existence of the solution of the above non-linear equation for $\hat{\pi}(t)$ can always be ensured. To see this, note that for $\hat{\pi}(t) = 0$ the left-hand side is always bigger than the right-hand one. Also the derivative of the right-hand side with respect to $\hat{\pi}(t)$ is strictly positive. In the case of $\pi^* \le 1/k^*$ we have a pole at $\hat{\pi}(t) = \pi^*$ where the right hand side equals $+\infty$. Thus, there must be a unique value for $\hat{\pi}(t)$ such that the right-hand side attains the value of the left-hand one. In the case of $\pi^* > 1/k^*$ a similar argument (but now with a pole at $\hat{\pi}(t) = 1/k^*$) yields the existence and uniqueness assertion.

To highlight the behavior and the performance of the 'equilibrium strategy' $\hat{\pi}(t)$ we also compute the best constant portfolio strategy in the crash setting.

Proposition 5 'Best constant portfolio under the threat of a crash'

In the log-utility case the best constant portfolio strategy for our worst-case problem is given by

$$\tilde{\pi} = \left(\tfrac{1}{2} \left(\frac{\mu - r}{\sigma^2} + \frac{1}{k^*} \right) - \sqrt{\tfrac{1}{4} \left(\frac{\mu - r}{\sigma^2} - \frac{1}{k^*} \right)^2 + \frac{1}{\sigma^2 T}} \right)^+ .$$

In particular, for $\mu > r$ we have

$$\tilde{\pi} \stackrel{T \to \infty}{\longrightarrow} \pi^* \quad \text{in case of} \quad \pi^* \leq k^*$$

and

$$\tilde{\pi} \stackrel{T \to \infty}{\longrightarrow} \pi^* \quad \text{in case of} \quad \pi^* > k^*.$$

Remark

The above limiting results deserve a closer look: if the time horizon is big and the optimal investment in the crash free model does not lead to the possibility of a negative wealth in the crash setting then it is close to the best constant portfolio under the threat of a crash. If it bears the possibility of a negative wealth after a crash (i.e. if we are in the case of $\pi^* > k^*$) then with a growing horizon the investor approaches the highest possible risk of a portfolio, i.e. attaining a value close to k^*. More precisely, he takes the risk of big crash losses for attaining a high growth rate of his holdings.

Proof

First note that if an investor follows a constant portfolio strategy then the worst time for a crash (of course of maximum size k^*) is just before the time horizon. To see this note that after a crash the investor is able to switch to the best possible constant portfolio strategy, π^*. Thus the earlier the crash happens, the longer the investor can take advantage of investing according to π^*. Given that the crash now happens immediately before the time horizon T the expected log-utility of an investor using the constant portfolio process π is given by

$$E\left(\log\left(X^{\pi}(T)\right)\right) = \log(x) + rT + \left(\pi\,(\mu - r) - \tfrac{1}{2}\pi^2\sigma^2\right) T + \log\left(1 - \pi k^*\right).$$

Differentiating the right hand side of this expression with respect to π and setting the derivative equal to zero (note that as a function of π the right hand side is concave) yields

$$\tilde{\pi} = \tfrac{1}{2} \left(\frac{\mu - r}{\sigma^2} + \frac{1}{k^*} \right) - \sqrt{\tfrac{1}{4} \left(\frac{\mu - r}{\sigma^2} - \frac{1}{k^*} \right)^2 + \frac{1}{\sigma^2 T}}$$

as the only zero of the derivative which is smaller than $\min\left(1/k^*, \mu - r/\sigma^2\right)$. But this value can only yield the active worst-case bound if it is non negative. Otherwise, the no-crash case would deliver the worst-case bound, a case where the pure bond investment has the best worst-case bound under all non-positive portfolio strategies. But it is easy to see that we have

$$\tilde{\pi} > 0 \Leftrightarrow \mu - r > \frac{k^*}{T}.$$

Finally, the remaining convergence assertions for $T \to \infty$ follows from the explicit form of $\tilde{\pi}$.

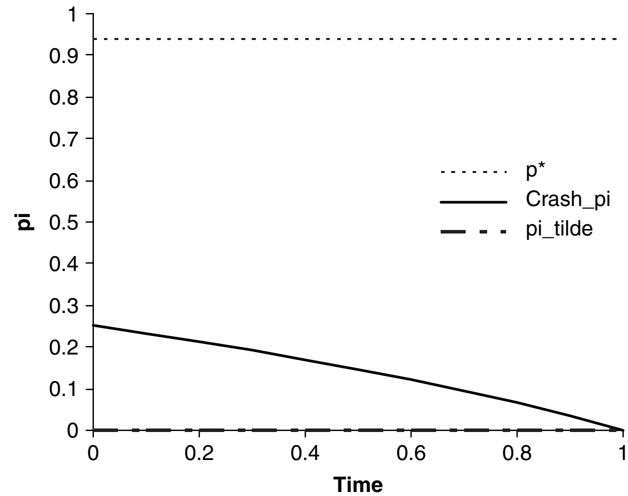


Figure 67.1 Optimal portfolios with and without crash possibility (small time horizon).

Numerical examples

For the following examples we solved the above non-linear equation for $\hat{\pi}(t)$ numerically (of course under the additional requirement of $0 \leq \hat{\pi}(t) < 1/k^*$). Figure 67.1 shows $\hat{\pi}(t)$ as a function of time for the choice of $\mu = 0.2$, $r = 0.05$, $\sigma = 0.4$, $k^* = 0.2$ and $T = 1$. Note that even at the initial time $t = 0$ the optimal portfolio process in view of a crash is considerably below the optimal portfolio π^* in the 'crash-free' standard model which is $\pi^* = 0.9375$. However, even the small values of the crash optimal portfolio process are much bigger than that of the best constant portfolio process in the crash setting which equals zero. These small values can be explained by the fact that if the time horizon is close then the crash risk dominates the possibilities of obtaining a better return via stock investment.

If, however, the time horizon is far away then the picture changes significantly. For the same data but now with $T = 10$ the resulting optimal portfolio processes are given in Figure 67.2. Now the optimal crash portfolio is much higher (at least at times much smaller than 10) and the optimal constant portfolio in the crash setting even exceeds it. The interpretation of this behavior is obvious. The longer the time to the trading horizon the more attractive it is to invest in the stock, and even a 'moderate crash' is no real threat. If however the final time is near then it is good to save the gains (i.e. reduce stock investment) as then there is not enough time to compensate the effect of a crash via an optimal stock investment afterwards.

67.3 MAXIMIZING GROWTH RATE UNDER THE THREAT OF A CRASH: *n* STOCKS

We are now considering a market that consists of one riskless bond and n stocks. The prices of the stocks are assumed to follow geometric Brownian motions in 'normal' times, i.e. they

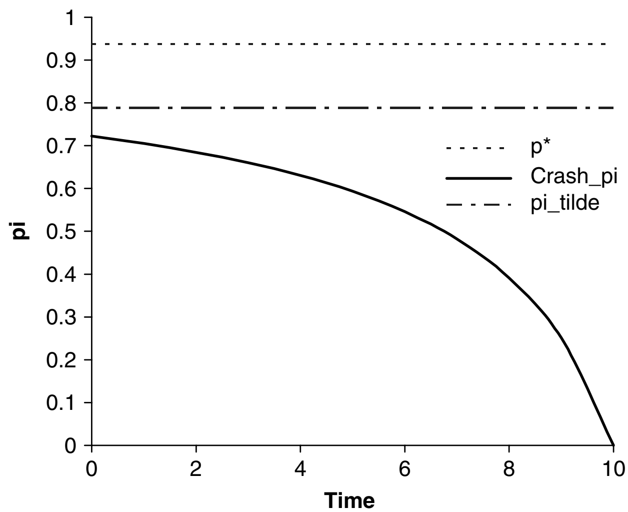


Figure 67.2 Optimal portfolios with and without crash possibility (large time horizon).

are given by

$$dP_0(t) = P_0(t)r\,dt, \quad P_0(0) = 1,$$

$$dP_i(t) = P_i(t)\left(\mu_i\,dt + \sum\nolimits_{j=1}^{n} \sigma_{ij}\,dW_j(t)\right), \quad P_i(0) = p_i, \quad i = 1, \ldots, n$$

as long as there is no crash. At the time of a crash we take on the view of Hua and Wilmott (1997) who assume that all stock prices become highly correlated and all fall at the same time as a certain index. The absolute values of all these falls are then given as suitable multiples k_i (the so-called crash coefficients) of the percentage jump of the index. As in the one-stock case we assume that there occurs at most one crash and that the crash sizes in the assets are in the intervals $[0, k_i]$. So for simplicity we assume that the jump in the index lies in the unit interval.

As in the closing part of the preceding section we here restrict ourselves to the use of the log-utility function. The main difference to the one-stock setting is that now there can exist more than one equilibrium strategy (i.e. portfolio processes with a worst-case bound which is determined simultaneously by all future time points and events). To see this note that one can obtain equilibrium strategies by simply restricting to the sub-markets made up of the bond and one arbitrary of the n stocks. However, it is then natural to conjecture that the best equilibrium strategy (i.e. the one delivering the highest worst-case bound) solves our worst-case bound portfolio problem.

Theorem 6

Assume that we are in the market setting as given above. Then the optimal portfolio process is given as the deterministic portfolio process which has the highest worst-case bound under all deterministic portfolio processes satisfying the 'equilibrium condition'

$$v_0\left(t, x\left(1 - \pi(t)'k^*\right)\right) = v_\pi\left(t, x\right) \tag{67.4}$$

Here, $v_0(t, x)$ is the value function in the crash-free setting and $v_\pi(t, x)$ denotes the expected log-utility of terminal wealth from following the portfolio process $\pi(.)$. That is, the above optimal strategy is determined as the solution of the problem

$$\sup_{\pi|[t,T]} \left(v_\pi(t, x) - v_0\left(t, x\left(1 - \pi(t)'k^*\right)\right)\right) = 0 \tag{67.5}$$

where the supremum is only taken over all such deterministic portfolio processes on $[t, T]$ that satisfy the equilibrium constraint (67.4).

Proof

To see the above claim we collect some facts:

i) Let $f(\pi) := \log\left(1 - \pi'k^*\right)$, $\pi \in \prod_{i=1}^{n} \left[0, 1/k_i^*\right)$, $\pi'k^* < 1$. Then $f(\pi)$ is concave for all such admissible vectors π. Hence, for each admissible portfolio process $\pi(t)$ we obtain

$$E\left(\log\left(1 - \pi(t)'k^*\right)\right) \le \log\left(1 - (E(\pi(t)))'k^*\right)$$

Also, it is easy to see that we can again restrict ourselves to portfolio vectors π having non-negative components as an optimal portfolio process in the sense of our worst-case problem has to be non-negative (at least $P \otimes l - a.s.$).

ii) Let $h(\pi) := \pi'\left(\mu - r\underline{1}\right) - \frac{1}{2}\pi'\sigma\sigma'\pi$. This function is also concave yielding

$$\begin{aligned} E\left(\pi(t)'\left(\mu - r\underline{1}\right) - \tfrac{1}{2}\pi(t)'\sigma'\pi(t)\right) &\le (E(\pi(t)))'\left(\mu - r\underline{1}\right) \\ &- \tfrac{1}{2}(E(\pi(t)))'\sigma\sigma'(E(\pi(t))). \end{aligned}$$

iii) As in the one-dimensional case it can now be shown that every portfolio process $\hat{\pi}(.)$ which satisfies the equilibrium condition (67.4) also satisfies

$$E\left(v_0\left(t, \hat{X}(t)\left(1 - \hat{\pi}(t)'k^*\right)\right)\right) = E\left(\hat{v}\left(t, \hat{X}(t)\right)\right) = \hat{v}(0, x) = E\left(\hat{v}\left(T, \hat{X}(T)\right)\right)$$

where $\hat{v}\left(t, \hat{X}(t)\right)$ denotes the expected log-utility of terminal wealth from following the portfolio process $\hat{\pi}(.)$.

iv) By the explicit form of

$$E\left(\log\left(\tilde{X}^\pi(T)\right)\right) = \log(x) + rT + E\left(\int_0^T \left(\pi(t)'(\mu - r) - \tfrac{1}{2}\pi(t)'\sigma\sigma'\pi(t)\right)dt\right),$$

a portfolio process $\pi(.)$ with a higher final expected log-utility than a deterministic 'equilibrium process' $\hat{\pi}(.)$ if no crash occurs at all has to satisfy

$$\hat{\pi}(t)'\left(\mu - r\underline{1}\right) - \tfrac{1}{2}\hat{\pi}(t)'\sigma\sigma'\hat{\pi}(t) \le E(\pi(t))'\left(\mu - r\underline{1}\right) - \tfrac{1}{2}(E(\pi(t)))'\sigma\sigma'(E(\pi(t)))$$

for some t.

v) Let us now prove optimality of the equilibrium portfolio process $\hat{\pi}(.)$ that admits the highest expected log-utility of terminal wealth $\hat{v}(t, x)$ in the crash free situation under all deterministic equilibrium strategies. Therefore, consider a portfolio process $\pi(.)$ with a higher final expected log-utility than $\hat{\pi}(.)$. Such a process can only yield a higher worst-case bound than $\hat{\pi}(.)$ if it also satisfies

$$\hat{\pi}(0)' k^* > \pi(0)' k^*.$$

vi) Due to the definition of $\hat{\pi}(.)$ it attains the minimum value of $\pi' k^*$ among all those vectors π that are at the same level set of $h(\pi)$ as $\hat{\pi}(.)$ (at least for almost all $t \in [0, T]$, because otherwise one can construct a better deterministic equilibrium strategy). Consequently, as long as we have

$$\hat{\pi}(t)' k^* > (E(\pi(t)))' k^*$$

we also have

$$E\left(\log\left(X^{\pi}(t)\right)\right) \leq E\left(\log\left(\hat{X}(t)\right)\right).$$

However, due to iv) there must be a first time t where we still have the above inequality between the expected log-wealth but also

$$\hat{\pi}(t)'\left(\mu - r\underline{1}\right) - \tfrac{1}{2}\hat{\pi}(t)'\sigma\sigma'\hat{\pi}(t) \leq E(\pi(t))'\left(\mu - r\underline{1}\right) - \tfrac{1}{2}(E(\pi(t)))'\sigma\sigma'(E(\pi(t))).$$

But due to those two relations and to ii) we then have

$$E\left(v_0\left(t, X^{\pi}(t)\left(1 - \pi(t)' k^*\right)\right)\right) < E\left(v_0\left(t, \hat{X}(t)\left(1 - \hat{\pi}(t)' k^*\right)\right)\right) = \hat{v}(0, x)$$

which proves optimality of $\hat{\pi}(.)$.

Remark

At first sight the optimization problem (67.5) seems to be very hard to solve. However, as by the explicit forms of both $v_0(t, x)$ and $v_{\pi}(t, x)$ the function over which the supremum is taken does not depend on the underlying stochastic process $X^{\pi}(t)$ one is at least able to get a numerical solution via backwards induction starting with $\pi(T) = 0$.

67.4 MAXIMIZING GROWTH RATE UNDER THE THREAT OF A CRASH: AN ARBITRARY NUMBER OF CRASHES AND OTHER REFINEMENTS

67.4.1 Arbitrary Upper Bound for the Number of Crashes

So far the maximum number of crashes was limited to one. However the extension to an arbitrary (but fixed) upper bound is straightforward. In fact the extension is something like a backward induction principle. If one has determined the best strategy given the maximum number of crashes is $n - 1$ then one can imitate the above proof of Theorem 3 (or Theorem 6, respectively) to get the optimal strategy in the case of the upper bound of n. Simply note that the role of $v_0(t, x)$ is then taken over by $v_{n-1}(t, x)$, the value function for the $n - 1$ case. For completeness we give the corresponding theorem in the one stock case. The n stock case is then similar to Theorem 6 but lacks the explicit formula that we can give in the single stock case:

Theorem 7 'Best portfolio under the threat of at most *n* crashes'

The optimal strategy in the log-utility case if at most n crashes of size $k \in [0, k^*]$ can occur is given by the unique solution(s) of the differential equation(s)

$$\dot{\hat{\pi}}_j(t) = \frac{1}{k^*}\left(1 - \hat{\pi}_j(t)k^*\right)\left(\hat{\pi}_j(t)\left(\mu - r\right) - \tfrac{1}{2}\left(\hat{\pi}_j(t)^2\sigma^2 + \hat{\pi}_{j-1}(t)^2\sigma^2\right)\right),$$

$$\hat{\pi}_j(T) = 0$$

with

$$0 \leq \hat{\pi}_j(.) < \frac{1}{k^*}, \quad j = 1, \ldots, n, \quad \hat{\pi}_0(t) = \frac{\mu - r}{\sigma^2}$$

where $\hat{\pi}_j(.)$ is the optimal portfolio process for the worst-case problem if at most j crashes can occur.

Proof

We give an induction on n, the maximum number of crashes. For the case of $n = 1$ (and also $n = 0$) all the claims follow from Theorem 3. Now that the above assertions are already proved for $n - 1$ and that $\hat{\pi}_{n-1}(.)$ is the corresponding equilibrium strategy (i.e. all future time instant yield binding constraints simultaneously for the worst case problem), then, by induction, $v_{n-1}(t, x)$ the corresponding value function is given by

$$v_{n-1}(t, x) = \log(x) + r\,(T - t) + \log\left(1 - \hat{\pi}_{n-1}(t)k^*\right) + \int_t^T \left(\left(\mu - r\right)\hat{\pi}_{n-2}(s) - \tfrac{1}{2}\hat{\pi}_{n-2}(s)^2\sigma^2\right) ds$$

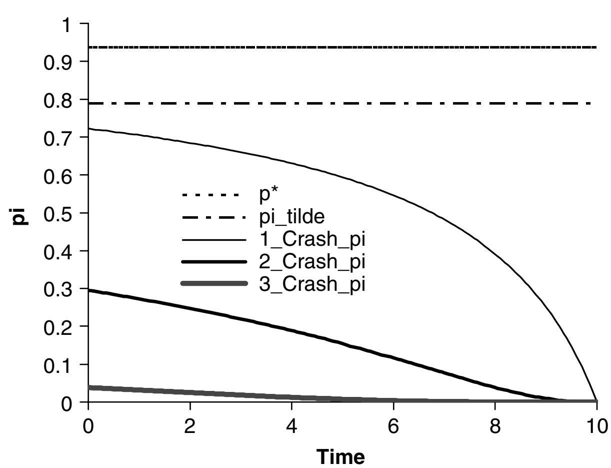


Figure 67.3 Optimal portfolios with and without crash possibility (large time horizon, at most three crashes).

(where this only holds for $n - 1 > 0$). The rest of the proof is now totally similar to that of Theorem 3 with an obvious change of notation.

In Figure 67.3 we illustrate the n-crash-situation, showing the situation for the 3-crash-situation with the same data as used for producing Figure 67.2. Of course, the more crashes possibly to come the less is the optimally invested fraction of wealth into the risky stock.

67.4.2 Changing Volatility After a Crash

It is a common phenomenon that after a crash has happened the volatility has the tendency to increase. For our worst-case problem this has the consequence that the 'starting value function' $v_0(t, x)$ has to be computed with a different value of σ. In particular, in the n-crash case it might be necessary to calculate all value functions with different values of σ where they are valid.

67.4.3 Further Possible Refinements

There are still a lot of possible problems in the above setting which are worth considering and which might be the subject of future research:

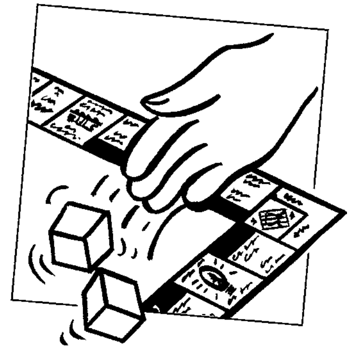
- inclusion of the possibility for consumption;
- explicit solution of the problem for a general utility function;
- inclusion of liquidity constraints;
- additional consideration of derivatives for portfolio insurance.

67.5 SUMMARY

It is particularly pleasing when a model produces results that tie in with intuition. We see that in this model since it recommends reducing exposure to the risky investment as our 'retirement' approaches.

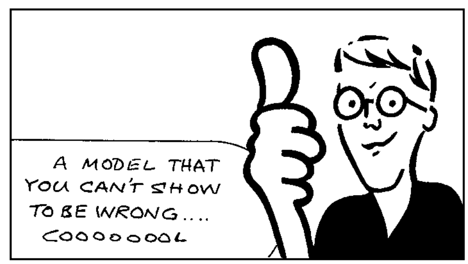
FURTHER READING

- See Korn (1997) for the maths of optimal portfolios.
- For further work in this area see Korn (2003, 2005), Korn & Menkens (2004a,b).



CHAPTER 68
interest-rate modeling without probabilities

In this Chapter...



- the Epstein–Wilmott model, a non-Brownian motion model for interest rates:
- how to value interest-rate products in a worst-case scenario
- the Yield Envelope
- optimal static hedging

68.1 INTRODUCTION

The two main classical approaches to pricing and hedging fixed income products may be termed 'yield-based' and 'stochastic.' The former (see Chapter 13) assumes that interest rates are constant for each product, which, of course, is inconsistent across products. These ideas are used a great deal for the simpler, 'linear' products. The latter approach (see Chapters 30 and 35) assumes that interest rates are driven by a number of random factors. It is used for 'non-linear' contracts, contracts having some form of optionality. In the stochastic models an equation for the short-term rate will give as an output the whole yield curve.

Both of these approaches can be criticized. The yield-based ideas are not suited to complex products and the popular stochastic models are inaccurate. For the latter, it is extremely difficult to estimate parameters, and after estimating them, they are prone to change, making a mockery of the underlying theory. One of the main problems is the assumption of a finite number of factors. From such an assumption it follows that you can delta hedge any contract with this same number of simpler contracts. For example, in a one-factor world you can hedge one part of the yield curve with any other part, something which is clearly not possible in practice. Is it acceptable to hedge a six-month option on a one-year bond with a ten-year bond? Although practitioners use common sense to get around this (they would hedge the option with the one-year bond), this common sense is not reflected in the modeling.

In this chapter I address the problem from a new perspective, by assuming as little as possible about the process underlying the movement of interest rates. I will model a short-term interest rate and price a portfolio of cashflows in a worst-case scenario, using the short rate as the rate for discounting. One of the key features of the model in this chapter is that delta hedging plays no important role. The resulting problem is non linear and thus the value of a contract then depends on what it is hedged with. This approach necessarily correctly prices traded instruments; no fitting is necessary. I also describe the Yield Envelope. This is a sophisticated version of the yield curve. We find a yield spread at maturities for which there are no traded instruments.

68.2 WHAT DO I WANT FROM AN INTEREST RATE MODEL?

Here is my list of properties of my ideal interest rate model.

- As few factors as possible, but able to model any realistic yield curve
- Easy to price many products quickly
- Insensitivity of results to hard-to-measure parameters, such as volatilities and correlations
- Robustness in general
- Sensible fitting to data
- Strategy for hedging

This list has been built up after conversations with many practitioners. In this chapter I describe a model that delivers all of these and more. In fact, we won't be seeing any mention of volatilities or correlations, or delta hedging. The only hedging will be entirely static.

68.3 A NON-PROBABILISTIC MODEL FOR THE BEHAVIOR OF THE SHORT-TERM INTEREST RATE

Motivated by a desire to model the behavior of the short-term interest rate, r, with as much freedom as possible, I assume only the following constraints on its movement:

$$r^- \leq r \leq r^+ \tag{68.1}$$

and

$$c^- \leq \frac{dr}{dt} \leq c^+. \tag{68.2}$$

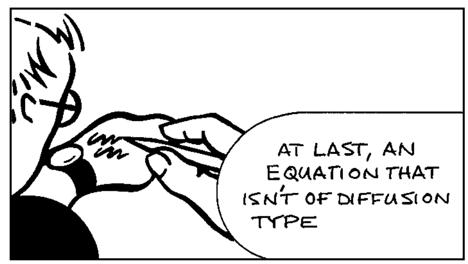
Equation (68.1) says that the interest rate cannot move outside the range bounded below by the rate r^- and above by the rate r^+. Equation (68.2) puts similar constraints on the *speed* of movement of r. The constraints can be time-dependent and, in the case of the speed constraints, functions of the spot interest rate, r.

There is an obvious difference between the classical stochastic models of the spot rate, with their Brownian motion evolution of r and locally unbounded growth, and the model I am now presenting. I can justify this on several grounds: (i) We are perhaps trying to model a long-term behavior for which we are less concerned about the very short-term movements (a weak justification); (ii) The Brownian models can also be criticized. It is still an open question exactly what the stochastic process is that underlies the evolution of financial quantities; Brownian motion is often chosen for its nice mathematical properties (a slightly better justification);

(iii) A combination of the model here together with bands, jumps etc. discussed later will, in practice, be indistinguishable from the real process (excellent justification, if it's true).

The worst-case scenario that we will be addressing is hard to criticize as long as it gives decent prices. Why then don't we simply present an interest rate version of the uncertain volatility model of Chapter 52. This has been done by Lewicki & Avellaneda (1996) in a Heath, Jarrow & Morton framework. It is not entirely satisfactory because of the usual problem: What we model and what we trade are two different things. If we are going to use a delta-hedging argument, then we have to assume that what we hedge with is perfectly correlated with our contract; this can never be the case in the fixed-income world. Different points on the yield curve may be correlated but they are certainly not perfectly correlated for all time. In the model we present here *there is no delta hedging* and we do not depend on any correlation between different parts of the yield curve.

68.4 WORST-CASE SCENARIOS AND A NON-LINEAR EQUATION



In this section we derive the equation governing the worst-case price of a fixed-income portfolio, first presented by Epstein & Wilmott (1997). Let $V(r, t)$ be the value of our portfolio when the short-term interest rate is r and the time is t. We consider the change in the value of this portfolio during a time step dt.

Using Taylor's theorem to expand the value of the portfolio for small changes in its arguments, we find that

$$V(r + dr, t + dt) = V(r, t) + \frac{\partial V}{\partial r}\,dr + \frac{\partial V}{\partial t}\,dt + \cdots.$$

Note that there is no second r-derivative term because the process is not Brownian. We want to investigate this change under our worst-case assumption. This change is given by

$$\min_{dr}(dV) = \min_{dr}\left(\frac{\partial V}{\partial r}\,dr + \frac{\partial V}{\partial t}\,dt\right).$$

Since the rate of change of r is bounded according to (68.2), we find that

$$\min_{dr}(dV) = \min_{dr}\left(\frac{\partial V}{\partial r}\,dr + \frac{\partial V}{\partial t}\,dt\right) = \left(c\left(\frac{\partial V}{\partial r}\right)\frac{\partial V}{\partial r} + \frac{\partial V}{\partial t}\right)dt$$

where

$$c(x) = \begin{cases} c^+ & \text{for } x < 0 \\ c^- & \text{for } x > 0. \end{cases}$$

We shall require that, in the worst case, our portfolio always earns the risk-free rate of interest. This gives us

$$\boxed{\frac{\partial V}{\partial t} + c\left(\frac{\partial V}{\partial r}\right)\frac{\partial V}{\partial r} - rV = 0.} \qquad (68.3)$$

This is a first-order non-linear hyperbolic partial differential equation. For an instrument having a known payoff at maturity we will know the final data $V(r, T)$. Also, if there is a cash flow K at time t_i, then we have

$$V(r, t_i^-) = V(r, t_i^+) + K.$$

Thus we solve backwards in time from T to the present, applying jump conditions when necessary.

In addition to the worst-case scenario, we can find the value of our portfolio in a best-case scenario. This is equivalent to a worst-case scenario where we are short the portfolio. We therefore have

$$-V_{\text{best}} = (-V)_{\text{worst}}.$$

68.4.1 Let's See That Again in Slow Motion

Now we've done the math, let's see that again with some numbers.

Suppose we hold a contract that, depending on the evolution of interest rates over the next dt time step, will be worth 1.03, 1.02, 1.015, 1.01 or 1.02. This is illustrated in Figure 68.1. The short-term interest rate is currently 8%. What is the worst that can happen? Clearly, the worst is for the value of 1.01 to be realized, since that is the lowest of all the possible values after the next time step. So we assume that's what happens. In that case, what is the value of the contract now? PVing the future value gives

$$? = \frac{1.01}{1 + 0.08\,dt}.$$

If the time step is one day this is just 1.0097.

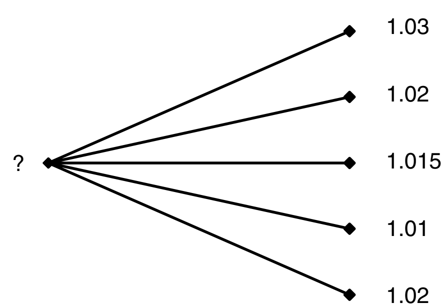


Figure 68.1 Five possible outcomes.

Example

Our first 'proper' example will be the simplest possible: We will value a zero-coupon bond in the worst and best cases according to the model. Results are shown in Figure 68.2 In this example I have valued a zero-coupon bond in the two scenarios, best- and worst-cases, and plotted the yield for different maturities. The spot rate is initially 6%, is allowed to grow or decrease at 4% p.a. at most, and cannot go outside the range 3–20%. The important point to

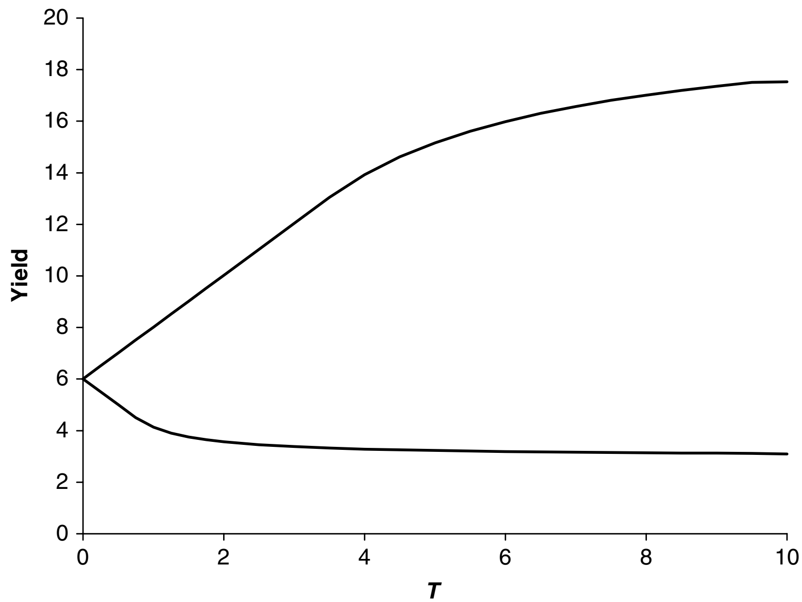


Figure 68.2 Yields in worst- and best-case scenarios for a zero-coupon bond.

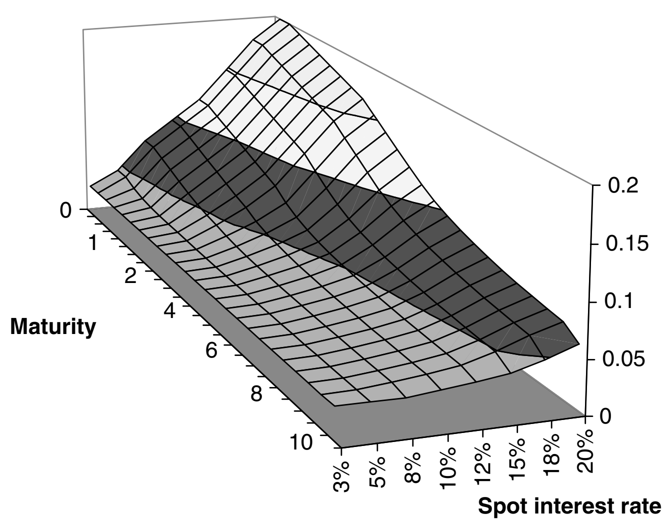


Figure 68.3 Yields in worst-case scenario for a zero-coupon bond.

note about this figure is how wide the yield spread is. If the model always gave such extreme results then it would be of no practical use. Fortunately we can reduce this dramatically by static hedging, and we will see examples of this shortly. This example is shown again in Figures 68.3 and 68.4. In these are shown the yields in the two cases but now against both the spot rate and maturity.

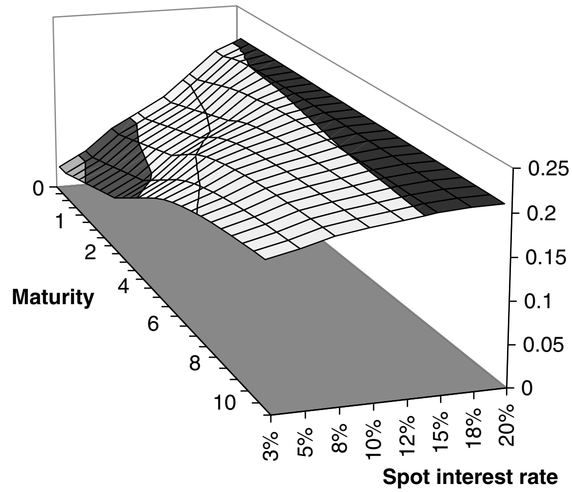


Figure 68.4 Yields in best-case scenario for a zero-coupon bond.

68.5 EXAMPLES OF HEDGING: SPREADS FOR PRICES

Since our uncertain interest rate model solely places bounds on the short-term interest rate, it is not surprising that the best we can do is to find bounds for the value of a contract. We have derived a partial differential equation for the value of a contract in a worst-case scenario, a lower bound, and for the value in the best case, an upper bound. We therefore find a spread for the possible price of a contract. Consequently, long and short positions in a contract have different values.

Finding a spread for prices is not necessarily a disadvantage of the model. After all, the market itself has such a property (the bid-offer spread). In some sense, spreads are therefore a more realistic result than a single price. However, it becomes a disadvantage when the spreads are so large that the result becomes meaningless. We require a method to reduce large spreads to more sensible levels; this is the process of static hedging.

Typically, we find that the spread between the worst- and best-case values is larger than the market spread. To reduce the spread, we hedge with market-traded instruments. Hedging is shown schematically in Figure 68.5. At the top of this figure is our original, or 'target' contract. Some of the cashflows are known amounts and some are floating. Below this are shown the contracts that are available for hedging; to keep things simple I've shown them all as zero-coupon bonds. How many of each of these available hedging bonds should we buy or sell to give our target cashflow the highest value in the worst-case scenario?

There is an optimal static hedge for which the worst-case value of the bond is as high as possible, and another for which the best-case value of the bond is as low as possible. This optimization technique was described in detail in another context in Chapter 60. To find this optimal static hedge in the worst-case scenario, we maximize the value of our zero-coupon bond with respect to the hedge quantities of the hedging instruments. In the best-case scenario, we minimize with respect to the hedge quantities.

We expect that the market price of a hedging instrument is contained in the spread of values for the instrument generated by our model. If this were not the case, we could make an arbitrage

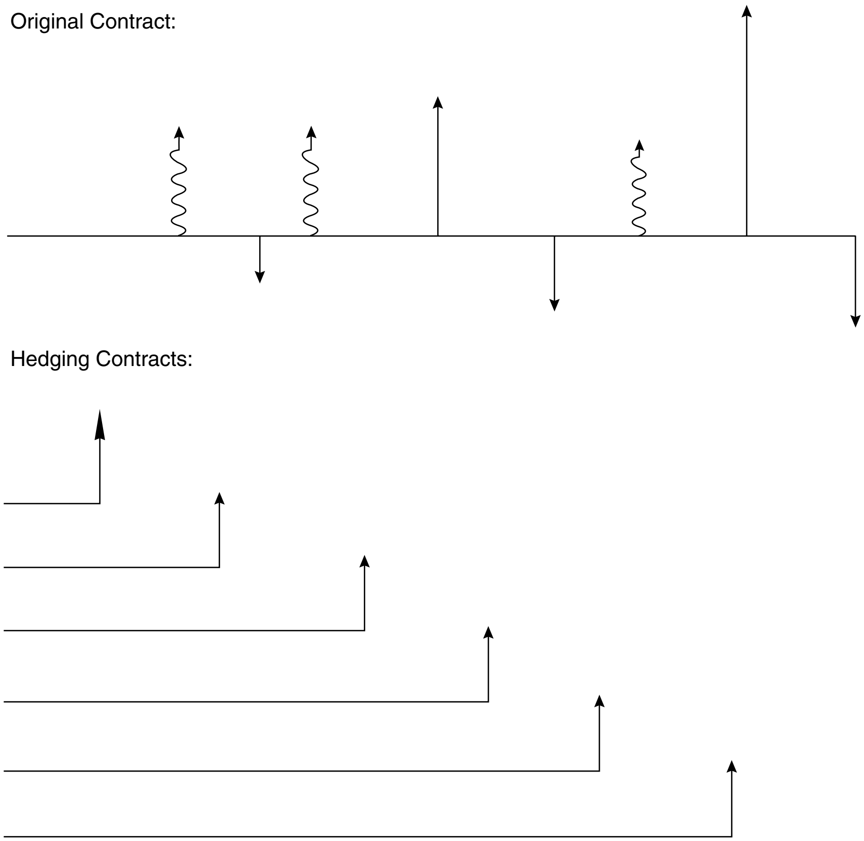


Figure 68.5 A schematic diagram of the hedging problem.

profit by selling (buying) the instrument at a price above (below) its maximum (minimum) possible value, assuming that the interest rate moves within the constraints of our model.

Observe how the nonlinearity in the model means that the value of a portfolio depends on what it is hedged with. This means that the 'fitting' that we saw in Chapter 31 becomes irrelevant. In fact, we are not concerned with the market prices of traded instruments except in so far as we exploit these instruments for hedging. We never say that we 'believe' market prices are 'correct' only that they tell us how much we must pay to get a particular cashflow.

68.5.1 Hedging with One Instrument

Consider a contract consisting of a set of cash flows. We wish to value this contract in a worst-case scenario. Suppose that there exists a market-traded instrument, with known market price (a zero-coupon bond, for instance). We hedge with this instrument and price the resulting portfolio in a worst-case scenario. The value of the overall portfolio is

$$\text{VALUE(contract + hedging instrument)}.$$

where 'VALUE' means the solution of the non-linear partial differential equation, Equation (68.3), with relevant final and jump conditions.

The cost of setting up this static hedge is equal to the current market value of the hedging instrument. The marginal value of our hedged contract is therefore the value of the overall portfolio minus the cost of the static hedge,

$$\text{MARGINAL VALUE(hedged contract)} = \text{VALUE(contract + hedging instrument)} \\ - \text{cost of hedge.}$$

Example 1

Let's try to hedge a five-year zero-coupon bond with a one-year zero-coupon bond. The latter has a market price of \$0.90484, the spot rate is currently 10%, the interest rate range is 3–20%, and the spot rate cannot change at a rate faster than 4% p.a., either up or down. In Figure 68.6 are shown the worst- and best-case values for the five-year bond as a function of the number of one-year bonds with which it is hedged. First look at the lower curve, the worst-case value. Note that it has a maximum. To get the most from the five-year bond in the worst case it must be hedged with -2.15 of the one-year bond. If you hold the five-year bond short then to get the best out of it you must hedge it with 2.29 of the one-year bond.

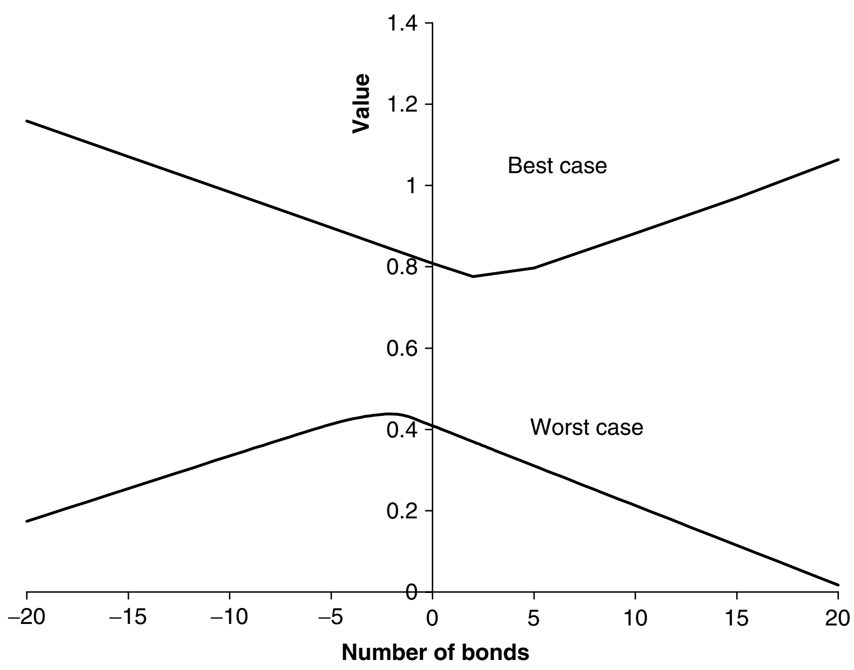


Figure 68.6 The value of the hedged five-year bond in the worst and best cases as a function of the number of one-year bonds with which it is hedged.

68.5.2 Hedging with Multiple Instruments

Suppose that there exist M instruments in the market, and each has a known market price C_j, for $1 \le j \le M$. We hedge our original contract with these instruments. We then price the resulting portfolio in a worst-case scenario.

We hedge with λ_j of the jth instrument. The value of the overall portfolio is

$$\text{VALUE}\left(\text{original contract} + \sum_{j=1}^{M} \lambda_j \ \text{ hedging instrument}_i\right).$$

The cost of setting up this static hedge is equal to the current market value of the hedging instruments,

$$\sum_{j=1}^{M} \lambda_j C_j.$$

The marginal value of our hedged contract is therefore the value of the overall portfolio minus the cost of the static hedge,

$$\text{MARGINAL VALUE}\left(\text{original contract} + \sum_{j=1}^{M} \lambda_j \ \text{ hedging instrument}_i\right) - \sum_{j=1}^{M} \lambda_j C_j.$$

To avoid arbitrage opportunities, we must again assume that the market value of each hedging instrument is contained within the best- and worst-case scenario values for that instrument, obtained from our model.

There will be an optimal static hedge, for which we obtain the maximum possible worst-case scenario value. To find this, we maximize the value of the portfolio with respect to the hedge quantities, λ_j:

$$\max_{\lambda_j}\left(\text{VALUE}\left(\text{original contract} + \sum_{j=1}^{M} \lambda_j \ \text{ hedging instrument}_i\right) - \sum_{j=1}^{M} \lambda_j C_j\right).$$

Similarly, there will be an optimal static hedge, for which we obtain the minimum possible best-case scenario value. To find this, we minimize the value of the portfolio with respect to the hedge quantities, λ_j:

$$\min_{\lambda_j}\left(\text{VALUE}\left(\text{original contract} + \sum_{j=1}^{M} \lambda_j \ \text{ hedging instrument}_i\right) - \sum_{j=1}^{M} \lambda_j C_j\right),$$

where ‘VALUE’ now denotes the solution of the best-case scenario equation.

To include a bid-offer spread in the market price for a hedging instrument, we simply make C_j dependent on the sign of λ_j. If $\lambda_j > 0$ then we are long the bond and the market price is the offer price. If $\lambda_j < 0$ then we are short the bond and the market price is the bid price, i.e.

$$C_j(\lambda_j) = \begin{cases} C_j^+ & \text{if } \lambda_j > 0 \\ C_j^- & \text{if } \lambda_j < 0, \end{cases}$$

where C^+ is the offer price and C^- is the bid price.

By hedging with market-traded instruments, we can significantly decrease the spread between the worst- and best-case valuations for our contract.

Example 2

We hedge a four-year zero-coupon bond with zero-coupon bonds of maturities 0.5, 1, 2, 3, 5, 7, 10 years. These hedging bonds are shown in Table 68.1 together with their market prices (we are assuming for the moment that the bid and ask prices are the same). The spot rate is currently 6% and we will take $r^+ = 0.2$, $r^- = 0.03$, with $c^+ = -c^- = 0.04$. Thus the spot rate is not allowed to grow or decrease faster than 4% p.a.

The results of the valuation, with and without the optimal static hedge, are shown in Table 68.2.

The optimal static hedges for the worst- and best-case valuations of the four-year bond are shown in Table 68.3.

The short-term interest rate paths for the optimally-hedged worst- and best-case scenarios are shown in Figure 68.7 and 68.8 respectively.

We observe that the hedge quantities for worst- and best-case scenario valuations differ and that the optimal static hedge will not necessarily contain all of the hedging instruments; in neither case is the 10-year bond used for hedging. Most of the hedging is with the three- and five-year maturities. Hedging has reduced the spread in the value of the four-year bond from $0.878 - 0.579 = 0.299$ to $0.758 - 0.730 = 0.028$, a factor of more than ten. This is still a vast spread, but considering how few instruments we have used to hedge it, and that the hedge

Table 68.1 Choice of hedging instruments.

Hedging bond	Maturity (years)	Market price
A	0.5	0.970
B	1	0.933
C	2	0.868
D	3	0.805
E	5	0.687
F	7	0.579
G	10	0.449

Table 68.2 Results of the valuation, with and without the optimal static hedge.

	Worst case	Best case
No hedge	0.579	0.878
Optimally hedged	0.730	0.758

Table 68.3 The optimal static hedges for the worst- and best-case valuations of the four-year bond.

Hedging bond	Worst-case hedge quantity	Best-case hedge quantity
A	0.000	0.002
B	−0.004	−0.004
C	0.169	0.117
D	−0.699	−0.653
E	−0.468	−0.481
F	0.020	0.000
G	0.000	0.000

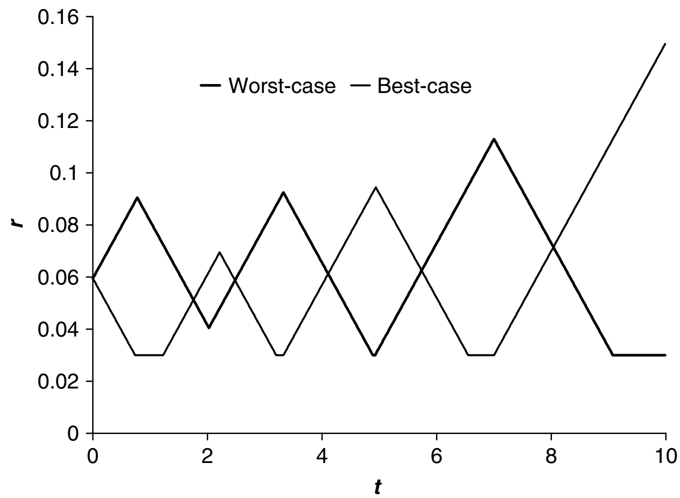


Figure 68.7 Interest rate paths for the four-year bond optimally hedged in the worst-case scenario.

is *perfect*, it's very good. It will be reduced further if we are allowed to hedge with more instruments.

68.6 GENERATING THE 'YIELD ENVELOPE'

We continue with our philosophy of finding spreads for prices by generating the **Yield Envelope**. As we found in the above example of hedging a four-year bond, at a maturity for which there are no traded instruments we find a yield spread. So, we find a highest and lowest value for the yield at all maturities for which there are no traded instruments. The plot of the highest and lowest yields against maturity is called the Yield Envelope. Note that at a maturity for which there is a traded instrument, the Yield Envelope converges to the observed yield, the highest and lowest values coincide.[1]

[1] This is not the case if there are distinct bid and ask prices, but this is a simple generalization of the concept.

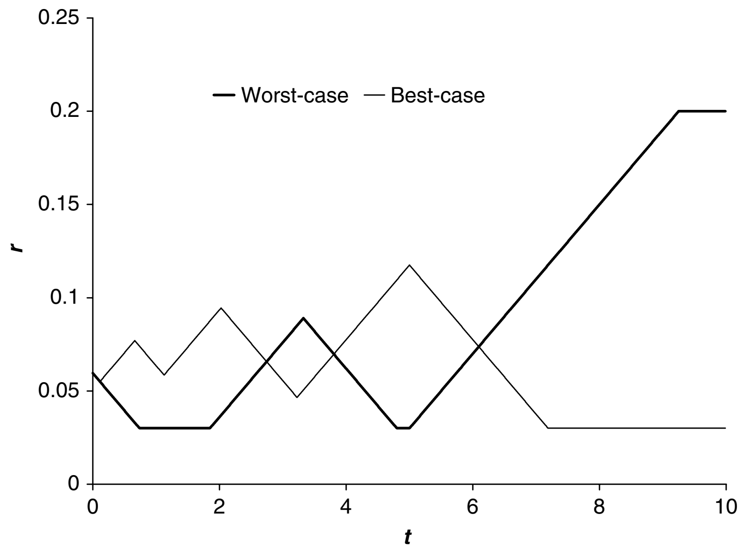


Figure 68.8 Interest rate paths for the four-year bond optimally hedged in the best-case scenario.

We calculate the worst- and best-case values of a zero-coupon bond with principal 1 and maturity T. We then calculate the maximum and minimum yields possible, using

$$Y = -\frac{\log Z}{T - t}.$$

To reduce the yield spread, we again hedge our zero-coupon bond with market-traded zero-coupon bonds.

Example

We hedge our zero-coupon bonds with the bonds in Table 68.4. This table shows the available zero-coupon bonds and their market values. The spot rate is 6%. The other parameters are as before. The results are shown in Figure 68.9 and in Table 68.5. In the latter table the numbers in bold represent the traded bonds. In the figure are four curves. The outer two are exactly

Table 68.4 Available zero-coupon bonds and their market values.

Hedging bond	Maturity (years)	Market price
A	0.5	0.950
B	1	0.899
C	2	0.803
D	3	0.712
E	5	0.566
F	7	0.448
G	10	0.304

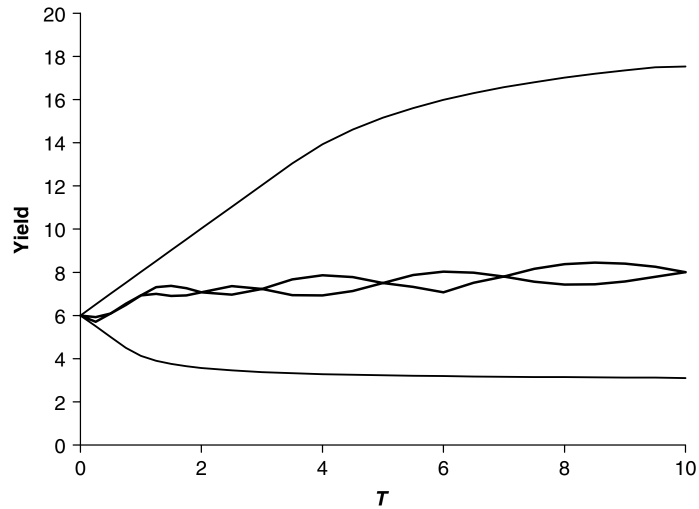


Figure 68.9 Yields in worst and best cases, with and without hedging.

Table 68.5 The optimally-hedged 'yield envelope.'

Maturity (years)	Yield in worst case	Yield in best case
0	6.000	6.000
0.25	5.919	5.706
0.5	**6.092**	**6.092**
0.75	6.545	6.485
1.0	**6.935**	**6.935**
1.25	7.313	7.002
1.5	7.378	6.908
1.75	7.269	6.933
2.0	**7.078**	**7.078**
2.5	7.357	6.969
3.0	**7.230**	**7.230**
3.5	7.668	6.942
4.0	7.864	6.927
4.5	7.777	7.138
5.0	**7.508**	**7.508**
5.5	7.877	7.329
6.0	8.029	7.079
6.5	7.986	7.517
7.0	**7.806**	**7.806**
7.5	8.159	7.558
8.0	8.381	7.438
8.5	8.454	7.450
9.0	8.407	7.571
9.5	8.259	7.786
10.0	**8.007**	**8.007**

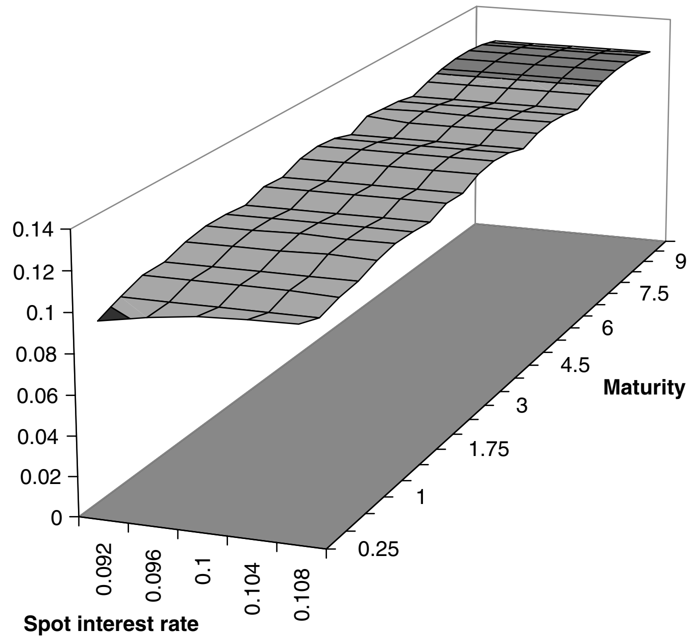


Figure 68.10 Yields in the worst-case scenario against spot rate and maturity, with optimal hedging.

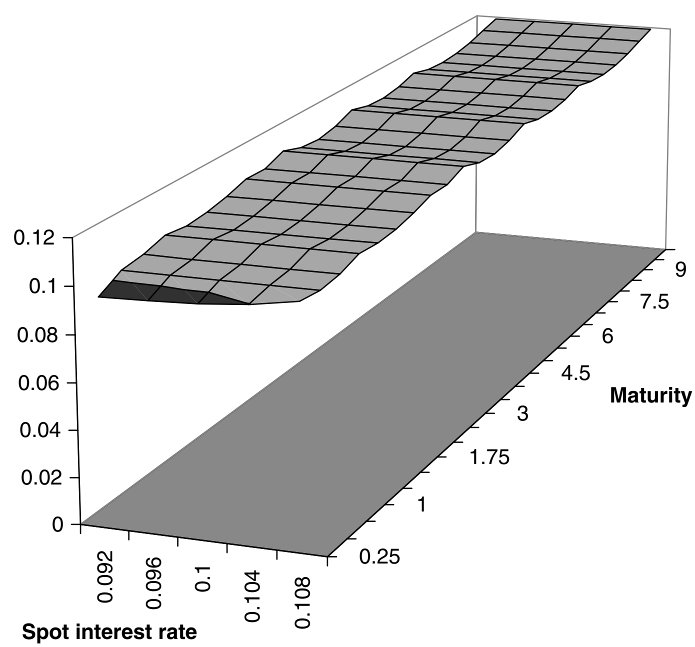


Figure 68.11 Yields in the best-case scenario against spot rate and maturity, with optimal hedging.

the same as in Figure 68.2. They are the yields in the best and worst cases, without any hedging, assuming a current spot rate of 6%, a growth of ±4%, and a range of 3–20%. The inner curves (the 'string of sausages') form the Yield Envelope. These are the best and worst cases for the yields, with the same parameters, but now with optimal static hedging. Note the dramatic decrease in the yield spreads at all maturities. In particular, at maturities with traded zero-coupon bonds the spread disappears, and the theoretical price becomes the market price.

At maturities for which there are traded instruments, the yields in both worst- and best-case scenarios equal the market yield. This is because we can fully hedge our zero-coupon bond with the market-traded instrument and this is the optimal static hedge in both scenarios.

Figures 68.10 and 68.11 show the yields in the worst- and best-case scenarios, respectively, with varying spot rate and maturity, again when we hedge optimally with the bonds shown in the table.

68.7 SWAPS

With the methodology that we have built up, we can price and hedge a variety of simple fixed-income products. To value any instrument whose cashflows solely depend on the short-term interest rate, we simply solve our partial differential equation and include these cashflows as jump conditions. I now show how to incorporate swaps into this approach, and in the next section, apply the method to caps and floors.

Consider a swap with a τ-period reference rate (measured τ before the payment date), an annualized fixed rate r_f, maturity at time T, a principal of 1 and swaplet dates every τ. There are two approaches to the pricing of this contract. Market practice is to decompose the swap into a portfolio of zero-coupon bonds, as shown in Chapter 14. These can then be incorporated into our model for the short-term interest rate.

We solve Equation (68.3), for the worst-case scenario value, or the equivalent equation for the best-case scenario value, with a final condition of,

$$V(r,T) = -\left(1+\tau r_f\right),$$

jump conditions of,

$$V(r,t_s^-) = V(r,t_s^+) - \tau r_f,$$

at each swaplet date before maturity, t_s, and a jump condition of,

$$V(r,t_0^-) = 1 + V(r,t_0^+),$$

at a time τ before the first swaplet date, t_0, say.

An alternative method, more often seen in academia, is to approximate the τ-period rate by substituting our annualized short-term interest rate (which is instantaneously compounding) in its place. We can then price the cashflows directly using our model since they now only depend on our short-term interest rate. The method is only valid when the period τ is short enough to make this a reasonable assumption.

To solve the problem using this latter approach, we solve Equation (68.3) with a final condition of,

$$V(r,T) = \tau(r - r_f),$$

and the jump condition,

$$V(r, t_s^-) = V(r, t_s^+) + \tau(r - r_f),$$

at each swaplet date, t_s, before maturity.

Example

We price and hedge a two-year swap, on the three-month interest rate, with a principal of 1. Swaplets occur every three months and the last payment is three years from today. The spot short-term interest rate is 6%. The interest rate bounds are 3% and 20% and the growth rate is bounded by −4% pa and 4% pa.

We use the yield curve determined from the hedging bonds of Table 68.1 and the decomposition approach to find the 'fair value' for r_f. We linearly interpolate to construct the yield curve, as shown in Figure 68.12. We can then price the relevant zero-coupon bonds and find that the fixed interest rate should be 7.44% per annum.

First value the swap using the decomposition into zero-coupon bonds. The results of this valuation, with and without hedging, are shown in Table 68.6. The optimal static hedges for the worst- and best-case valuations are shown in Table 68.7 (where we hedge with the bonds in Table 68.1).

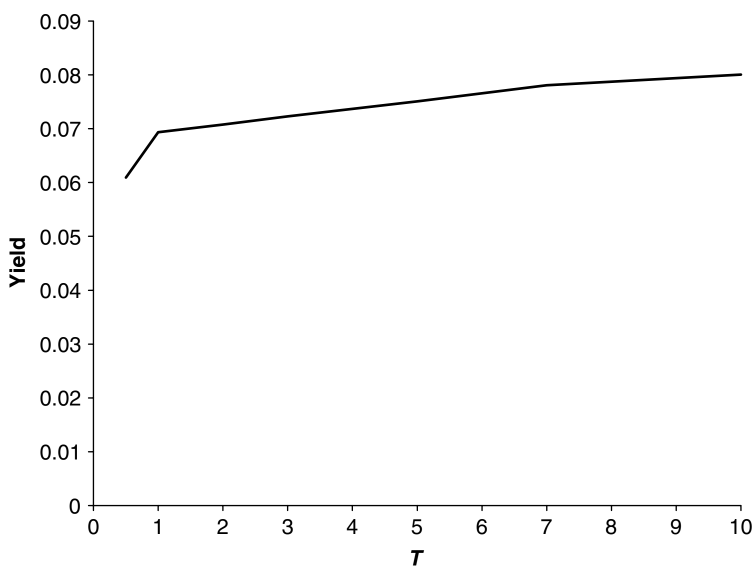


Figure 68.12 The yield curve for the bonds from Table 68.1.

Table 68.6 Value of the decomposed swap.

	Worst case	Best case
No hedge	−0.0822	0.1056
Optimally hedged	−0.0002	0.0002

Table 68.7 The optimal static hedges for the decomposed swap.

Hedging bond	Worst case hedge quantity	Best case hedge quantity
A	0.000	0.000
B	−0.977	−0.983
C	0.092	0.097
D	1.037	1.035
E	0.000	0.000
F	0.000	0.000
G	0.000	0.000

We also value the swap using the short-term interest rate to approximate the three-month rate. The results of the valuation, with and without hedging, are shown in Table 68.8. The optimal static hedges for the worst- and best-case valuations are shown in Table 68.9.

The value of the swap before hedging is similar for both the decomposition and approximation methods. The small difference between the two prices could be due to the choice of fixed rate via the former rather than the latter approach. We could therefore remark that the latter method is a reasonable approximation to make.

However, the effect of hedging is far more effective for the decomposition method, where the spread is reduced from 0.1878 to 0.0004, than for the approximation method, where the spread is reduced from 0.1919 to 0.0096. (Even so, hedging has still had a significant effect in the latter case.) A plausible explanation for the greater spread reduction could be that, in the first case, we are hedging cashflows with cashflows of the same form (both are zero-coupon bonds) whereas in the second case, the cashflows to be hedged are of a different form and consequently, the hedge will not be as effective.

In choosing which instruments to use in the setting up of a static hedge, we should identify those products most similar in form to our contract as the most suitable hedging instruments. (For example, when we come to price the convertible bond, in Chapter 69, zero-coupon bonds

Table 68.8 Value of the approximated swap.

	Worst case	Best case
No hedge	−0.0824	0.1095
Optimally hedged	−0.0060	0.0036

Table 68.9 The optimal static hedges for the approximated swap.

Hedging bond	Worst case hedge quantity	Best case hedge quantity
A	−0.050	1.625
B	−0.746	−1.863
C	−0.312	−0.036
D	1.237	1.163
E	0.000	0.000
F	0.000	0.000
G	0.000	0.000

are a natural choice of hedging instrument, but when we price the index amortizing rate swap we choose to hedge with swaps.)

68.8 CAPS AND FLOORS

We consider an interest rate agreement with a τ-period reference rate, an annualized strike rate r_s, with maturity at time T, a principal of 1 and cashflow dates every τ. As with the swap valuation, there are two approaches to the pricing of these contracts. Market practice is to decompose the contract into a portfolio of bond options, as shown in Section 32.4. We discuss the valuation of such options in the next chapter.

Alternatively, we can again approximate the τ-period rate by substituting our short-term interest rate in its place. We then price the cashflows directly using our model. The method is still only valid when the period τ is short enough to make this a reasonable assumption.

To solve the problem using this latter approach, we solve Equation (68.3), for the worst-case scenario value, with jump and final conditions dependent on the particular specification of the interest rate agreement.

For a cap, we apply a final condition of

$$V(r, T) = \tau \max(r - r_s, 0),$$

and the jump condition

$$V(r, t_c^-) = V(r, t_c^+) + \tau \max(r - r_s, 0),$$

at each caplet date, t_c, before maturity. We must again divide by τ because the short-term interest rate, r, and the strike rate, r_s, are annualized rates, and we require rates that hold for a period of τ only.

For a floor, we apply a final condition of

$$V(r, T) = \tau \max(r_s - r, 0),$$

and the jump condition

$$V(r, t_f^-) = V(r, t_f^+) + \tau \max(r_s - r, 0),$$

at each floorlet date before maturity, t_f.

Example

We price and hedge two-year contracts, on the three-month interest rate, with a principal of 1. Cashflows occur every three months and the last payment is two years from today. The spot short-term interest rate is 6%. The interest rate bounds are 3% and 20% and the growth rate is bounded by −4% pa and 4% pa.

We hedge the contracts with the zero-coupon bonds in Table 68.1. The results for the valuation are shown in Tables 68.10 and 68.11.

Note that hedging appears to be more effective at reducing the best-case price than raising the worst-case price. A cause of this could be that, in the worst-case, the interest rate will move to ensure that the caplets or floorlets have zero value. To have an effect hedging must either move the worst-case interest rate path so that these cashflows have value, or take advantage of the worst-case path to make a profit on the hedging bonds. On the other hand, to reduce the best-case price, the hedging only has to counteract the positive value of the cashflows.

Table 68.10 Value of a cap with varying strike.

Cap (5% strike)	Worst case	Best case
No hedge	0.000	0.096
Optimally hedged	0.035	0.046

Cap (6% strike)	Worst case	Best case
No hedge	0.000	0.078
Optimally hedged	0.018	0.036

Cap (7% strike)	Worst case	Best case
No hedge	0.000	0.060
Optimally hedged	0.003	0.022

Table 68.11 Value of a floor with varying strike.

Floor (5% strike)	Worst case	Best case
No hedge	0.000	0.031
Optimally hedged	0.000	0.005

Floor (6% strike)	Worst case	Best case
No hedge	0.000	0.050
Optimally hedged	0.000	0.011

Floor (7% strike)	Worst case	Best case
No hedge	0.000	0.069
Optimally hedged	0.002	0.020

68.9 APPLICATIONS OF THE MODEL

There are a number of different ways in which the theory and techniques that we have developed can be applied in the marketplace.

68.9.1 Identifying Arbitrage Opportunities

Assuming that the movements of the interest rate conform to our constraints, the realized value of any contract must lie within our bounds. We have therefore obtained a spread for the possible price of a contract. If we find a contract whose value lies outside of these bounds, then we have identified an arbitrage opportunity. We should obviously buy the contract if its current market value is below this range, and sell the contract if its value is above the range.

68.9.2 Establishing Prices for the Market Maker

If we are making a market in a contract, then setting our bid price at the low end and our offer price at the high end of the spread range guarantees that we cannot lose money on any deal. This technique is particularly appropriate in the OTC contracts business where spreads are usually much higher due to the often exotic nature and illiquidity of the product.

68.9.3 Static Hedging to Reduce Interest Rate Risk

In both of the above applications, we can create a static hedge, using market-traded instruments, to reduce the spread in our prices and make them more competitive. It is important to note that these new prices are only valid as long as we create the static hedge in reality (on trading the contract).

We can statically hedge with any product in the market. It does not have to be similar to the contract we are pricing. However, the most effective static hedges are likely to be those which are made up of contracts similar in form to that being priced. Hence, the more exotic a contract, the fewer similar contracts that will be available and the larger the spread, as we would expect.

This form of hedging is 'static' in that, once we have set up a hedge, we should leave it until maturity of our contract. However, if the market changes significantly, then a rehedging of the contract may yield a different static hedge, with a higher worst-case (or lower best-case) price. Subject to transaction costs, updating our static hedge guarantees a better worst-case (or best-case) price and 'locks in' an additional profit of the difference between this price and the old one.

68.9.4 Risk Management: A Measure of Absolute Loss

Conventional risk management models produce results of the form 'a maximum loss of X with probability Y.' Since our spreads for prices are absolute, they give us a good measure of the total risk in a portfolio, and we can predict the absolute maximum loss possible.

Furthermore, by static hedging, we can generate an optimal static hedge to reduce this interest rate risk. This would not be possible with a linear model calibrated to the yield curve. This is because the price of a contract would be invariant to the addition of hedging instruments if they had originally been used to construct the yield curve. Hence, a further use of the model is to construct static hedges for portfolios, regardless of which interest rate model we choose to finally price them.

68.9.5 A Remark on the Validity of the Model

In Chapter 70 I describe various extensions to our model. These extensions allow us to construct tighter bounds for the model which give reasonable spreads for prices. We can still be confident that actual interest rate movements will definitely lie within our bounds. With these additions, the applications that we have described above can become feasible in the marketplace.

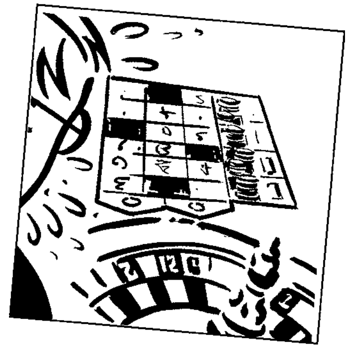
68.10 SUMMARY

In this chapter I have presented a model for valuing and hedging interest rate securities. The framework is that of 'worst-case scenarios' and we have derived a non-linear first-order partial differential equation for the value of a portfolio of products. Since the equation is nonlinear we find that the value of a product depends on the rest of the portfolio. This nonlinearity gives the model both advantages and disadvantages compared with other, more traditional approaches. The main disadvantage is that to obtain the full benefit of the model one must solve the equation for the entire portfolio. The principal advantage is that optimal hedges can be found, maximizing the portfolio's value. A further advantage of the model is that one can be fairly confident in the accuracy of the model parameters. In the next two chapters we look at the pricing and hedging of derivatives and at extensions to the model.

FURTHER READING

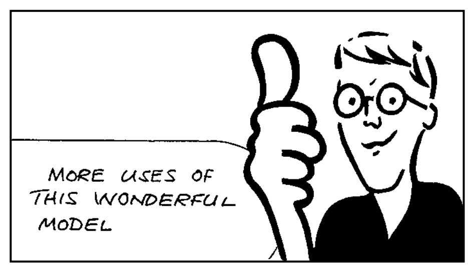
- See Epstein & Wilmott (1997, 1998) for examples of pricing other fixed-income instruments.
- See Lewicki & Avellaneda (1996) for an interest rate model in the HJM framework but assuming uncertain volatility.

CHAPTER 69
pricing and optimal hedging of derivatives, the non-probabilistic model cont'd



In this Chapter...

- bond options
- index amortizing rate swaps
- convertible bonds



69.1 INTRODUCTION

This chapter starts with an examination of the pricing and hedging of a real-life leasing portfolio. We then look at using the non-probabilistic methodology for the pricing and optimal static hedging of a variety of derivatives. In particular, we look at bond options, both European and American, options with complex decisions, index amortizing rate swaps and convertibles. Between them these contracts have a wide variety of interesting exotic features.

69.2 A REAL PORTFOLIO

In Figure 69.1 are shown the sterling-denominated cashflows of Dresdner, Kleinwort, Benson arising from a leasing portfolio. Let's see how we might use the model to hedge this portfolio. First of all, we will calculate the present value of all the cashflows using the yield curve today, 8th January 1998. This yield curve is shown in Figure 69.2. The present value is found to be −£3,539,362. To test the sensitivity of this value to shifts in the yield curve, I show in Table 69.1 the present value of the cashflows assuming instantaneous parallel shifts in the yield curve of various magnitudes. These give a traditional way of determining whether or not a position is hedged against moves in the yield curve. If we are being pessimistic then a 2% downward shift in the yield curve reduces the portfolio's value to −£4,432,153.

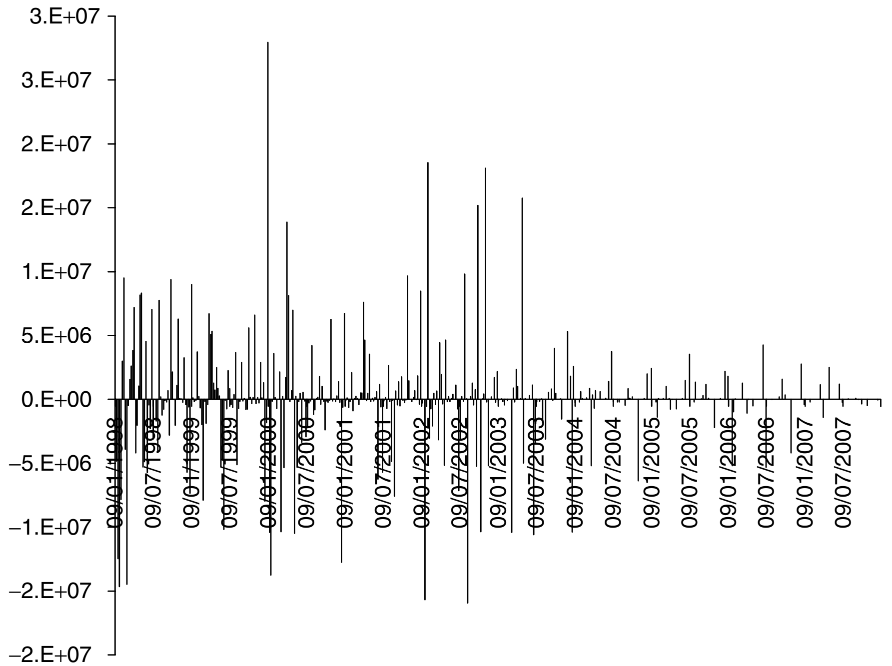


Figure 69.1 Cashflows of Dresdner, Kleinwort, Benson.

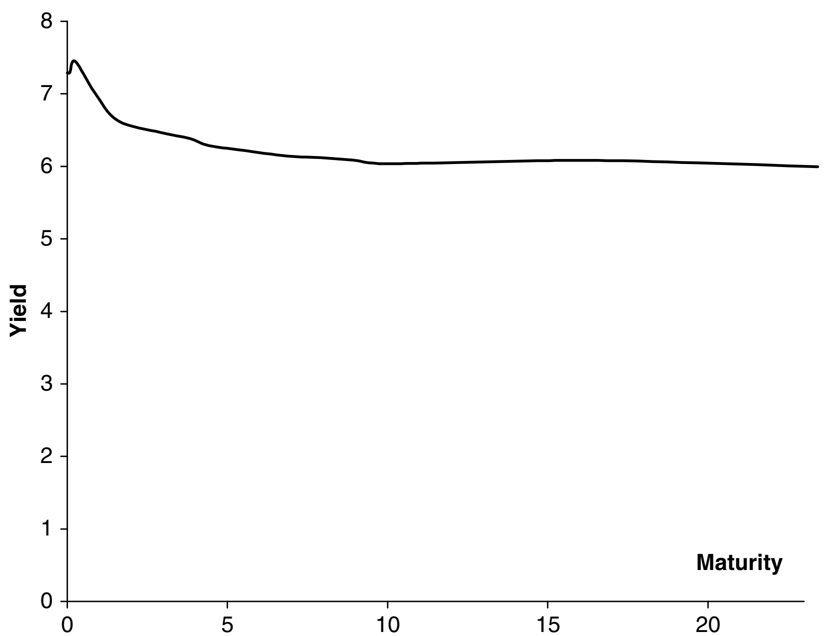


Figure 69.2 Current yield curve.

Table 69.1 Present value of portfolio with parallel shifts in the yield curve.

Shift in curve	Present value (£)
2%	−2,620,568
1%	−3,083,010
0%	−3,539,362
−1%	−3,989,193
−2%	−4,432,153

Table 69.2 Benchmark bonds.

	Coupon	Maturity	Market price (£)
1M T-BILL	0	04 FEB 1998	99.46
3MB T-BILL	0	08 APR 1998	98.15
1Y GILT	12	20 NOV 1998	104.125
2YB GILT	6	10 AUG 1999	99.125
3Y GILT	8	07 DEC 2000	104.03125
4Y GILT	7	06 NOV 2001	102.09375
5YB GILT	7	07 JUN 2002	102.75
6Y GILT	8	10 JUN 2003	108.09375
7Y GILT	6.75	26 NOV 2004	103.40625
8Y GILT	8.5	07 DEC 2005	114.75
9Y GILT	7.5	07 DEC 2006	109.625
10YB GILT	7.25	07 DEC 2007	108.96875
15Y GILT	8	27 SEP 2013	119.21875
20YB GILT	8.75	25 AUG 2017	130.75
25YB GILT	8	07 JUN 2021	125.03125

The yield curve was estimated from the market values of benchmark bonds. These bonds are shown in Table 69.2. We are going to use these bonds, which are coupon bearing except for the first two, for hedging purposes.

Using the Epstein–Wilmott model, with the same parameters as before, we find that the worst-case value for the portfolio as it stands is −£6,135,878, and the best-case value is −£1,898,173. If we are being pessimistic then a value of −£6,135,878 is perfectly possible. This is significantly lower than the present value of −£3,539,362 calculated off the yield curve today. It is also significantly lower than the value we calculated above using a 2% parallel shift in the yield curve. Is a 2% shift in the yield curve realistic? Remember that the cashflows go out to the year 2008. The price calculated using the non-linear model is far more robust and reliable.

Now we will hedge the portfolio using the benchmark bonds so as to optimize the worst-case value. The quantities of each of the hedging bonds giving the optimal hedge is shown in Table 69.3. Note the large quantity of the three-year bond sold short. The cashflows of these bonds are shown in Figure 69.3. The worst-case value for the original portfolio hedged with these bonds is now −£4,009,581. Again this is a very robust pessimistic value for the portfolio, when optimally hedged. It is hedged against a far broader set of future scenarios than a simple yield curve shift and is consequentially far more reliable. Moreover, it is only −£470,219 less than

Table 69.3 Optimal hedge for the leasing portfolio.

Hedging bond	Hedge quantity
1M	762
3M	556
1Y	745
2Y	−7112
3Y	−123013
4Y	7064
5Y	−30844
6Y	1689
7Y	3135
8Y	−6944
9Y	−285
10Y	11674
15Y	−122
20Y	1456
25Y	−1299

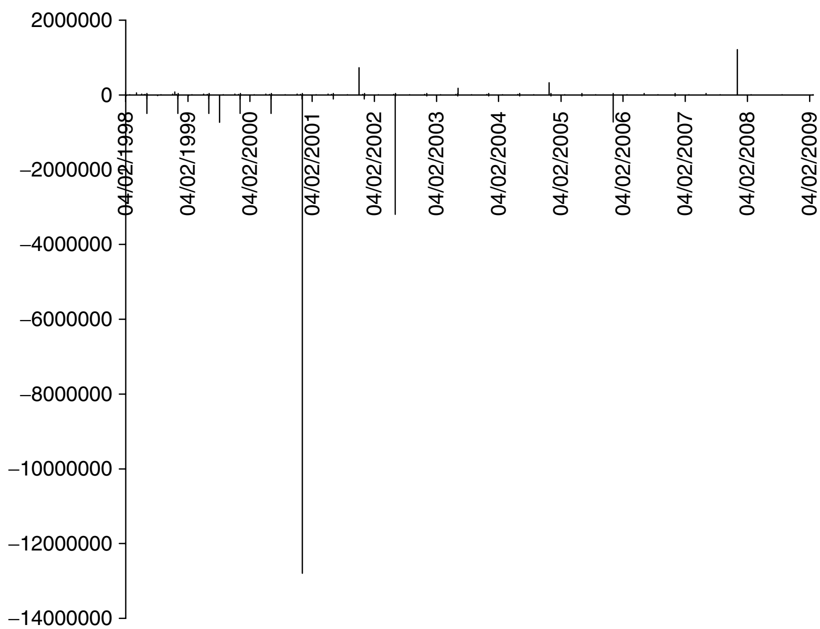


Figure 69.3 Optimal hedging portfolio cashflows.

the simple present value off the yield curve. By optimally hedging the worst-case has improved from −£6,135,878 to −£4,009,581, simply by hedging with benchmark bonds. (As a matter of interest, the best-case value when hedged with these same bonds is −£3,112,337.)

A summary of the prices using different models is shown in Table 69.4.

Table 69.4 Summary of the prices using different models.

Model	Value (£)
Deterministic PV	−3,539,362
2% downward shift	−4,432,153
Worst case, unhedged	−6,135,878
Best case, unhedged	1,898,173
Worst case, hedged	−4,009,581

69.3 BOND OPTIONS

Now consider the bond option. This contract includes optionality in its specification. I will show how to price this in the Epstein–Wilmott world.

69.3.1 Pricing the European Option on a Zero-coupon Bond

We consider a European option with a zero-coupon bond as the underlying. The bond has a principal payment of 1 at time T_Z. The option expires at time $T_O < T_Z$ and has payoff $\Lambda(Z)$ e.g. for a long call option,

$$\Lambda(Z) = \max(Z - E, 0),$$

where E is the exercise price of the option.

We will develop a general approach to the pricing of contracts with optionality by considering the cases in which we exercise and do not exercise the option separately, i.e. we consider all of our options individually and then choose the appropriate course of action. The drawback to this approach will be that for each instance of 'either/or' optionality, we double the number of cases to be considered.

Let Π_0 be the overall portfolio of cashflows that we would have if we chose to exercise the option at expiry. This consists of the cashflows due to the hedging instruments plus the cashflows that we would receive if we were to exercise the option. In the case of a call option, the latter cashflows would be those of the underlying bond (for a put option, they would be the cashflows for the short bond). We also let Π_1 be the portfolio of cashflows that we would have if we did not exercise the option (i.e. just those due to the hedging instruments). So, subscript 0 means exercised and subscript 1 means not exercised.

We solve Equation (68.3) with the appropriate final and jump conditions (dependent on the nature of the hedging instruments and option payoff) to find the value of the portfolio in a worst-case scenario at expiry, when we do exercise the option. This is $\Pi_0^-(r, T_0)$. We also solve Equation (68.3) with the appropriate final and jump conditions to find the value of the portfolio in a worst-case scenario at expiry, when we do not exercise the option. This is $\Pi_1^-(r, T_0)$. Again, superscript'−'denotes worst case.

Since we are long the option, we have control over whether or not to exercise, and so we set the value of the portfolio at expiry of the option to be the most valuable of the two courses of action, where we take the exercise price into account. For a call option, this is the maximum of the value of the portfolio when we do exercise minus the exercise price and the value when we do not exercise, i.e.

$$\Pi^-(r, T_O) = \max\left(\Pi_0^-(r, T_0) - E, \Pi_1^-(r, T_0)\right).$$

(For a put option, we would add rather than subtract the exercise price, since the holder of the option receives the exercise price at expiry.)

We then solve Equation (68.3) with $\Pi^{-}(r, T_O)$ as final data and appropriate jump conditions (for the hedging instruments) to find the current worst-case scenario value of the portfolio. To find the marginal worst-case value of the option, we then subtract the cost of the static hedge. Finally, we can maximize with respect to the hedge quantities to find the optimal worst-case scenario value for the option.

We can also find the value of the option in a best-case scenario. We solve the relevant equation with appropriate final and jump conditions to find the best-case values of Π_0 and Π_1 at expiry of the option, $\Pi_0^{+}(r, T_0)$ and $\Pi_1^{+}(r, T_0)$ respectively. We then set the value of the portfolio at expiry of the option to be the most valuable course of action. For example, for a call option,

$$\Pi^{+}(r, T_O) = \max\left(\Pi_0^{+}(r, T_0) - E, \Pi_1^{+}(r, T_0)\right).$$

We then use $\Pi^{+}(r, T_O)$ as final data and appropriate jump conditions to find the current best-case scenario value of the portfolio. We then subtract the cost of the static hedge to find the marginal best-case value of the option. Again, we can optimize the result and minimize with respect to the hedge quantities to find the optimal best-case value.

Example

Price vanilla European call and put options with expiry in one year, on a zero-coupon bond with principal 1 and maturity in five years. The current market price of the bond is 0.687. The spot short-term interest rate is 6% and the parameters of our model are the usual. We hedge with the hedging bonds of Table 68.1.

The results when we price a call option, with exercise price 0.5, are shown in Tables 69.5 and 69.6 and those when we price a put option, with exercise price 0.9, are shown in Tables 69.7 and 69.8. Figure 69.4 shows the value of the put option in a worst-case scenario under the various hedging strategies.

The static hedges still include approximately one of the underlying bond (short for the call, long for the put) although the specific quantities of this bond have altered slightly. The hedges also include sizable amounts of the shorter-dated bonds. With these extra hedging instruments it has been possible to reduce the spread to a level which is of the same magnitude as the bid-offer spread seen in practice.

From Figure 69.4 we can see that although the extra hedging instruments have not had a particularly noticeable effect in raising the worst-case price (over and above that when we hedged with the underlying), they have 'flattened out' the curve and this must correspond to a significant decrease in the interest rate risk in the portfolio.

Table 69.5 Value of the optimally-hedged European call option.

Call, $E = 0.5$	Worst case	Best case
No hedge	0.009	0.372
Optimally hedged	0.220	0.221

Table 69.6 The optimal static hedges for the European call option.

Hedging bond	Maturity (years)	Worst case hedge quantity	Best case hedge quantity
A	0.5	0.012	0.082
B	1	0.488	0.456
C	2	0.001	0.002
D	3	−0.002	−0.008
E	5	−1.004	−0.994
F	7	0.006	0.000
G	10	0.000	0.000

Table 69.7 Value of the optimally-hedged European put option.

Put, $E = 0.9$	Worst case	Best case
No hedge	0.012	0.361
Optimally hedged	0.152	0.153

Table 69.8 The optimal static hedges for the European put option.

Hedging bond	Maturity (years)	Worst case hedge quantity	Best case hedge quantity
A	0.5	−0.136	−0.034
B	1	−0.826	−0.874
C	2	−0.003	−0.003
D	3	0.008	0.008
E	5	0.996	0.995
F	7	−0.001	−0.001
G	10	0.000	0.000

69.3.2 Pricing and Hedging American Options

For the European option, we compared the value of two portfolios at expiry of the option and just picked the course of action that had the higher value. However, for the American option, we may also exercise at earlier times. We must therefore be aware that at any time before expiry, it may be optimal to exercise rather than hold the option. This presents itself as a constraint on the value of the portfolio in which we continue to hold the option.

We again consider two portfolios, one containing the cashflows we would have if we were to exercise the option, the other containing those we would have if we continued to hold the option. We consider an option with the same specification as before, with the one exception that the holder now has the right to exercise the option at any time before T_O. We let $\Pi_0(r, t)$ be the overall portfolio of cashflows that we would have at time t if we were to exercise the option at time t and $\Pi_1(r, t)$ be the portfolio of cashflows that we would have if we continued to hold the option at time t (this does not include any cashflow due to the option payoff at expiry).

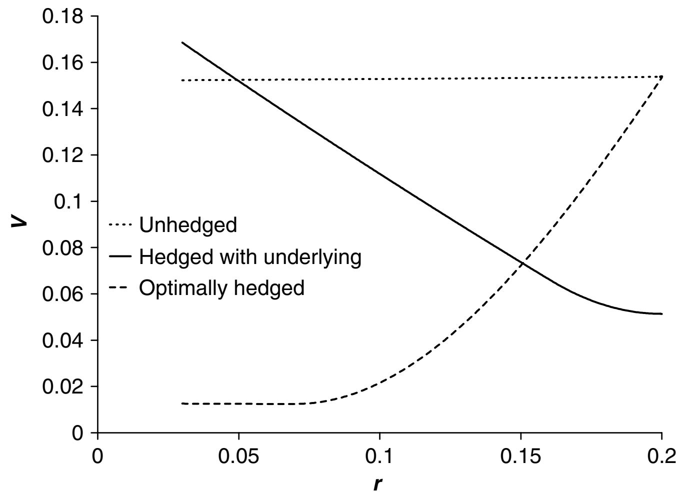


Figure 69.4 European put option value in a worst-case scenario.

We solve Equation (68.3) with the appropriate final data and jump conditions to find the worst-case value of the portfolio when we do exercise at time t, $\Pi_0^-(r, t)$. This tells us what our portfolio payoff would be if we decided to exercise at time t.

We then solve Equation (68.3) with the appropriate final data and jump conditions to find the worst-case value of the portfolio when we continue to hold the option at time t, $\Pi_1^-(r, t)$. In the absence of arbitrage, we would exercise the option if the value of the portfolio, with the exercise price taken into account, were greater than the current value of the portfolio when we continue to hold the option. For the call option, this gives us the additional constraint

$$\Pi_1^-(r, t) \geq \Pi_0^-(r, t) - E,$$

during the period in which we are allowed to exercise the option (in this case, for $t \leq T_O$). For the put option, the constraint would be

$$\Pi_1^-(r, t) \geq \Pi_0^-(r, t) + E.$$

The marginal worst-case value of the option at time t is then the value of the portfolio in which we still hold the option, $\Pi_1^-(r, t)$, minus the cost of the static hedge. We can maximize with respect to the hedge quantities to find the optimal marginal value.

We can also find the best-case scenario value of the option.

Example

We price the vanilla American put option with expiry in one year and exercise price E, on a zero-coupon bond with principal 1 and maturity in five years. The current market price of the bond is 0.687. The spot short-term interest rate is 6% and the parameters of our model are the usual.

Initially, we solely hedge with the underlying zero-coupon bond. The results for the option valuation, without hedging and with the optimal static hedges for both worst- and best-case valuations, are shown in Table 69.9.

Table 69.9 Value of an American put option hedged with the underlying.

$E = 0.8$	Worst case	Best case
No hedge	0.000	0.328
Optimal hedged	0.112	0.142
$E = 0.9$	Worst case	Best case
No hedge	0.048	0.428
Optimal hedge on worst-case	0.212	0.213
$E = 1.0$	Worst case	Best case
No hedge	0.148	0.528
Optimal hedge on worst-case	0.312	0.313

Table 69.10 Value of the optimally-hedged American put option.

Put, $E = 0.9$	Worst case	Best case
No hedge	0.048	0.428
Optimally hedged on worst-case	0.212	0.213

Table 69.11 The optimal static hedges for the American put option.

Hedging bond	Maturity (years)	Worst case hedge quantity	Best case hedge quantity
A	0.5	−0.002	−0.036
B	1	0.002	−0.001
C	2	−0.002	−0.003
D	3	0.008	0.009
E	5	0.994	0.994
F	7	0.000	−0.001
G	10	0.000	0.000

We then hedge the option with the bonds of Table 68.1. The results when we price a put option, with exercise price 0.9, are shown in Tables 69.10 and 69.11. Figure 69.5 shows the value of the unhedged option in a worst-case scenario.

The further an option is 'in the money' (whether American or European) the more likely it is to be exercised and the nearer the quantity of the underlying, in the static hedge, is to unity. The spread in price also decreases because we are effectively valuing the exercised option and we can hedge this very efficiently with the underlying bond.

The spreads for the American option are therefore smaller than those for the European option and the hedge quantities of the underlying are larger. This is because the American option has more exercise opportunities and is consequently more likely to be exercised.

If we were to price the American call option then we would find that it had the same value as the European call option. This is because the underlying is a zero-coupon bond. It is an equivalent result to the equality of American and European call options on equities which do

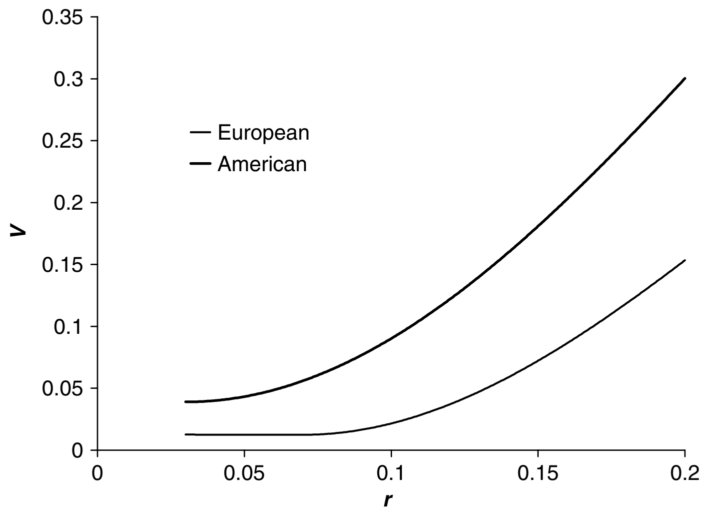


Figure 69.5 American put option value in a worst-case scenario.

not pay a dividend. If we were to price a call option on a coupon bond (where the coupon was paid before expiry of the option) then the American option would be worth more than the European, since the holder of the option would not receive the coupon, whereas the holder of the bond would.

We remark that we can also use this approach to value Bermudan options (options with exercise allowed only on or between specified dates). To value such an option, we proceed as for the American option. However, when we come to solve our partial differential equation for Π_1, we only include the relevant constraint at (or between) times when exercise is allowed.

Unfortunately, there is a drawback to the approach described above. For each instance of either/or optionality, we must double the number of cases to be considered. This means that to value a portfolio which includes N instances of optionality (N vanilla options, for example), we must price 2^N separate portfolios. This can quickly become computationally intensive.

69.4 CONTRACTS WITH EMBEDDED DECISIONS

The flexiswap requires the holder to make a series of decisions during the life of the contract. The swap has M possible cashflow dates, on each of which, the holder may choose to exchange a floating rate payment for a fixed rate payment. Moreover, the holder must do so on exactly m occasions (where $m \leq M$).

The fixed rate payments are the interest that would be due on some principal due to a predetermined fixed rate. The floating rate payments are the interest that would be due on the principal due to some designated interest rate, the reference rate.

We solve the problem, in a worst-case scenario, by approximating this designated interest rate using the short-term interest rate, r. This is because we must make a decision on each cashflow date and we cannot do this without knowledge of the value of the entire cashflow. We will then hedge the contract with vanilla swaps, and to be consistent, we will also value these using the approximation approach. As I note later, it is possible to value the contract using the swap decomposition approach, but it can be impractical to do so.

We assume that the M cashflow dates are on the τ-period interest rate and are τ apart. We designate these dates by T_j for $j = 1, 2, \ldots, M$. If the holder chooses to exchange interest payments on one of these dates, then he receives a cashflow of the form,

$$\tau(r - r_f),$$

where the principal is assumed to be 1 and r_f is the fixed rate.

To price the contract, we must consider separately the cases when there are i cashflows still to be taken, for $i = 0 \ldots m$. We therefore introduce $m + 1$ functions, $V(r, t, i)$. The index i represents the number of cashflows that the holder has left to take before maturity and V is the subsequent value of the contract.

Solve Equation (68.3) with suitable final, jump conditions and constraints, to find the value of each of these functions today. The value of our contract in a worst-case scenario is then $V(r, t, m)$.

Clearly

$$V(r, t, 0) = 0,$$

since there are no cashflows left to be taken.

At each possible swaplet date, we must apply either a jump condition, if we are forced to take the cashflow, or a constraint, if we are not obliged to take the cashflow and only do so in the case that it is the optimal decision.

If there are i cashflows left to choose between, and we still have to take a total of i cashflows, then we must take every cashflow left. Mathematically, we represent this as,

$$V(r, T_{M-i+1}, i) = V(r, T_{M-i+1}, i - 1) + \tau(r - r_f),$$

for $i = 1, 2, \ldots, m$.

On the other hand, if we are at a cashflow date, and there are more cashflow dates left than the number of cashflows that we are obliged to take, then we only take the cashflow if it is optimal to do so. This is the case when the value of the contract is less than the value of the contract with one less cashflow left to take plus the value of the cashflow. We therefore have the constraint,

$$V(r, T_j, i) \geq V(r, T_j, i - 1) + \tau(r - r_f),$$

for $M - j + 1 > i$.

To hedge the contract, we just include the appropriate jump conditions for the hedging cashflows, price the overall portfolios $\Pi(r, t, i)$, for $i = 1, 2, \ldots, m$ and then subtract the cost of the static hedge to find the marginal contract value. In this case, $\Pi(r, t, 0)$ no longer has zero value, but represents the value of the hedging instruments when there is no swap contract.

Example

We price an eight-choice swap. The swap has a principal of \$1,000,000 and a fixed rate of 7%. The holder must take four of the eight possible cashflows, which are every six months, with the first cashflow in one year's time.

The spot short-term interest rate is 7.4% and the parameters of our model are the usual.

We hedge the contract with the market-traded swaps from Table 69.12. They are par swaps with a principal of \$100,000 and payment dates biannually until their maturity.

Table 69.12 The hedging par swaps

Hedging Swap	Maturity (years)	Fixed Rate (%)
A	2	7.10
B	3	6.85
C	4	6.69
D	5	6.58
E	6	6.49
F	7	6.42
G	8	6.39
H	9	6.36
I	10	6.34

Table 69.13 Value of an eight-choice swap.

	Worst case	Best case
No hedge	−68784.4	108726.3
Optimally hedged	−26655.1	3774.1

Table 69.14 The optimal static hedges for the eight-choice swap.

Hedging swap	Worst case hedge quantity	Best case hedge quantity
A	0.302	1.725
B	0.859	4.968
C	−6.215	−10.009
D	0.000	0.000
E	0.000	0.000
F	0.000	0.000
G	0.000	0.000
H	0.000	0.000
I	0.000	0.000

The results for the worst- and best-case valuations are shown in Tables 69.13 and 69.14.

It is also possible to price this contract using the swap decomposition approach. However, if we decompose the swap cash flows into sets of zero-coupon bonds, then we will not be able to make the decision on whether or not to take a cashflow at the cashflow date. This is because one of the bond cashflows will be before this date. To price the swap, we must use the same approach to optionality as for the general bond option problem. At each possible cashflow date, we consider the cases when we do and do not choose to take the cashflow separately.

69.5 THE INDEX AMORTIZING RATE SWAP

We saw the IAR swap in Chapter 32, here we will see how to value it using the model of this chapter. This is a particularly good example of the use of this model because static hedging of the IAR swap with vanilla swaps can be done in an optimal fashion.

As before we must introduce a new state variable. This new state variable is the current level of the principal, again denoted by P. The value of the IAR swap is $V(r, P, t)$.

The variable P is deterministic and jumps to its new level at each resetting. Since P is piecewise constant, the governing differential equation for the value of the swap is, in the present model,

$$\frac{\partial V}{\partial t} + c\left(\frac{\partial V}{\partial r}\right)\frac{\partial V}{\partial r} - rV = 0.$$

At each reset date there is an exchange of interest payments and an amortization of the principal. If we use t_i to denote the reset dates and r_f for the fixed interest rate, then the swap jumps in value by an amount $(r - r_f)P$. Subsequently, the principal P becomes $g(r)P$ where the function $g(r)$ is the representation of the amortizing schedule. This gives us the jump condition

$$V(r, P, t_i^-) = V(r, g(r)P, t_i^+) + (r - r_f)P.$$

At the maturity of the contract there is one final exchange of interest payments, thus

$$V(r, P, T) = (r - r_f)P.$$

The problem is nonlinear, and must be solved numerically. The structure of this particular IAR swap is such that there is a similarity reduction; just look for a solution of the form

$$V(r, P, t) = PH(r, t).$$

We are lucky that the similarity reduction is not affected by the nonlinearity.

The problem for H is

$$\frac{\partial H}{\partial t} + c\left(\frac{\partial H}{\partial r}\right)\frac{\partial H}{\partial r} - rH = 0,$$

i.e. H also satisfies Equation (68.3), with the jump condition

$$H(r, t_i^-) = g(r)H(r, t_i^+) + (r - r_f),$$

over a payment date, and the final condition

$$H(r, T) = r - r_f.$$

Example

We value an index amortizing rate swap with a maturity of $5\frac{1}{2}$ years, fixed interest rate of 6.5% and initial principal \$1,000,000. Payment dates occur once a year until maturity, with the first payment $3\frac{1}{2}$ years from today. The amortizing schedule is shown in Figure 69.6.

The spot short-term interest rate is 7.40%. We set the parameters of our model to be the usual.

We hedge the contract with the market-traded swaps from Table 69.12 (these were the actual market prices on 14th May 1998). They are par swaps with a principal of \$100,000 and payment dates biannually until their maturity. The yield curve for these swaps is shown in Figure 69.7.

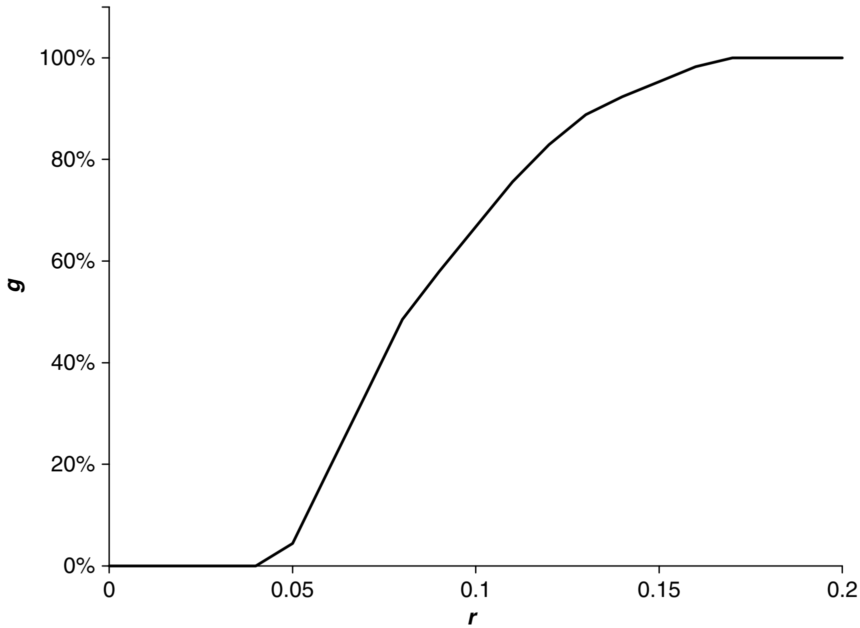


Figure 69.6 The amortizing schedule.

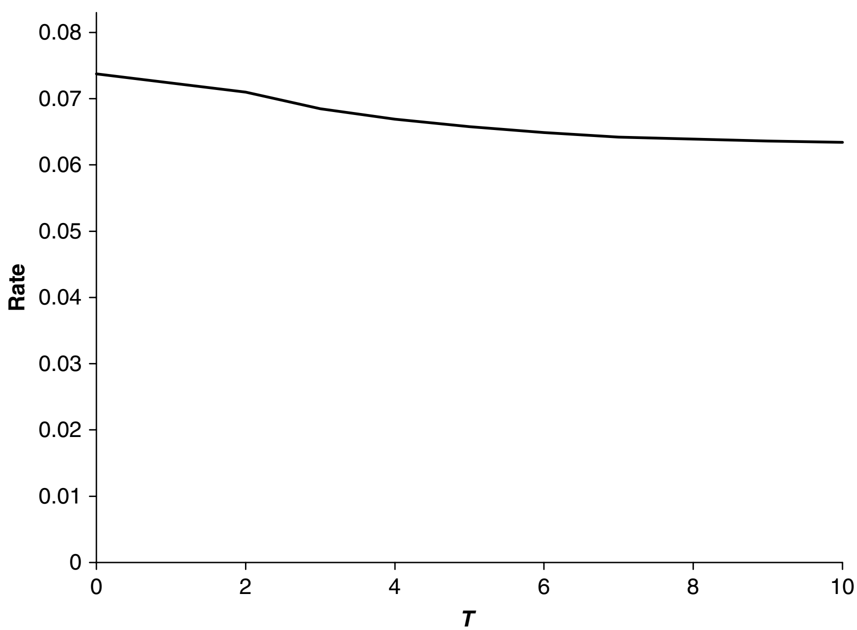


Figure 69.7 The swap curve.

Table 69.15 Value of the index amortizing rate swap.

	Worst case	Best case
No hedge	−30820.2	216336.6
Optimally hedged	−11753.2	143129.6

Table 69.16 The optimal static hedge for the index amortizing rate swap.

Hedging swap	Worst case hedge quantity
A	−0.008
B	9.433
C	−9.528
D	−0.009
E	0.001
F	0.000
G	−0.001
H	0.001
I	0.000

Table 69.17 Sensitivity of the index amortizing rate swap value to the yield curve.

Yield Shift (%)	Swap Value
+2	21552.8
+1	8297.5
0	−2440.3
−1	−10908.0
−2	−18786.6
−3	−28219.6

The results of the valuation for the index amortizing rate swap, without hedging and with the optimal static hedge for the worst-case scenario, are shown in Table 69.15. The optimal static hedge for this valuation is shown in Table 69.16.

The use of a static hedge has significantly improved the value of the index amortizing rate swap in a worst-case scenario. The optimal hedge does not consist of all of the possible contracts, and is essentially composed of the few contracts that are most similar in form to the cashflows of our swap.

We can also price the index amortizing rate swap directly from the yield curve of Figure 69.7, using linear interpolation between points. In this case, we find that the swap is worth −2440.3. We can then test the sensitivity of the valuation to shifts in the yield curve. The results are shown in Table 69.17. We note that the hedged worst-case scenario price is similar in value to the swap price with a −1% yield shift (a comparable result to that found for the leasing portfolio in Section 69.2).

69.6 CONVERTIBLE BONDS

The model can easily be applied to contracts with other asset-price dependence such as convertible bonds. In this case the pricing equation is

$$\frac{\partial V}{\partial t}+\tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+rS\frac{\partial V}{\partial S}+c\left(\frac{\partial V}{\partial r}\right)\frac{\partial V}{\partial r}-rV=0,$$

where the bond value is a function of S, r and t: $V(S,r,t)$. The usual payoff condition and constraints apply as discussed in Chapter 33.

Worst-case results are shown in Figure 69.8 for a two-year CB with conversion allowed at any time. In Figure 69.9 is plotted the difference between the best and the worst cases.

Example

In this example we value a convertible bond using three different models. The convertible bond matures in 18 months, is convertible into 0.5 of the underlying at any time and there are 3% coupons paid twice *per annum*. The underlying asset is currently 1.5, and has a dividend yield of 3% and a volatility of 25%.

- The simplest model is the stochastic asset model having a constant interest rate. With an interest rate of 7% this model gives the CB value of 1.058.
- The second model is Vasicek with

$$dr=(a(t)-br)\,dt+\nu\,dX_2$$

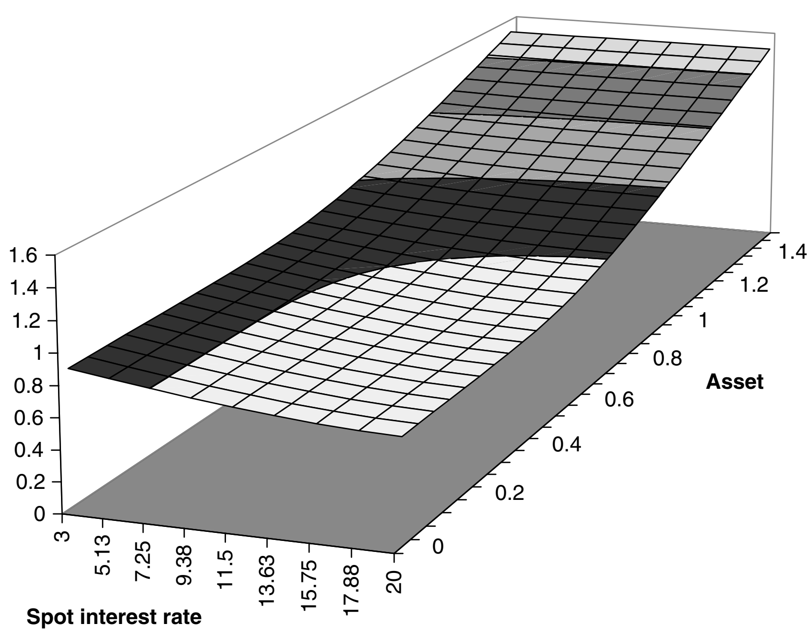


Figure 69.8 The worst-case value of a convertible bond using the Epstein–Wilmott interest rate model.

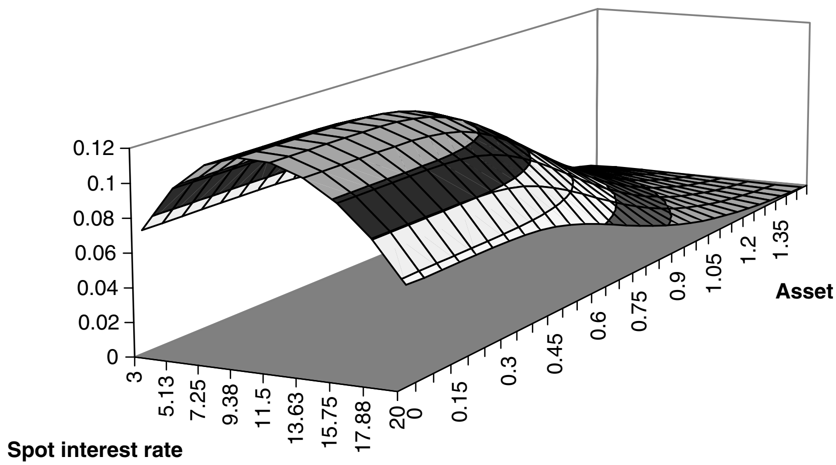


Figure 69.9 The spread in values of a convertible bond using the Epstein–Wilmott interest rate model.

with $b = 0.1$, $\nu = 0.02$ and a correlation of 0.1. This has been *fitted* to a flat 7% yield curve. This model gives a value of 1.059. The results from these two models are very close. But are these values realistic?

- The final example is of the Epstein–Wilmott model worst- and best-case analyses. In the worst case we get a value of 1.042 and in the best case 1.089. This spread is very large, yet it is a very robust estimate of the bond's range of values. Can we improve these values by static hedging?

Well, how good is Vasicek first of all? If we increase the parameter b to 0.11 the bond value increases by 0.001, if ν is increased to 0.25 the bond value increases by 0.002. It's very easy to get quite a wide range of Vasicek prices by playing around with the parameters, and this still assumes that the fitting is 'correct.'

Now let's see if we can improve the Epstein–Wilmott values by hedging.

Table 69.18 shows three zero-coupon bonds with which we can hedge and their market prices. In the third column are the optimal quantities of each of the bonds that maximize the worst-case value of the convertible bond. With this hedge in place the worst case is improved from 1.042 to 1.052. This price is far less sensitive to the interest rate model than the other prices we have obtained with the other two models. When optimally hedged the best-case value is 1.067. The spread has been reduced from $1.089 - 1.042 = 0.047$ to $1.067 - 1.052 = 0.015$, a factor of three.

Table 69.18 Universe of hedging bonds, market prices and optimal hedge quantities.

Maturity	Market price	Hedge (worst)
0.5	0.9656	−0.7350
1	0.9323	−0.1639
2	0.8693	−0.4882

Table 69.19 Summary of CB prices for different models.

Model	Value
Constant interest rate	1.058
Vasicek	1.059
Worst case, unhedged	1.042
Best case, unhedged	1.089
Worst case, hedged	1.052

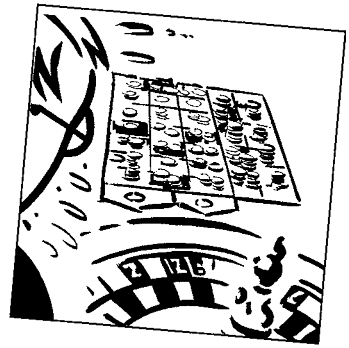
A summary of the prices using different models is shown in Table 69.19.

We have applied our non-linear, non-probabilistic interest rate model to the pricing of a convertible bond. The resulting worst-case scenario valuation produced results which were far lower in price than those found using typical interest rate models. However, unlike the worst-case scenario approach, these models were found to be quite sensitive to their parameters, the values of which can often be uncertain.

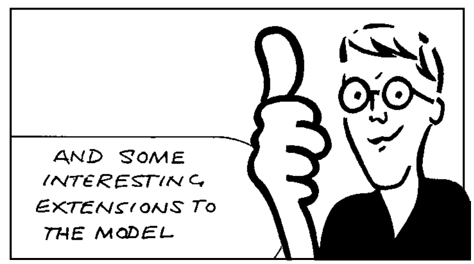
Through the process of static hedging, we were able to increase the worst-case scenario price significantly for the convertible bond. In addition, the process found an optimal static hedge. This hedge could be applied to a convertible bond portfolio to reduce the inherent interest rate risk, regardless of which model were chosen to price the resulting portfolio.

69.7 SUMMARY

For obvious reasons, I am particularly fond of this model. Being nonlinear it sometimes presents computational problems. These are not because the pde is hard to solve numerically, it is not, but because we must lump together all contracts in our portfolio to achieve the best marginal values and the best hedges.



CHAPTER 70
extensions to the non-probabilistic interest-rate model



In this Chapter...

- more on the EW model

70.1 INTRODUCTION

There is no end to the number of bells and whistles that can be attached to financial models. We can incorporate many of the features that we have seen in other chapters into the present model. The obvious improvement, which will not take much more computational effort, is to include jumps of some form. It is common experience that rates do jump. In some countries these jumps are orchestrated by governments. The crash models of Chapter 58 are a nice addition to the model since they are also based on a worst-case scenario analysis.

We also consider other extensions to the non-probabilistic model. For example, we explore the possibility of modeling economic cycles. Then we introduce the uncertainty band. This allows us to model more accurately real interest rate movements. We then examine past data to choose a sensible width for the band. Finally, we consider the impact of liquidity.

70.2 FITTING FORWARD RATES

As a first thought, consider taking the Epstein–Wilmott model explained in the previous chapters and making it closer in spirit to both the classical stochastic models and the duration type of models. Suppose that we believe interest rates to be completely deterministic and governed by the equation

$$\frac{dr}{dt} = a(b(t) - r). \tag{70.1}$$

That is, the rates are mean reverting to some level $b(t)$ at a rate a. This is similar to a fitted Vasicek, but without the random element. If this is true and markets price according to this model, then the forward rate curve today $F(t^*;T)$ must satisfy

$$\frac{d}{dt}F(t^*;t) = a(b(t) - F(t^*;t)).$$

It follows that

$$b(t) = F(t^*;t) + \frac{1}{a}\frac{d}{dt}F(t^*;t). \tag{70.2}$$

So our interest rate model is just (70.1) with $b(t)$ given by (70.2). Now add a 'margin of error' to this model so that it becomes

$$a(b(t) - r) - c < \frac{dr}{dt} < a(b(t) - r) + c.$$

This puts the model firmly into the Epstein–Wilmott framework. This model combines the best of all possible model worlds. The margin of error c can be found by 'fitting' the function $b(t)$ many, many times using historical forward-rate data, and then determining by how much this function is in error.

70.3 ECONOMIC CYCLES

There is evidence that interest rates follow economic cycles with a period of around five to ten years. Can we incorporate this observation into our model for the short-term interest rate?

A starting point for such a model could be to say that an economic cycle could be represented by simple harmonic motion, i.e. by the solution of

$$\frac{d^2r}{dt^2} = a - \omega^2 r.$$

The period of this cycle is $2\pi/\omega$, with the cycle centered on $r = a/\omega^2$.

Of course, we cannot be sure about either the period or the mean value of r; cycles have been observed but they are far from being predictable.

Accordingly, we modify the above equation for the evolution of r to allow for that uncertainty:

$$a^- - {\omega^-}^2 r \le \frac{d^2r}{dt^2} \le a^+ - {\omega^+}^2 r.$$

More generally we could write

$$a^-(r,s) \le \frac{d^2r}{dt^2} \le a^+(r,s),$$

where

$$s = \frac{dr}{dt}.$$

The governing equation for the worst-case value, $V(r,s,t)$, of a product is then

$$\frac{\partial V}{\partial t} + s\frac{\partial V}{\partial r} + a\left(r, s, \frac{\partial V}{\partial s}\right)\frac{\partial V}{\partial s} - rV = 0,$$

where

$$a(r, s, x) = \begin{cases} a^-(r, s) & \text{for} \quad x \geq 0 \\ a^+(r, s) & \text{for} \quad x < 0. \end{cases}$$

Constraints on r and s will result in proscribed regions of r, s, t space.

70.4 INTEREST RATE BANDS

Interest rates do have a stochastic nature that is not quite captured by the present model. This stochastic nature may or may not be Brownian motion, and it may or may not matter in practice. We can easily extend the model to allow for interest rate movements that are practically indistinguishable from the real behavior of rates. This extension is the introduction of **uncertainty bands**. We subtly redefine r to be some estimate of the spot interest rate that is always within a distance ϵ of the *real* short-term rate. Thus

$$|\text{real spot rate} - r| \leq \epsilon.$$

The details of the worst-case analysis are obvious, we must discount at either $r + \epsilon$ or $r - \epsilon$ depending on the sign of the value V:

$$\frac{\partial V}{\partial t} + c\left(\frac{\partial V}{\partial r}\right)\frac{\partial V}{\partial r} - (r + \epsilon(V))V = 0,$$

where

$$\epsilon(x) = \begin{cases} \epsilon & \text{for} \quad x > 0 \\ -\epsilon & \text{for} \quad x < 0. \end{cases}$$

Figure 70.1 shows a possible evolution of r and of r'. The latter, the spot interest rate, is the volatile line, although the word ‘volatile’ does not have its usual, or indeed any precise, meaning here. I have deliberately plotted a rather bizarre evolution, demonstrating the rich structure that the model allows. Working from left to right, we see (a) a steady rate increase followed by (b) a jump, then (c) a further smooth increase followed by (d) a period during which the rate jumps discontinuously every day from one extreme to another. There is a period (e) where the rate is constant, followed by (f) a Brownian-looking spell with an upward trend. There is then another Brownian-looking period but with a downward trend, (h) another rising period, followed by (i) a low volatility period. There is then (j) a very volatile time with sinusoidal periodicity, followed by (k) a calmer spell.

Since the cashflows depend on r' rather than r, we minimize their value to the holder (as we are in a worst-case scenario), when we write them in terms of r.

We apply the last cashflow as final data for the equation.

$$V(r, T) = C_N(r'),$$

which gives us that, in the worst-case,

$$V(r, T) = \min_{-\epsilon \leq e \leq \epsilon}(C_N(r + e)). \tag{70.3}$$

Similarly, a cashflow of $C_i(r')$ at time T_i gives the jump condition that, in the worst-case,

$$V(r, T_i^-) = V(r, T_i^+) + \min_{-\epsilon \leq e \leq \epsilon}(C_i(r + e)). \tag{70.4}$$

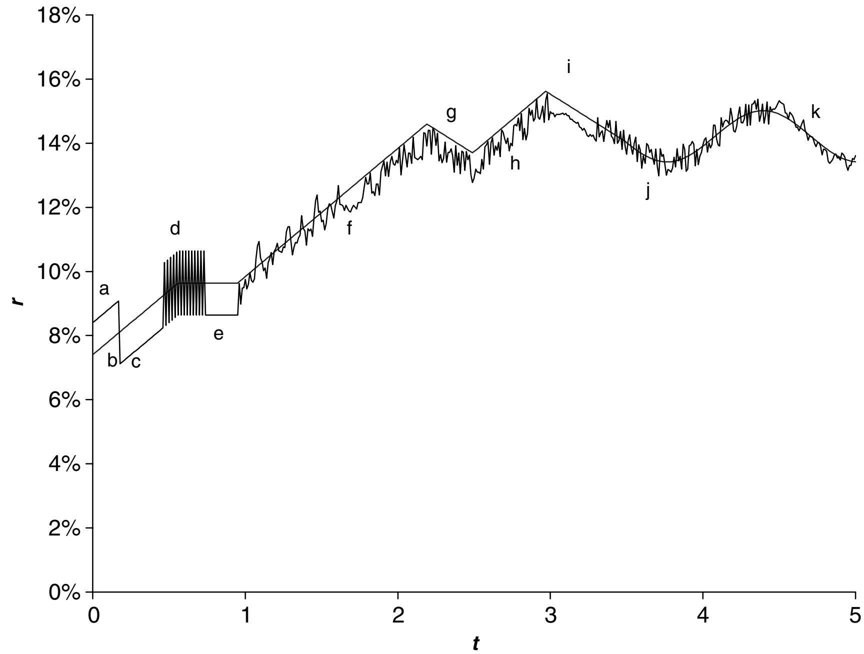


Figure 70.1 A possible evolution of the short-term interest rate. See text for an explanation.

Example

Price and hedge a four-year zero-coupon bond with principal 1. The hedging instruments are the zero-coupon bonds from Table 68.1. The spot short-term interest rate is 6% and the parameters of our model are the usual.

Table 70.1 shows the results for the zero-coupon bond value in a worst-case scenario with various values for ϵ without hedging and with the optimal static hedge for the worst-case scenario. Table 70.2 shows the hedge quantities for these static hedges. Figure 70.2 shows the value of the zero-coupon bond with varying ϵ.

It is clear that the wider the uncertainty band, the lower the worst-case price. As ϵ increases, we also see that the quantity of lower maturity hedging instruments in the portfolio decreases and the quantity of higher maturity instruments increases. The effect of this hedging strategy is that the worst-case price decreases less quickly with increasing ϵ.

70.4.1 Estimating ϵ from Past Data

We can use past interest rate data to choose a sensible value for ϵ. We use data for a longer period rate, the t_d period rate say, which is more readily available, and relate this to our

Table 70.1 Worst-case value of a four-year zero-coupon bond.

ϵ	0.00	0.01	0.02	0.03
No hedge	0.575	0.534	0.497	0.463
Optimally hedged	0.730	0.722	0.714	0.707

Table 70.2 The optimal static hedges for a four-year zero-coupon bond.

Hedging bond	$\epsilon = 0.00$ hedge quantity	$\epsilon = 0.01$ hedge quantity	$\epsilon = 0.02$ hedge quantity	$\epsilon = 0.03$ hedge quantity
A	0.000	0.002	0.002	0.002
B	−0.004	−0.004	−0.004	−0.001
C	0.169	0.091	0.091	0.010
D	−0.699	−0.569	−0.563	−0.439
E	−0.468	−0.511	−0.506	−0.527
F	0.020	0.018	0.018	0.000
G	0.000	0.000	0.000	0.000

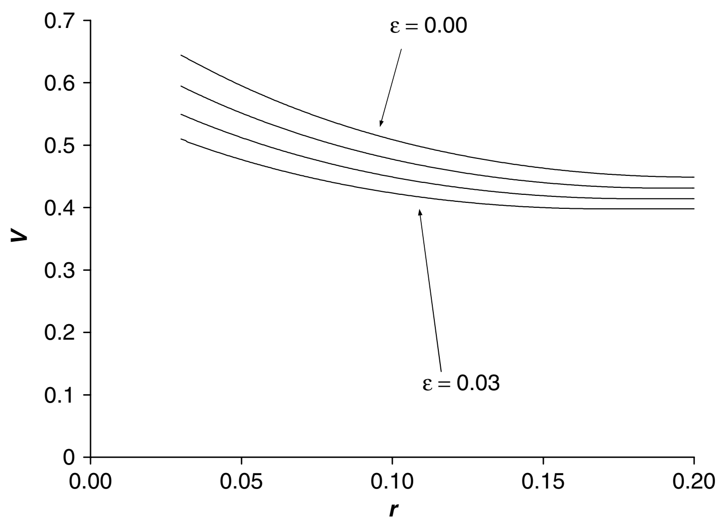


Figure 70.2 Value of a four-year zero-coupon bond with varying ϵ.

short-term interest rate. We can then examine the data and find the lowest value of ϵ for which the actual interest rate movements are consistent with our model.

We choose our parameters to be

$$r^- = 3\%,\ r^+ = 20\%,\ c^- = -4\% \text{ p.a. and } c^+ = 4\% \text{ p.a.}$$

and examine daily, one-month US interest rate data from 21st October 1986 to 25th April 1995.

The minimum value for ϵ for which the model is consistent with the data is

$$\epsilon = 0.005754.$$

If we use a value of ϵ slightly larger than this, then we can be fairly certain that any interest rate movements seen in the market would be allowed under our model. The spreads for prices predicted by our model, as well as all the other applications that we have discussed, would then be realistic in practice.

70.5 CRASH MODELING

To model further the precise nature of actual interest rate movements, we may wish to include the possibility of jumps or crashes in the interest rate. With these in place, we would find an even smaller minimum value for ϵ when we examine interest rate data.

I present two different approaches to the modeling of a crash. Both are instantaneous changes in the interest rate. In the first approach, the interest rate can crash on, at most, a specified number of occasions over the time horizon. In the second approach, the interest rate can crash an unlimited number of times, but can only do so when a specified length of time has passed since the previous crash. In both cases, we assume that the crashes occur at the worst possible time, since we are pricing the contract in a worst-case scenario. Note that, under these circumstances, it may be optimal for there to be no crash at all.

70.5.1 A Maximum Number of Crashes

First consider the situation when there can be at most one crash before the maturity of our contract. We value the contract, V, in a worst-case scenario and model the crash as an instantaneous movement of the short-term interest rate from r to $r-k$, for some specified k.

Introduce the subscript 0 to denote the value of the portfolio when there is no crash allowed and 1 to denote the value when the interest rate is allowed to crash once. Thus V_0 is the usual worst-case value and is the solution of Equation (68.3) with suitable final and jump conditions for the contract in question.

To value V_1, we also solve Equation (68.3), with the same final and jump conditions as before. However, the interest rate is now allowed to crash if that would lower the value of the contract. If the interest rate does not crash, then the contract is worth $V_1(r,t)$ since a crash is still allowed in the future. On the other hand, if the interest rate does crash from r to $r-k$, then the contract is worth $V_0(r-k,t)$, since the interest rate cannot crash again. In a worst-case scenario, a crash will only occur if that would give the contract a lower value. We therefore have the constraint,

$$V_1(r,t) \le V_0(r-k,t).$$

The value of the contract today is then $V_1(r,t)$.

We can generalize the model by allowing a range for the size of any possible crash. We model the crash as an instantaneous movement from r to $r-k$, where

$$k^- \le k \le k^+,$$

for some specified k^- and k^+. The effect of this generalization is seen in the constraint. In a worst-case scenario, the interest rate will always crash to the level that would minimize the value of the contract. The constraint therefore becomes

$$V_1(r,t) \le \min_{k^- \le k \le k^+} \left(V_0(r-k,t)\right).$$

We can also generalize the model to allow a number of crashes to occur before the maturity of the contract. We consider the situation where there are a maximum number of crashes, N say, before maturity. To price the contract under these circumstances, we must introduce $N+1$ functions, V_i, for $i = 1, 2, \ldots, N$. V_i is the value of the contract when the total number of crashes still allowed is i. As before, each of the V_i satisfies the usual equation with the appropriate final and jump conditions for the contract in question. Rather than a single constraint, we now have a set of constraints linking the $N+1$ functions, of the form,

$$V_i(r,t) \le \min_{k^- \le k \le k^+} \left(V_{i-1}(r-k,t)\right),$$

for $i = 1, 2, \ldots, N$. The value of the contract, today, in a worst-case scenario is then $V_N(r,t)$.

Example

We price and hedge a four-year zero-coupon bond with principal 1. The hedging instruments are the zero-coupon bonds from Table 68.1. The spot short-term interest rate is 6% and the parameters of our model are

$$r^- = 3\%,\ r^+ = 20\%,\ c^- = -4\% \text{ p.a. and } c^+ = 4\% \text{ p.a.}$$

Table 70.3 shows the results for the zero-coupon bond value in a worst-case scenario with various numbers of crashes allowed. These crashes have a maximum magnitude of 1% (i.e. $k^- = -1\%$ and $k^+ = 1\%$). Results are shown for the unhedged bond and for the bond optimally hedged in a worst-case scenario. Table 70.4 shows the relevant hedge quantities for some of these scenarios. Figure 70.3 shows the value of the zero-coupon bond when various numbers of crashes are allowed.

The addition of crashes lowers the value of the unhedged zero-coupon bond in a worst-case scenario. However, it has less effect on the value of the optimally hedged bond because almost the entire interest rate risk of the bond can be hedged away. However, to hedge away this risk, increasing amounts of the hedging instruments are required as the number of crashes allowed increases.

Table 70.3 Worst-case value of a four-year zero-coupon bond with crashes allowed.

No. of crashes	0	1	2	3	4
No hedge	0.575	0.558	0.542	0.528	0.516
Optimally hedged	0.730	0.728	0.726	0.724	0.722

Table 70.4 The optimal static hedges for a four-year zero-coupon bond with crashes allowed.

Hedging bond	No crashes hedge quantity	1 crash hedge quantity	2 crashes hedge quantity
A	0.000	0.000	−0.002
B	−0.004	−0.004	0.001
C	0.169	0.204	0.239
D	−0.699	−0.753	−0.816
E	−0.468	−0.453	−0.425
F	0.020	0.027	0.028
G	0.000	0.000	0.000

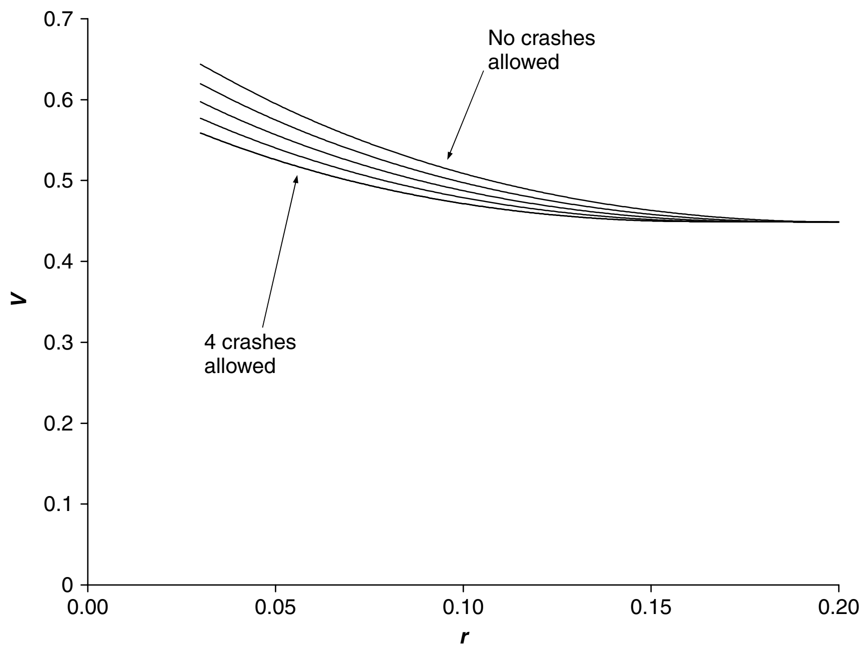


Figure 70.3 Value of a four-year zero-coupon bond with crashes allowed.

For the long zero-coupon bond, the worst-case scenario interest rate path will always be as high as possible. Since there is an upper bound for the short-term interest rate the addition of a crash does not have any effect at high interest rates, as the rate cannot crash to a higher value than the upper bound. This is illustrated in Figure 70.3 where the values converge at high interest rates.

70.5.2 A Maximum Frequency of Crashes

As an alternative, we can constrain the frequency of the crashes, instead of the total number allowed. We model crashes from r to $r - k$, as before. However, the interest rate can only crash again when a specified time ω has passed since the previous crash.

To value a contract under this model, we must introduce another variable, τ, the time since the last crash. Also introduce two functions: $V_0(r, t, \tau)$ is the value of the contract when the last crash was τ ago, and $V_1(r, t)$ is the value of the contract when a crash is allowed. The value of the contract today is then $V_1(r, t)$.

Since τ and t increase at the same rate when a crash is not allowed, the pricing equation for V_0 is

$$\frac{\partial V_0}{\partial \tau} + \frac{\partial V_0}{\partial t} + c\left(r, \frac{\partial V_0}{\partial r}\right)\frac{\partial V_0}{\partial r} - rV_0 = 0.$$

When $\tau = \omega$, another crash is allowed. We therefore have the final condition in τ,

$$V_0(r, t, \omega) = V_1(r, t).$$

The final condition in t, and any jump conditions, are dependent on the cashflows of the contract in question.

V_1 satisfies Equation (68.3), with suitable final and jump conditions dependent on the contract. In a worst-case scenario, a crash occurs if that would lower the value of the contract, and so we have the constraint,

$$V_1(r, t) \le \min_{k^- \le k \le k^+} \left(V_0(r - k, t, 0) \right).$$

With these two approaches to crash modeling, we can formulate a wide variety of crash events. For instance, we could combine the two models to allow a limited number of large crashes (1%, say) plus a larger number of smaller jumps (0.1%, say) to which we could assign a frequency. We must examine interest rate data to choose a sensible option for these parameters.

Example

We price and hedge a four-year zero-coupon bond with principal 1. The hedging instruments are the zero-coupon bonds from Table 68.1. The spot short-term interest rate is 6% and the parameters of our model are

$$r^- = 3\%,\ r^+ = 20\%,\ c^- = -4\% \text{ p.a. and } c^+ = 4\% \text{ p.a.}$$

One crash of maximum size 1% is allowed over the time horizon. Smaller crashes of at most 0.1% are also allowed, but can occur, at most, once a month.

The worst-case scenario value of the four-year bond without hedging is 0.558. When we optimally hedge the bond, the worst-case value increases to 0.728. The static hedge for this valuation is shown in Table 70.5.

Table 70.5 The optimal static hedges for a four-year zero-coupon bond with crashes allowed.

Hedging bond	Worst case hedge quantity
A	0.000
B	−0.001
C	0.202
D	−0.770
E	−0.415
F	0.004
G	0.000

70.5.3 Estimating ϵ from Past Data

When we performed the data analysis of Section 70.4.1, we effectively identified the point at which the largest jump in the short-term interest rate occurred. We then used this jump to find the smallest possible value for ϵ under the model. This is because any smaller movement would also be allowed, if this movement were allowed by the model.

Table 70.6 The largest changes in the one-month daily interest rate data.

Date	Size of crash (%)
22 DEC 86	1.328
30 DEC 86	−1.203
27 NOV 87	1.047
30 DEC 87	−0.813
29 NOV 88	1.000
29 NOV 90	1.125
28 DEC 90	−1.500
27 NOV 92	0.938

By including the possibility of a crash in our model, we can exclude this largest jump (by making it the point where the crash occurred). We can then calculate the lowest value of ϵ for the next largest jump. This should lead to a smaller value for ϵ. We therefore examine the data to choose a sensible number and size for crashes and then calculate the corresponding ϵ.

Example

Choose our parameters to be

$$r^- = 3\%,\ r^+ = 20\%,\ c^- = -4\% \text{ p.a. and } c^+ = 4\% \text{ p.a.}$$

and examine daily, one-month US interest rate from 21st October 1986 to 25th April 1995, as before. The eight largest changes in the one-month rate are shown in Table 70.6.

To some extent these are spurious jumps in the one-month rate. They are due to unusual supply/demand at the end of the year as people liquidate assets.

When we exclude these points, we find that the new minimum value for ϵ, so that the data are consistent with our model, is dramatically reduced to

$$\epsilon = 0.001066.$$

70.6 LIQUIDITY

The liquidity of a market is a measure of the inherent difficulty encountered on entering or exiting the market. As a market becomes more illiquid, the spreads between bid and offer prices increase. These spreads are likely to increase further for trades of particularly large quantities. However, this feature is more often apparent on exit rather than entry into a trade.

We can gauge the effect of liquidity on our model by including a bid-offer spread in the price of a hedging instrument, as discussed in Section 68.5.2. We can increase this spread, simulating illiquidity in the market, and examine how the worst-case price of a hedged contract changes, as well as the effect of the spread on the make-up of the optimal static hedge.

To try and gain some insight into what is happening, I shall only make one particular instrument illiquid. All of the hedging instruments, however, will admit a bid-offer spread. As the instrument becomes more illiquid, we would expect the static hedge to adjust to include less of the illiquid instrument and larger quantities of more liquid instruments. I shall increase the size of the bid-offer spread for the illiquid instrument. However, I will not make this spread dependent on the quantity of the instrument being traded.

Example

Include a bid-offer spread for the hedging bonds of Table 68.1. The resulting prices are shown in Table 70.7.

Consider different liquidities for the five-year bond, E, increasing the bid-offer spread from 0.683–0.691, to 0.647–0.727 and then to 0.607–0.767. The Yield Envelopes for these liquidity levels, are shown, in close-up, in Figure 70.4. It is clear that as the five-year bond becomes less liquid, the yield spread increases near five years maturity. There is still a maximum spread possible under the model (that of the contract hedged with all but the five-year bond) and the yield spread tends towards this as the liquidity decreases.

Table 70.7 The hedging bonds with a bid-offer spread.

Hedging bond	Maturity (years)	Bid price	Offer price
A	0.5	0.966	0.974
B	1	0.929	0.937
C	2	0.864	0.872
D	3	0.801	0.809
E	5	0.683	0.691
F	7	0.575	0.583
G	10	0.445	0.453

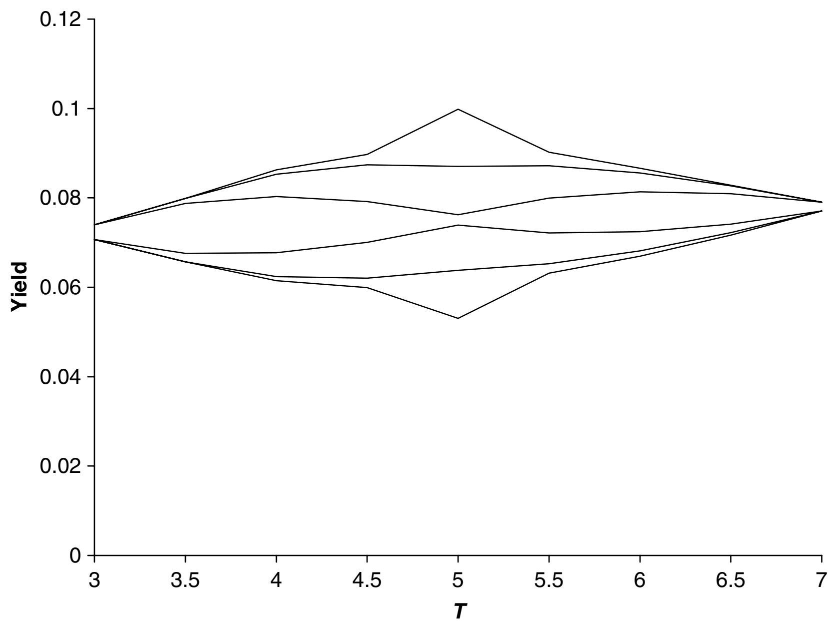


Figure 70.4 Illiquid Yield Envelope.

Table 70.8 Worst-case value of the four-year zero-coupon bond with an illiquid hedge.

Five-year spread	0.683–0.691	0.647–0.727	0.607–0.767
No hedge	0.575	0.575	0.575
Optimally hedged	0.725	0.711	0.708

Table 70.9 The illiquid optimal static hedges for the four-year zero-coupon bond.

Hedging bond	0.683–0.691 hedge quantity	0.647–0.727 hedge quantity	0.607–0.767 hedge quantity
A	0.000	0.000	0.000
B	0.000	0.000	0.000
C	0.105	0.370	0.577
D	−0.612	−1.047	−1.437
E	−0.479	−0.301	0.000
F	0.002	0.000	−0.105
G	0.000	0.000	0.000

To examine the effect of the liquidity changes on the optimal static hedge, consider the value of the four-year zero-coupon bond. The results of this worst-case scenario valuation are shown in Tables 70.8 and 70.9.

As we would expect, the value of the optimally-hedged four-year bond decreases as the liquidity of the hedging instrument decreases. The quantity of the five-year bond in the static hedge also decreases and the quantities of the bonds close in maturity to the five-year bond increase significantly to make up for this shortfall.

70.7 SUMMARY

Some of the new models presented in this chapter increase the spread between the best and worst cases. Some decrease it. There's clearly lots of work still to do to find the most realistic and most useful model of this type.

CHAPTER 71
modeling inflation

In this Chapter...

- an examination of inflation data
- comparison with interest rates
- a two-factor model
- calibration

71.1 INTRODUCTION

Index-linked bonds have coupons and principal that are increased in line with a measure of inflation, the Consumer Price Index (CPI) in the US and the Retail Price Index (RPI) in the UK. Thus income from these bonds is not eroded by inflation. In this chapter we will look at the various inflation-protected instruments available, and also examine inflation data to try to derive a statistical model for inflation that can be used for pricing complex inflation-linked products. We will focus on UK RPI data.

71.2 INFLATION-LINKED PRODUCTS

71.2.1 Bonds

Inflation-linked bonds have coupons and principal that are linked directly to the level of the inflation index.[1] If we denote the level of the index by I then the payoff of the bond will be just

$$\frac{I}{I_0},$$

that is inflation scaled with its initial level I_0. From now on I'll drop explicit mention of the scaling, or equivalently assume that $I_0 = 1$.

71.2.2 Inflation Caps and Floors

These contracts have call or put payoffs in I, with a maturity date and a strike price.

[1] There may be some time lag in the measurement of the data, such as the principal payment, depending on the index level eight months previously.

Example

A bond pays a coupon every year for 30 years. This coupon starts at \$1 but is increased by a specified percentage every year. That percentage increase is

$$R = \frac{I_i}{I_{i-1}} - 1$$

where I_i is the level of the index at the ith sampling (every year, say, in this example) but with a floor at 3% and a cap at 6%.

71.2.3 Swaps

Inflation swaps have a periodic exchange of a fixed rate (representing the market's estimate of average inflation) and a payment depending on the relative change in inflation during that period. There will be a series of payments of the form

$$\frac{I_i}{I_{i-1}} - 1 - I_f.$$

There are also inflation swaptions, options to enter into a swap.

71.2.4 Barriers

There are also barrier-like contracts which knock in or out depending on the level of inflation.

71.3 PRICING, FIRST THOUGHTS

The obvious way to approach the pricing of inflation-linked products is to treat the index I like an exchange rate or an equity price. We would write down a stochastic differential equation model for I,

$$dI = \cdots,$$

another for the interest rate r,

$$dr = u\,dt + w\,dX,$$

and, hey presto!, we have a pricing method. One subtlety here is that I is not traded so we would expect a market price of inflation risk term appearing.

By now you should all be able to fill in the details for yourself.[2] But this is what everyone does. So let's see a different approach, based on some data analysis, which is similar but subtly different.

[2] No, you should have been able to fill in the details 500 pages ago!

71.4 WHAT THE DATA TELL US

We will use the notation I to denote the index itself and R to denote the instantaneous relative change in I. Thus

$$I = \exp\left(\int^t R\,d\tau\right).$$

In Figure 71.1 we see the UK RPI index, I, and in Figure 71.2 the relative changes, R. Prior to 1960 the data look a bit odd because there are too few decimal places to distinguish small moves.

Since 1960 the data for R look like that in Figure 71.3. These data are more suggestive of a random walk than an independent, Normally distributed random number. Our 'first thoughts' above may have been easy, but not necessarily appropriate. Perhaps we should write

$$dR = a\,dt + b\,dX$$

as our model. The notation 'R' is meant to help parallels be drawn with an interest rate r. We are effectively smoothing the index, and putting r and R at the same level, both being stochastic, rather than r and I.

71.5 PRICING, SECOND THOUGHTS

The payoff for an index-linked bond is I. Its value also depends on the interest rate process r. With

$$dr = u\,dt + w\,dX_1,$$

$$I = \exp\left(\int^t R\,d\tau\right),$$

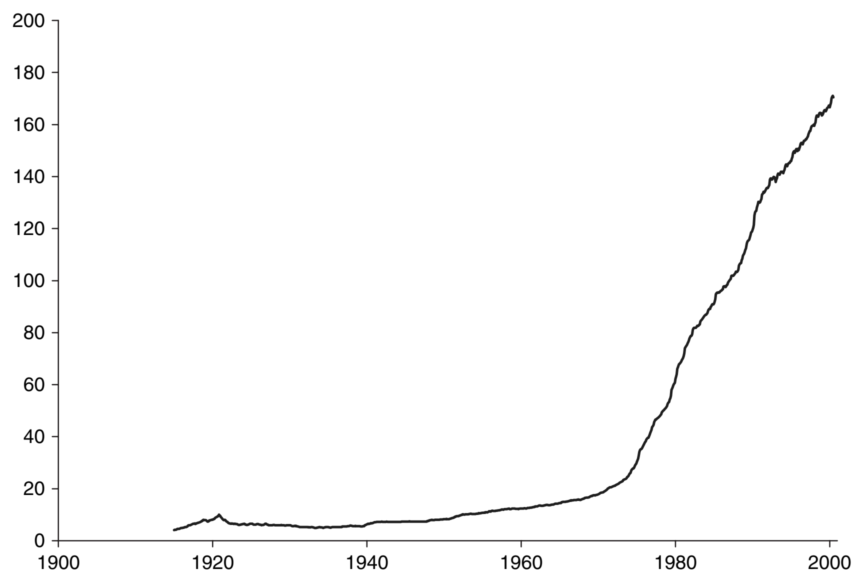


Figure 71.1 The UK RPI index (i.e. I).

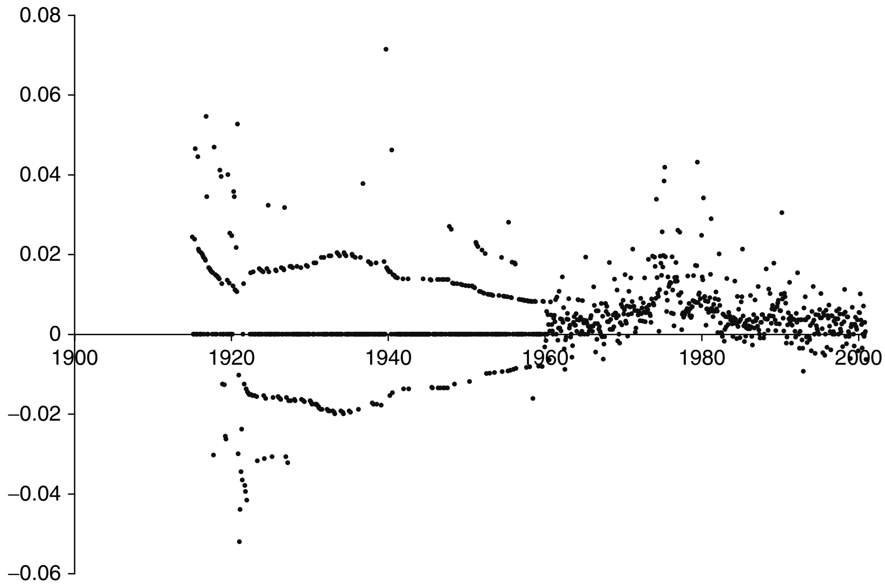


Figure 71.2 Relative changes in the monthly UK RPI index (i.e. R). (Extra decimal place from 1960.)

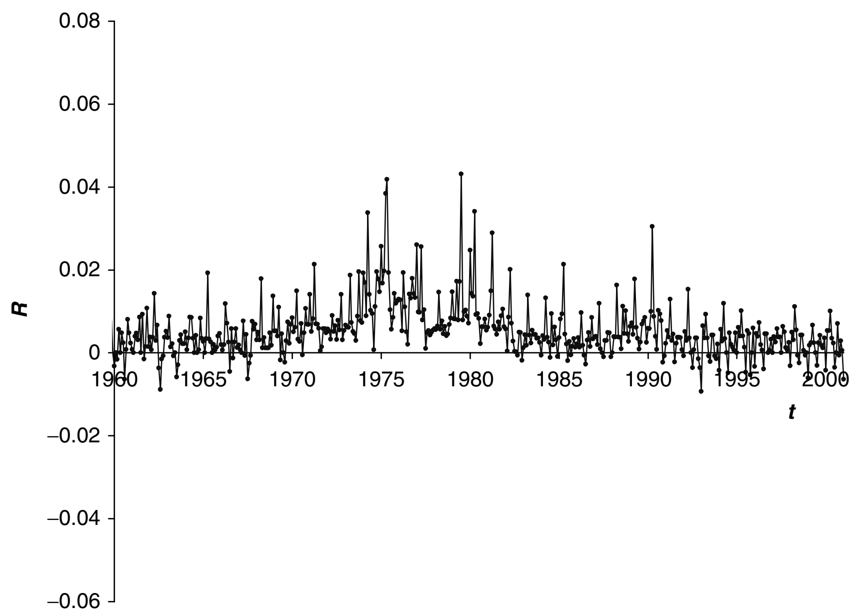


Figure 71.3 R since 1960.

so that

$$dI = RI\,dt$$

and

$$dR = a\,dt + b\,dX_2,$$

with a correlation of ρ, then the pricing equation is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda_r w)\frac{\partial V}{\partial r} + \rho b w\frac{\partial^2 V}{\partial r\,\partial R} + \tfrac{1}{2}b^2\frac{\partial^2 V}{\partial R^2}$$
$$+ (a - \lambda_R b)\frac{\partial V}{\partial R} + RI\frac{\partial V}{\partial I} - rV = 0$$

with

$$V(r, R, I, T) = I.$$

The λs are the market prices of risk.

There is a similarity solution $V(r, R, I, t) = IH(r, R, t)$ such that

$$\frac{\partial H}{\partial t} + \tfrac{1}{2}w^2\frac{\partial^2 H}{\partial r^2} + (u - \lambda_r w)\frac{\partial H}{\partial r} + \rho b w\frac{\partial^2 H}{\partial r\,\partial R} + \tfrac{1}{2}b^2\frac{\partial^2 H}{\partial R^2} + (a - \lambda_R b)\frac{\partial H}{\partial R} - (r - R)H = 0$$

with

$$H(r, R, T) = 1.$$

Choose your favorite interest rate model, and a suitable model for R, and solve. If you choose risk-neutral drifts that are linear in r and R and constant or square roots for the volatility terms then you should be able to find explicit solutions.

71.6 ANALYZING THE DATA

What if we look at the data, rather than choosing a tractable model?

In Figure 71.4 we plot the distribution of R, fitted to a Normal distribution. Like all financial distributions, this shows a skew. This, and the apparent limited downside, suggest that we try fitting a lognormal distribution to the monthly changes.

That fit is shown in Figure 71.5. And as is so often the case, the lognormal seems to be the better fit.

The best-fit lognormal has the following distribution

$$\frac{1}{\sqrt{2\pi}c(R - R^-)}\exp\left(-\frac{1}{2c^2}\left(\log\left(\frac{R - R^-}{\overline{R}}\right)\right)^2\right)$$

where $c = 0.2222$, $R^- = -0.0150$ and $\overline{R} = 0.0192$.

Now let's take a look at dR. Can this tell us anything about the volatility of R?

In Figure 71.6 is plotted the empirically calculated value of $E[dR^2]$ versus R. Is there any structure in this that suggests how the volatility of R varies with the level of R?

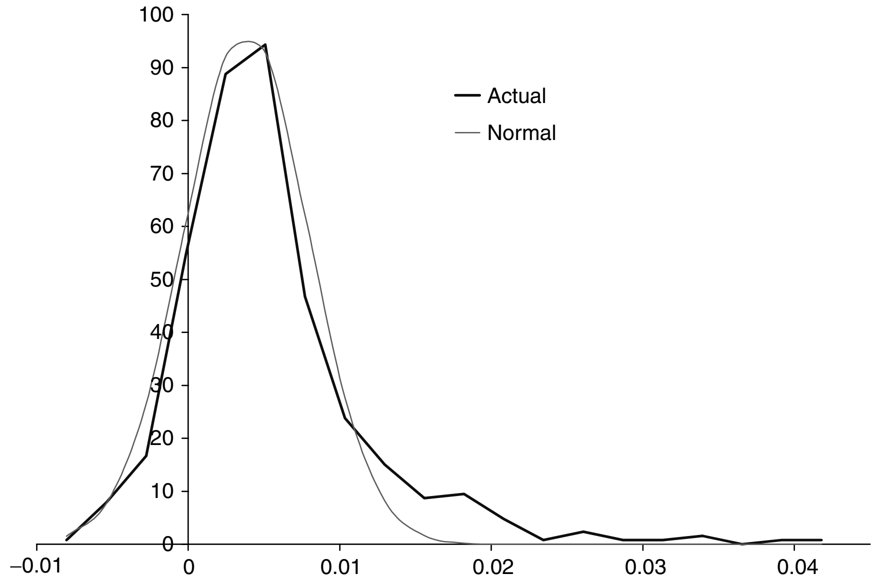


Figure 71.4 Distribution of relative changes in the monthly UK RPI index (i.e. *R*), post-1960 and best-fit Normal. The Normal has a mean of 0.0038 and standard deviation of 0.0041.

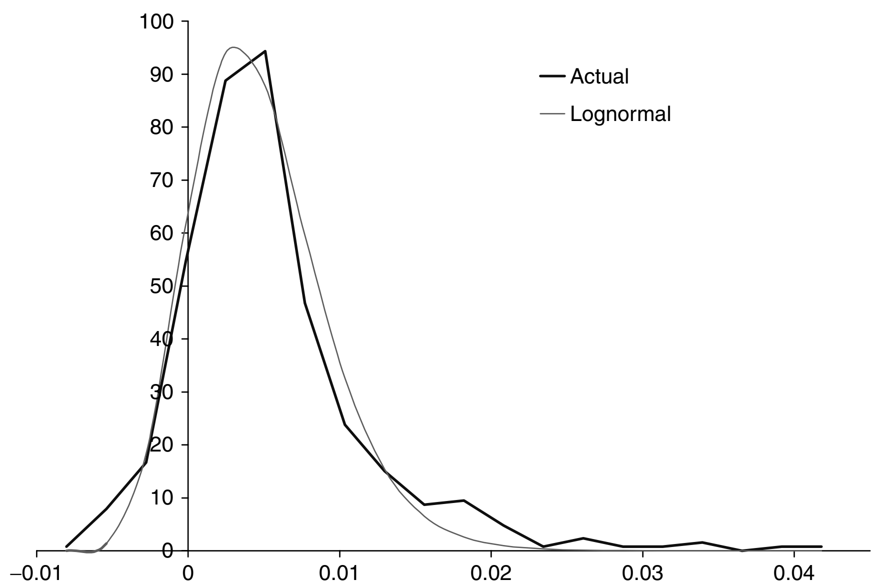


Figure 71.5 Distribution of relative changes in the monthly UK RPI index (i.e. *R*), post-1960 and best-fit Lognormal.

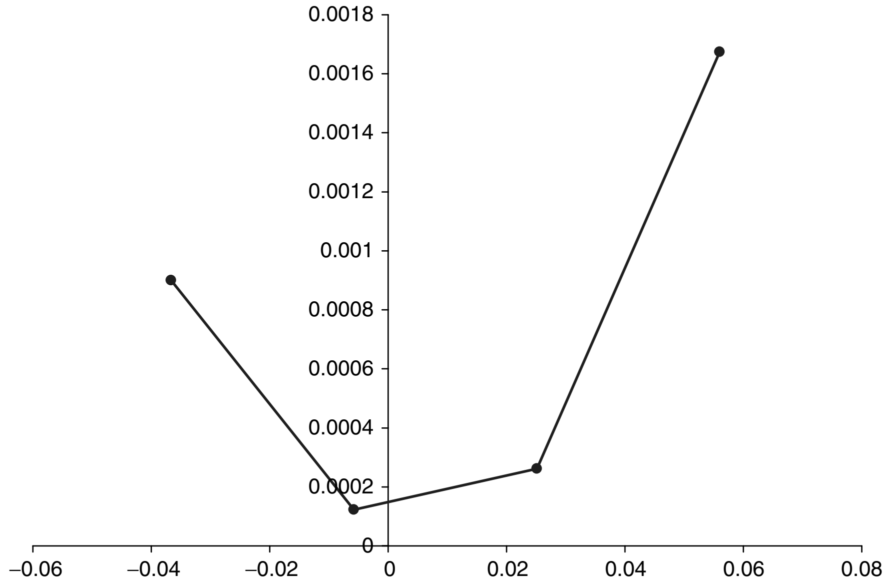


Figure 71.6 $E[dR^2]$ versus R.

It appears that the volatility of R is greatest for large R and for negative, deflationary, R. However, there is insufficient data for any concrete conclusions.

Staying within a 'classical' framework let's write

$$dR = a(R, t)\,dt + b(R, t)\,dX.$$

Given that there doesn't seem to be any volatility structure in R, for simplicity let us take as our model

$$dR = a(R)\,dt + b\,dX$$

where b is a constant and $a(R)$ is chosen to give a lognormal steady-state distribution for R.

Our data for the steady-state distribution suggest that we use the model

$$a(R) = -\frac{b^2}{2(R - R^-)}\left(1 + \frac{1}{c^2}\log\left(\frac{R - R^-}{\overline{R}}\right)\right)$$

and

$$b = 0.047.$$

71.7 CAN WE MODEL INFLATION INDEPENDENTLY OF INTEREST RATES?

Is such a model realistic? Take a look at the year-on-year inflation since 1960 (Figure 71.7). Should we really be using data from 1960 onwards in our model, or is the high inflation of the 1970s really dead? Famous-last-words territory!

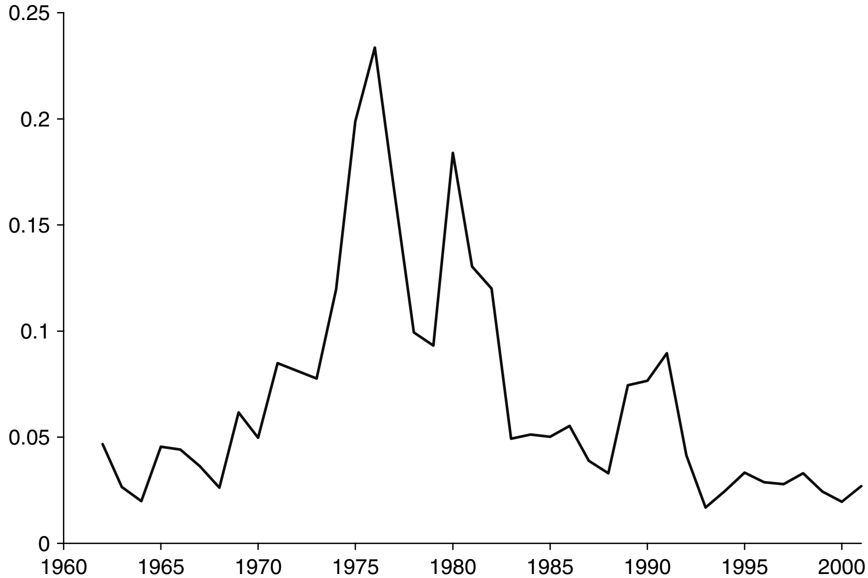


Figure 71.7 Year-on-year inflation since 1960.

For inflation-linked bonds the coupons and principal increase with R and therefore so does the present value. However, increasing interest rates decrease the present value. The current value of inflation-linked cashflows therefore depends on the difference between R and r (see Figure 71.8).

There is clearly an interesting and complex relationship between inflation and interest rates. Interest rates are used by governments to control inflation, and inflation affects people's use of money, their tendency to either save or spend. The causal relationship is not an easy one to model.

Ignoring the high inflation of the 1970s, what simple relationship between inflation and interest rates can we find?

It seems that $R = -0.8672 + 0.5596r$ is a good fit, where r is a short-term interest rate (Figure 71.9).

How can we get something like

$$R = \alpha + \beta r + \phi?$$

Let's assume that

$$d\phi = -f\phi\, dt + g\, dX_2.$$

The steady-state distribution for ϕ is Normal with mean zero and standard deviation

$$\sqrt{\frac{g^2}{2f}}.$$

An analysis of the data for $R - \alpha - \beta r$ gives us

$$f = 0.977 \quad \text{and} \quad g = 1.44.$$

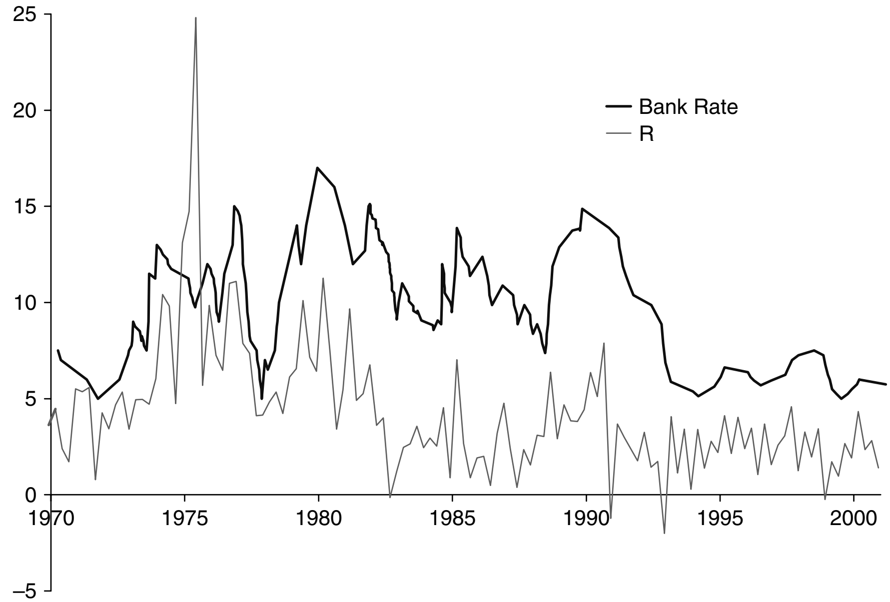


Figure 71.8 Interest rate and inflation.

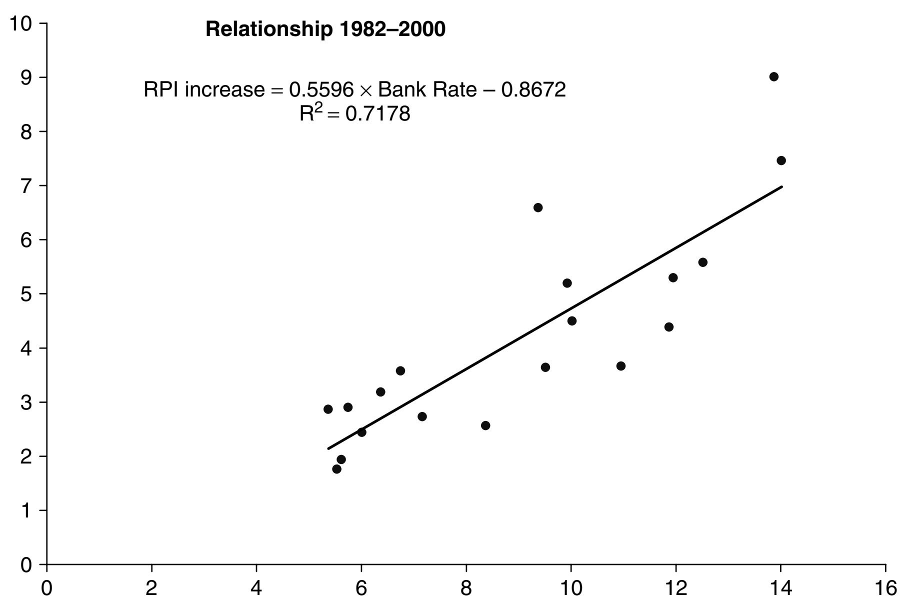


Figure 71.9 Interest rate and *R*, 1982–2000.

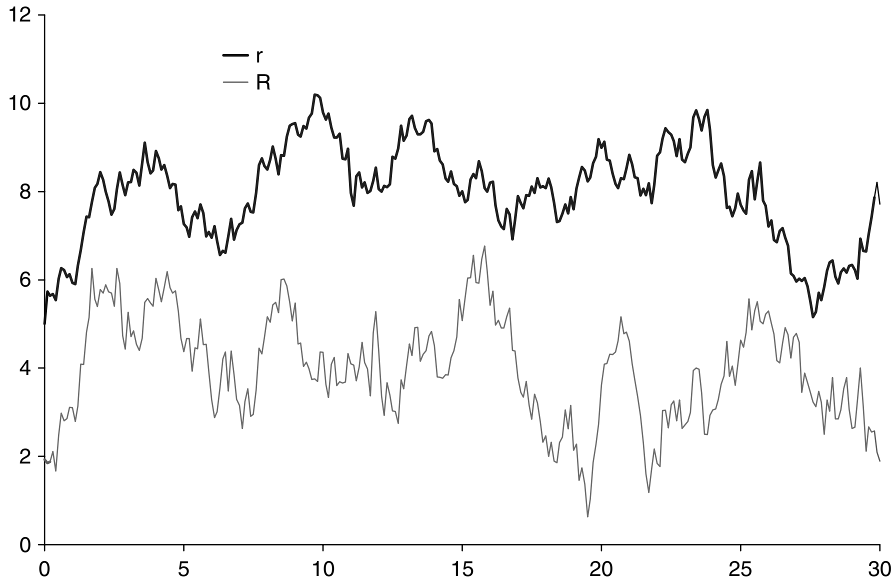


Figure 71.10 A simulation of r and R.

71.8 CALIBRATION AND MARKET PRICE OF RISK

Much of what we've done above has involved the statistical analysis of real inflation data (and potentially real interest rate data). But in the classical framework, the pricing of derivatives is more to do with the statistics of risk-neutral random walks than real ones. In this framework what we need for modeling purposes are the risk-neutral drifts for r and R (Figure 71.10).

Why go to all the trouble of analyzing the statistics of inflation and interest rates when all the work, except for the volatilities, gets ignored in the pricing model? The problem with modern pricing methodologies is that they assume that the market is in some sense 'correct.' There is always the desire to calibrate a pricing algorithm so that it gets theoretical prices to match market prices.

Whenever we model a non-traded quantity such as an interest rate or inflation we find ourselves having to model the market price of risk for that quantity. Unfortunately, this quantity is hard to observe and measure, and even harder to model since it is so unstable.

71.9 NON-LINEAR PRICING METHODS

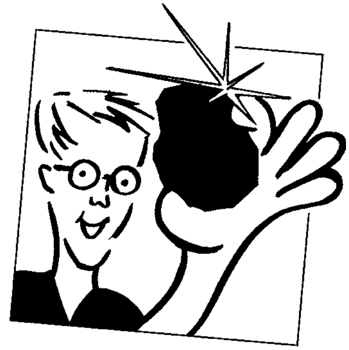
Modern pricing methods get around this problem by acknowledging it and confronting it. The obvious example in the fixed income world, and especially in the index-linked world, is the mean-variance model. In this model we accept that we cannot be certain of the price of any instrument. We cannot eliminate risk therefore pricing becomes more subtle than finding a single number. We must calculate both the *average* value (as if there were many parallel universes) and the variance about that average. These non-linear models do not require calibrating since they do not result in a single price; a range of prices is possible, some being more likely than others.

71.10 SUMMARY

Inflation modeling is still in its infancy. Most research on this topic seems to follow the standard approaches we have seen in modeling interest rates and risk of default.

FURTHER READING

- See Bloch (1995) for an empirical analysis of inflation data and a theoretical discussion of pricing index-linked bonds.
- For further work in this rapidly expanding area see Hughston (1998), Mercurio (2004) and Beletski & Korn (2005).



CHAPTER 72
energy derivatives

In this Chapter...

- why the energy markets are special
- the convenience yield
- the Pilopović two-factor model

72.1 INTRODUCTION

This chapter can only be a very brief introduction to the energy markets. The issues involved can be quite complex and would require a whole book to capture all of the subtleties. And, to be honest, I don't believe much of what I'm writing, and by the end of the chapter I hope you'll see why.

72.2 THE ENERGY MARKET

The energy market consists of commodities such as

- Natural gas
- Oil, of various types
- Electricity

These markets share some of the characteristics of other commodity markets, such as that for pork bellies. There is a spot market for immediate delivery and a futures market for delivery at some specified time in the future. Spot prices are typically very volatile.

As with other commodities there is a cost associated with the storage of the energy for future use or delivery. Sometimes this storage is simply a matter of physically holding onto the commodity in special locations, such as with oil. Sometimes, as with electricity, more complex storage arrangements must be made, and often aren't possible at all. For example, use the electricity to pump water into a reservoir, store the energy as potential energy for later release as hydroelectric energy. Generally speaking though, the opportunities for arbitrage by exploiting apparent mispricings are negligible; and because of storage difficulties, one day's delivery of electricity is a completely different product from the following day's.

There are two forms of delivery, physical and cash. As its name suggests, physical delivery is when the actual commodity is given to the buyer. The contract usually specifies the location

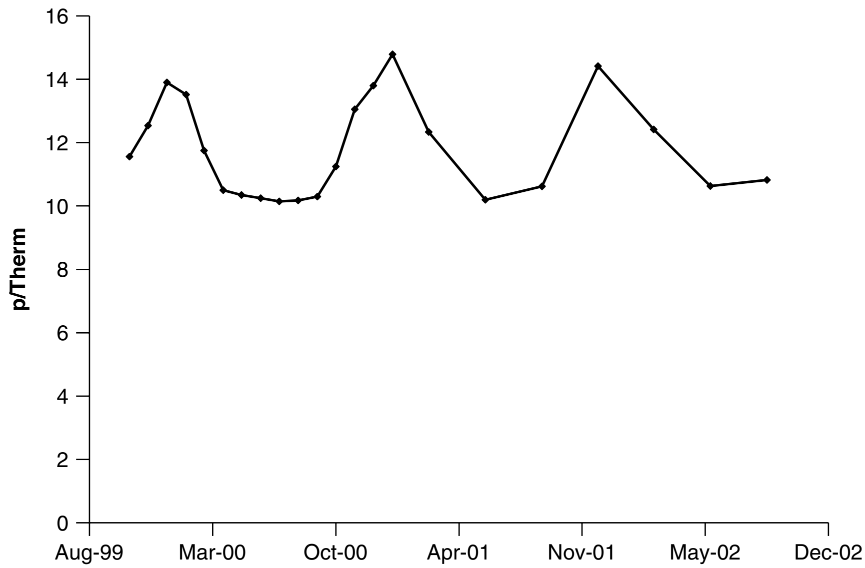


Figure 72.1 Future natural gas prices.

at which delivery will be made. Different commodities have different delivery locations or hubs, and there will be several such locations around the country. Cash delivery is when the equivalent amount is given in money.

A very important effect that is seen in energy prices is the seasonality effect. Prices have a strong seasonality component that is seen in the future price curve. Figure 72.1 shows future prices of natural gas. In the US, heating oil is more expensive in the winter months. Electricity is more expensive in the summer when it is used to power air conditioners.

Because of subtleties in the production of an energy commodity, such as oil, or in the point of delivery, there arises the problem of basis risk. Basis risk is when two almost identical commodities are not quite as identical as they may seem. For example, in the summer electricity will be more expensive in the Southern US than in the Northern US. It will not be possible to perfectly hedge electricity to be delivered at one hub with electricity to be delivered at another hub, even though both products are identical.

There are many differences to other commodity markets and these are what make the market interesting.

72.3 WHAT'S SO SPECIAL ABOUT THE ENERGY MARKETS?

Figure 72.2 shows a time series of daily spot electricity prices for immediate delivery. This shows the very typical wild price swings common in spot energy prices. Future prices will also show large price moves for maturities of a couple of weeks. Longer maturities are much more well behaved. Figure 72.3 shows the same data, but now on a logarithmic scale. Even on this scale the spikes are enormous.

In Table 72.1 is shown a time series of spot electricity prices during one day.

The reasons for these sudden price hikes stem from the lack of storability for electricity. Demand shows large spikes at certain times during the day, during ad breaks on the TV, for

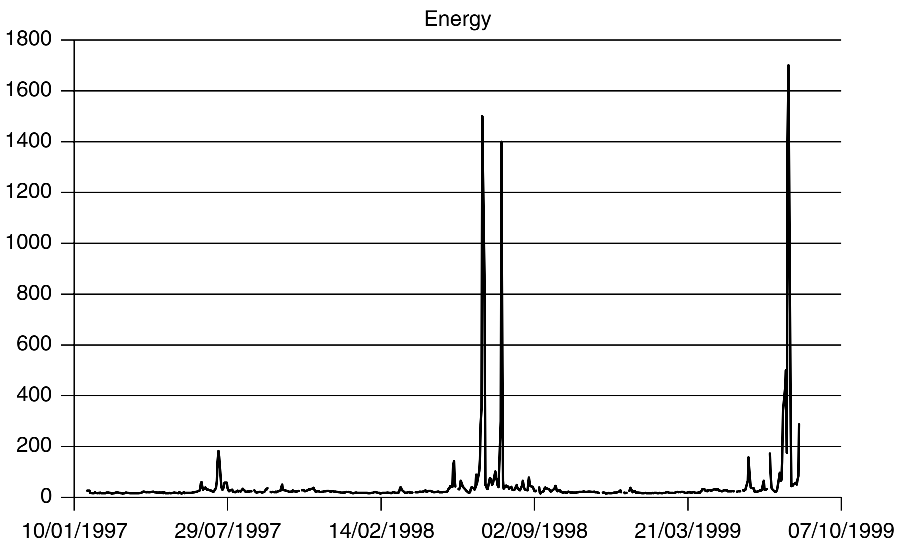


Figure 72.2 Daily spot electricity prices.

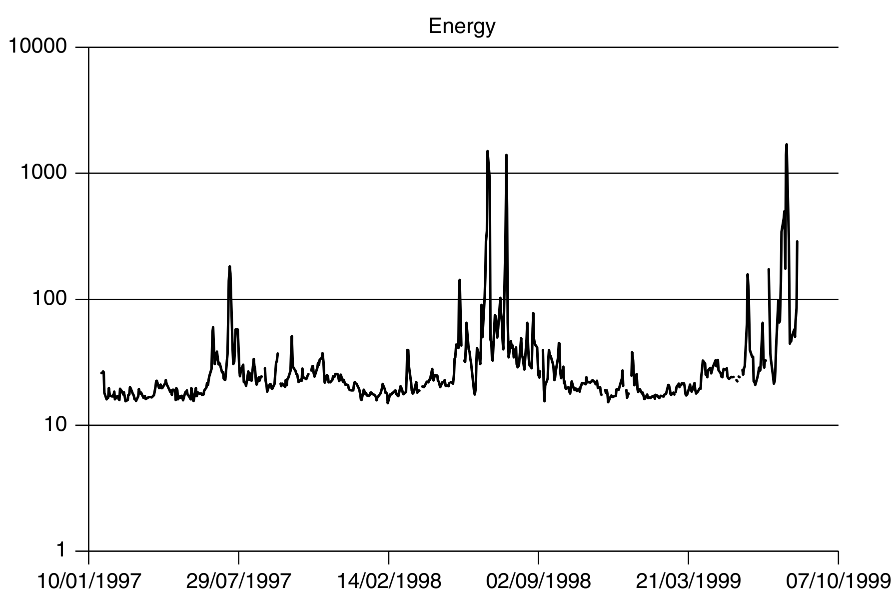


Figure 72.3 Daily spot electricity prices, logarithmic vertical scale.

example. On August 11th 1999 there was a total solar eclipse in the Southern UK. Just before the eclipse there was a drop in power demand as everyone went outside to 'see' the eclipse and switched off lights, TVs etc. On returning a few minutes later, they switched everything back on, leading to a surge in demand. If electricity could be easily stored the rapid fluctuations in demand wouldn't necessarily lead to rapid fluctuation in price. The same is true of natural

Table 72.1 Spot electricity prices.

Time	$/MWh
0:00	14.6
1:00	13.7
2:00	14.3
3:00	14.1
4:00	13.7
5:00	14.3
6:00	15.8
7:00	34.2
8:00	42.8
9:00	47.7
10:00	55.2
11:00	89.9
12:00	468.0
13:00	900.0
14:00	900.0
15:00	900.0
16:00	709.1
17:00	189.4
18:00	92.8
19:00	31.3
20:00	58.1
21:00	51.5
22:00	31.3
23:00	19.7

gas, which can also rise many times in price quite suddenly. Other energy commodities are relatively well behaved.

It's not even obvious that we should use a continuous-price process as a model. Certainly, the traditional lognormal asset price model is not going to be appropriate. Although the spikes are extreme, prices do fall back to earlier levels equally dramatically, usually within a day or two. We could model this as a jump process but it may have to be path-dependent to capture the price falls. Maybe a mean-reverting process would be a reasonable approximation, which would fit in well with our modeling experience so far. We'll pursue this route later on.

In Table 72.2 are shown prices of US mid-continent electricity prices on July 28th 1999. These prices, in $/MWh, are taken from Bloomberg whose headline was 'Heat Boosts Next-Day Midwest Electricity Prices Near 1998 Highs.' Two days later the headline was 'Midwest Electricity Prices Plummet on Cooler Monday Weather Forecast.' Prices are shown in Table 72.3. What a crazy market.

To summarize, energy prices have the following characteristics:

- Basis risk due to location of delivery (difficulty in transportation);
- Basis risk due to time of delivery (difficulty in storage);
- Large jumps and mean reversion;
- Seasonality effects.

Table 72.2 Bloomberg mid-continent electricity prices July 28th 1999.

Index	$/MWh	Daily change %	Low	High
ECAR	1780	+1586	600	2500
East	1966	+1757	1500	2400
AEP	1544	+1357	800	2100
West	1800	+1618	1800	1800
Central	1972	+1792	600	2500
Cinergy	1972	+1792	600	2500
South	1700	+1529	1000	2100
North	1700	+1475	1400	2000
Main	1981	+1777	1500	2200
Com-Ed	1762	+1554	1500	2100
Lower	2200	+2000	2200	2200
MAPP	1425	+1237	325	2000
North	1850	+1684	325	2000
Lower	1000	+792	800	1200

Table 72.3 Bloomberg mid-continent electricity prices July 30th 1999.

Index	$/MWh	Daily change %	Low	High
ECAR	281	−1496	200	325
East	275	−1585	250	300
AEP	282	−1618	225	282
West	300	−1200	275	325
Central	281	−1469	225	325
Cinergy	281	−1469	225	325
South	298	−1452	200	300
North	250	−1650	225	275
Main	247	−1353	150	862
Com-Ed	225	−1475	150	275
Lower	269	−1231	200	325
MAPP	95	−1321	78	150
North	90	−1360	90	150
Lower	100	−1283	78	100

72.4 WHY CAN'T WE APPLY BLACK–SCHOLES THEORY TO ENERGY DERIVATIVES?

The high volatility at the short end of the forward curve and the stability of the long end rule out the use of the basic Black–Scholes model. The actual volatility of the spot price would give unrealistically large option prices. In Figure 72.4 are shown the implied at-the-money forward volatilities for Brent crude oil. This rapid decay of volatility is very typical of energy underlyings, and represents the wild swings seen in spot prices and prices for early delivery.

We need something more sophisticated, not only because of this high volatility but also because, as I've said before, but it can bear repeating, it's hard to store electricity and virtually impossible to hedge.

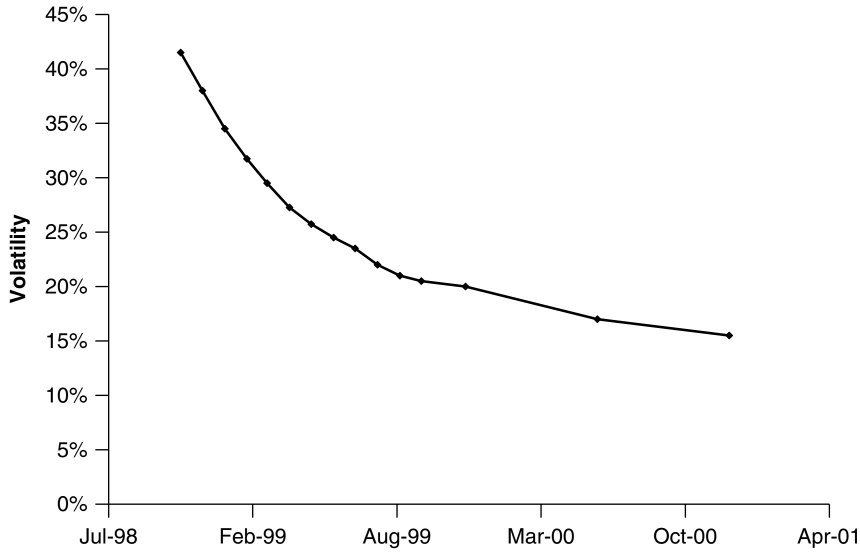


Figure 72.4 Forward volatilities of at-the-money Brent crude oil options.

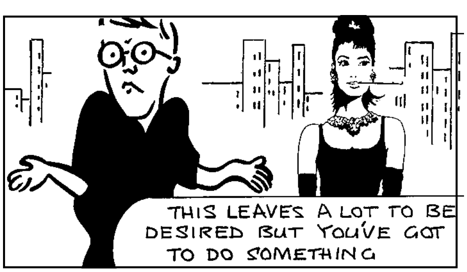
72.5 THE CONVENIENCE YIELD

The **convenience yield** is to energy what dividend yield is to stocks. Specifically, it measures the net benefit less cost of holding the energy. Users of energy are willing to pay a premium for the ability to get the energy they require, when they require it. On the other hand, there is the cost of energy storage to take into account. The benefits minus the costs become the convenience yield which can be quantified by examining the future prices of energy. If the convenience yield were constant then we would expect the usual relationship between forward and spot prices,

$$F = Se^{(r-q)(T-t)},$$

where q is the convenience yield.

Because the convenience yield plays such an important role in shaping the forward price curve, it is usual to use a much more sophisticated model for it than a simple constant.



72.6 THE PILOPOVIĆ TWO-FACTOR MODEL

The Pilopović two-factor model takes the form

$$dS = \alpha(L - S)\,dt + \sigma S\,dX_1$$

and

$$dL = \mu L\,dt + \xi L\,dX_2$$

where dX_1 and dX_2 are uncorrelated. S is the spot price and L is a long-term price or an 'equilibrium' price. In this model the spot price reverts to L, while L grows lognormally. The speed of mean reversion is much faster than the long-term growth so that

$$\alpha \gg \mu.$$

We want to relate the spot and equilibrium price processes, together with the convenience yield and any market prices of risk, to the shape of the forward price curve. To do this we must first write down the equation governed by the forward price $F(S, L, t)$.

This equation is

$$\frac{\partial F}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} + \tfrac{1}{2}\xi^2 L^2 \frac{\partial^2 F}{\partial L^2} + (r - q)S\frac{\partial F}{\partial S} + (\mu - \lambda\xi)L\frac{\partial F}{\partial L} = 0. \tag{72.1}$$

This has the final condition

$$F(S, L, T) = S.$$

Note that there is no $-rF$ term because the forward price is paid at expiry, not at the initiation of the contract. λ is the market price of risk for the equilibrium price. Also note that I have assumed risk-neutrality in that the coefficient of $\partial F/\partial S$ is the risk-adjusted drift of the spot price.[1]

We need to model the convenience yield, perhaps as a function of S, L and t. As we often do, let's choose a functional form so that we get a nice convenient explicit solution for F.

Working backwards, let's suppose

$$F(S, L, t) = a(t)S + b(t)L.$$

Plugging this into (72.1), and assuming that λ is a constant, we get

$$q = q_0(t) + \frac{q_1(t)L}{S},$$

with

$$q_0(t) = \frac{\dot{a}}{a} + r$$

and

$$q_1(t) = \frac{\dot{b} + (\mu - \lambda\xi)b}{a}.$$

Here the dots over a and b mean differentiation with respect to time.

Final conditions are

$$a(T) = 1 \quad \text{and} \quad b(T) = 0.$$

There is plenty of freedom (i.e. time dependence) here to fit the forward rate curve and also the volatility structure of the curve.

---

[1] Since electricity is so difficult to store this is a bad assumption, but one that most people probably make. Combine the impossibility of hedging with the rapid price fluctuations and you'll see that it might be better to use a mean-variance model, such as those discussed in Chapter 59. As I said, this chapter is only the briefest of intros to a complex subject.

The governing equation for the value of non-path-dependent energy derivatives, under the Pilopović two-factor model, is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \tfrac{1}{2}\xi^2 L^2 \frac{\partial^2 V}{\partial L^2} + ((r - q_0)S - q_1 L)\frac{\partial V}{\partial S} + (\mu - \lambda\xi)L\frac{\partial V}{\partial L} - rV = 0. \quad (72.2)$$

72.6.1 Fitting

From the above we get

$$\log(a(t;T)) = -r(T-t) + \int_t^T q_0(\tau)\,d\tau$$

and

$$b(t;T) = e^{(\mu-\lambda\xi)(T-t)} \int_t^T \exp\left(-(\mu-\lambda\xi)(T-s) - r(T-s) + \int_s^T q_0(\tau)\,d\tau\right) ds.$$

If we know the forward curve, and S and L, at time t^*, $F(S^*, L^*, t^*; T)$, then

$$F(S^*, L^*, t^*; T) = a(t^*; T)S^* + b(t^*; T)L^*$$

is one equation for $q_0(t)$ and $q_1(t)$. (Note that the qs aren't allowed to be functions of T.) Another equation could come from the volatility structure:

$$\text{Volatility}(S^*, L^*, t^*; T) = \sqrt{a(t^*; T)^2\sigma^2 S^{*2} + b(t^*; T)^2\xi^2 L^{*2}}.$$

As already mentioned, seasonality plays an important role in the modeling of energy prices. So we might want to add an oscillatory (Fourier series) term or two to the basic model for S.

72.7 ENERGY DERIVATIVES

The types of derivatives seen in the energy markets are not that different from the many types we have seen already in other markets.

72.7.1 One-day Options

One-day options on electricity are very popular in the US. Because of the possibility of enormous price fluctuations these contracts are very hard to price. This effect is slightly mitigated by a timewise averaging that is sometimes part of these contracts.

72.7.2 Asian Options

Asian options are very popular in the energy markets. Two forms of averaging are seen, averaging over several realized spot prices that have settled, and averaging over unsettled forward prices.

The former are easier to value, in the path-dependent framework we've already seen.

The latter require a very accurate forward curve model. If we think of forward contracts as derivatives then the latter type of contract is a second-order contract. The contract is, however, not path-dependent.

72.7.3 Caps and Floors

Caps and floors exist in the energy markets. These restrict the price that must be paid for delivery of the energy. There is often the added, path-dependent, complication that some form of timewise averaging takes place.

72.7.4 Cheapest to Deliver

This contract allows the delivering party to deliver the same energy source to one of two delivery points. Presumably he will choose to deliver the cheaper. The correlation between the prices of the underlying at the two delivery points is a crucial parameter in the valuation of this contract. It is a multi-asset contract.

72.7.5 Basis Spreads

A **basis spread** is a contract on the difference or spread between two very similar, but different underlyings, two different types of oil for example. Again, this is a multi-asset contract and the correlation between the two assets is of paramount importance.

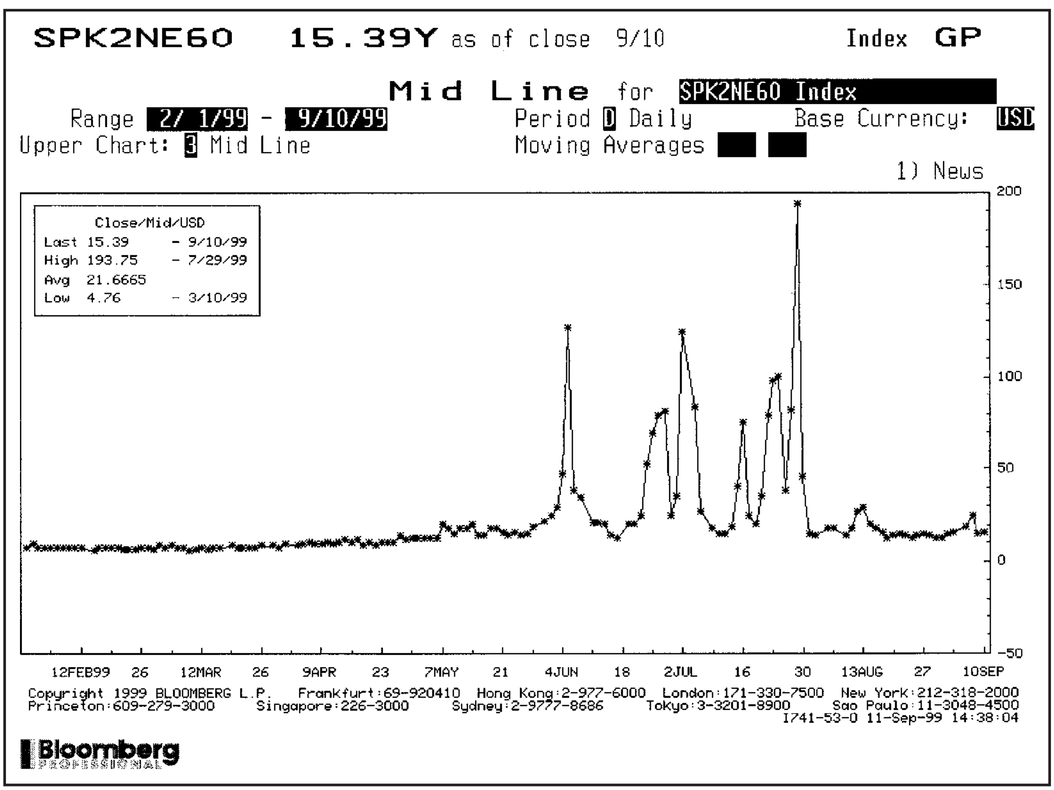


Figure 72.5 Time series of electricity spark spreads. Source: Bloomberg L.P.

72.7.6 Swing Options

Swing options allow the energy user to vary his energy delivery to vary between set limits. Perhaps also the amount is only allowed to vary a set number of times. Perhaps the user can take any ten days' of electricity during the summer, with the choice of precise days left up to the option holder. These contracts are complex, having characteristics similar to the passport option contracts seen in Chapter 27 and requiring the same mathematics of stochastic control.

72.7.7 Spread Options

Spread options are options on the spread between the price of the fuel used to produce the electricity and the price of that electricity. They are used to hedge the electricity production costs. Figure 72.5 shows a time series of electricity spark spreads, the relative efficiency of buying gas for making electricity versus buying the electricity itself.

72.8 SUMMARY

Energy derivatives are a fascinating subject, and still relatively new. Modeling the underlying is tricky because of the enormous spikes in prices. Modeling derivatives is difficult because hedging is often impossible. I think that there's a long way to go before the models become satisfactory.

FURTHER READING

- Gabillon (1995) was the first to present a two-factor model applied to the oil market.
- Pilopović (1998) is the best, by virtue of being the only affordable, book on the energy market and derivatives.
- See Ahn *et al*. (2002) for an analysis of the costs of storage of energy.

CHAPTER 73
real options

In this Chapter...

- the application of derivatives theory outside of finance
- optimal investment

73.1 INTRODUCTION

We've seen the word 'option' used to describe a financial contract containing some element of choice. And that choice is made complicated, and interesting, by the randomness underpinning the financial markets. But high finance is not the only area in which randomness and choice play important roles, every decision in life could be interpreted as trying to make the best choice, given an unknown future.

In this chapter we will see how many of the ideas of derivatives theory can be applied to other walks of life, and we'll also see a few more ideas as well. This is the subject of **Real options**.

73.2 FINANCIAL OPTIONS AND REAL OPTIONS

The key points that relate financial option theory and Real option theory are as follows:

- Randomness concerning the future introduces the idea of examining probability distributions for outcomes;
- Decisions should be made optimally, there is some question over the timing;
- Decisions may be partially or wholly irreversible.

In this chapter we will slowly build up the concepts and math used in Real option theory. In particular, we look at project valuation, optimal entry into and exit from a business, and optimally and sequentially investing. Most of the math is very similar to what we've seen already.

73.3 AN INTRODUCTORY EXAMPLE: ABANDONMENT OF A MACHINE

You own a factory that produces goods with a profitability of P which is realized continuously. So that $P\,dt$ is the profit made between times t and $t + dt$. The factory has a natural life span of T.

If P is a constant then the present value of all the future profits $V(t)$ satisfies

$$\frac{dV}{dt} - rV + P = 0 \tag{73.1}$$

with

$$V(T) = 0.$$

Notice that I have discounted at the risk-free rate; we may be discounting at a different rate later on.

The solution of this is just

$$V = \frac{P}{r}\left(1 - e^{-r(T-t)}\right).$$

As long as $P > 0$ then V is always positive and there would be no reason to close down the factory.

If P is a deterministic function of time then we can still solve (73.1). If P goes negative for sufficiently long then it might be worthwhile closing the factory down. Since this is a completely deterministic problem you just wait until $V = 0$. You could even open up the factory later on if conditions improve. Let's not worry about this until later in the chapter.

More realistically, suppose that the profit fluctuates due to general market conditions such as supply and demand, and so we'll assume that it follows the random walk

$$dP = a\,dt + b\,dX$$

with a and b constant, for the sake of argument. In a mature market you may expect $a < 0$ reflecting the general decline in profit as the product becomes outmoded. It is certainly possible that the profit could become negative, again a likely outcome in practice.

If there is some randomness in the dynamics of P, because b is non-zero then we must solve the diffusion problem

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}b^2\frac{\partial^2 V}{\partial P^2} + a\frac{\partial V}{\partial P} - rV + P = 0 \tag{73.2}$$

and now V is a function of P and t. By solving this problem, subject to conditions below, we will be finding out information about the present value of expected quantities. Observe that (73.2) is similar to the equation for the transition probability density function; the only differences are in the discounting to get the present value and the cashflow term, the last on the left.

The final condition is

$$V(P, T) = 0.$$

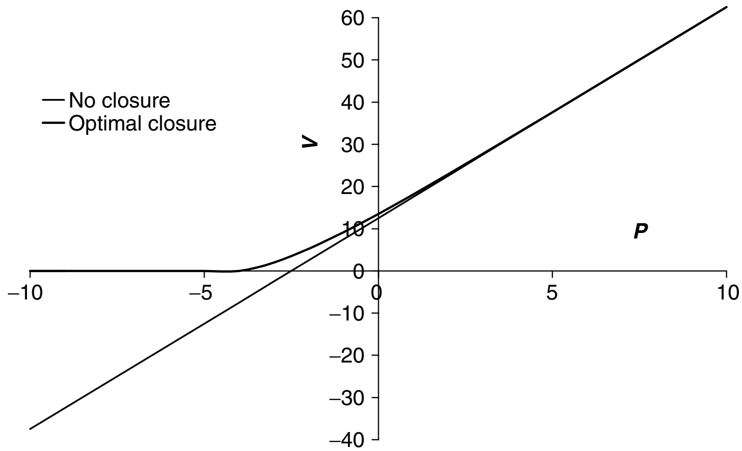


Figure 73.1 The value of a perpetual factory, with and without closure.

The solution of this problem tells us the present value of the expected total profit *provided that the factory stays open until* $t = T$.

Now we are at the point where we can ask about how it is best to run the factory: Should we actually keep it open even if it starts to lose money?

The optimal time at which to close down the factory, if we want to maximize the present value of accumulated expected profit, comes from solving the free boundary problem, Equation (73.2), subject to the final condition and

$$V(P, t) \geq 0$$

with continuity of $\partial V/\partial P$. This is the same in principle as the American option problem.

In Figure 73.1 are shown the theoretical values for the present value of the accumulated profits in two cases, with and without the option to close the factory down. Obviously the former is greater and the difference between the two is the value of the option of closure. In the example, T is taken to be infinity so that the factory has no natural lifespan.

The important point to note about the above analysis is that by postponing the closure of the factory until the optimal time we add significant value to the factory. If we had had to decide at the start whether to close the factory immediately, or continue forever, then we would not be fully exploiting the opportunity offered by waiting for new information (the realization of P) before making the big decision.

This model is also applicable to other interesting situations, such as modeling the 'rewards' from marriage. Some change of notation would be needed, for instance P would have to mean something other than profitability. T could be the expected lifespan of a man after marriage. The question is then when to separate.

73.4 OPTIMAL INVESTMENT: SIMPLE EXAMPLE #2

Here's another example, equally straightforward. When is the optimal time to invest a given fixed amount in return for a product whose value evolves randomly? This could be the purchase of our factory.

This problem sounds very much like an option problem, so let's use familiar notation. We use E to denote the fixed amount and S the value of the product, and we'll even use

$$dS = \mu S\,dt + \sigma S\,dX$$

although the exact specification would depend on the nature of the product.

If the investment opportunity has value $V(S, t)$ then we must solve

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} - rV = 0$$

subject to

$$V > S - E$$

and continuity of V and its S derivative. There will be a final condition of the form

$$V(S, T) = 0$$

if the opportunity has a finite lifespan, otherwise the problem will be time-independent. What can you say about the existence of an optimal investment time when there is an infinite time horizon? How does it depend on the parameters?

73.5 TEMPORARY SUSPENSION OF THE PROJECT, COSTLESS

We've seen how to model the closure of a factory, permanently, and the optimal time to invest. What if we can *costlessly* stop and start production depending on the prevailing market conditions? This must add even more value to our factory. I'll go back to using P to denote the instantaneous profit at any time and which is evolving randomly,

$$dP = a\,dt + b\,dX.$$

There's no point in specifying a and b since they will be very problem-dependent; generally they can be functions of P and t.

We will only operate the factory when $P > 0$ and suspend its operation when $P < 0$. Because there is no cost associated with stopping and starting we simply get all positive cashflows, thus

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}b^2 \frac{\partial^2 V}{\partial P^2} + a \frac{\partial V}{\partial P} - rV + \mathcal{H}(P) = 0$$

where $\mathcal{H}(\cdot)$ is the Heaviside function. There will be the obvious final condition if there is a finite time horizon.

73.6 TEMPORARY SUSPENSION OF THE PROJECT, COSTLY

Here's a real option problem that has features of some of the exotic options that we saw in Part Two. It's like the problem above but now it costs something to start or suspend the factory, C to start and K to stop.

To value this opportunity we must introduce two functions: The values of the opportunity while the firm is idle, $V_0(P,t)$, and while it is active, $V_1(P,t)$.

It's easy enough to write down the two pdes satisfied by these two functions. While the factory is idle we have

$$\frac{\partial V_0}{\partial t} + \tfrac{1}{2}b^2\frac{\partial^2 V_0}{\partial P^2} + a\frac{\partial V_0}{\partial P} - rV_0 = 0.$$

There is no source term since there is no profit being made. While the factory is active it is making money:

$$\frac{\partial V_1}{\partial t} + \tfrac{1}{2}b^2\frac{\partial^2 V_1}{\partial P^2} + a\frac{\partial V_1}{\partial P} - rV_1 + P = 0.$$

Now to incorporate the switch from idle to active and vice versa.

In (re-)opening the factory we lose an amount C. Thus

$$V_1(P,t) \geq V_0(P,t) - C.$$

This just says that the difference between active and idle is the cost.

Similarly (re-)closing the factory gives

$$V_0(P,t) \geq V_1(P,t) - K.$$

If the time/profit at which the factory changes state were given then the above would be sufficient. However, we want to change state in the optimal way ... and this gives a free boundary problem: We must also insist that

$$\frac{\partial V_0}{\partial P} \quad \text{and} \quad \frac{\partial V_1}{\partial P} \quad \text{are continuous.}$$

A numerical example of this is given in Figure 73.2 for an infinite time horizon.

How would this model change if there were an ongoing cost associated with the dormant factory?

73.7 SEQUENTIAL AND INCREMENTAL INVESTMENT

Faced with a project that comes in discrete stages we can use ideas from compound options for its valuation. For example, if one stage must be completed before moving on, at a cost, to the next stage we effectively have a call option on a call (see Chapter 22 for the details).

More interesting is the project for which the investment must be made continuously, but can be stopped at any time. This brings together some of the math from path-dependent option pricing as well as control theory.

Let's be more specific about the nature of the project.

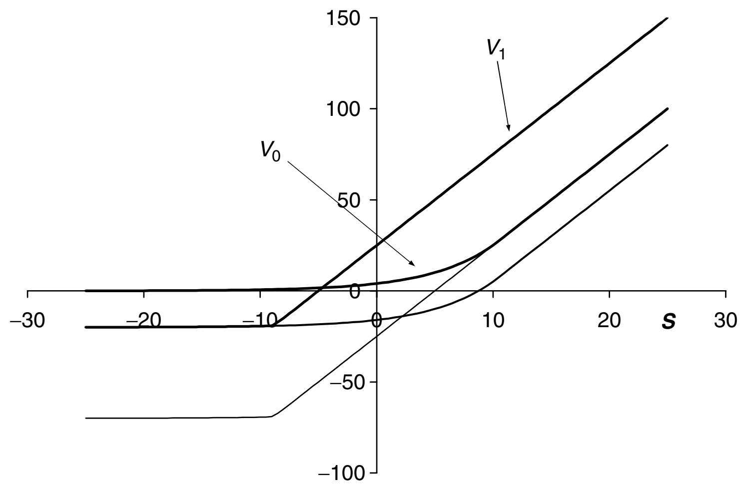


Figure 73.2 Costly entry and exit.

The firm invests at a rate I, which is somewhere between zero and a maximum possible rate: $0 \le I \le k$. The total remaining to be spent is K so that

$$dK = I\,dt.$$

The firm does not make any money until the project has been completed, when $K = 0$. On completion the project becomes worth an amount P where

$$dP = a\,dt + b\,dX.$$

The firm earns nothing before this time.

The equation to be solved for the real expected worth of the firm $V(P, K, t)$ is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}b^2\frac{\partial^2 V}{\partial P^2} + a\frac{\partial V}{\partial P} - I\frac{\partial V}{\partial K} - rV - I = 0.$$

The two terms of interest in this are both proportional to I. One is a derivative with respect to K which looks similar to what we've seen before with some path-dependent contracts. The other term is a sink of money as it gets spent.

We will solve the infinite horizon, time-independent problem for $V(P, K)$

$$\tfrac{1}{2}b^2\frac{\partial^2 V}{\partial P^2} + a\frac{\partial V}{\partial P} - I\frac{\partial V}{\partial K} - rV - I = 0. \tag{73.3}$$

We also have

$$V(P, 0) = P$$

representing the firm's value on completion of the project. Are there any other conditions?

Since we are not given I we must choose it to maximize the firm's expected value. So, by examining (73.3) we find that

$$I = k \quad \text{when} \quad \frac{\partial V}{\partial K} + 1 < 0 \quad \text{otherwise} \quad I = 0.$$

This is an example of **bang-bang control**, as the rate of expenditure goes from one extreme to the other.

73.8 ASHANTI: GOLD MINE CASE STUDY

Ashanti is a company that mines gold. It hedges against the price of gold with options. In the late 1990s Ashanti may have overhedged, so that an investment in Ashanti was an investment in a short gold position. Let's see how the ideas of real options can be applied.

Gold is a commodity unlike any other: Romanticized by poets, valued higher than human life. So, nothing like pork bellies, then. But what causes its value to fluctuate? Supply, demand? Well, they affect gold prices in the usual manner, but there are special factors that must be taken into consideration. Not only does price depend on production but also, vice versa, production depends on price. When the price is high, mines open up and it becomes economical to extract lower grade ore. When the price is low, some mines close and others concentrate on the higher quality ore; Ashanti estimates that it costs US$218 to get the gold out of the ground, and on top of that are corporate costs. This makes gold prices mean reverting to some degree; the higher the price, the more gold is produced, and so the price eventually falls. This is not unlike most storable commodities such as oil and other metals. However, this mean-reversion effect is less for gold because it is rarely destroyed during its use. So the world stock of gold is ever increasing. This makes the gold price more like an equity price. See Figures 73.3 and 73.4.

Most important for the behavior of the gold price is the attitude of central banks, currently holding one quarter of the world's total supply. Bulk selling by these bodies has been responsible for most of the dramatic price falls in recent years. Argentina, Australia, Switzerland, the UK and the EC have all affected gold prices one way or the other, depending on whether they did or didn't sell.

What does all this mean for the modeling of the gold price? As always, it depends on your timescale. If you are interested in the long timescale, five years and over, say, you will need to take into account the mean reversion and the relationship to inflation. For shorter timescales the mean reversion is irrelevant but the importance of bank action soars. Beliefs about future

Figure 73.3 Gold price in real terms.

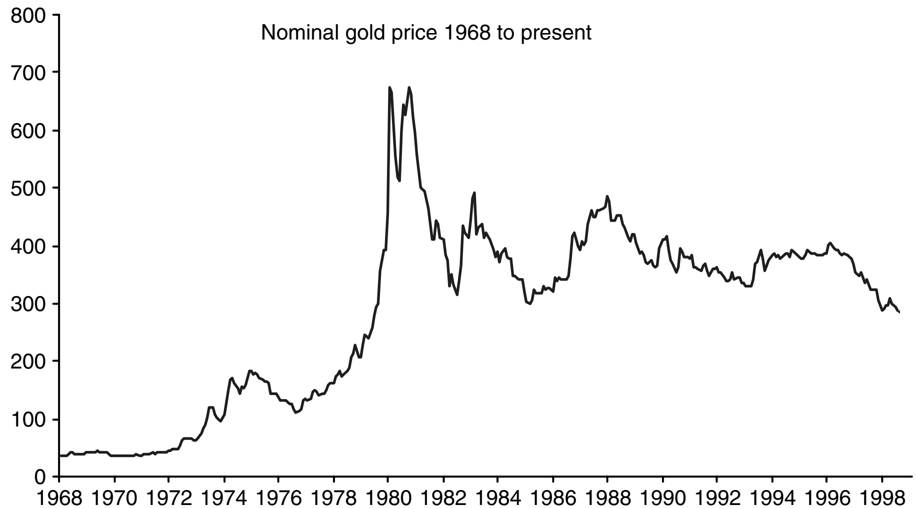


Figure 73.4 More recent gold price history.

bank action drive the price, down when they sell, up when they threaten to but change their minds.

Using gold price data from 1985 I've calibrated a simple mean-reverting model to historical data (nothing to do with calibrating to option data):

$$dS = \nu^2 S^{2\gamma-1}\left(\gamma - \frac{1}{2} - \frac{1}{2a^2}\log(S/\overline{S})\right) dt + \nu S^\gamma \, dX$$

with $\nu = 16.2$, $\gamma = 0.42$, $\overline{S} = 366$ and $a = 0.138$. This fits the volatility structure of the gold price, its dependence on the level of the gold price, and the long-term distribution.

One of the most important aspects in the Ashanti story is the effect of the derivatives portfolio on the share price. In particular, the relationship between the price of gold and the Ashanti share price is not what might be expected from a gold mining company ... and if you are an investor in a gold mining company this is something you should be aware of. The question is: Did Ashanti shareholders know that an investment in Ashanti was to some extent a bet on the gold price falling?

To see how this works, I've constructed a very simple model of the gold price/share price relationship using basic real option theory. Model assumptions are

- above stochastic model for price of gold;
- cost of \$235 per ounce cost of extraction, this includes both physical and corporate costs;
- simple expectation valuation with optimal closure;
- the company's profit comes solely from selling gold.

The results of this analysis are shown in the following figure, Figure 73.5. I've 'calibrated' the share price so that in July 1999 the theoretical share price, the faint line, and actual, the bold line, are the same, given the gold price. As you'd expect, as the gold price rises, from July to September, the Ashanti share price also rises. The theoretical price is very close to the actual. Now see what happened when the derivatives situation became known. From September

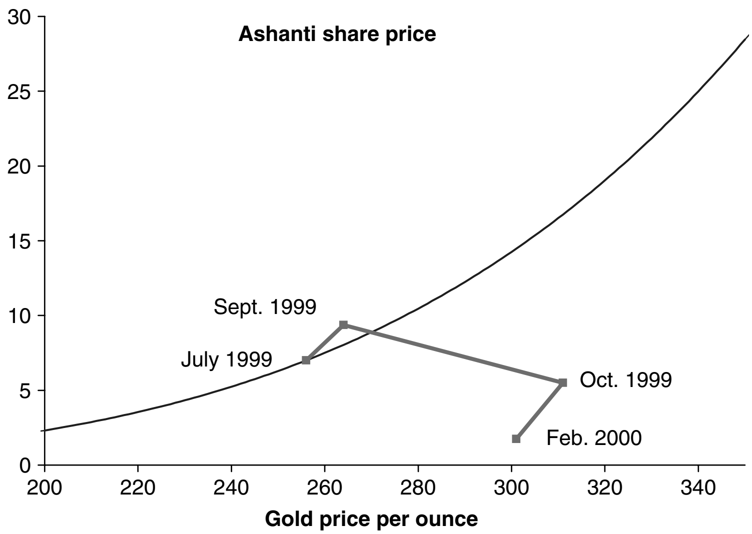


Figure 73.5 Ashanti share price.

to October the gold price rose significantly. Had Ashanti not been 'hedging' we might have expected an even greater rise in its share price. But this was not to be. The relationship between the gold price and the Ashanti share price became inverted as seen in the bold line.

Another output from this model is the gold price at which Ashanti should cease operations. We find this to be US$210 per ounce, a level fortunately not seen since the late 1970s.

This model can be made as sophisticated as you want. Putting derivatives into the picture is straightforward enough. This will change the share price versus gold price structure. For example, buying puts would raise the company value for low gold price, perhaps even making closure suboptimal. This would come at some cost, lowering the company's current value. Alternatively, the purchase of puts could be funded by writing calls. Again this would change the pricing structure. At this point the question arises as to what is best, what is trying to be achieved, perhaps what is 'optimal.'

73.9 SUMMARY

The theory of real options has been applied to marriage, d-i-v-o-r-c-e and suicide. It is harder to model these because the benefits are rarely perceived in simple quantitative terms, such as monetary value ... except by dumb blondes. However, there is little that cannot be mathematically modeled. The following is taken *verbatim* from Dixit & Pindyck (1994) p. 24 since it is so well written, and so upbeat about the possibilities in math modeling.

> "Perhaps the most extreme example of economics applied to sociological phenomena is the Beckerian theory of suicide developed by Hamermesh and Soss (1974). According to them, an individual will end his or her own life when the expected present value of the utility of the rest of the life falls short of a benchmark or cutoff standard 'zero.' Most people react by saying that the model gives an excessively rational view of what is inherently an irrational action. Our theory [Dixit & Pindyck] suggests exactly the opposite. Whatever its merits or demerits

as descriptive theory, the Hamermesh–Soss model is *not rational enough* from the prescriptive viewpoint, because it forgets the option value of staying alive. Suicide is the ultimate irreversible act [and unhedgeable], and the future has a lot of ongoing uncertainty. Therefore the option of waiting to see if 'something will improve' should be very large. The circumstances must be far more bleak than the cutoff standard of the Hamermesh–Soss model to justify pulling the trigger. This is true even if the expected direction of life is still downward; all that is needed is some positive probability on the upside.

"Now return to the argument that most suicides are irrational, and ask exactly how they fail to be rational. There are several possibilities, but one seems especially pertinent. Suicides project the bleak present into an equally bleak future, ignoring uncertainty, and thereby ignoring the option value of life. Then religious or social proscriptions against suicide serve a useful function as measures to compensate for this failure of rationality [an irrationality compensating for a lack of rationality?]. By raising the perceived cost of the act, these taboos lower the threshold of quality of life that leads to suicide when the option value is ignored. This can correct the failure of the individuals' forethought, and bring their threshold in conformity with the optimal rule that recognizes the option value."

FURTHER READING

- For details of the model is Section 73.4 see McDonald & Siegel (1986).
- The classic text on Real option theory with well-explained math is by Dixit & Pindyck (1994).
- The continuous investment project was first developed by Majd & Pindyck (1987)
- More recent, though generally less technical, texts are Trigeorgis (1998) and Amram & Kulatilaka (1999)
- For modeling of suicide see Hamermesh & Soss (1974).

CHAPTER 74
life settlements and viaticals

In this Chapter...

- life
- sex
- death

74.1 INTRODUCTION

And now for something completely ... morbid.

Life settlements and **viaticals** are contracts associated with death. The two expressions can be used interchangeably, although 'viatical' is often associated with a life with an already diagnosed terminal illness. With these contracts expiration means precisely that.

Life settlements are a secondary market for the life insurance policies held by individuals. These individuals may, typically later in life, want to sell their policy. The common reasons are

- they can no longer afford to pay the life-insurance premium;
- as part of estate planning;
- they need the money to pay for medical treatment;
- they no longer need the life insurance.

The policy is usually worth a lot more than its surrender value.

Many of these life insurance policies are then usually packaged together and sold as one product. To the quant, the question is how to model and price, and hedge, individual policies and portfolios of policies.

74.2 LIFE EXPECTANCY

In estimating the chances of an individual dying at different ages you typically have non-specific and specific pieces of information. The first comes from actuarial tables for populations as a whole. Such tables may be specific to the 'type' of person selling their policy, in the sense that it may refer to their sex, whether or not they smoke, the illness(es) they suffer from. But it won't be specific to the individual person. For that sort of information the policy purchaser will require the individual to be examined by an approved **life expectancy provider**. This will be a medical examiner (ME) who will examine the policyholder to determine a **life expectancy** for that individual. Although termed a 'life expectancy' (LE) the number given by the ME will actually be an estimate for the 'median' of the person's life, and not its expected value or average. The ME will often include in their estimation an allowance for expected new cures for terminal illnesses. Figure 74.1 shows a typical LE certificate from a LE provider.

[redacted]

Life Expectancy Certificate

[redacted]

Insured: [redacted]
Soc. Sec. No.: 911-21-4321
DoB: 8/4/1924
Gender: M
Smoking Non smoker

Life expectancy estimate: 76 months

[redacted]

Summary of relevant medical history:

The insured has a diagnosis of stage 0 chronic lymphocytic leukaemia with sustained lymphocytosis. His CLL is stable and requires no treatment at this time. No overt cardiovascular issues are present although there is a history of elevated cholesterol. [redacted]

Family medical history, social history, functional status:

He is a non smoker. A history of left knee and hip replacements was noted. However, he is able to complete activities of daily living without assistance.

Years	Lives	Deaths	Accum. deaths	Years	Lives	Deaths	Accum. deaths
1	895	105	105	13	2	2	998
2	772	123	228	14	1	1	999
3	604	168	396	15	0	1	1000
4	470	134	530	16	0	0	1000
5	352	118	648	17	0	0	1000
6	242	110	758	18	0	0	1000
7	152	90	848	19	0	0	1000
8	93	59	907	20	0	0	1000
9	46	47	954	21	0	0	1000
10	21	25	979	22	0	0	1000
11	9	12	991	23	0	0	1000
12	4	5	996	24	0	0	1000

Figure 74.1 Life expectancy certificate.

74.2.1 Sex

One of the dominant factors affecting LE is sex. No, not how often you get it, although I'm sure that has an impact, but your gender. As every man knows, their life expectancy is invariably shorter than that for a woman, all things being equal. American males born now have a LE of 74 years; American females 79.5. Japanese woman have a LE of 83 years, the longest in the world. In the US, men smoke more cigarettes and drink more alcohol than women. Men are three times as likely as women to die from accidents and four times more likely to be murdered.

74.2.2 Health

The World Health Organization (WHO) recently designed a new system of measuring life expectancy that takes into account disability, and so measures 'healthy lifespan.' Under this system Japan still comes top of the LE table, with an average healthy lifespan of 75 years. Sierra Leone comes in bottom with 26 years. Under this system the US ranks 24th. Christopher Murray, Director of WHO's Global Programme on Evidence for Health Policy, said: 'Basically, you die earlier and spend more time disabled if you're an American rather than a member of most other advanced countries.' The WHO's reasons for the low ranking of the US are, amongst others, as follows.

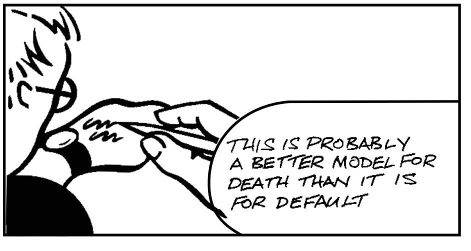
- Some ethnic groups and the inner city poor have bad health, more like that of the poor from a developing country rather than a rich country.
- HIV causes a higher proportion of death and disability to young and middle-aged Americans than in most other advanced countries.
- High incidence of cancer related to tobacco.
- High coronary heart disease rate.
- High levels of violence, especially homicides, compared to other industrial countries.

74.3 ACTUARIAL TABLES

Actuaries quantify probability of death in their famous tables. These tables give probabilities of dying within the year, say, according to a person's age, sex, smoking or non-smoking, and various health factors. A section of such a table is shown in Figure 74.2. The numbers in the cells are probabilities.

We will use such information but convert it into a form that we are already used to, that of the probability of default. Thus we shall interpret death as default, with no prospect of 'recovery.'

74.4 DEATH SEEN AS DEFAULT



Let us introduce some simple notation: $p(a)$ will be the probability of dying at age a. More precisely, the probability of dying between ages a and $a + da$ is

$$\int_a^{a+da} p(s)\, ds,$$

Male Nonsmoker 2001 Valuation Basic Table

Issue Age						Duration			
	1	2	3	4	5	6	7	8	9	10
0	0.00090	0.00049	0.00032	0.00020	0.00014	0.00014	0.00015	0.00015	0.00015	0.00016
1	0.00040	0.00029	0.00019	0.00014	0.00014	0.00014	0.00015	0.00015	0.00015	0.00016
2	0.00028	0.00019	0.00014	0.00014	0.00014	0.00014	0.00015	0.00015	0.00015	0.00016
3	0.00017	0.00013	0.00013	0.00014	0.00014	0.00014	0.00015	0.00015	0.00015	0.00018
4	0.00013	0.00013	0.00013	0.00014	0.00014	0.00014	0.00015	0.00015	0.00018	0.00023
5	0.00013	0.00013	0.00013	0.00014	0.00014	0.00014	0.00015	0.00017	0.00022	0.00027
6	0.00013	0.00013	0.00013	0.00014	0.00014	0.00014	0.00017	0.00022	0.00026	0.00038
7	0.00013	0.00013	0.00013	0.00014	0.00014	0.00017	0.00021	0.00026	0.00037	0.00057
8	0.00013	0.00013	0.00013	0.00014	0.00016	0.00020	0.00025	0.00037	0.00056	0.00068
9	0.00013	0.00013	0.00013	0.00016	0.00019	0.00025	0.00036	0.00056	0.00068	0.00076
10	0.00013	0.00013	0.00015	0.00019	0.00025	0.00035	0.00056	0.00067	0.00076	0.00080
11	0.00013	0.00015	0.00018	0.00024	0.00034	0.00056	0.00067	0.00076	0.00079	0.00079
12	0.00014	0.00017	0.00024	0.00033	0.00056	0.00067	0.00076	0.00078	0.00078	0.00078
13	0.00016	0.00023	0.00032	0.00056	0.00067	0.00076	0.00077	0.00077	0.00077	0.00077
14	0.00022	0.00031	0.00056	0.00067	0.00076	0.00077	0.00077	0.00077	0.00077	0.00077
15	0.00030	0.00056	0.00067	0.00076	0.00076	0.00076	0.00076	0.00076	0.00076	0.00076
16	0.00056	0.00067	0.00075	0.00075	0.00075	0.00075	0.00075	0.00075	0.00075	0.00076
17	0.00067	0.00074	0.00074	0.00074	0.00074	0.00074	0.00074	0.00074	0.00075	0.00077
18	0.00073	0.00073	0.00073	0.00073	0.00073	0.00073	0.00073	0.00073	0.00075	0.00075
19	0.00070	0.00070	0.00070	0.00070	0.00070	0.00070	0.00070	0.00072	0.00072	0.00072
20	0.00066	0.00066	0.00066	0.00066	0.00066	0.00067	0.00067	0.00068	0.00068	0.00068
21	0.00060	0.00060	0.00060	0.00060	0.00061	0.00062	0.00063	0.00063	0.00064	0.00065
22	0.00055	0.00055	0.00055	0.00056	0.00058	0.00058	0.00059	0.00060	0.00060	0.00063
23	0.00049	0.00049	0.00051	0.00053	0.00055	0.00056	0.00057	0.00058	0.00059	0.00062
24	0.00043	0.00044	0.00048	0.00051	0.00054	0.00055	0.00056	0.00058	0.00060	0.00063
25	0.00035	0.00040	0.00046	0.00051	0.00053	0.00055	0.00057	0.00058	0.00062	0.00066
26	0.00032	0.00039	0.00046	0.00052	0.00054	0.00056	0.00057	0.00061	0.00064	0.00069
27	0.00030	0.00039	0.00046	0.00053	0.00056	0.00057	0.00061	0.00064	0.00068	0.00074
28	0.00030	0.00039	0.00047	0.00053	0.00057	0.00060	0.00063	0.00068	0.00073	0.00079
29	0.00029	0.00039	0.00048	0.00054	0.00060	0.00063	0.00068	0.00073	0.00078	0.00087
30	0.00028	0.00039	0.00047	0.00055	0.00062	0.00068	0.00073	0.00078	0.00086	0.00095
31	0.00027	0.00038	0.00047	0.00055	0.00063	0.00071	0.00078	0.00084	0.00091	0.00099
32	0.00026	0.00036	0.00047	0.00055	0.00064	0.00072	0.00080	0.00087	0.00094	0.00102
33	0.00027	0.00038	0.00048	0.00058	0.00067	0.00076	0.00085	0.00091	0.00098	0.00107
34	0.00029	0.00039	0.00051	0.00060	0.00071	0.00080	0.00090	0.00097	0.00105	0.00116
35	0.00031	0.00041	0.00052	0.00063	0.00073	0.00084	0.00094	0.00104	0.00115	0.00130
36	0.00032	0.00043	0.00054	0.00065	0.00077	0.00087	0.00099	0.00114	0.00129	0.00147
37	0.00034	0.00044	0.00056	0.00067	0.00078	0.00091	0.00107	0.00126	0.00145	0.00163
38	0.00037	0.00049	0.00061	0.00071	0.00083	0.00098	0.00118	0.00140	0.00161	0.00179
39	0.00040	0.00054	0.00066	0.00077	0.00090	0.00108	0.00132	0.00154	0.00176	0.00194
40	0.00043	0.00058	0.00071	0.00083	0.00099	0.00121	0.00146	0.00170	0.00189	0.00209

Figure 74.2 A section of an actuarial table.

or, for small da, approximately

$$p(a)\, da.$$

This function we will get from actuarial tables. In practice the function will vary from person to person; we will deal with this later. It will also be a function of time, not just age. The probability of a 60-year old dying before their 61st birthday is different now from what it was 50 years ago. So we should write $p(a, t)$. Again, we'll come back to this point later on. A typical such $p(a)$ is shown in Figure 74.3.

If we write $P(a; a_0)$ as the probability of still being alive at age a given that you were definitely alive at age a_0 then

$$\frac{dP}{da} = -p(a)P,$$

so that

$$P(a; a_0) = \exp\left(-\int_{a_0}^{a} p(s)\, ds\right).$$

See Figure 74.4.

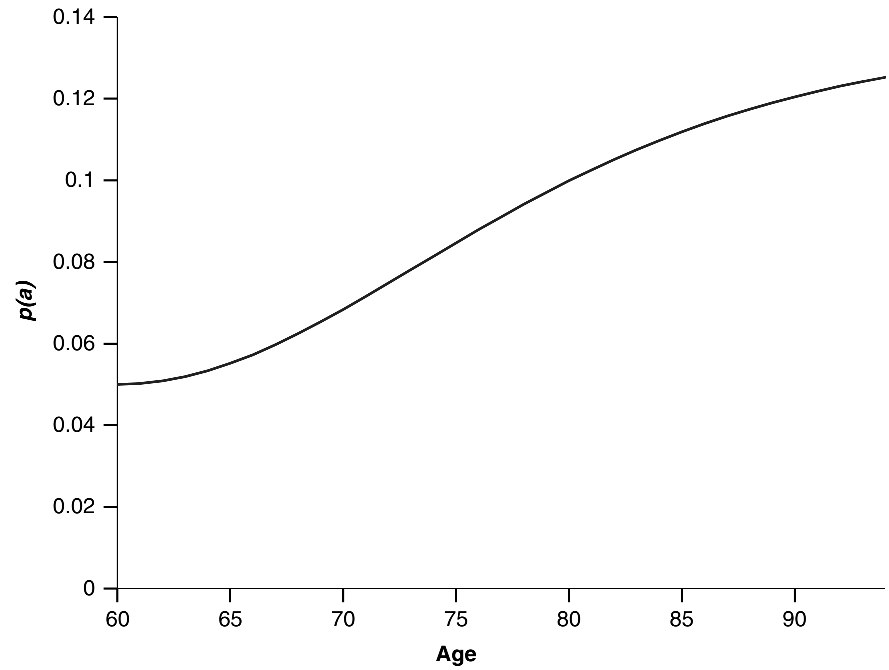


Figure 74.3 $p(a)$.

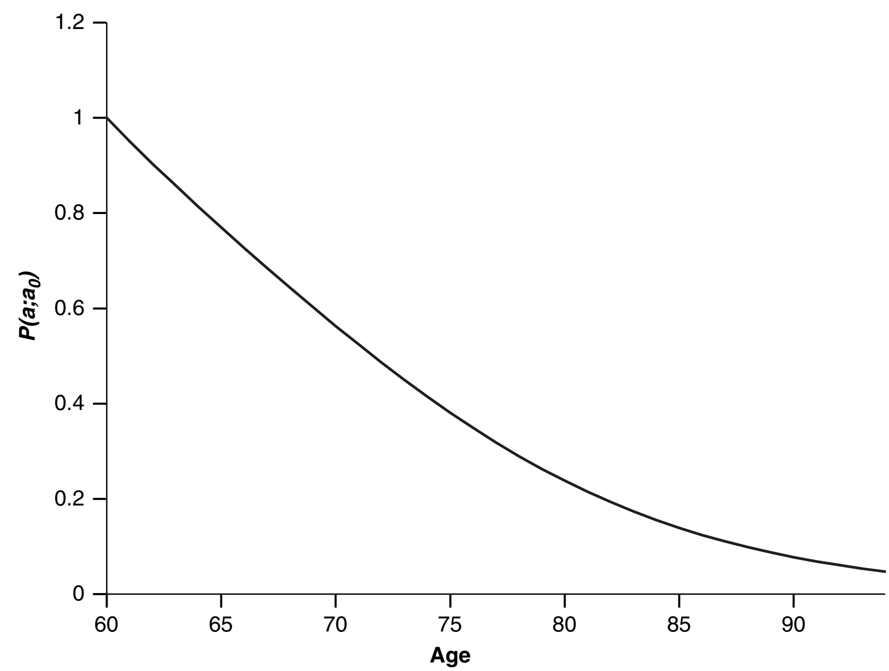


Figure 74.4 $P(a; a_0)$.

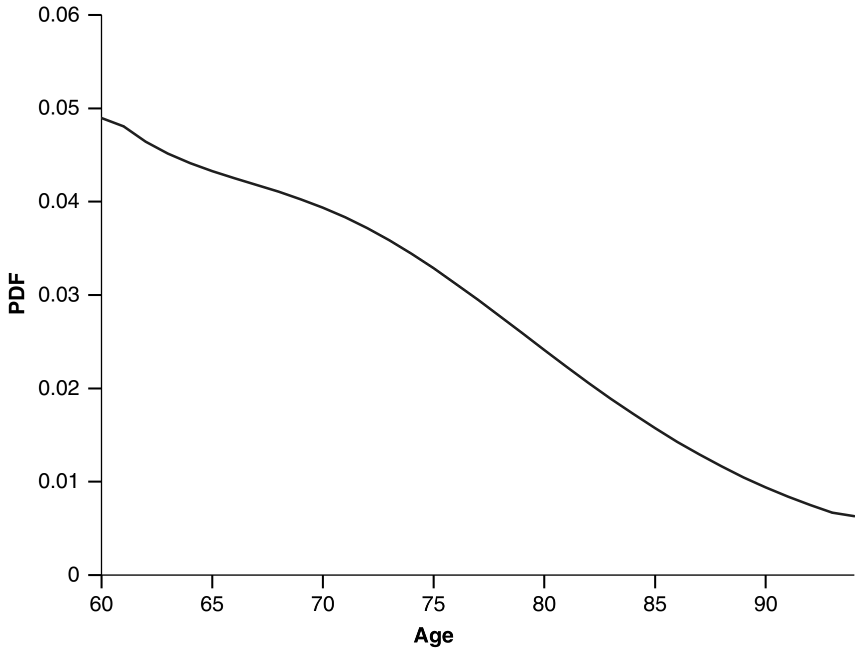


Figure 74.5 Probability function for age at death.

The median, or life expectancy, is that a such that

$$\exp\left(-\int_{a_0}^{a} p(s)\,ds\right) = \tfrac{1}{2}.$$

The probability density function for age at death is

$$-\frac{dP}{da} = p(a)\,\exp\left(-\int_{a_0}^{a} p(s)\,ds\right).$$

See Figure 74.5. Therefore, the expected age at death is

$$\int_{a_0}^{\infty} a\;p(a)\;\exp\left(-\int_{a_0}^{a} p(s)\,ds\right) da.$$

74.5 PRICING A SINGLE POLICY

We now move on to pricing of a single life policy. Here is an example:

Age	70
Principal	\$1,000,000
Monthly premiums	\$2,083
Policy purchase price	\$375,000
Life expectancy	5.2 years

This means that the policy holder is aged 70 when the policy is sold. It was sold for $375,000. The purchaser must take over payment of the monthly insurance premiums $2,083 while the insured is still alive. On his/her death the purchaser will receive $1,000,000. The policy holder has a life expectancy of 5.2 years.

To value this policy I have taken a published table of mortality rates, and adjusted them to give approximately the correct LE. (Of course, this can be done in any number of ways.)

Adjust1 = 0.02
Adjust2 = 0.025
If Age > 70 Then ProbabilityOfDeath = 0.01
If Age > 75 Then ProbabilityOfDeath = 0.02
If Age > 80 Then ProbabilityOfDeath = 0.036
If Age > 85 Then ProbabilityOfDeath = 0.07
If Age > 90 Then ProbabilityOfDeath = 0.1
If Age > 95 Then ProbabilityOfDeath = 0.13
If Age > 100 Then ProbabilityOfDeath = 0.19
If Age > 105 Then ProbabilityOfDeath = 0.2
If Age > 110 Then ProbabilityOfDeath = 100000
ProbabilityOfDeath = ProbabilityOfDeath + Adjust1 + Adjust2 × (Age − 70)

The resulting distribution of ages at death is given in Figure 74.6. This plot and all results that follow are based on simple Monte Carlo simulations of ages at death, not on any analytic calculations. The standard deviation of ages at death is 3.58.

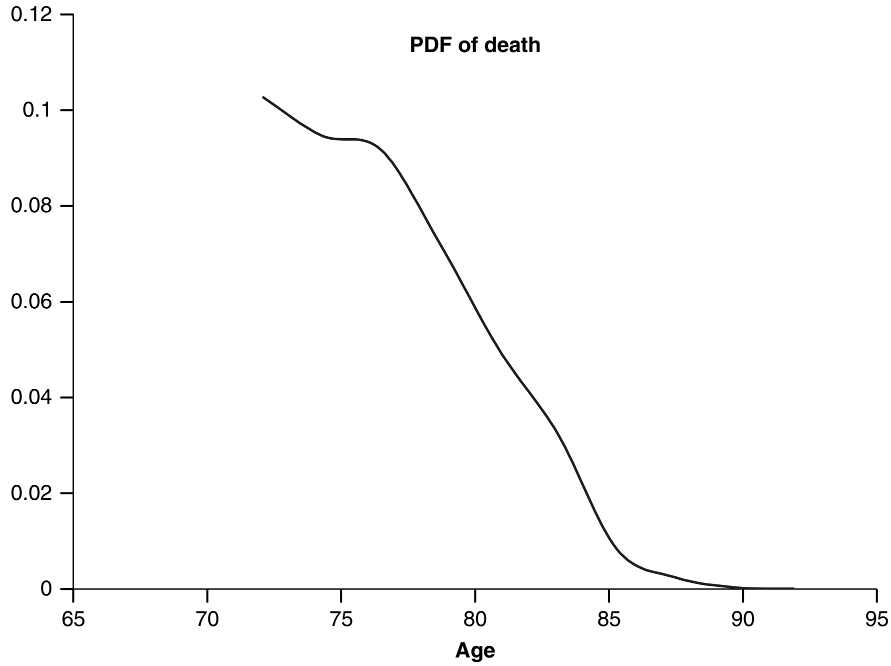


Figure 74.6 Probability function for age at death, 10,000 Monte Carlo simulations.

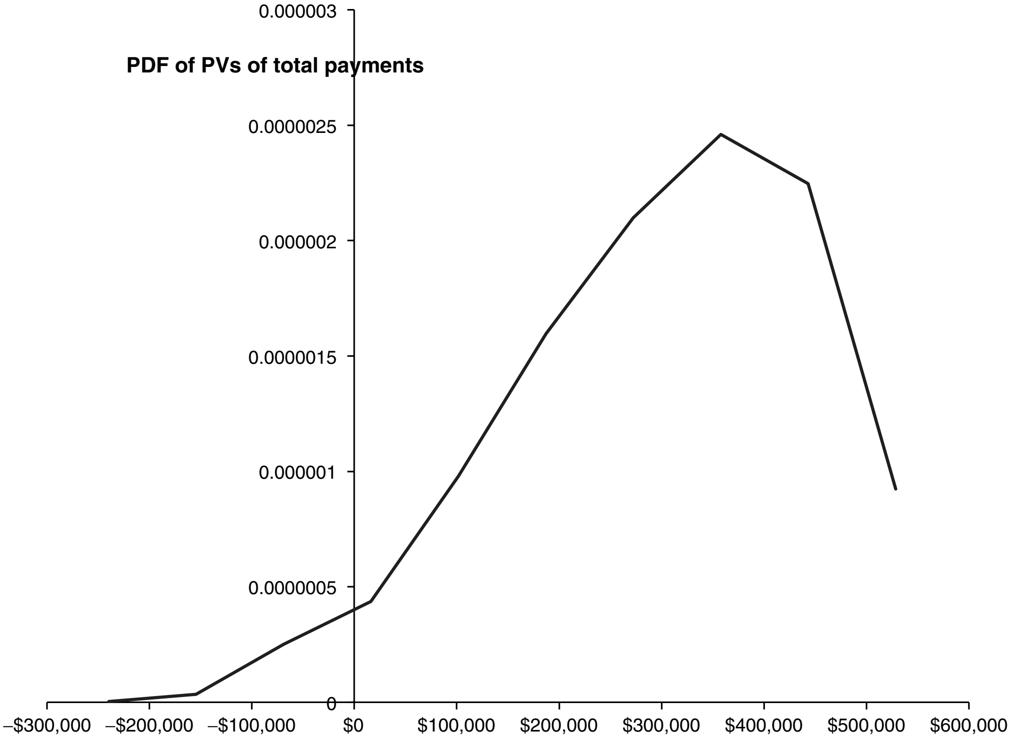


Figure 74.7 Probability distribution of present value of all cashflows, 10,000 Monte Carlo simulations.

Associated with this policy are various cashflows as far as the purchaser is concerned: The purchase payment; the monthly premiums; the final principal. For each of the simulations of death we can present value all of these cashflows and so get a distribution of present values. This is shown in Figure 74.7, using a 3% interest rate. The mean is \$310,120, the standard deviation \$157,401. So, on average you will almost double your money if you buy this policy for the \$375,000.

In Figure 74.8 you can see how this present value varies with age at death. After 84 years you start to lose money on this investment.

74.5.1 Internal Rate of Return

Most investors (i.e. purchasers) in these products do not think in terms of present value. Instead they like to work in terms of internal rate of return, or yield. So for a policy such as the above, rather than know the expected PV of all cashflows, they would like to know the expected IRR. Figure 74.9 shows the distribution of IRR. The average is 25.7% and the standard deviation is 36.5%.

One of the problems with using the IRR as a measure of the 'value' of this policy is that the calculation is very sensitive to age of death, especially at the short end. When someone dies immediately after the policy has been purchased you will find the IRR is infinite. So IRR is not necessarily that informative.

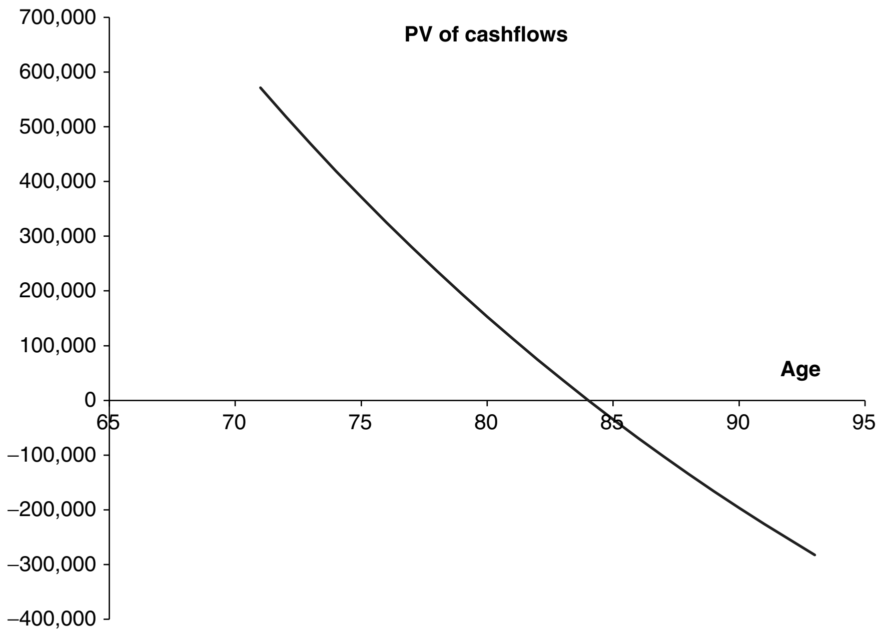


Figure 74.8 Present value of all cashflows versus age at death.

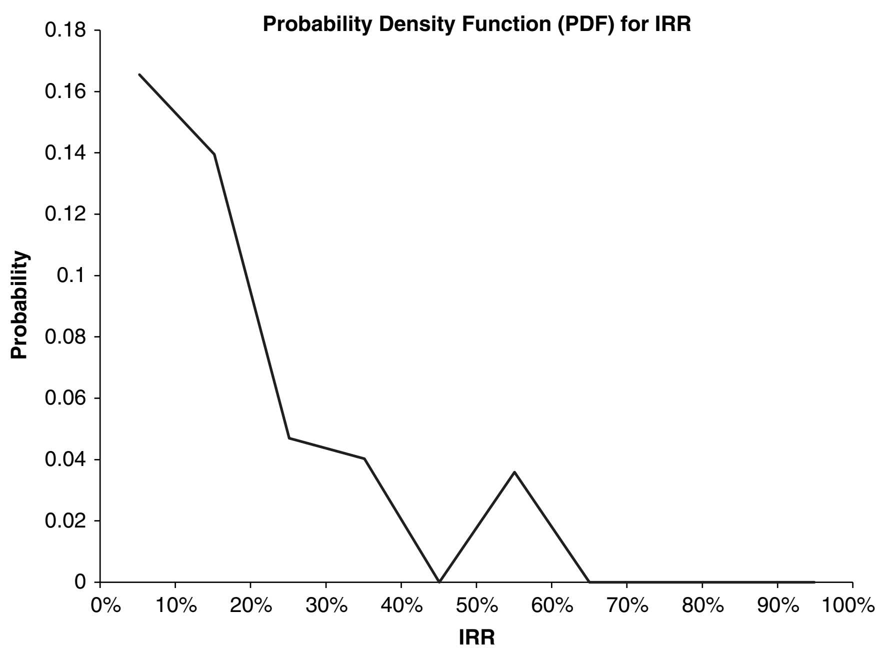


Figure 74.9 Probability distribution of IRR.

74.6 PRICING PORTFOLIOS

People invest in these portfolios by buying many, dozens or hundreds, at a time. This reduces exposure to individual lives. The process of pricing, calculating risks and IRRs is exactly the same for a portfolio of many policies as it is for a single policy, except that you must simulate all of the underlying lives simultaneously. At least we can reasonably assume that deaths are not correlated.

Let's crunch the numbers for many policies, all identical to the one analyzed above. In practice, of course, the policies and actuarial tables will vary from policy to policy. But this is just a detail needed in programming, not in concept.

In Figures 74.10–74.12 are the probability density functions for the present value of all cashflows and for the IRRs, with the number of identical policies being five, 20 and 100.

The statistics for these portfolios are shown in the following table.

No. of policies	SD of PV ($)	Av. IRR (%)	SD of IRR (%)
1	157,401	25.7	36.5
5	70,554	15.4	7.44
10	50,302	14.6	4.76
20	34,467	13.9	2.88
100	15,531	14.0	1.30
500	7,240	13.8	0.60

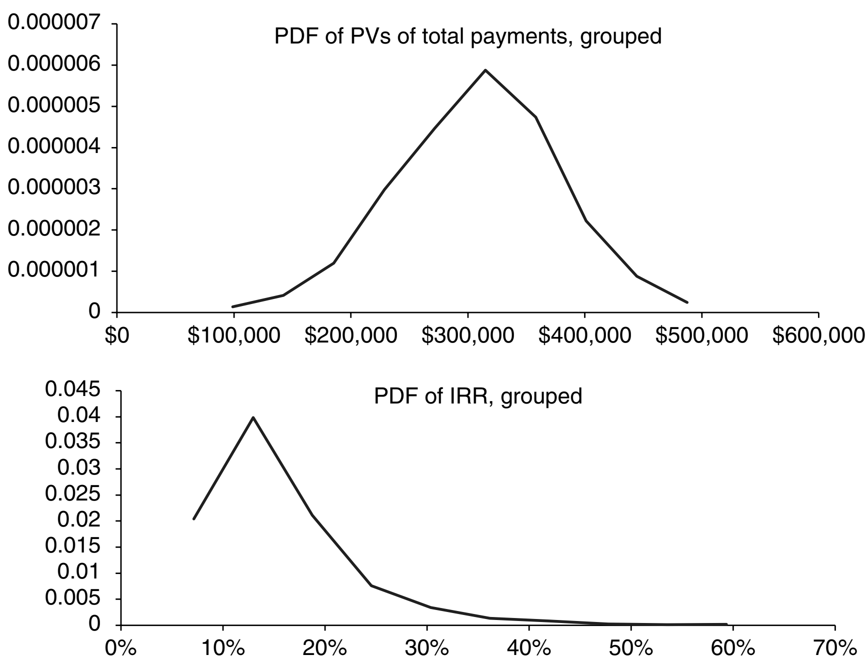


Figure 74.10 Probability distributions, five policies.

Figure 74.11 Probability distributions, 20 policies.

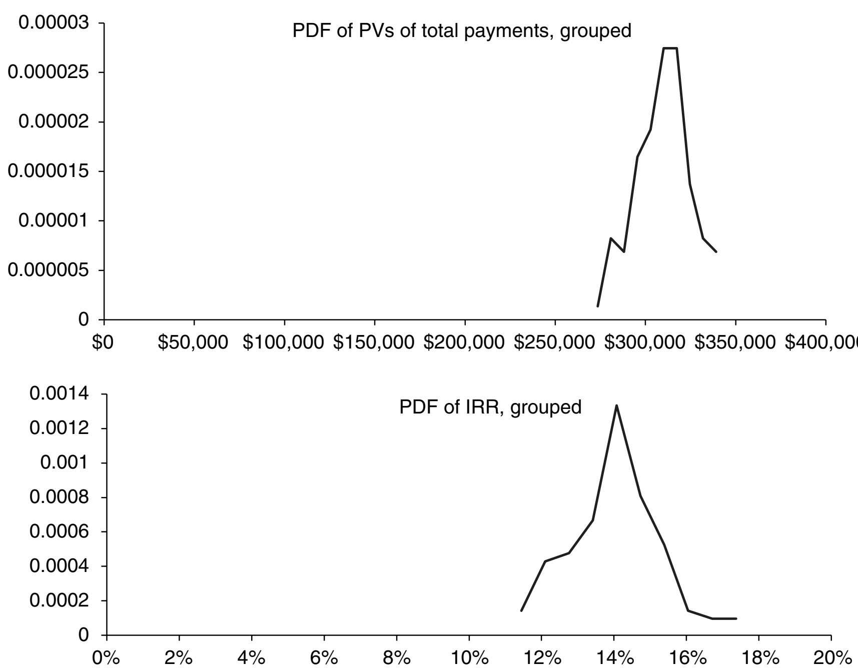


Figure 74.12 Probability distributions, 100 policies.

74.6.1 Extension Risk

If you buy one of these policies or portfolios, the last thing you want is for the insured to have a long life. Obviously, each life is going to be difficult to forecast, but when you look at portfolios you hope that some form of the Central Limit Theorem will be working. One of your major exposures, however, is to a systematic extension to lifespans generally, so called **extension risk**. The unexpected cure for otherwise terminal illnesses, for example. On a positive note—positive for the owner of the contract if not the insured persons—there are factors that may systematically decrease lifespans.

Factors systematically increasing lifespans:

- Health care improvements
- Health education improvement
- Improvements in safety
- Improvements in nutrition
- Therapeutic advances
- Genetic research
- Early detection
- Improved access to medical care
- Globalization—more research and advances
- Environmental changes
- Shifting demographics
- Diet
- Telecommuting
- Cell phone usage
- Technology

Factors systematically decreasing lifespans:

- Catastrophes
- Terrorist activities
- War
- Globalization—prone to more diseases
- Natural disasters
- Nuclear accidents
- Global warming
- Outbreak of disease
- Stress
- Drug-resistant germs
- Pollution
- Smoking
- Side effects of new drugs
- Shifting demographics
- Increasing wealth disparity
- Cell phone usage
- Lack of sleep

(Source: Klein, 2002)

It appears that lately the probability of dying at any age is decreasing by a factor of about 2% per annum. So that if the probability of dying at age a today is p then the probability of someone aged a next year, dying next year is 0.98 p. Probability of death is not time homogeneous. And this is where the more general function $p(a, t)$ comes in; it is the hazard rate for a person aged a at time t. For most people in most countries this function is slowly decreasing with time as health generally improves. However, there are some exceptions to this even in developed countries.

74.7 SUMMARY

This is a subject that is becoming popular with quants, and as such we should expect a little bit more in the way of sophistication of the modeling in the future. This will be driven in the main by the increasing amount of securitization of these life products, portfolios of policies lumped together and sold as a package.

FURTHER READING

- See Stephen Jay Gould's *The median isn't the message* (1985) about his diagnosis of and fight against abdominal mesothelioma.
- See Klein (2002) for a discussion of the future of mortality forecasting.

APPENDIX: THE AGE OF QUANTS

Here's a fun bit of mathematics, not totally unrelated to the above, and which shows you a little bit more of the mathematics of life and death. It concerns the probability distribution of the age of quants and the age at which people first become a quant. Of course, the same idea can be applied to other problems as well.

Suppose $N(a, t)$ is the number of quants in the world who are aged a at time t. There is a steady stream of newbies into the business, so that in a time step dt a number $n(a)\,dt$ of people of age a become quants. (I have kept this independent of time for simplicity.)[1]

The increment in number of quants aged a in a time step dt is caused by newcomers joining, and quants aging (an amount dt as well). So

$$N(a, t + dt) = N(a - dt, t) + n(a)\,dt.$$

From Taylor series we get

$$\frac{\partial N}{\partial t} + \frac{\partial N}{\partial a} = n(a).$$

The general solution of this is

$$N(a, t) = f(a - t) + \int_0^a n(\tau)\,d\tau,$$

[1] I have also been a bit loose with the definition of N, it should also contain some 'bucket' size.

where $f(\cdot)$ is an arbitrary function. Let's suppose that at time $t = 0$ there were no quants, $N(a, 0) = 0$. This means we can find f and we end up with

$$N(a,t) = \int_{a-t}^{a} n(\tau)\,d\tau.$$

Furthermore suppose that all quants retire at age a_R, then the probability density function for the age of quants is just

$$P(a,t) = \frac{N(a,t)}{\displaystyle\int_0^{a_R} N(a,t)\,da} = \frac{\displaystyle\int_{a-t}^{a} n(\tau)\,d\tau}{\displaystyle\int_0^{a_R}\int_{a-t}^{a} n(\tau)\,d\tau\,da}.$$

After a long time this settles down to the steady-state distribution

$$P_\infty(a) = \frac{\displaystyle\int_0^{a} n(\tau)\,d\tau}{\displaystyle\int_0^{a_R} (a_R - \tau)\,n(\tau)\,d\tau}.$$

If we know $n(a)$ we can find $P_\infty(a)$ or vice versa.

One of the unrealistic bits in the above is that all quants retire at the same age. Suppose instead that a proportion of quants retire (or change job, or die, etc.), and let's make the coefficient of proportionality a function of age, $\alpha(a)$. With N and n having the same meanings as before we now get

$$\frac{\partial N}{\partial t} + \frac{\partial N}{\partial a} = n(a) - \alpha(a)\,N.$$

The general solution of this is

$$N(a,t) = f(a-t)\,e^{-\int_0^a \alpha(s)\,ds} + e^{-\int_0^a \alpha(s)\,ds}\int_0^a n(\tau)\,e^{\int_0^\tau \alpha(s)\,ds}d\tau,$$

where $f(\cdot)$ is again an arbitrary function. The solution having $N(a, 0) = 0$ is

$$N(a,t) = e^{-\int_0^a \alpha(s)\,ds}\int_{a-t}^a n(\tau)\,e^{\int_0^\tau \alpha(s)\,ds}d\tau.$$

We can then scale this to find the probability density function. The steady-state limit is

$$P_\infty(a) = \frac{e^{-\int_0^a \alpha(s)\,ds}\displaystyle\int_0^a n(\tau)\,e^{\int_0^\tau \alpha(s)\,ds}d\tau}{\displaystyle\int_0^\infty e^{-\int_0^a \alpha(s)\,ds}\int_0^a n(\tau)\,e^{\int_0^\tau \alpha(s)\,ds}\,d\tau\,da}.$$

In the simple case that α is constant, $\overline{\alpha}$, this becomes

$$P_\infty(a) = \frac{\overline{\alpha}e^{-\overline{\alpha}a}\displaystyle\int_0^a e^{\overline{\alpha}\tau}n(\tau)\,d\tau}{\displaystyle\int_0^\infty n(\tau)\,d\tau}.$$

CHAPTER 75
bonus time

In this Chapter...

- modeling the bonus compensation of traders
- incorporating a skill factor into valuation models

75.1 INTRODUCTION

Traders are compensated by their bonus, in addition to their basic salary. A typical bonus structure might be to pay out five or ten percent of the profit they have made for the bank in the previous year. This type of compensation has a major drawback. Assuming that the trader wants to maximize his expected bonus, this structure encourages him to take silly risks. For example, a trader with a negative trading account would take large risks in the hope of getting his trading account into the black.

Of course, the story is not quite as simple as this. Traders have limits imposed on their available capital and on the gearing and type of instruments they are allowed to trade. Nevertheless little is known about how to reward a good trader optimally. Most of the academic work has taken a game-theoretic approach but hasn't seen much use in practice.

In this chapter we are going to build up a framework for the study of this problem and explore a variety of possible compensation structures. The mathematics of this chapter is similar to that required in the analysis of the passport option, Chapter 27.

75.2 ONE BONUS PERIOD

We've seen the analysis of the problem of the passport option in Chapters 27 and 63. Well, that's pretty much the same as the trader bonus problem when the bonus is a fraction of the (positive part of the) profit. Things get more interesting, and sensible perhaps, if the bonus depends also on the realized Sharpe ratio. Why should a bank want to compensate a trader depending on his Sharpe ratio? Because that way they exert some control over the risk that the trader takes.

75.2.1 Bonus Depending on the Sharpe Ratio

As always S is the asset price following

$$dS = \mu S\,dt + \sigma S\,dX.$$

Let π be the value of the trading account. It follows that

$$d\pi = r\pi\, dt + q(dS - rS\, dt),$$

where q is the position of the trader. He will have some restriction such as $|q| \leq C$, C is the position limit. Let I be the variance of the trading account. It follows that

$$I = \int_0^t q^2\sigma^2 S^2\, dt$$

and

$$dI = q^2\sigma^2 S^2\, dt.$$

At time $t = 0$, $\pi = I = 0$ and the trader begins to trade the underlying asset. At the end of the year, the bank gives to the trader a bonus depending on the profit made, $\pi(T)$ and the Sharpe ratio

$$\frac{\pi}{\sqrt{I}}.$$

I'm ignoring the risk-free interest rate that should be in the Sharpe ratio.

We will assume that the trader is trying to maximize his expected bonus. If $V(S, \pi, I, t)$ is that expectation then we get the pricing equation from

$$\max_{|q|\leq 1} E[dV - rV\, dt] = 0.$$

This equation says that the trader tries to maximize his expected bonus over the risk-free rate.

The equation for $V(S, \pi, I, t)$ is therefore

$$\begin{aligned}&\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} + r\pi \frac{\partial V}{\partial \pi} - rV \\ &\quad + \max_{|q|\leq 1}\left(q\left(S(\mu - r)\frac{\partial V}{\partial \pi} + \sigma^2 S^2 \frac{\partial^2 V}{\partial S \partial \pi}\right) + q^2\left(\tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial \pi^2} + \sigma^2 S^2 \frac{\partial V}{\partial I}\right)\right) = 0\end{aligned}$$

with final condition

$$V(S, \pi, I, T) = \max(\pi, 0) P\left(\frac{\pi}{\sqrt{I}}\right).$$

A suitable form for the function $P(\cdot)$ would be monotonically increasing from zero to a constant $c < 1$.

Note that we can assume that the trading limit $C = 1$ without loss of generality because $V_{C=C^*}(S, \pi, I, t) = V_{C=1}(C^* S, \pi, I, t)$.

This equation can be simplified. If we introduce the variables $z_1 = \pi/S$ and $z_2 = I/S^2$, then

$$V(S, \pi, I, t) = S\Phi(z_1, z_2, t),$$

where Φ satisfies

$$\frac{\partial\Phi}{\partial t}+(\mu-r)\Phi-(\mu-r)z_1\frac{\partial\Phi}{\partial z_1}+\max_{|q|\le 1}\left(q(\mu-r)\frac{\partial\Phi}{\partial z_1}+q^2\sigma^2\frac{\partial\Phi}{\partial z_2}+\tfrac{1}{2}\sigma^2(z_1-q)^2\frac{\partial^2\Phi}{\partial z_1^2}\right)=0$$

with

$$\Phi(z_1,z_2,T)=\max(z_1,0)P\left(\frac{z_1}{\sqrt{z_2}}\right).$$

75.2.2 Numerical Results

In the following numerical examples, the bonus is paid after one year. The real drift of the underlying is 10%, volatility 20% and zero risk-free rate. We have

$$P(x)=0.1\frac{x^4}{x^4+1}.$$

First we calculate the theoretical Φ and the theoretical strategy q, assuming that the trader wants to maximize his expected bonus. In Figure 75.1 is shown the function Φ. This is plotted one year before the bonus is paid. In Figure 75.2 we see the trading strategy q at time $t = T - 0.1$ At the start of the option, $t = 0$, this picture would have shown the very boring plane $q = 1$.

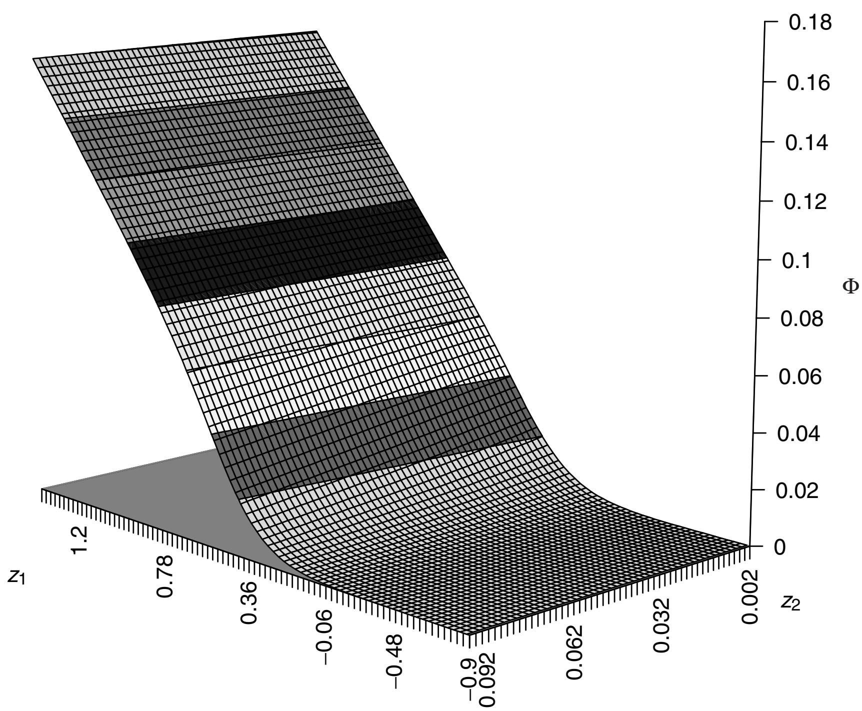


Figure 75.1 Value of Φ as a function of z_1 and z_2 at time $t = 0$.

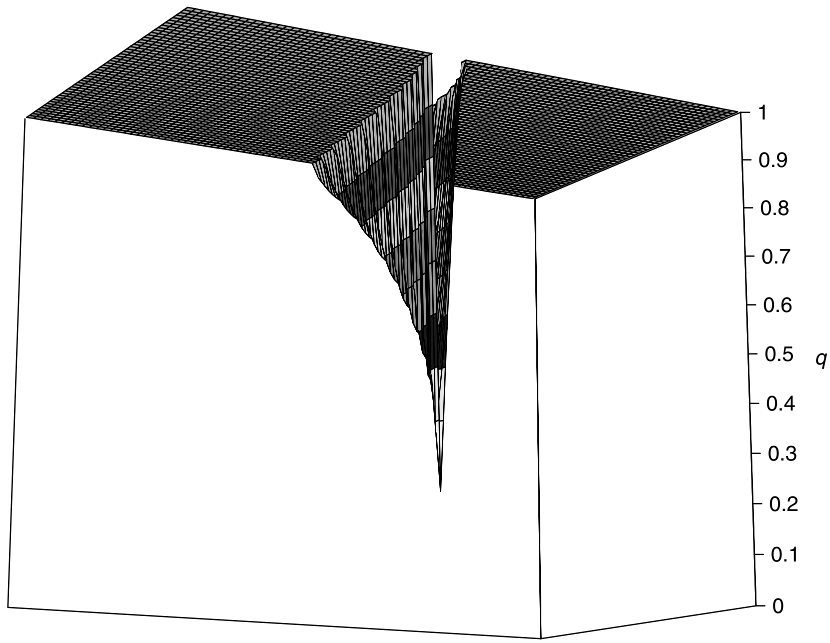


Figure 75.2 The trading strategy q at time $t = 0.9$ which maximizes the expected bonus.

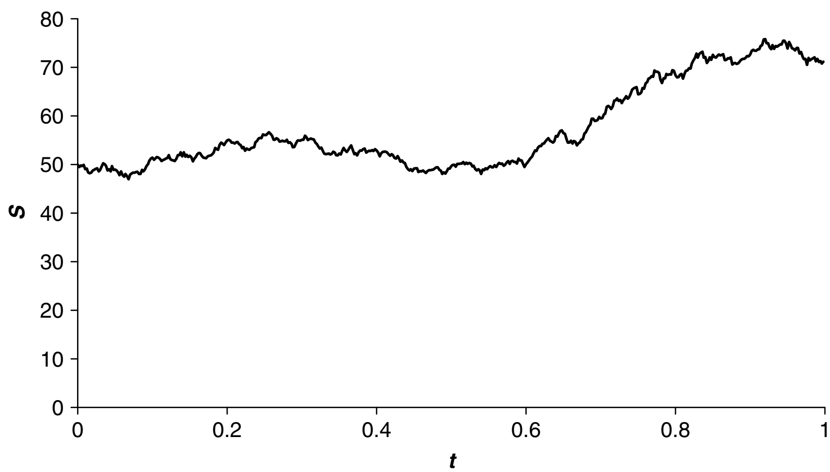


Figure 75.3 A simulation of the asset price S.

Now we'll look at one simulation of an asset price S, of the trading strategy of a trader q, of the trading account π, and of the Sharpe ratio $\pi/\sqrt{I}$. In Figure 75.3 is shown one simulation, one realization, of the asset S. The corresponding q is then shown in Figure 75.4. Observe that it is only towards the bonus time, the end of the year, that the trader stops holding

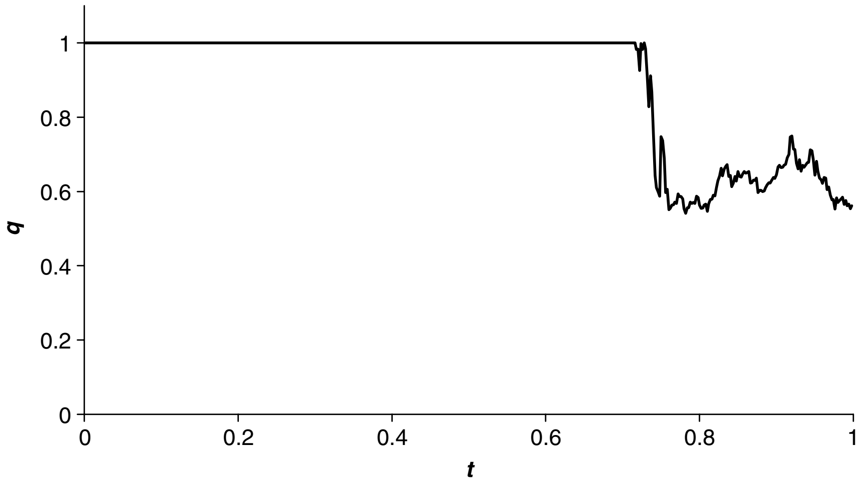


Figure 75.4 The strategy q, as realized.

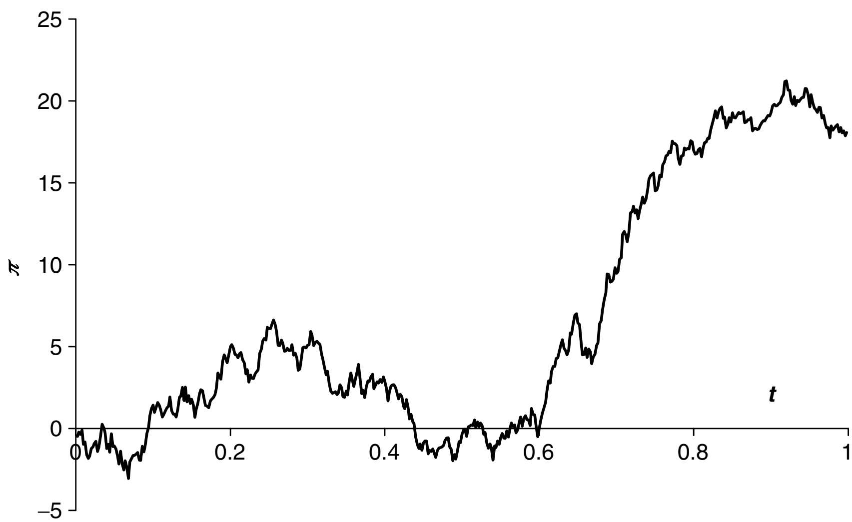


Figure 75.5 The evolution of the trading account.

the maximum long position. He is trying to reduce his volatility to increase his Sharpe ratio and get the highest bonus. The trading account is shown in Figure 75.5. This looks very like the asset path, except near the end where it is less volatile. The Sharpe ratio is then shown in Figure 75.6.

Now we perform 10,000 Monte Carlo simulations and plot the distributions of the profit made by the trader and by the bank. These are shown in Figures 75.7 and 75.8.

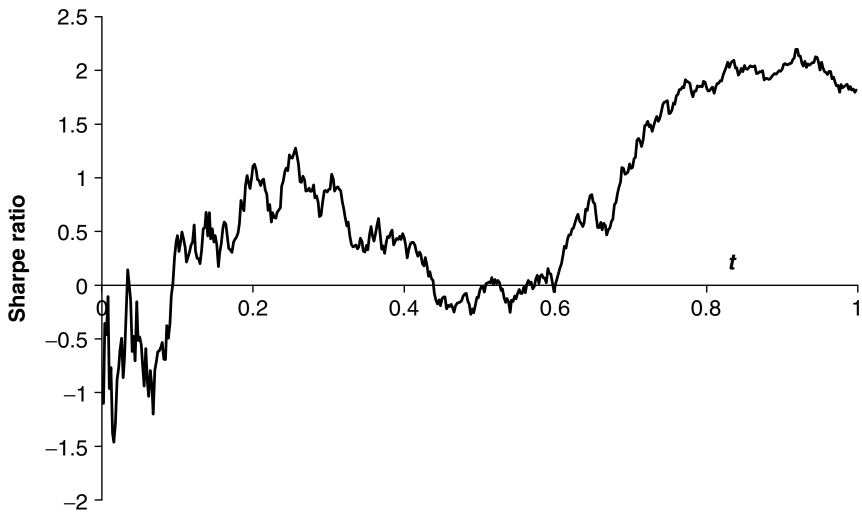


Figure 75.6 The evolution of the Sharpe ratio.

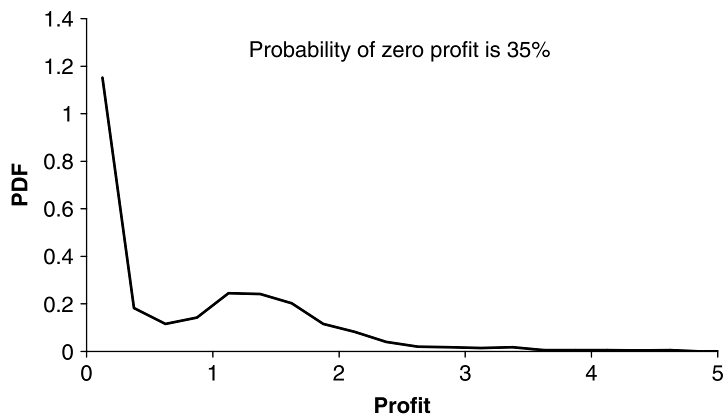
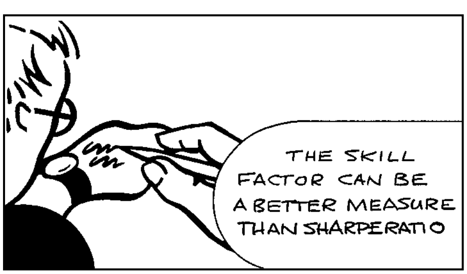


Figure 75.7 The distribution of the trader's profit: mean = 0.5143 and standard deviation = 0.7943.



75.3 THE SKILL FACTOR

There is little point in rewarding traders, or even hiring them, if they do not possess some skill above that of a chimpanzee. Often 'skill' is quantified by the Sharpe ratio. Here I want to suggest something a little more complex, but more realistic. I will model a possible way in which traders act, incorporating a **skill factor** that quantifies how much correct information they receive. I'll then back out an implied skill factor from the time series of the P&L of a trader friend. Finally, we'll put the skill factor into the bonus problem.

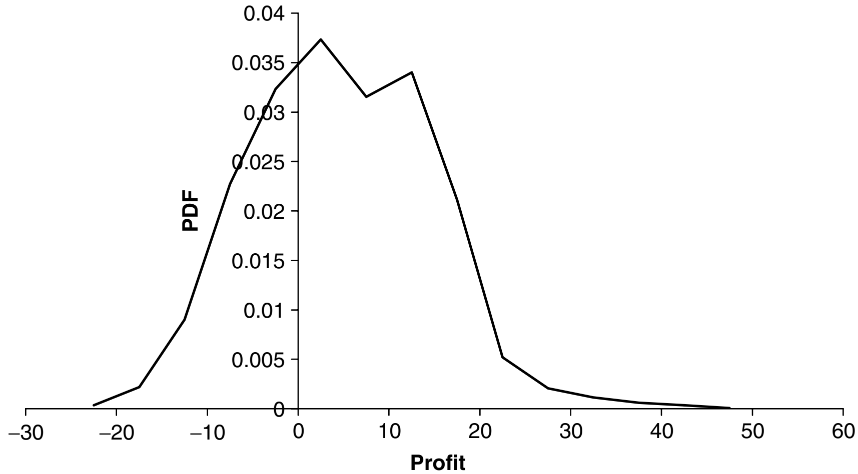


Figure 75.8 The distribution of the bank's profit: mean = 4.8498 and standard deviation = 9.8967.

I'm going to assume that our trader gets important and correct information about the direction of the market a fraction p of the time. If the information is that the market will rise, she[1] buys to a set limit, if the information is that the market will fall she sells to the same limit. The remaining $1-p$ she trades the same volume but randomly buying or selling.

Assuming Normal distributions of returns, ϕ, she makes a profit of

$$C\phi \quad \text{for} \;\; 1-p \;\; \text{of the time}$$

and

$$C|\phi| \quad \text{for} \;\; p \;\; \text{of the time.}$$

The latter represents the certain profit made when information is received. The factor C includes trade size, leverage and volatility of the underlying.

What is the distribution of profit and loss, ψ? The pdf of the distribution is

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\psi^2} \left(1 - p + 2p\mathcal{H}(\psi)\right).$$

The mean of this distribution is

$$Cp\sqrt{\frac{2}{\pi}}$$

and the standard deviation

$$C\sqrt{1 - \frac{2p^2}{\pi}}.$$

The probability of being correct is $p + \frac{1}{2}(1-p) = \frac{1}{2} + \frac{1}{2}p$.

[1] Do not worry, dear reader, I haven't gone all PC... in this very real example the data are from a female trader friend.

The next stage is to deduce from a trader's P&L time series the skill factor p. To do this properly requires a knowledge of all the quantities affecting C, such as currency, leverage, volume and volatility. In the example below I make rather sweeping assumptions about these.

Figures 75.9 and 75.10 show the accumulated profit (suitably scaled to hide the actual profits and with the risk-free return taken out) made by my trader friend in 1998 and half of 1999 respectively. Her Sharpe ratios for these years were 3.7 and 4.7. In Figure 75.11 are shown the pdfs for my friend's P&L for 1998, scaled to have a standard deviation of one, of course, and the theoretical distribution using the above model. I have fitted the two by matching means and

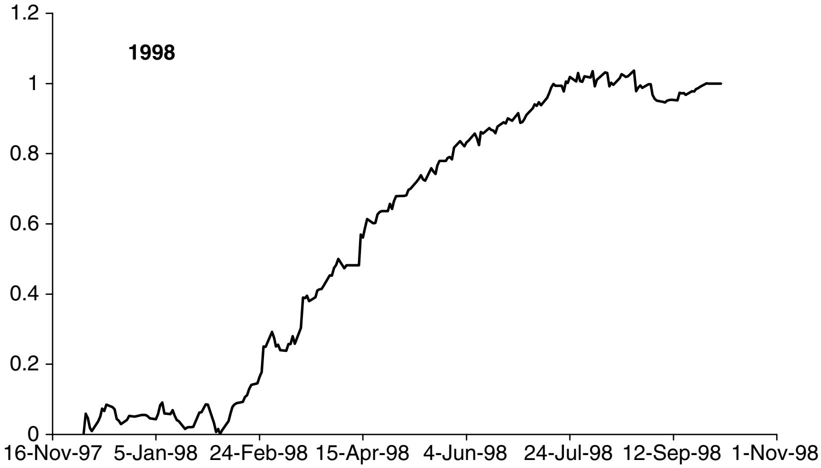


Figure 75.9 Trader friend's scaled profit for 1998.

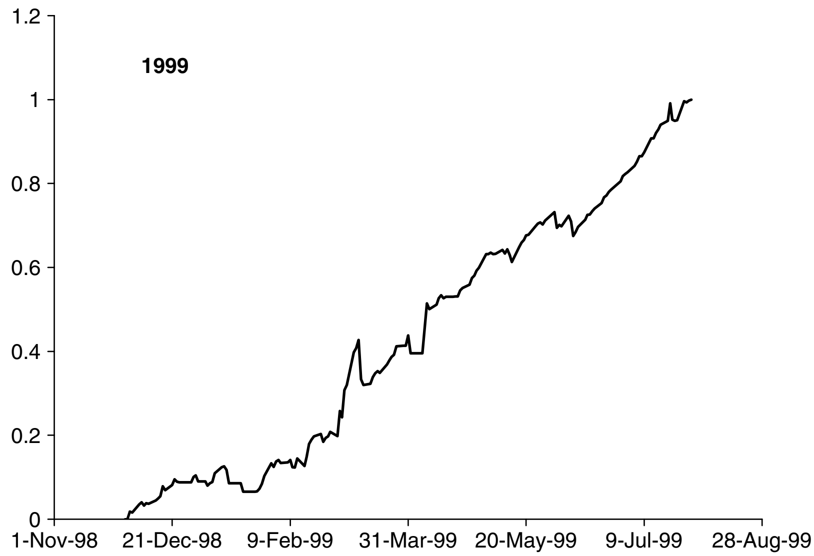


Figure 75.10 Trader friend's scaled profit for 1999.

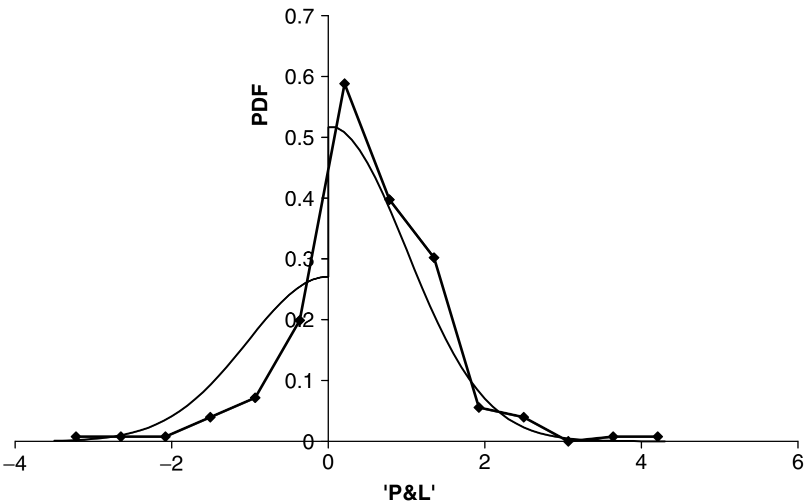


Figure 75.11 Trader friend's scaled daily P&L distribution and theoretical curve, 1998.

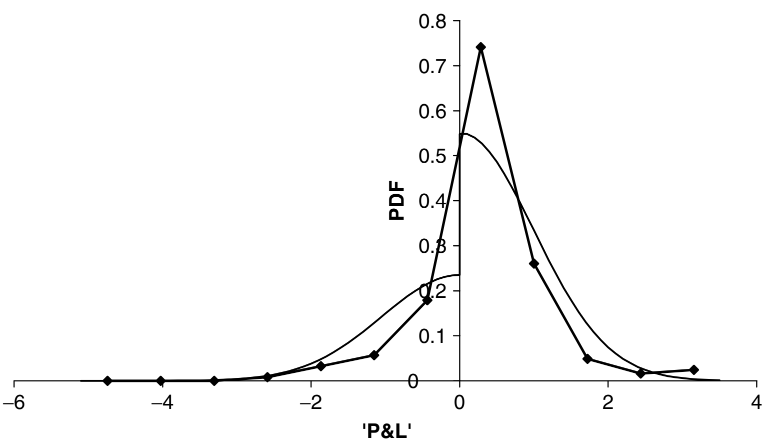


Figure 75.12 Trader friend's scaled daily P&L distribution and theoretical curve, 1999.

standard deviations. I could have simply found p by looking at how often her trade is correct. Instead I'll use this to 'confirm' the results.

In Figure 75.12 are the same plots but using 1999 data. Her skill factor in 1998 was an impressive 30%, but in 1999 an awesome 38%. Theoretically this would mean that she got the market direction correct $38 + 0.5 \times (100 - 38)\% = 69\%$ of the time. The actual number was 68%. That is talent. Her P&L for 1999 had a few very extreme data points, possibly due to non-Normal market moves. Taking these extremes out and refitting gives an even better fit and a higher p.

If any reader wants me to calculate their skill factor, please email me a (small, please) Excel spreadsheet of your P&L. If you can, also tell me how much trading capital you use and what are your markets.

75.4 PUTTING SKILL INTO THE EQUATION

Let $p = \alpha\sqrt{dt}$ be the probability that the trader gets 'told' the direction of the market's next move. If she gets told, she trades up to the position limit. Note the necessary scaling of p with the time step. We get

$$d\pi = r\pi\, dt + (1-p)q(dS - rS\, dt) + pC|dS - rS\, dt|$$

This leads to

$$E[d\pi] = (r\pi - qrS + q\mu S + \alpha C\sigma S)\, dt.$$

Here $\omega = \sqrt{2/\pi} = 0.7978846\ldots$, with the 'other' pi $= 3.1415926\ldots$. The pde for V is therefore

$$\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + (r\pi + \alpha\sigma\omega S)\frac{\partial V}{\partial \pi} - rV + \tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}$$
$$+ \max_{|q|\le 1}\left(q\left(S(\mu - r)\frac{\partial V}{\partial \pi} + \sigma^2 S^2\frac{\partial^2 V}{\partial S\partial \pi}\right) + q^2\left(\tfrac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial \pi^2} + \sigma^2 S^2\frac{\partial V}{\partial I}\right)\right) = 0$$

There is a similarity solution Φ satisfying

$$\frac{\partial \Phi}{\partial t} + (\mu - r)\Phi - (\mu - r)z_1\frac{\partial \Phi}{\partial z_1} + \alpha\sigma\omega\frac{\partial \Phi}{\partial z_1}$$
$$+ \max_{|q|\le 1}\left(q(\mu - r)\frac{\partial \Phi}{\partial z_1} + q^2\sigma^2\frac{\partial \Phi}{\partial z_2} + \tfrac{1}{2}\sigma^2(z-q)^2\frac{\partial^2 \Phi}{\partial z_1^2}\right) = 0.$$

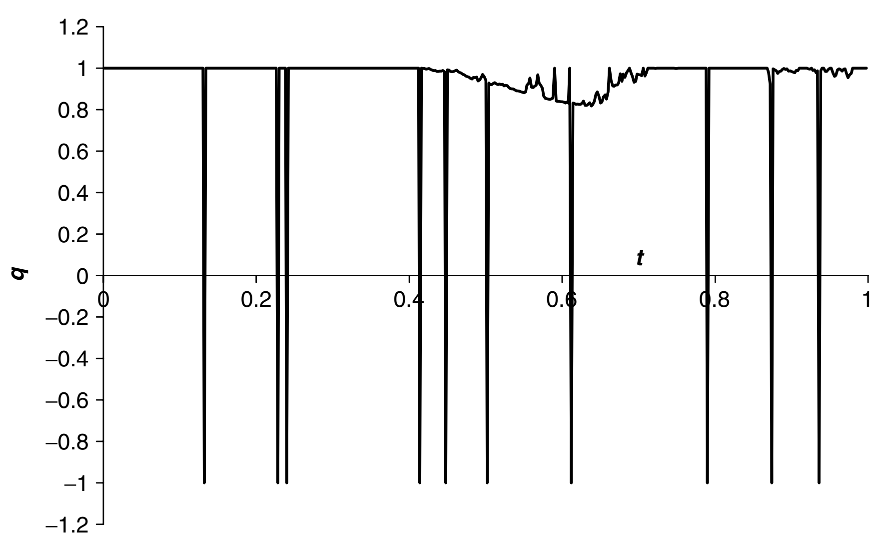


Figure 75.13 Trading strategy for a skilled trader.

75.4.1 Example

In the next examples we take exactly the same data and simulations as in the previous examples but the trader now possesses some skill.

Figure 75.13 shows the trading strategy of the skilled trader. Notice how q occasionally jumps down to -1. This corresponds to the arrival of correct information about the market's next move. The equivalent trading account is shown in Figure 75.14. Without skill the trader made 17, with skill this has gone up to 25. This is a much better performance.

The distributions of profit for the trader and bank are shown in Figures 75.15 and 75.16.

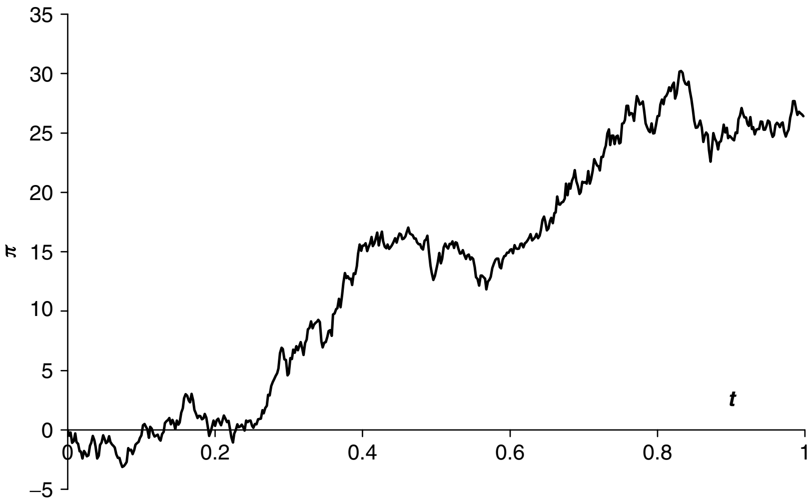


Figure 75.14 Trading account for the skilled trader.

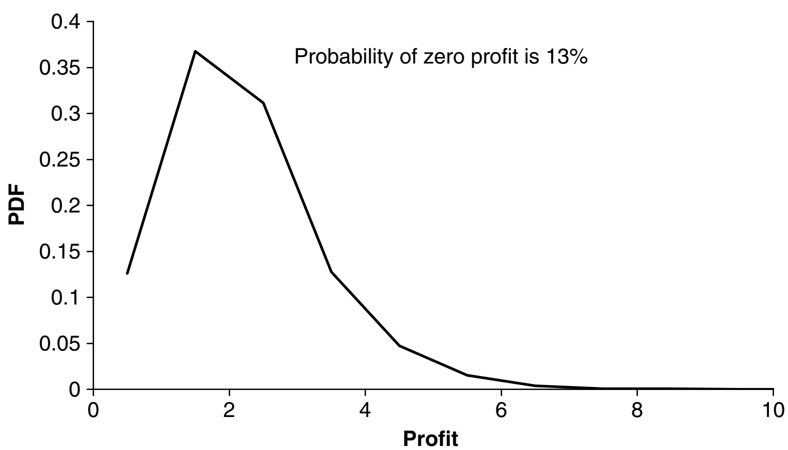


Figure 75.15 The distribution of the skilled trader's profit: mean = 1.1295 and standard deviation = 1.1238.

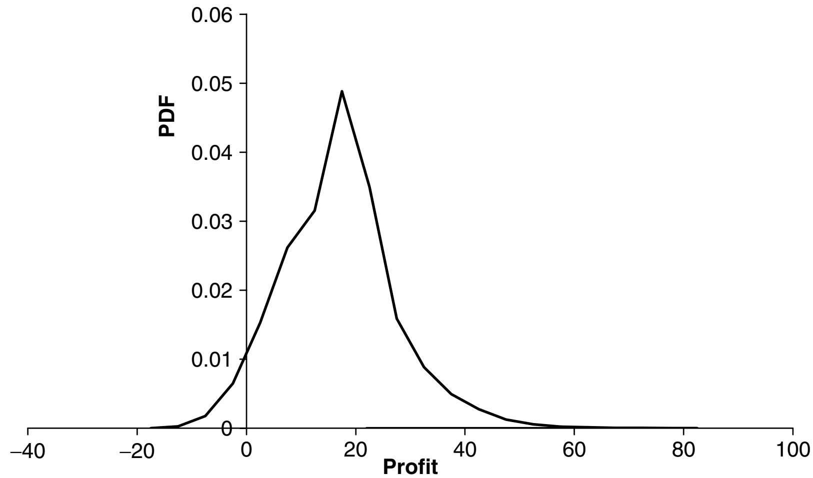


Figure 75.16 The distribution of the bank's profit: mean = 11.9403 and standard deviation = 10.3443.

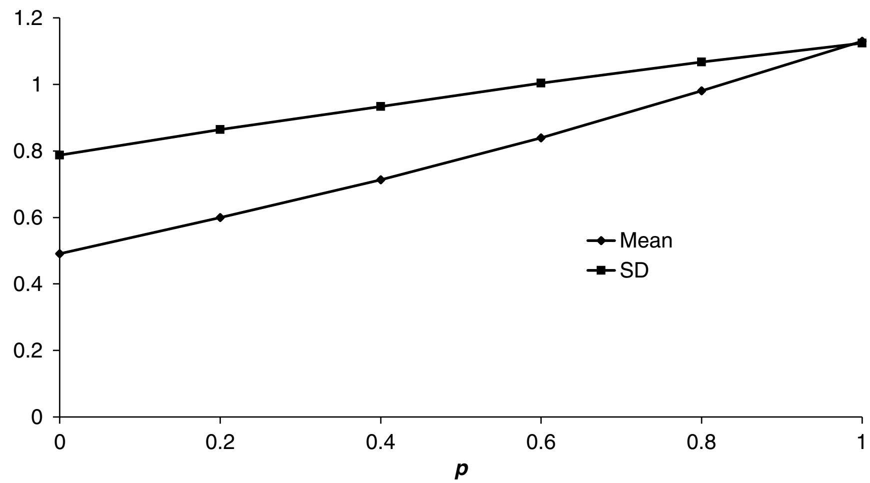


Figure 75.17 The dependence of the trader's mean and standard deviation of profit on p.

Finally in Figures 75.17 and 75.18 are shown the dependence of means and standard deviations of profit for the trader and the bank on the skill of the trader.

75.5 SUMMARY

There is clearly much scope for playing around with various bonus compensation packages to see what is best for everyone. Other issues that should be addressed are the possibility of firing the trader should the trading account get too negative. This would discourage excessive risk taking at times when the trader would otherwise have nothing to lose. The trader can also walk

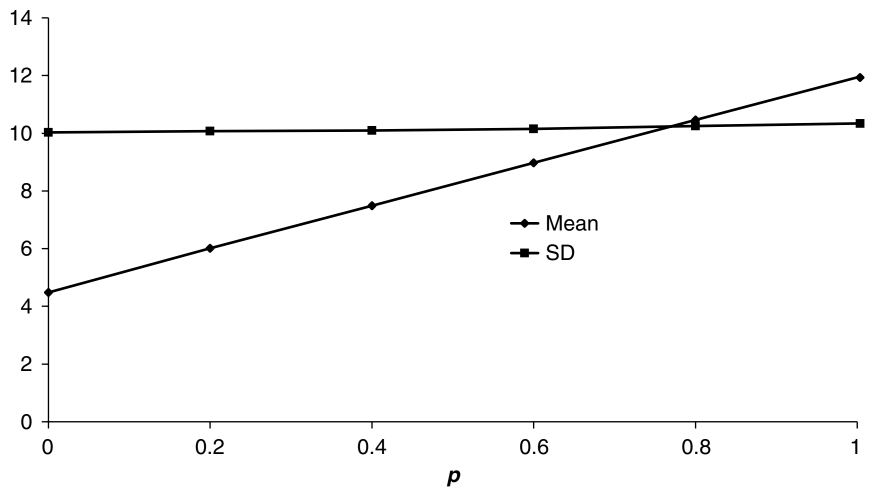


Figure 75.18 The dependence of the bank's mean and standard deviation of profit on *p*.

away after receiving her bonus. The compensation for one year could be locked up and only paid out on good performance the following year. Hedge funds are often incentivized in such a path-dependent manner.

FURTHER READING

- See Ahn, Hua, Penaud & Wilmott (1999) for more details, and further work.

PART SIX
numerical methods and programs

The final part of the book concerns the numerical methods needed for implementing the models described in the rest of the book. Many of the techniques have grown up in other branches of applied mathematics or physics and are now seeing successful use in financial problems. This part is necessarily just an overview, albeit quite detailed and self-contained, of possible techniques. The reader is encouraged to read widely on this subject.

Chapter 76: Overview of Numerical Methods A chapter to set the scene, and make suggestions for a program of self study.

Chapter 77: Finite-difference Methods for One-factor Models Most of the models we have seen in the book have led to some form of partial differential equation, usually, but not always, parabolic. The parabolic partial differential equation is a very easy differential equation to solve numerically. This chapter describes the very basics of finite-difference meshes and the approximation of derivatives. It ends with a description of the explicit method.

Chapter 78: Further Finite-difference Methods for One-factor Models More sophisticated than the explicit method are the implicit methods. These are harder to program but this effort usually pays off in faster and more accurate results. Other techniques are explored. The methods are applied to American and path-dependent options.

Chapter 79: Finite-difference Methods for Two-factor Models When we have two stochastic factors we end up with a partial differential equation with second derivatives in two variables. We can still use the explicit method, which is described, but for speed we may want to employ one of the available implicit methods.

Chapter 80: Monte Carlo Simulation For some path-dependent problems, or when we have a high-dimensional problem, we may want to use simulation methods for pricing.

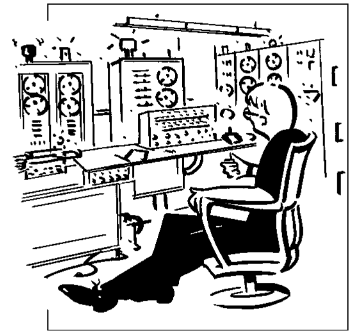
Chapter 81: Numerical Integration Related to the Monte Carlo methods for simulation of paths are the methods for numerical integration that are very useful in very special circumstances.

Chapter 82: Finite-difference Programs This chapter contains several Visual Basic programs demonstrating many of the numerical techniques that have been described.

Chapter 83: Monte Carlo Programs The book concludes with some Monte Carlo programs written in Visual Basic.

CHAPTER 76

overview of numerical methods



In this Chapter...

- finite-difference methods
- Monte Carlo simulation
- numerical integration
- programs of study

76.1 INTRODUCTION

We are in the home straight now, only numerical methods separate us from the bibliography. The aim of this chapter, before we get to the meat of the numerical methods, is to put these techniques into context, to explain what they are used for, why, and give you some initial thoughts on how they are implemented. I'll also explain how efficient these techniques are in terms of computational time, and suggest a program of self study to help you build up your experience of implementation.

76.2 FINITE-DIFFERENCE METHODS

Finite-difference methods are designed for finding numerical solutions of differential equations. Since we work with a mesh, not unlike the binomial method, we will find the contract value at all points is stock price-time space. In quantitative finance that differential equation is almost always of diffusion or parabolic type, as explained. The only real difference between the partial differential equations are the following:

- Number of dimensions
- Functional form of coefficients
- Boundary/final conditions

- Decision features
- Linear or non linear

Number of dimensions

Is the contract an option on a single underlying or many? Is there any strong path dependence in the payoff? Answers to these questions will determine the number of dimensions in the problem, as we saw in Chapter 24. At the very least we will have two dimensions: S or r, and t. Finite-difference methods cope extremely well with smaller number of dimensions, up to four, say. Above that they get rather time consuming. We'll look at finite-difference methods in up to three dimensions, meaning time plus two others, although the explicit finite-difference method can be easily extended to as many dimensions as you want.

Functional form of coefficients

The main difference between an equity option problem and a single-factor interest rate option problem is in the functional form of the drift rate and the volatility. These appear in the governing partial differential equations as coefficients. The standard model for equities is the lognormal model, but there are many more 'standard' models in fixed income. Does this matter? No, not if you are solving the equations numerically, only if you are trying to find a closed-form solution in which case the simpler the coefficients the more likely you are to find a closed-form solution. When we look at the details of the finite-difference methods we won't be assuming any particular functional forms at all.

Boundary/final conditions

In a numerical scheme the difference between a call and a put is in the final condition. You tell the finite-difference scheme how to start. And in finite-difference schemes in finance we start, strangely, at expiration and work towards the present. Boundary conditions are where we tell the scheme about things like knock-out barriers. When we write our code we'd like it to be as general and reusable as possible. That means writing it so that it doesn't have to be changed too much in going from one contract or model to another. So we might put things like final conditions in some external function, to be changed easily.

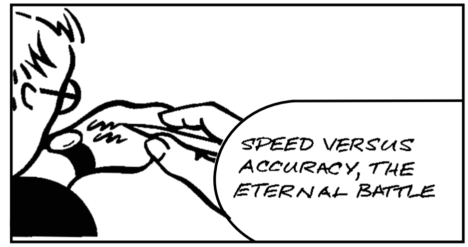
Decision features

Early exercise, instalment premiums, chooser features, are all examples of embedded decisions seen in exotic contracts. Coping with these numerically is quite straightforward using finite-difference methods, making these numerical techniques the natural ones for such contracts. The difference between a European and an American option is about three lines of code in a finite-difference program and less than a minute's work.

Linear or non linear

Almost all finance models are linear, so that you can solve for a portfolio of options by solving each contract at a time and adding. Some more modern models are non linear, and we've seen a few in this book. Linear or non linear doesn't make that much difference when you are solving by finite-difference methods. So choosing this method gives you a lot of flexibility in the type of model you can use.

76.2.1 Efficiency



Finite differences are very good at coping with low dimensions, and are the method of choice if you have a contract with embedded decisions. They are excellent for non-linear differential equations.

Here I will just quote an expression for the calculation time of a typical finite-difference scheme used to price a portfolio of options. You will see where this comes from in Chapter 79. The time taken to price an option, and calculate the sensitivities to underlying(s) and time, will be

$$O\left(M\epsilon^{-1-d/2}\right),$$

where M are the number of options in the portfolio and we want an accuracy of ϵ, and d is the number of dimensions other than time.[1]

76.2.2 Program of Study

If you are new to numerics and you really want to study them to implement the models I describe then you need a program of study. Here I make some suggestions for how you should approach the numerical methods you'll be seeing.

- **Explicit method/European calls, puts and binaries:** To get started you should learn the explicit method as applied to the Black–Scholes equation for a European option. This is very easy to program and you won't make many mistakes.
- **Explicit method/American calls, puts and binaries:** Not much harder is the application of the explicit method to American options.
- **Crank–Nicolson/European calls, puts and binaries:** Once you've got the explicit method under your belt you should learn the Crank–Nicolson implicit method. This is harder to program, but you will get a better accuracy.
- **Crank–Nicolson/American calls, puts and binaries:** There's not much more effort involved in pricing American-style options than in the pricing of European-style options.
- **Explicit method/path-dependent options:** By now you'll be quite sophisticated and it's time to price a path-dependent contract. Start with an Asian option with discrete sampling, and then try a continuously-sampled Asian. Finally, try your hand at lookbacks.
- **Interest rate products:** Repeat the above program for non-path-dependent and then path-dependent interest rate products. First price caps and floors and then go on to the index amortizing rate swap.
- **Two-factor explicit:** To get started on two-factor problems price a convertible bond using an explicit method, with both the stock and the spot interest rate being stochastic.
- **Two-factor implicit:** The final stage, using methods described in this book, is to implement the implicit two-factor method as applied to the convertible bond.

[1] So if we have a non-path-dependent option on a single underlying then $d = 1$.

76.3 MONTE CARLO METHODS

Monte Carlo methods simulate the random behavior underlying the financial models. So, in a sense they get right to the heart of the problem. Always remember, though, that when pricing you must simulate the risk-neutral random walk(s), the value of a contract is then the expected present value of all cashflows. When implementing a Monte Carlo method look out for the following:

- Number of dimensions
- Functional form of coefficients
- Boundary/final conditions
- Decision features
- Linear or non linear

again!

Number of dimensions

For each random factor you will have to simulate a time series. It will obviously take longer to do this, but the time will only be proportional to the number of factors, which isn't so bad. This makes Monte Carlo methods ideal for higher dimensions when the finite-difference methods start to crawl.

Functional form of coefficients

As with the finite-difference methods it doesn't matter so much what the drift and volatility functions are in practice, since you won't be looking for closed-form solutions.

Boundary/final conditions

These play a very similar role as in finite differences. The final condition is the payoff function and the boundary conditions are where we implement trigger levels etc.

Decision features

When you have a contract with embedded decisions the Monte Carlo method becomes cumbersome. This is easily the main drawback for simulation methods. You'll see that when we use the Monte Carlo method we are only finding the option value at today's stock price and time. But to price an American option correctly, say, we need to know what the option value *would be* at every point in stock price-time space. We don't typically find this as part of the Monte Carlo solution.

Linear or non linear

Simulation methods also cope poorly with non-linear models. Some models just don't have a useful interpretation in terms of probabilities and expectations so you wouldn't expect them to be amenable to solution by methods based on random simulations.

76.3.1 Efficiency

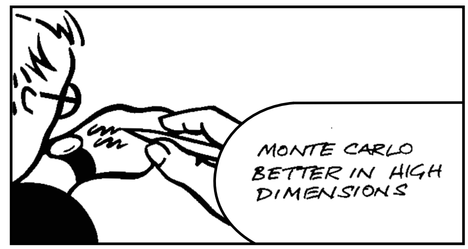
If we want an accuracy of ϵ and we have d underlyings then the calculation time is

$$O\left(d\epsilon^{-3}\right).$$

It will take longer to price the greeks, but, on the positive side, we can price many options at the same time.

76.3.2 Program of Study

Here is a program of study for the Monte Carlo path-simulation methods.



- **European calls, puts and binaries on a single equity:** Simulate a single stock path, the payoff for an option, or even a portfolio of options, calculate the expected payoff and present value to price the contract.
- **Path-dependent option on a single equity:** Price a barrier, Asian, lookback etc.
- **Options on many stocks:** Price a multi-asset contract by simulating correlated random walks. You'll see how time taken varies with number of dimensions.
- **Interest rate derivatives, spot rate model:** This is not that much harder than equities. Just remember to present value along each realized path of rates *before* taking the expectation across all paths.
- **HJM model:** Slightly more ambitious is the HJM interest rate model. Use a single factor, then two factors etc.
- **BGM model:** A discrete version of HJM.

76.4 NUMERICAL INTEGRATION

Occasionally one can write down the solution of an option-pricing problem in the form of a multiple integral. This is because you can interpret the option value as an expectation of a payoff, and an expectation of the payoff is mathematically just the integral of the product of that payoff function and a probability density function. This is only possible in special cases. The option has to be European, the underlying stochastic differential equation must be explicitly integrable (so the lognormal random walk is perfect for this) and the payoff shouldn't usually be path-dependent. So if this is possible then pricing is easy... you have a formula. The only difficulty comes in turning this formula into a number. And that's the subject of numerical integration or quadrature. Look out for the following.

- Can you write down the value of an option as an integral?

That's it in a nutshell.

76.4.1 Efficiency

There are several numerical quadrature methods we'll describe in Chapter 81. But the two most common are based on random number generation again. One uses Normally distributed numbers and the other uses what are called low-discrepancy sequences. The low-discrepancy numbers are clever in that they appear superficially to be random but don't have the inevitable clustering that truly random numbers have.

Using the simple Normal numbers, if you want an accuracy of ϵ and you are pricing M options the time taken will be

$$O\left(M\epsilon^{-2}\right).$$

If you use the low-discrepancy numbers the time taken will be

$$O\left(M\epsilon^{-1}\right).$$

You can see that this method is very fast, unfortunately it isn't often applicable.

76.4.2 Program of Study

Here is a program of study for the numerical quadrature methods.

- **European calls, puts and binaries on a single equity using Normal numbers:** Very simple. You will be evaluating a single integral.
- **European calls, puts and binaries on several underlying lognormal equities, using Normal numbers:** Very simple again. You will be evaluating a multiple integral.
- **Arbitrary European, non-path-dependent payoff, on several underlying lognormal equities, using Normal numbers:** You'll only have to change a single function.
- **Arbitrary European, non-path-dependent payoff, on several underlying lognormal equities, using low-discrepancy numbers:** Just change the source of the random numbers in the previous code.

76.5 SUMMARY

Subject	FD	MC	Quad.
Low dimensions	Good	Inefficient	Good
High dimensions	Slow	Excellent	Good
Path dependent	Depends	Excellent	Not good
Greeks	Excellent	Not good	Excellent
Portfolio	Inefficient	Very good	Very good
Decisions	Excellent	Poor	V. poor
Non linear	Excellent	Poor	V. poor

If you get to the end of this program successfully then you will have reached a very high level of sophistication.

FURTHER READING

- Finite differences and Monte Carlo aren't the only numerical methods we can use. See Topper (2005) for a description of the finite-element method, a technique commonly used in the hard sciences and now finding its way into finance.

CHAPTER 77
finite-difference methods for one-factor models

In this Chapter...

- finite-difference grids
- how to approximate derivatives of a function
- how to go from the Black–Scholes partial differential equation to a *difference* equation
- the explicit finite-difference method, a generalization of the binomial method

77.1 INTRODUCTION

Rarely can we find closed-form solutions for the values of options. Unless the problem is very simple indeed we are going to have to solve a problem numerically. In an earlier chapter I described the binomial method for pricing options. This used the idea of a finite tree structure branching out from the current asset price and the current time right up to the expiry date. One way of thinking of the binomial method is as a method for solving a partial differential equation. Finite-difference methods are no more than a generalization of this concept, although we tend to talk about **grids** and **meshes** rather than 'trees.' Once we have found an equation to solve numerically then it is much easier to use a finite-difference grid than a binomial tree, simply because the transformation from a differential equation (Black–Scholes) to a difference equation is easier when the grid/mesh/tree is nice and regular. Moreover, there are many, many ways the finite-difference method can be improved upon, making it faster and more accurate. The binomial method is not so flexible. And finally, there is a great deal in the mathematical/numerical analysis literature on these and other methods; it would be such a shame to ignore this. The main difference between the binomial method and the finite-difference methods is that the former contains the diffusion, the volatility, in the tree structure. In the finite-difference methods the 'tree' is fixed but parameters change to reflect a changing diffusion.

To those of you who are new to numerical methods let me start by saying that you will find the parabolic partial differential equation very easy to solve numerically. If you are not new to these ideas, but have been brought up on binomial methods, now is the time to wean yourself off them. On a personal note, for solving problems in practice I would say that I use finite-difference methods about 75% of the time, Monte Carlo simulations 20%, and the rest would be explicit formulae. Those explicit formulae are almost always just the regular Black–Scholes formulae for calls and puts, never for barriers for which it is highly dangerous to use constant volatility. Only once have I ever seriously used a binomial method, and that was more to help with modeling than with the numerical analysis.

In this chapter I'm going to show how to approximate derivatives using a grid and then how to write the Black–Scholes equation as a difference equation using this grid. I'll show how this can be done in many ways, discussing the relative merits of each. In the next chapter I'll also show how to extend the ideas to price contracts with early exercise and to price exotic options.

When I describe the numerical methods I often use the Black–Scholes equation as the example. *But the methods are all applicable to other problems, such as stochastic interest rates.* I am therefore assuming a certain level of intelligence from my reader, that once you have learned the methods as applied to the equity, currency, commodity worlds you can use them in the fixed-income world. I am sure you won't let me down.

77.2 OVERVIEW

This chapter is the first devoted to the finite-difference family of methods. We will kick off with an introduction to grids, followed by how to approximate derivatives (as in '$\partial \cdot /\partial \cdot$,' not as in 'options') using a discrete set of data points, that are given by values on the mesh points of our grid.

After we have looked at differentiation we'll see how to incorporate already known information about an option's value. By this I mean the option's payoff function, and how the option value behaves for large and small S. We'll use this information as our final and boundary conditions.

Last, but most importantly, we'll see how the Black–Scholes partial differential equation, which is obviously valid in continuous time and continuous S, can be approximated by a difference equation, valid in discrete time and discrete S, i.e. valid on our grid.

Then we solve for the option's value and the greeks.

77.3 GRIDS

Figure 77.1 is the binomial tree figure from Chapter 15. This is the structure commonly used for pricing simple non-path-dependent contracts. The idea was explained in Chapter 15. In the world of finite differences we use the grid or mesh shown in Figure 77.2. In the former figure the **nodes** are spaced at equal time intervals and at equal intervals in $\log S$. The finite-difference grid usually has equal **time steps**, the time between nodes, and either equal S steps or equal $\log S$ steps. If we wanted, we could make the grid any shape we wanted.

I am only going to describe finite-difference methods with constant time and asset step. There are advantages and disadvantages in this. If we are solving the Black–Scholes equation there is something appealing in having a grid with constant $\log S$ steps, after all, the underlying is following a lognormal random walk. But if you want to use constant $\log S$ steps then it is conceptually simpler to change variables, to write the Black–Scholes equation in terms of the new variable $x = \log S$. Once you have done this then constant $\log S$ step size is equivalent to constant x step size. You could even go so far as to transform to the much neater heat equation

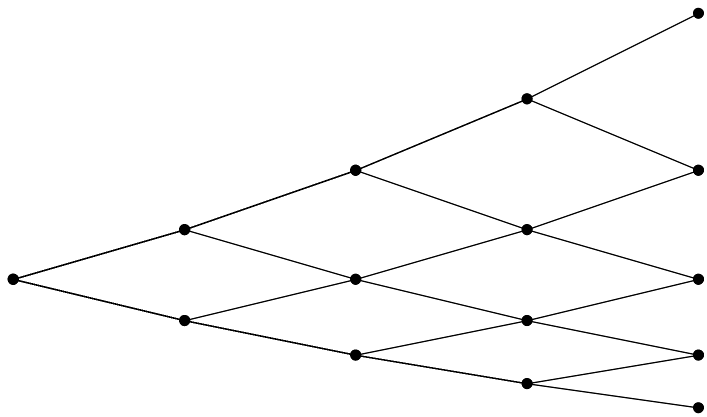


Figure 77.1 The binomial tree.

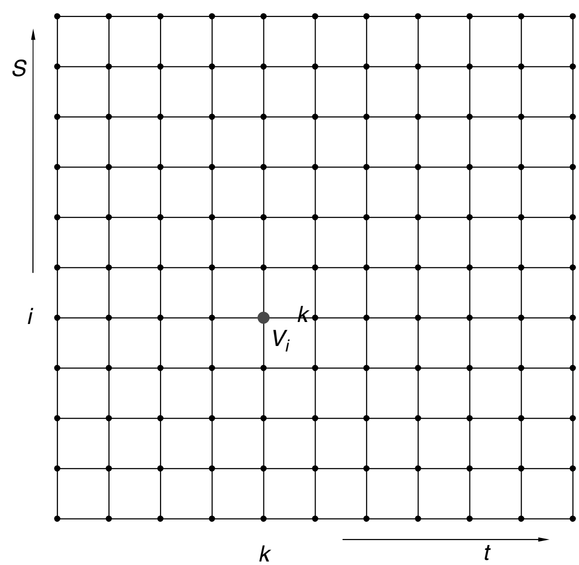


Figure 77.2 The finite-difference grid.

as in Chapter 6. One downside to such a transformation is that equal spacing in $\log S$ means that a lot of grid points are spread around small values of S where there is usually not very much happening. The main reason that I rarely do any transforming of the equation when I am solving it numerically is that I like to solve in terms of the real financial variables since terms of the contract are specified using these real variables: Transforming to the heat equation could cause problems for contracts such as barrier options. For other problems such a transformation to something nicer is not even possible. Examples would be an underlying with asset- and time-dependent volatility, or an interest rate product.

I'm also going to concentrate on backward parabolic equations. Every partial differential equation or numerical analysis book explains methods with reference to the forward equation.

But of course, the difference between forward and backward is no more than a change of the sign of the time (but make sure you apply initial conditions to forward equations and final conditions to backward).

77.4 DIFFERENTIATION USING THE GRID

Let's introduce some notation. The time step will be δt and the asset step δS, both of which are constant. Thus the grid is made up of the points at asset values

$$S = i\,\delta S$$

and times

$$t = T - k\,\delta t$$

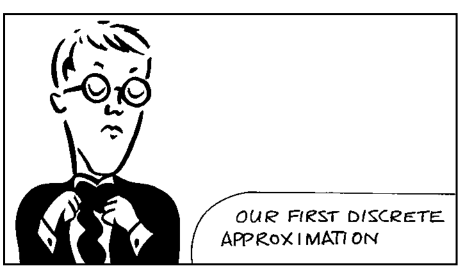
where $0 \leq i \leq I$ and $0 \leq k \leq K$. This means that we will be solving for the asset value going from zero up to the asset value $I\,\delta S$. Remembering that the Black–Scholes equation is to be solved for $0 \leq S < \infty$ then $I\,\delta S$ is our approximation to infinity. In practice, this upper bound does not have to be too large. Typically it should be three or four times the value of the exercise price, or more generally, three or four times the value of the asset at which there is some important behavior. In a sense barrier options are easier to solve numerically because you don't need to solve over all values of S; for an up-and-out option there is no need to make the grid extend beyond the barrier.

I will write the option value at each of these grid points as

$$V_i^k = V(i\,\delta S, T - k\,\delta t),$$

so that the superscript is the time variable and the subscript the asset variable. Notice how I've changed the direction of time; as k increases so real time decreases.

Suppose that we know the option value at each of the grid points, can we use this information to find the derivatives of the option value with respect to S and t? That is, can we find the terms that go into the Black–Scholes equation?



77.5 APPROXIMATING θ

The definition of the first time-derivative of V is simply

$$\frac{\partial V}{\partial t} = \lim_{h \to 0} \frac{V(S, t+h) - V(S, t)}{h}.$$

It follows naturally that we can approximate the time derivative from our grid of values using

$$\frac{\partial V}{\partial t}(S, t) \approx \frac{V_i^k - V_i^{k+1}}{\delta t}. \tag{77.1}$$

This is our approximation to the option's theta. It uses the option value at the two points marked in Figure 77.3.

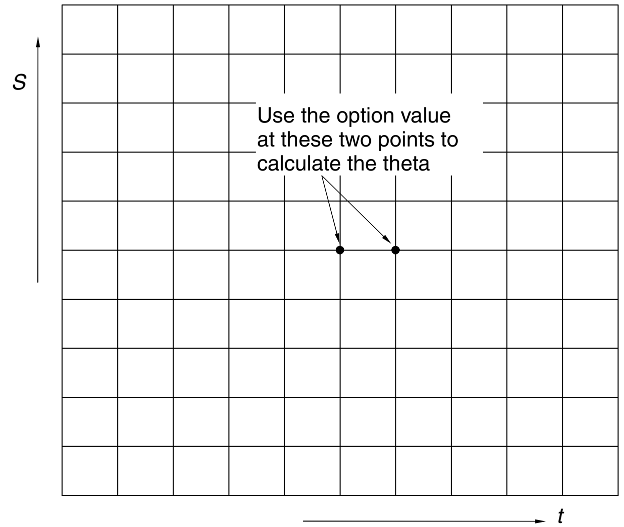


Figure 77.3 An approximation to the theta.

How accurate is this approximation? We can expand the option value at asset value S and time $t - \delta t$ in a Taylor series about the point S, t as follows.

$$V(S, t - \delta t) = V(S, t) - \delta t \frac{\partial V}{\partial t}(S, t) + O(\delta t^2).$$

In terms of values at grid points this is just

$$V_i^k = V_i^{k+1} + \delta t \frac{\partial V}{\partial t}(S, t) + O(\delta t^2).$$

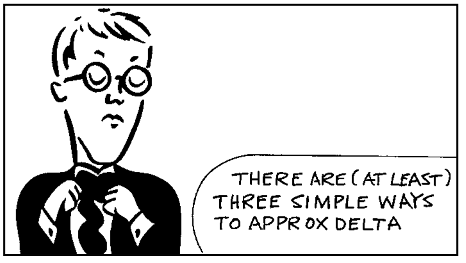
Which, upon rearranging, is

$$\frac{\partial V}{\partial t}(S, t) = \frac{V_i^k - V_i^{k+1}}{\delta t} + O(\delta t).$$

Our question is answered, the error is $O(\delta t)$. It is possible to be more precise than this; the error depends on the magnitude of the second t derivative. But I won't pursue the details here.

There are other ways of approximating the time derivative of the option value, but this one will do for now.

77.6 APPROXIMATING Δ



The same idea can be used for approximating the first S derivative, the delta. But now I *am* going to present some choices.

Let's examine a cross section of our grid at one of the time steps. In Figure 77.4 is shown this cross section. The figure shows three things: The function we are approximating (the curve), the values of the function at the grid points (the dots)

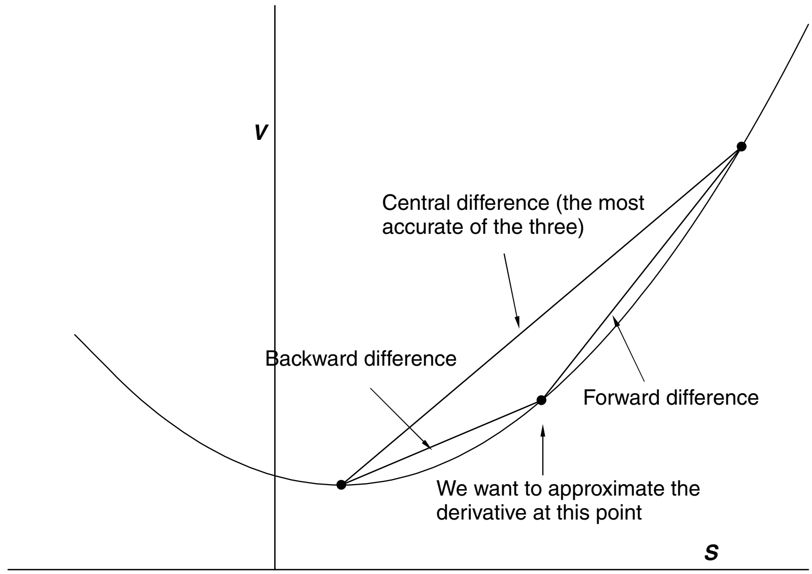


Figure 77.4 Approximations to the delta.

and three possible approximations to the first derivative (the three straight lines). These three approximations are

$$\frac{V_{i+1}^k - V_i^k}{\delta S}, \quad \frac{V_i^k - V_{i-1}^k}{\delta S} \quad \text{and} \quad \frac{V_{i+1}^k - V_{i-1}^k}{2\,\delta S}.$$

These are called a **forward difference**, a **backward difference** and a **central difference** respectively.

One of these approximations is better than the others, and it is obvious from the diagram which it is. From a Taylor series expansion of the option value about the point $S + \delta S$, t we have

$$V(S + \delta S, t) = V(S, t) + \delta S \frac{\partial V}{\partial S}(S, t) + \tfrac{1}{2}\delta S^2 \frac{\partial^2 V}{\partial S^2}(S, t) + O(\delta S^3).$$

Similarly,

$$V(S - \delta S, t) = V(S, t) - \delta S \frac{\partial V}{\partial S}(S, t) + \tfrac{1}{2}\delta S^2 \frac{\partial^2 V}{\partial S^2}(S, t) + O(\delta S^3).$$

Subtracting one from the other, dividing by $2\,\delta S$ and rearranging gives

$$\frac{\partial V}{\partial S}(S, t) = \frac{V_{i+1}^k - V_{i-1}^k}{2\,\delta S} + O(\delta S^2).$$

The central difference has an error of $O(\delta S^2)$ whereas the error in the forward and backward differences are both much larger, $O(\delta S)$. The central difference is that much more accurate because of the fortunate cancelation of terms, due to the symmetry about S in the definition of the difference.

77.6.1 One-sided Differences

The central difference calculated at S requires knowledge of the option value at $S + \delta S$ and $S - \delta S$. However, there will be occasions when we do not know one of these values, for example, if we are at the extremes of our region i.e. at $i = 0$ or $i = I$. Then there are times when it may be beneficial to use a one-sided derivative for reasons of stability, an important point which I will come back to. If we do need to use a one-sided derivative, must we use the simple forward or backward difference or is there something better?

The simple forward and backward differences use only two points to calculate the derivative; if we use three points we can get a better order of accuracy. To find the best approximations using three points we need to use Taylor series again.

Suppose I want to use the points S, $S + \delta S$ and $S + 2\,\delta S$ to calculate the option's delta, how can I do this as accurately as possible? First, expand the option value at the points $S + \delta S$ and $S + 2\,\delta S$ in a Taylor series:

$$V(S + \delta S, t) = V(S, t) + \delta S \frac{\partial V}{\partial S}(S, t) + \tfrac{1}{2}\delta S^2 \frac{\partial^2 V}{\partial S^2}(S, t) + O(\delta S^3)$$

and

$$V(S + 2\,\delta S, t) = V(S, t) + 2\,\delta S \frac{\partial V}{\partial S}(S, t) + 2\,\delta S^2 \frac{\partial^2 V}{\partial S^2}(S, t) + O(\delta S^3).$$

If I take the combination

$$-4V(S + \delta S, t) + V(S + 2\,\delta S, t)$$

I get

$$-3V(S, t) - 2\,\delta S \frac{\partial V}{\partial S}(S, t) + O(\delta S^3),$$

since the second derivative, $O(\delta S^2)$, terms both cancel. Thus

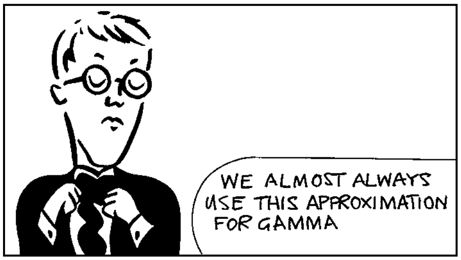
$$\frac{\partial V}{\partial S}(S, t) = \frac{-3V(S, t) + 4V(S + \delta S, t) - V(S + 2\,\delta S, t)}{2\,\delta S} + O(\delta S^2).$$

This approximation is of the same order of accuracy as the central difference, but of better accuracy than the simple forward difference. It uses no information about V for values below S.

If we want to calculate the delta using a better *backward* difference then we would choose

$$\frac{\partial V}{\partial S}(S, t) = \frac{3V(S, t) - 4V(S - \delta S, t) + V(S - 2\,\delta S, t)}{2\,\delta S} + O(\delta S^2).$$

For most of the time I will use the central difference as the approximation to the delta, but there will be times when I need to use one of the one-sided approximations.



77.7 APPROXIMATING Γ

The gamma of an option is the second derivative of the option with respect to the underlying. One way of calculating it is to estimate the delta using a forward difference, then estimate the delta using a backward difference ... and then use the difference between these two estimates (divided by the distance between nodes) as the estimate of gamma.

The forward difference is

$$\frac{V_{i+1}^k - V_i^k}{\delta S},$$

which you can think of as an accurate central difference at the point $S + \frac{1}{2}\delta S$. The backward difference is

$$\frac{V_i^k - V_{i-1}^k}{\delta S},$$

which is also just another accurate central difference but now at the point $S - \frac{1}{2}\delta S$.

Take the difference between these two deltas, divide by δS, the distance between the two mid-points, to get the natural approximation to gamma being

$$\frac{\partial^2 V}{\partial S^2}(S,t) \approx \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}.$$

Again, we can check this by looking at Taylor series expansions. The error in this approximation is also $O(\delta S^2)$. I'll leave the demonstration of this as an exercise for the reader.

77.8 EXAMPLE

In Figure 77.5 are shown some option values on the grid. The time step is 0.1 and the asset step is 2.

From these numbers we can estimate the theta as

$$\frac{12 - 13}{0.1} = -10.$$

The delta is approximately

$$\frac{15 - 10}{2 \times 2} = 1.25.$$

And the gamma is approximately

$$\frac{15 - 2 \times 12 + 10}{2 \times 2} = 0.25.$$

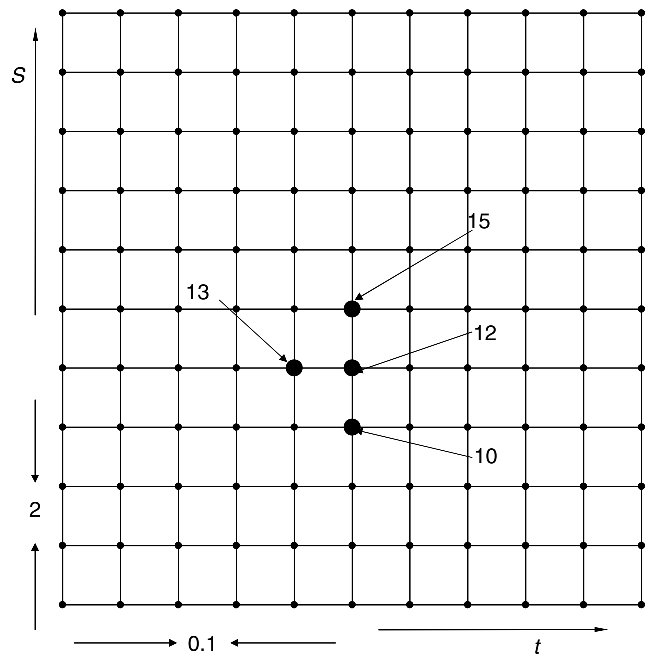


Figure 77.5 Calculation of the greeks.

77.9 FINAL CONDITIONS AND PAYOFFS

We know that at expiry the option value is just the payoff function. This means that we don't have to solve anything for time T. At expiry we have

$$V(S, T) = \text{Payoff}(S)$$

or, in our finite-difference notation,

$$V_i^0 = \text{Payoff}(i\,\delta S).$$

The right-hand side is a known function. For example, if we are pricing a call option we put

$$V_i^0 = \max(i\,\delta S - E, 0).$$

This final condition will get our finite-difference scheme started. It will be just like working down the tree in the binomial method.

77.10 BOUNDARY CONDITIONS

When we come to solving the Black–Scholes equation numerically in the next section, we will see that we must specify the option value at the extremes of the region. That is, we must prescribe the option value at $S = 0$ and at $S = I\,\delta S$. What we specify will depend on the type of option we are solving. I will give some examples.

Example 1

Suppose we want to price a call option. At $S = 0$ we know that the value is always zero, therefore we have

$$V_0^k = 0.$$

Example 2

For large S the call value asymptotes to $S - Ee^{-r(T-t)}$ (plus exponentially small terms). Thus our upper boundary condition could be

$$V_I^k = I\,\delta S - Ee^{-rk\,\delta t}.$$

This would be slightly different if we had a dividend.

Example 3

For a put option we have the condition at $S = 0$ that $V = Ee^{-r(T-t)}$. This becomes

$$V_0^k = Ee^{-rk\,\delta t}.$$

Example 4

The put option becomes worthless for large S and so

$$V_I^k = 0.$$

Example 5

The most useful boundary condition to apply at $S = 0$ for most contracts (including calls and puts) is that the diffusion and drift terms 'switch off.' This means that on $S = 0$ the payoff is guaranteed, resulting in the condition

$$\frac{\partial V}{\partial t}(0, t) - rV(0, t) = 0.$$

Numerically, this becomes

$$\boxed{V_0^k = (1 - r\,\delta t)V_0^{k-1}.}$$

Example 6

When the option has a payoff that is at most linear in the underlying for large values of S then you can use the upper boundary condition

$$\frac{\partial^2 V}{\partial S^2}(S, t) \to 0 \quad \text{as } S \to \infty.$$

Almost all common contracts have this property. The finite-difference representation is

$$\boxed{V_I^k = 2V_{I-1}^k - V_{I-2}^k.}$$

This is particularly useful because it is independent of the contract being valued meaning that your finite-difference program does not have to be too intelligent.[1]

77.10.1 Other Boundary Conditions

Often there are natural boundaries at finite, non-zero values of the underlying, which means that the domain in which we are solving either does not extend down to zero or up to infinity. Barrier options are the most common form of such contracts.

By way of example, suppose that we want to price an up-and-out call option. This option will be worthless if the underlying ever reaches the value S_u. Clearly,

$$V(S_u, t) = 0.$$

If we are solving this problem numerically how do we incorporate this boundary condition?

The ideal thing to do first of all is to choose an asset step size such that the barrier $S = S_u$ is a grid point i.e $S_u/\delta S$ should be an integer. This is to ensure that the boundary condition

$$V_I^k = 0$$

is an accurate representation of the correct boundary condition. Note that we are no longer solving over an asset price range that extends to large S. The upper boundary at $S = S_u$ may be close to the current asset level. In a sense this makes barrier problems easier to solve; the solution region is always smaller than the region over which you would solve a non-barrier problem.

Sometimes it is not possible to make your grid match up with the barrier. This would be the case if the barrier were moving, for example. If this is the case then you are going to have to find an approximation to the boundary condition. There is something that you must not do, and that is to set V equal to zero at the nearest grid point to the barrier. Such an approximation is very inaccurate, of $O(\delta S)$, and will ruin your numerical solution. The trick that we can use to overcome such problems is to introduce a **fictitious point**. This is illustrated in Figure 77.6.

The point $i = I - 1$ is a real point, within the solution region. The point $i = I$ is just beyond the barrier.

Example 7

Suppose that we have the condition that

$$V(S_u, t) = f(t).$$

If we have an 'out' option then f would be either zero or the value of the rebate. If we have an 'in' option then f is the value of the option into which the barrier option converts.

[1] Sometimes I even use this condition for small values of S, not taking the grid down to $S = 0$.

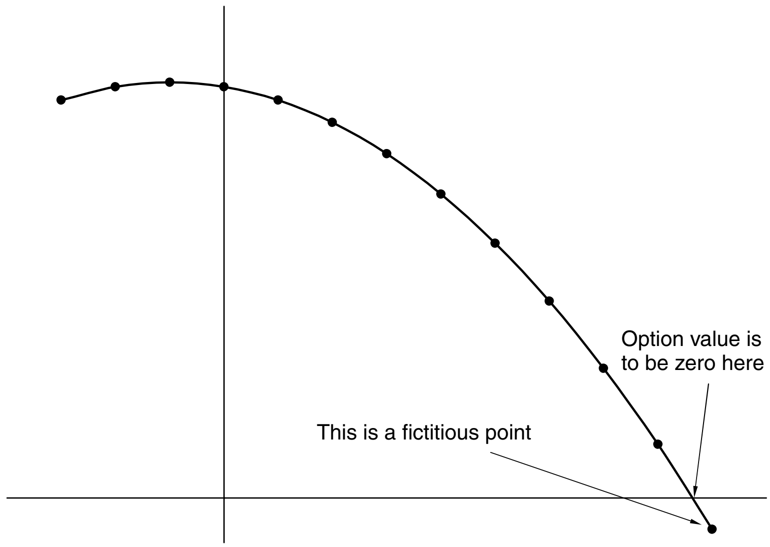


Figure 77.6 A fictitious point, introduced to ensure accuracy in a barrier option boundary condition.

This condition can be approximated by ensuring that the straight line connecting the option values at the two grid points straddling the barrier has the value f at the barrier. Then a good discrete version of this boundary condition is

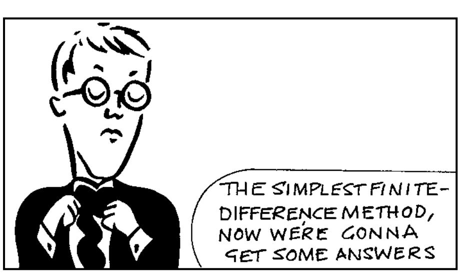
$$V_I^k = \frac{1}{\alpha}\left(f - (1-\alpha)V_{I-1}^k\right)$$

where

$$\alpha = \frac{S_u - (I-1)\delta S}{\delta S}.$$

This is accurate to $O(\delta S^2)$, the same order of accuracy as in the approximation of the S derivatives.

I have set up all the foundations for us to begin solving some equations. Remember, there has been nothing difficult in what we have done so far, everything is a simple application of Taylor series.



77.11 THE EXPLICIT FINITE-DIFFERENCE METHOD

The Black–Scholes equation is

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

I'm going to write this in a more general form as

$$\frac{\partial V}{\partial t} + a(S,t)\frac{\partial^2 V}{\partial S^2} + b(S,t)\frac{\partial V}{\partial S} + c(S,t)V = 0$$

to emphasize the wide applicability of the finite-difference methods. We won't only be using it to solve for options on lognormal stocks. The only constraint we must pose on the coefficients is that if we are solving a backward equation, i.e. imposing final conditions, we must have $a > 0$.

I'm going to take the approximations to the derivatives, explained above, and put them into this equation:

$$\begin{aligned} &\frac{V_i^k - V_i^{k+1}}{\delta t} \\ &+ a_i^k \left(\frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2} \right) \\ &+ b_i^k \left(\frac{V_{i+1}^k - V_{i-1}^k}{2\,\delta S} \right) \\ &+ c_i^k V_i^k = O(\delta t, \delta S^2). \end{aligned}$$

I have used a different line for each of the terms in the original equation.

Points to note:

- The time derivative uses the option values at 'times' k and $k+1$, whereas the other terms all use values at k.
- The gamma term is a central difference, in practice one never uses anything else.
- The delta term uses a central difference. There are often times when a one-sided derivative is better. We'll see examples later.
- The asset- and time-dependent functions a, b and c have been valued at $S_i = i\,\delta S$ and $t = T - k\,\delta t$ with the obvious notation.
- The error in the equation is $O(\delta t, \delta S^2)$.

I am going to rearrange this **difference equation** to put all of the $k+1$ terms on the left-hand side:

$$V_i^{k+1} = A_i^k V_{i-1}^k + (1 + B_i^k) V_i^k + C_i^k V_{i+1}^k \tag{77.2}$$

where

$$\begin{aligned} A_i^k &= \nu_1 a_i^k - \tfrac{1}{2}\nu_2 b_i^k, \\ B_i^k &= -2\nu_1 a_i^k + \delta t\, c_i^k \end{aligned}$$

and

$$C_i^k = \nu_1 a_i^k + \tfrac{1}{2}\nu_2 b_i^k$$

where

$$\nu_1 = \frac{\delta t}{\delta S^2} \quad \text{and} \quad \nu_2 = \frac{\delta t}{\delta S}.$$

The error in this is $O(\delta t^2, \delta t\, \delta S^2)$; I will come back to this in a moment. The error in the approximation of the differential equation is called the **local truncation error**.

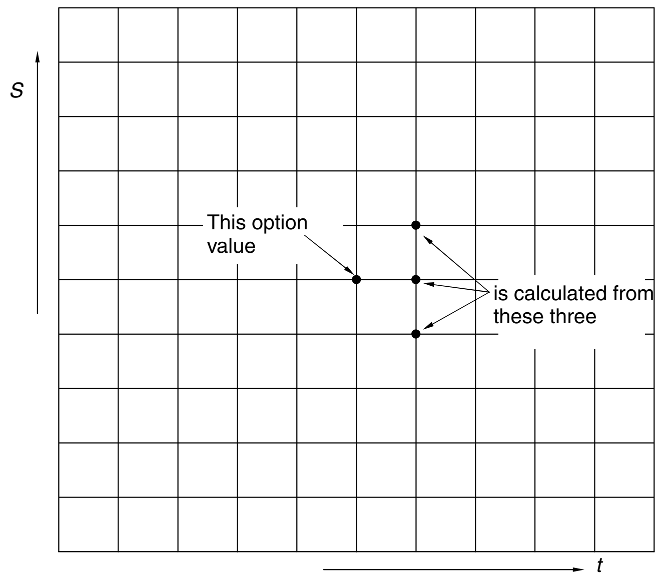


Figure 77.7 The relationship between option values in the explicit method.

Equation (77.2) only holds for $i = 1, \ldots, I-1$, i.e. for interior points, since V_{-1}^k and V_{I+1}^k are not defined. Thus there are $I-1$ equations for the $I+1$ unknowns, the V_i^k. The remaining two equations come from the two boundary conditions on $i=0$ and $i=I$. The two end points are treated separately.

If we know V_i^k for all i then Equation (77.2) tells us V_i^{k+1}. Since we know V_i^0, the payoff function, we can easily calculate V_i^1, which is the option value one time step before expiry. Using these values we can work step by step back down the grid as far as we want. Because the relationship between the option values at time step $k+1$ is a simple function of the option values at time step k this method is called the **explicit finite-difference method**. The relationship between the option values in Equation (77.2) is shown in Figure 77.7.

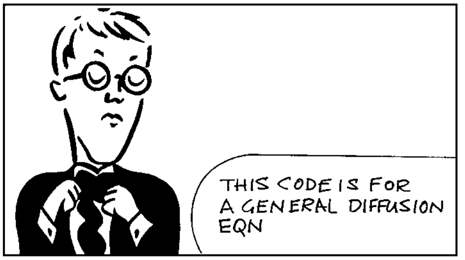
Equation (77.2) is only used to calculate the option value for $1 \leq i \leq I-1$ since the equation requires knowledge of the option values at $i-1$ and $i+1$. This is where the boundary conditions come in. Typically we either have a V_i^k being prescribed at $i=0$ and $i=I$ or, as suggested above, we might prescribe a relationship between the option value at an end point and interior values. This idea is illustrated in the following Visual Basic code fragment. This code fragment does not have all the declarations etc. at the top, nor the return of any answers. I will give a full function shortly, but for the moment I want you to concentrate on setting up the final condition and the finite-difference time loop.

The array `V(i, k)` holds the option values. Unless we wanted to store all option values for all time steps this would not be the most efficient way of writing the program. I will describe a better way in a moment.

First set up the final condition, the payoff.

```
For i = 0 To NoAssetSteps
    S(i) = i * AssetStep
    V(i, 0) = CallPayoff(S(i)) ' Set up final condition
Next i
```

Now we can work backwards in time using the following time loop.




```
' Time loop
For k = 1 To NoTimesteps
    RealTime = Expiry - k * Timestep
    For i = 1 To NoAssetSteps - 1
        V(i, k + 1) = A(S(i), RealTime) * V(i - 1, k) _
                      + B(S(i), RealTime) * V(i, k) _
                      + C(S(i), RealTime) * V(i + 1, k)
    Next i
' BC at S=0
    V(0, k + 1) = 0
' BC at "infinity"
    V(NoAssetSteps, k + 1) = 2 * V(NoAssetSteps - 1, k + 1) _
                      - V(NoAssetSteps - 2, k + 1)
Next k
```

The explicit finite-difference algorithm is simple, the functions `A(S(i), RealTime)`, `B(S(i), RealTime)` and `C(S(i), RealTime)` would be defined elsewhere and are in terms of the asset price `S(i)` and the time `RealTime`. Since I am valuing a call option here the boundary condition at $S = 0$ is simply `V(0, k + 1) = 0` but the boundary condition I have used at the upper boundary `i = NoAssetSteps` is that the gamma is zero.

77.11.1 The Black–Scholes Equation

For the Black–Scholes equation with dividends the coefficients above simplify to

$$A_i^k = \frac{1}{2}\left(\sigma^2 i^2 - (r - D)i\right)\delta t,$$

$$B_i^k = -\left(\sigma^2 i^2 + r\right)\delta t$$

and

$$C_i^k = \frac{1}{2}\left(\sigma^2 i^2 + (r - D)i\right)\delta t.$$

This uses $S = i\ \delta S$. If the volatility, interest rate and dividend are constant then there is no time or k-dependence in these coefficients.

77.12 CONVERGENCE OF THE EXPLICIT METHOD

I can write the value of the option at any i point at the final time step K as

$$V_i^K = V_i^0 + \sum_{k=0}^{K-1} (V_i^{k+1} - V_i^k).$$

Each of the terms in this summation is in error by $O(\delta t^2, \delta t\ \delta S^2)$. This means that the error in the final option value is

$$O(K\delta t^2, K\delta t\ \delta S^2)$$

since there are K terms in the summation. If we value the option at a finite value of T then $K = O(\delta t^{-1})$ so that the error in the final option value is $O(\delta t, \delta S^2)$.

Although the explicit method is simple to implement it does not always converge. Convergence of the method depends on the size of the time step, the size of the asset step and the size of the coefficients a, b and c.

The technique often used to demonstrate convergence is quite fun, so I'll show you how it is done. The method, as I describe it, is not rigorous but it can be made so.

Ask the question, 'If a small error is introduced into the solution, is it magnified by the numerical method or does it decay away?' If a small error, introduced by rounding errors for example, becomes a large error then the method is useless. The usual way to analyze such stability is to look for a solution of Equation (77.2) of the form[2]

$$V_i^k = \alpha^k e^{2\pi i\sqrt{-1}/\lambda}. \tag{77.3}$$

In other words, I'm going to look for an oscillatory solution with a wavelength of λ. If I find that $|\alpha| > 1$ then there is instability. Note that I am not worried about how the oscillation gets started, I could interpret this special solution as part of a Fourier series analysis.

Substituting (77.3) into (77.2) I get

$$\alpha = \left(1 + c_i^k\,\delta t + 2a_i^k \nu_1(\cos(2\pi/\lambda) - 1)\right) + \sqrt{-1}\,b_i^k \nu_2 \sin(2\pi/\lambda).$$

It turns out that to have $|\alpha| < 1$, for stability, we require

$$c_i^k \le 0,$$

$$2\nu_1 a_i^k - \delta t\, c_i^k \le 1$$

and

$$\tfrac{1}{2}\nu_2 |b_i^k| \le \nu_1 a_i^k.$$

To get this result, I have assumed that all the coefficients are slowly varying over the δS lengthscales.

In financial problems we almost always have a negative c, often it is simply $-r$ where r is the risk-free interest rate. The other two constraints are what really limit the applicability of the explicit method.

Typically we choose ν_1 to be $O(1)$ in which case the second constraint is approximately

$$\nu_1 \le \frac{1}{2a}$$

(ignoring sub- and superscripts on a). This is a serious limitation on the size of the time step,

$$\delta t \le \frac{\delta S^2}{2a}.$$

If we want to improve accuracy by halving the asset step, say, we must reduce the time step by a factor of four. The computation time then goes up by a factor of *eight*. The improvement in accuracy we would get from such a reduction in step sizes is a factor of four since the explicit finite-difference method is accurate to $O(\delta t, \delta S^2)$.

[2] This is the only place in the book that I use $\sqrt{-1}$. Normally I'd write this as i but I need i for other quantities.

In the Black–Scholes equation this time step constraint becomes

$$\delta t \leq \frac{\delta S^2}{2a} = \frac{\delta S^2}{\sigma^2 S^2} = \frac{1}{\sigma^2}\left(\frac{\delta S}{S}\right)^2.$$

This constraint depends on the asset price. Since δt will be independent of S the constraint is most restrictive for the largest S in the grid. If there are I equally spaced asset grid points then the constraint is simply

$$\delta t \leq \frac{1}{\sigma^2 I^2}.$$

If the time step constraint is not satisfied, if it is too large, then the instability is obvious from the results. The instability is so serious, and so oscillatory, that it is easily noticed. It is unlikely that you will get a false but believable result if you use the explicit method.

The final constraint can also be a serious restriction. It can be written as

$$\delta S \leq \frac{2a}{|b|}. \tag{77.4}$$

If we are solving the Black–Scholes equation this restriction does not make much difference in practice unless the volatility is very small. It can be important in other problems though and I will show how to get past this restriction later.

77.13 THE CODE # 1: EUROPEAN OPTION

The following VB code will output the value of a call or put option (depending on whether `PType` is `C` or `P`). The time step is hard coded and depends on the asset step. It has been chosen to be just about as big as possible while still satisfying the constraint on ν_1, so that the scheme shouldn't be unstable.

A whole array of option value versus stock price and time is output in this example.

```
Function Option_Value_3D(Vol, Int_Rate, PType, Strike, Expiration, NAS)
' NAS is number of asset steps

ReDim S(0 To NAS) As Double ' Asset array

dS = 2 * Strike / NAS ' 'Infinity' is twice the strike
dt = 0.9 / Vol ^ 2 / NAS ^ 2 ' For stability
NTS = Int(Expiration / dt) + 1 ' Number of time steps
dt = Expiration / NTS ' To ensure that expiration is an integer number of time _
                      steps away

ReDim V(0 To NAS, 0 To NTS) As Double ' Option value array

q = 1
If PType = "P" Then q = -1 ' Test for call or put

For i = 0 To NAS
S(i) = i * dS ' Set up S array
V(i, 0) = Application.Max(q * (S(i) - Strike), 0) ' Set up payoff
Next i
```

```
For k = 1 To NTS ' Time loop
For i = 1 To NAS - 1 ' Asset loop. End points treated separately
Delta = (V(i + 1, k - 1) - V(i - 1, k - 1)) / 2 / dS ' Central difference
Gamma = (V(i + 1, k - 1) - 2 * V(i, k - 1) + V(i - 1, k - 1)) / dS / dS _
                           ' Central difference
Theta = -0.5 * Vol ^ 2 * S(i) ^ 2 * Gamma - _
                           Int_Rate * S(i) * Delta + Int_Rate * V(i, k - 1) _
                           ' Black-Scholes
V(i, k) = V(i, k - 1) - dt * Theta
Next i

V(0, k) = V(0, k - 1) * (1 - Int_Rate * dt) ' Boundary condition at S=0
V(NAS, k) = 2 * V(NAS - 1, k) - V(NAS - 2, k) ' Boundary condition at S=infinity

Next k

Option_Value_3D = V ' Output array

End Function
```

Results from this program for $\sigma = 0.2$, $r = 0.05$, $E = 100$, $T = 1$ and $\mathtt{NAS} = 20$ are shown in Table 77.1 and Figure 77.8 for a call option.

Table 77.2 and Figure 77.9 show the results for a put, same parameter values.

Table 77.1 Call option values output by the explicit finite-difference code. Stock price ranges from 60 to 100, time from 0 (expiration) to 1.

	0.000	0.111	0.222	0.333	0.444	0.556	0.667	0.778	0.889	1.000
60	0.000	0.000	0.000	0.000	0.001	0.004	0.011	0.023	0.041	0.066
70	0.000	0.000	0.001	0.008	0.028	0.067	0.126	0.207	0.308	0.430
80	0.000	0.000	0.037	0.141	0.310	0.534	0.799	1.096	1.416	1.754
90	0.000	0.128	0.592	1.182	1.812	2.450	3.080	3.700	4.306	4.899
100	0.000	2.253	3.819	5.054	6.109	7.054	7.925	8.741	9.515	10.255
110	10.000	10.671	11.587	12.535	13.455	14.337	15.180	15.990	16.770	17.523
120	20.000	20.555	21.159	21.826	22.529	23.247	23.967	24.683	25.391	26.089
130	30.000	30.555	31.109	31.680	32.272	32.882	33.504	34.134	34.768	35.402
140	40.000	40.555	41.106	41.658	42.213	42.777	43.348	43.926	44.509	45.095

Table 77.2 Put option values output by the explicit finite-difference code. Stock price ranges from 60 to 100, time from 0 (expiration) to 1.

	0.000	0.111	0.222	0.333	0.444	0.556	0.667	0.778	0.889	1.000
60	40.000	39.445	38.894	38.345	37.801	37.261	36.728	36.204	35.688	35.182
70	30.000	29.445	28.894	28.353	27.827	27.323	26.843	26.387	25.955	25.546
80	20.000	19.445	18.930	18.486	18.110	17.790	17.516	17.276	17.063	16.871
90	10.000	9.573	9.486	9.527	9.612	9.706	9.798	9.880	9.953	10.015
100	0.000	1.699	2.713	3.399	3.909	4.311	4.642	4.921	5.162	5.372
110	0.000	0.116	0.481	0.880	1.254	1.593	1.898	2.171	2.417	2.639
120	0.000	0.000	0.052	0.171	0.328	0.503	0.684	0.864	1.038	1.205
130	0.000	0.000	0.003	0.025	0.071	0.138	0.221	0.315	0.415	0.518
140	0.000	0.000	0.000	0.002	0.013	0.033	0.065	0.106	0.156	0.211

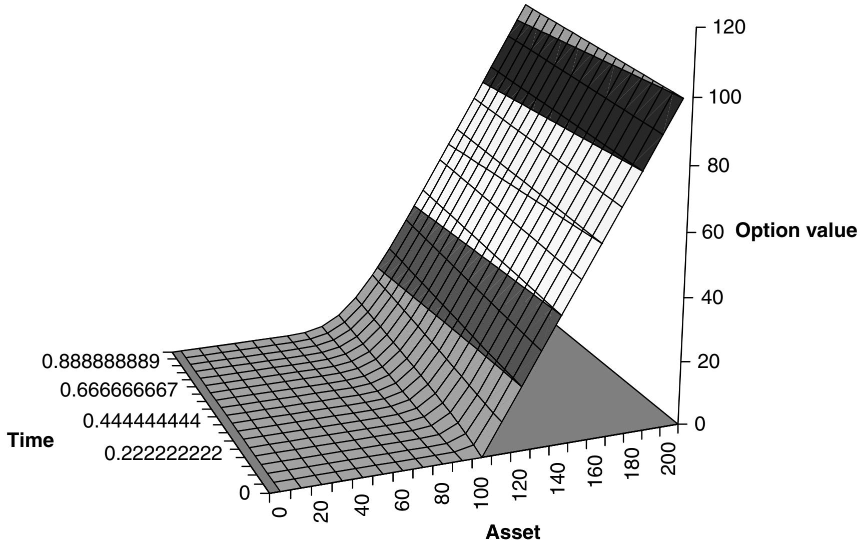


Figure 77.8 Call option values output by the explicit finite-difference code.

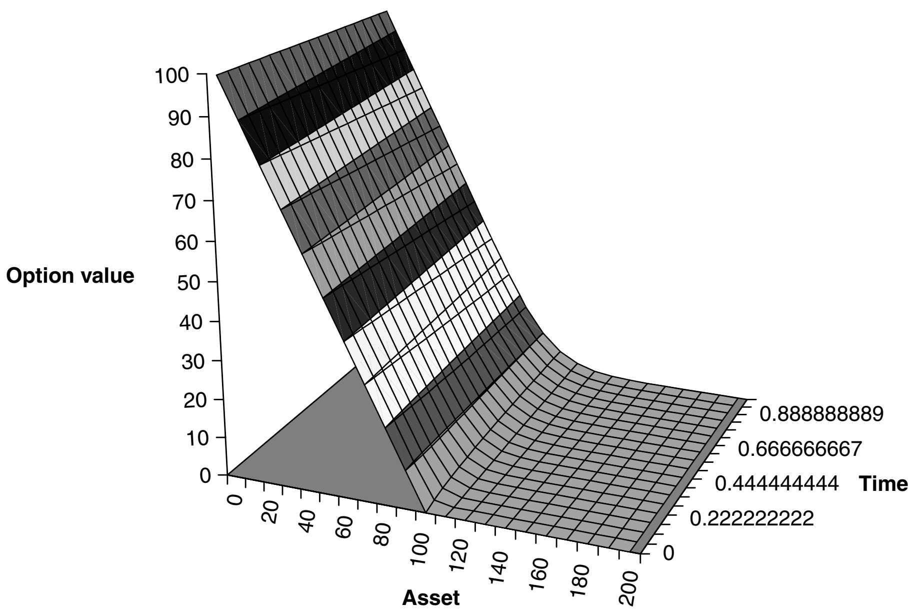


Figure 77.9 Put option values output by the explicit finite-difference code.

The error between the results of this explicit finite-difference program and the exact Black–Scholes formula as a function of the underlying is shown in Figure 77.10 for 50 asset points, a volatility of 20%, and interest rate of 10%, an expiry of one year and a strike of 100. The option is a call.

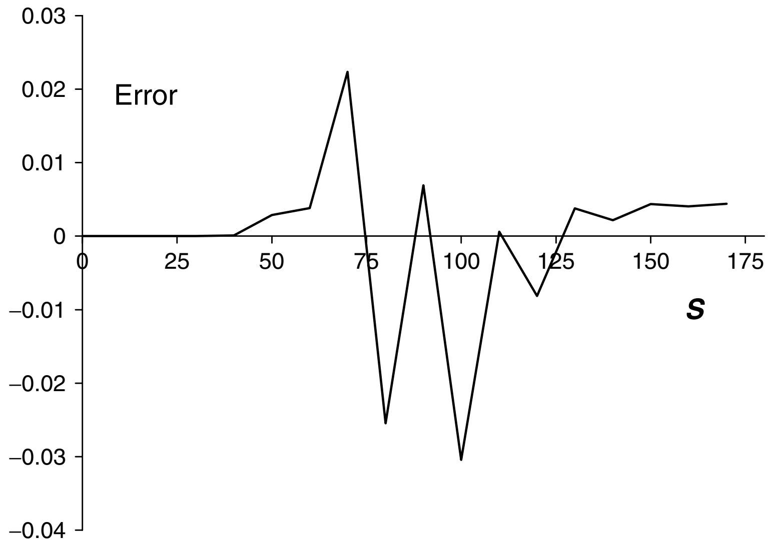


Figure 77.10 Error as a function of the underlying using the finite-difference scheme.

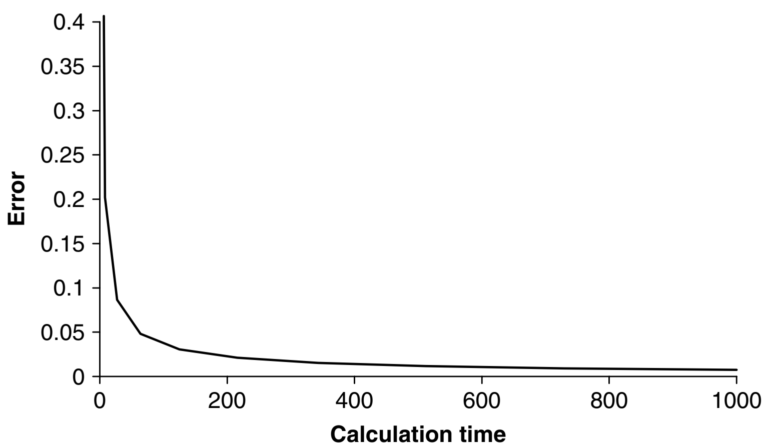


Figure 77.11 Log(Error) as a function of log(δS^2) using the finite-difference scheme.

The logarithm of the absolute error as a function of the logarithm of the square of the asset step size is shown in Figure 77.11. The $O(\delta S^2)$ behavior is very obvious. In these calculations I have kept ν_1 constant.

In Figure 77.12 is shown the error as a function of calculation time; time units will depend on your machine.

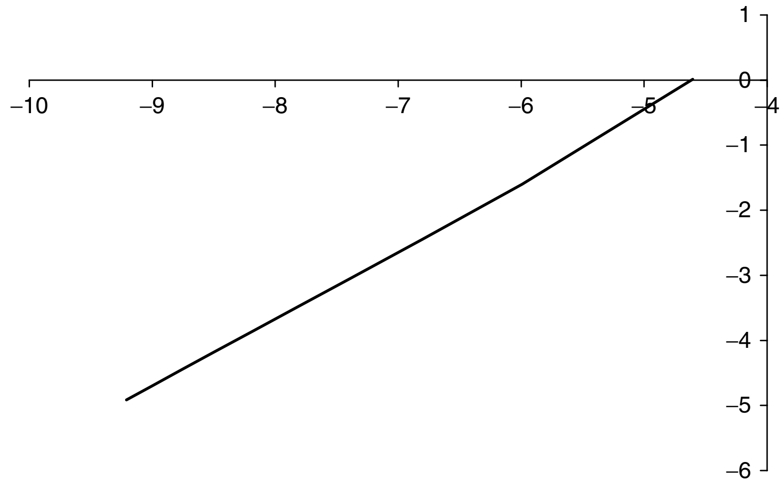


Figure 77.12 Absolute error as a function of the calculation time using the finite-difference scheme.

77.14 THE CODE # 2: AMERICAN EXERCISE

The code changes trivially when we have early exercise.

```
Function Option_Value_3D_US(Vol, Int_Rate, PType, Strike, Expiration, _
                           EType, NAS)
' NAS is number of asset steps

ReDim S(0 To NAS) As Double ' Asset array
ReDim Payoff(0 To NAS) As Double ' Payoff array

dS = 2 * Strike / NAS ' 'Infinity' is twice the strike
dt = 0.9 / Vol ^ 2 / NAS ^ 2 ' For stability
NTS = Int(Expiration / dt) + 1 ' Number of time steps
dt = Expiration / NTS ' To ensure that expiration is an integer number of _
                           time steps away

ReDim V(0 To NAS, 0 To NTS) As Double ' Option value array

q = 1
If PType = "P" Then q = -1 ' Test for call or put

For i = 0 To NAS
S(i) = i * dS ' Set up S array
V(i, 0) = Application.Max(q * (S(i) - Strike), 0) ' Set up payoff
Payoff(i) = V(i, 0) ' Store payoff
Next i

For k = 1 To NTS ' Time loop
For i = 1 To NAS - 1 ' Asset loop. End points treated separately
```

```
Delta = (V(i + 1, k - 1) - V(i - 1, k - 1)) / 2 / dS ' Central difference
Gamma = (V(i + 1, k - 1) - 2 * V(i, k - 1) + V(i - 1, k - 1)) / dS / dS _
                           ' Central difference
Theta = -0.5 * Vol ^ 2 * S(i) ^ 2 * Gamma - _
                           Int_Rate * S(i) * Delta + Int_Rate * V(i, k - 1) _
                           ' Black-Scholes
V(i, k) = V(i, k - 1) - dt * Theta
Next i

V(0, k) = V(0, k - 1) * (1 - Int_Rate * dt) ' Boundary condition at S=0
V(NAS, k) = 2 * V(NAS - 1, k) - V(NAS - 2, k) ' Boundary condition at S=infinity

If EType = "Y" Then ' Check for early exercise
For i = 0 To NAS
V(i, k) = Application.Max(V(i, k), Payoff(i))
Next i
End If

Next k

Option_Value_3D_US = V ' Output array

End Function
```

In Table 77.3 and Figure 77.13 are results for the American put, same parameters as above.

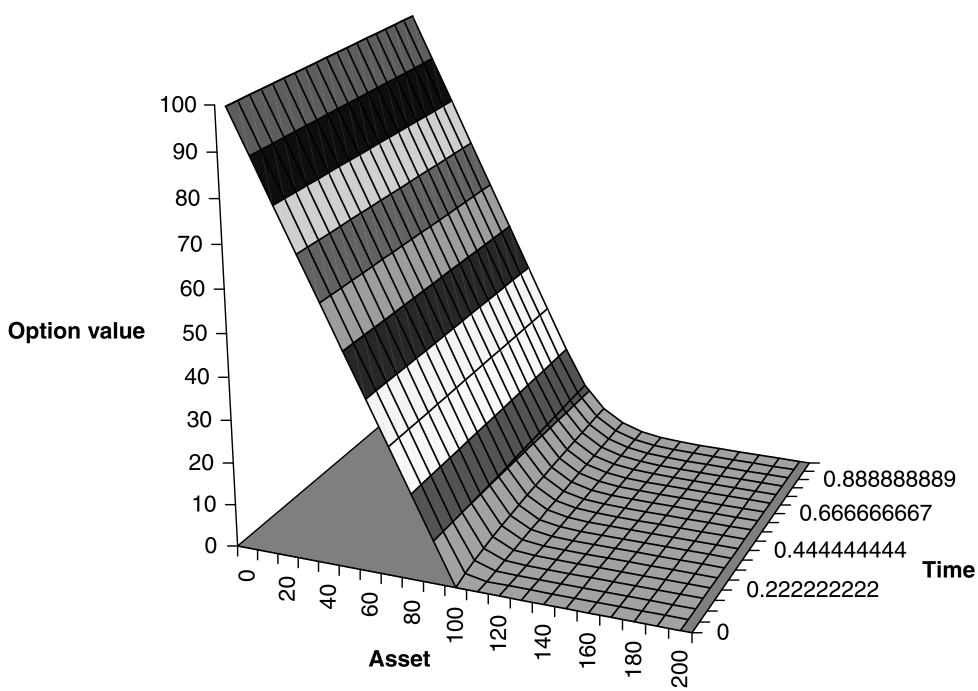


Figure 77.13 American put option values output by the explicit finite-difference code.

Table 77.3 American put option values output by the explicit finite-difference code. Stock price ranges from 60 to 100, time from 0 (expiration) to 1.

	0.000	0.111	0.222	0.333	0.444	0.556	0.667	0.778	0.889	1.000
60	40.000	40.000	40.000	40.000	40.000	40.000	40.000	40.000	40.000	40.000
70	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
80	20.000	20.000	20.000	20.000	20.000	20.000	20.000	20.000	20.000	20.000
90	10.000	10.000	10.000	10.060	10.226	10.434	10.653	10.869	11.074	11.267
100	0.000	1.726	2.810	3.549	4.102	4.551	4.935	5.275	5.581	5.859
110	0.000	0.116	0.491	0.909	1.305	1.666	1.994	2.294	2.571	2.827
120	0.000	0.000	0.053	0.174	0.338	0.523	0.714	0.905	1.093	1.276
130	0.000	0.000	0.003	0.025	0.073	0.143	0.230	0.328	0.434	0.544
140	0.000	0.000	0.000	0.002	0.013	0.034	0.067	0.110	0.162	0.221

77.15 THE CODE # 3: 2-D OUTPUT

The above code is fine if you want to output the entire function $V(S, t)$. But if you only want to know today's curve or value, it is very inefficient since it stores too much information. The following code only uses `VOld` and `VNew` to store enough information to calculate the curve today. The code also outputs the greeks.

```
Function Option_Value_2D_US(Vol, Int_Rate, PType, Strike, Expiration, _
                           EType, NAS)
' NAS is number of asset steps

ReDim S(0 To NAS) As Double ' Asset array
ReDim VOld(0 To NAS) As Double ' One option array
ReDim VNew(0 To NAS) As Double ' Second option array
ReDim Payoff(0 To NAS) As Double ' Payoff array
ReDim Dummy(0 To NAS, 1 To 6) As Double ' Used for storing data for output

dS = 2 * Strike / NAS ' 'Infinity' is twice the strike
dt = 0.9 / Vol ^ 2 / NAS ^ 2 ' For stability
NTS = Int(Expiration / dt) + 1 ' Number of time steps
dt = Expiration / NTS ' To ensure that expiration is an integer number of time _
                           steps away

q = 1
If PType = "P" Then q = -1 ' Test for call or put

For i = 0 To NAS
S(i) = i * dS ' Set up S array
VOld(i) = Application.Max(q * (S(i) - Strike), 0) ' Set up payoff
Payoff(i) = VOld(i) ' Store payoff
Dummy(i, 1) = S(i) ' First column of Dummy is the stock price
Dummy(i, 2) = Payoff(i) ' Second column of Dummy is the payoff
Next i

For k = 1 To NTS ' Time loop
For i = 1 To NAS - 1 ' Asset loop. End points treated separately
Delta = (VOld(i + 1) - VOld(i - 1)) / 2 / dS ' Central difference
Gamma = (VOld(i + 1) - 2 * VOld(i) + VOld(i - 1)) / dS / dS ' Central difference
```

```
Theta = -0.5 * Vol ^ 2 * S(i) ^ 2 * Gamma - _
                           Int_Rate * S(i) * Delta + Int_Rate * VOld(i) _
                           ' Black-Scholes
VNew(i) = VOld(i) - dt * Theta ' Update option value
Next i

VNew(0) = VOld(0) * (1 - Int_Rate * dt) ' Boundary condition at S=0
VNew(NAS) = 2 * VNew(NAS - 1) - VNew(NAS - 2) ' Boundary condition at S=infinity

For i = 0 To NAS ' Replace Old with New
VOld(i) = VNew(i)
Next i

If EType = "Y" Then ' Check for early exercise
For i = 0 To NAS
VOld(i) = Application.Max(VOld(i), Payoff(i))
Next i
End If

Next k

For i = 1 To NAS - 1
Dummy(i, 3) = VOld(i) ' Third column of Dummy is option value
Dummy(i, 4) = (VOld(i + 1) - VOld(i - 1)) / 2 / dS ' Fourth column of Dummy _
             is delta
Dummy(i, 5) = (VOld(i + 1) - 2 * VOld(i) + VOld(i - 1)) / dS / dS
                      ' Fifth column of Dummy is gamma
Dummy(i, 6) = -0.5 * Vol ^ 2 * S(i) ^ 2 * Dummy(i, 5) - _
             Int_Rate * S(i) * Dummy(i, 4) + Int_Rate * VOld(i) ' Sixth column _
             of Dummy is theta
Next i

Dummy(0, 3) = VOld(0)
Dummy(NAS, 3) = VOld(NAS)
Dummy(0, 4) = (VOld(1) - VOld(0)) / dS ' End points need to be treated _
             separately for delta, gamma and theta
Dummy(NAS, 4) = (VOld(NAS) - VOld(NAS - 1)) / dS
Dummy(0, 5) = 0
Dummy(NAS, 5) = 0
Dummy(0, 6) = Int_Rate * VOld(0)
Dummy(NAS, 6) = -0.5 * Vol ^ 2 * S(NAS) ^ 2 * Dummy(NAS, 5) - _
             Int_Rate * S(NAS) * Dummy(NAS, 4) + Int_Rate * VOld(NAS)

Option_Value_2D_US = Dummy ' Output array

End Function
```

Table 77.4 shows the results from this code, same parameters but now 40 asset steps.

77.16 BILINEAR INTERPOLATION

Suppose that we have an estimate for the option value, or its derivatives, on the mesh points; how can we estimate the value at points in between? The simplest way to do this is to use

Table 77.4 Call option values and greeks output by the explicit finite-difference code.

Stock	Payoff	Option	Delta	Gamma	Theta
0	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000
15	0.000	0.000	0.000	0.000	0.000
20	0.000	0.000	0.000	0.000	0.000
25	0.000	0.000	0.000	0.000	0.000
30	0.000	0.000	0.000	0.000	0.000
35	0.000	0.000	0.000	0.000	0.000
40	0.000	0.000	0.000	0.000	−0.001
45	0.000	0.001	0.000	0.000	−0.005
50	0.000	0.003	0.002	0.000	−0.023
55	0.000	0.016	0.005	0.001	−0.086
60	0.000	0.058	0.016	0.003	−0.256
65	0.000	0.173	0.038	0.006	−0.617
70	0.000	0.437	0.078	0.010	−1.235
75	0.000	0.953	0.139	0.015	−2.113
80	0.000	1.832	0.222	0.019	−3.168
85	0.000	3.174	0.321	0.021	−4.255
90	0.000	5.044	0.429	0.022	−5.224
95	0.000	7.461	0.536	0.021	−5.962
100	0.000	10.403	0.636	0.019	−6.425
105	5.000	13.816	0.723	0.016	−6.625
110	10.000	17.628	0.795	0.013	−6.614
115	15.000	21.764	0.852	0.010	−6.459
120	20.000	26.149	0.896	0.007	−6.225
125	25.000	30.722	0.928	0.005	−5.965
130	30.000	35.430	0.951	0.004	−5.712
135	35.000	40.235	0.968	0.003	−5.488
140	40.000	45.107	0.979	0.002	−5.301
145	45.000	50.023	0.986	0.001	−5.153
150	50.000	54.969	0.991	0.001	−5.039
155	55.000	59.935	0.994	0.001	−4.954
160	60.000	64.913	0.997	0.000	−4.892
165	65.000	69.900	0.998	0.000	−4.848
170	70.000	74.892	0.999	0.000	−4.817
175	75.000	79.886	0.999	0.000	−4.796
180	80.000	84.883	0.999	0.000	−4.781
185	85.000	89.881	1.000	0.000	−4.770
190	90.000	94.880	1.000	0.000	−4.761
195	95.000	99.878	1.000	0.000	−4.754
200	100.000	104.877	1.000	0.000	−4.754

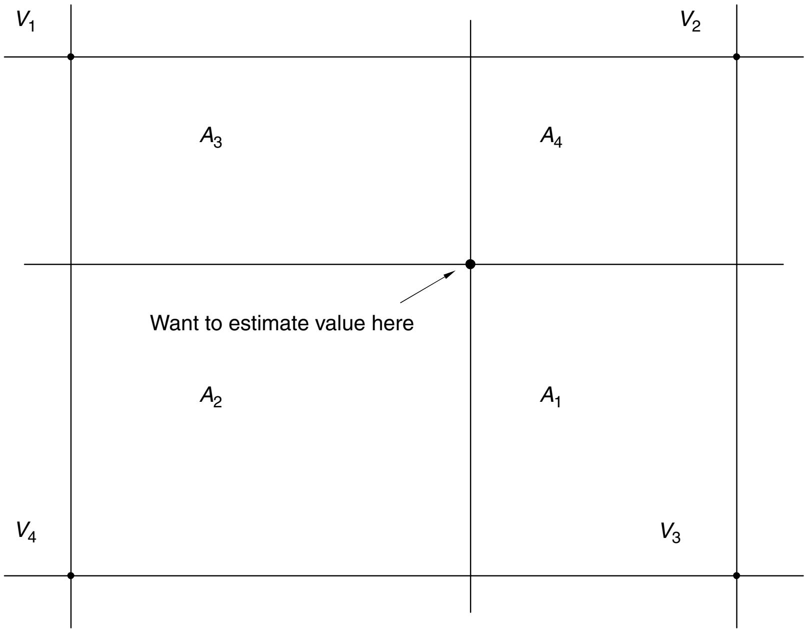
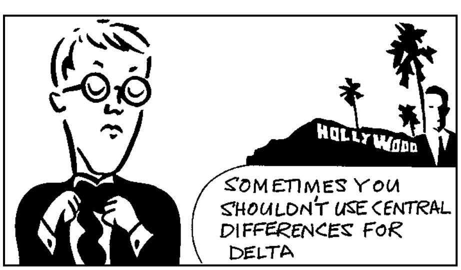


Figure 77.14 Bilinear interpolation.

a two-dimensional interpolation method called **bilinear interpolation**. This method is most easily explained via the schematic diagram in Figure 77.14.

We want to estimate the value of the option, say, at the interior point in the figure. The values of the option at the four nearest neighbors are called V_1, V_2, V_3 and V_4, simplifying earlier notation just for this brief section. The areas of the rectangles made by the four corners and the interior point are labeled A_1, A_2, A_3 and A_4. But note that the subscripts for the areas correspond to the subscripts of the option values at the corners *opposite*. The approximation for the option value at the interior point is then

$$\frac{\sum_{i=1}^{4} A_i V_i}{\sum_{i=1}^{4} A_i}.$$



77.17 UPWIND DIFFERENCING

The constraint (77.4) can be avoided if we use a one-sided difference instead of a central difference for the first derivative of the option value with respect to the asset. As I said in Chapter 6 the first S derivative represents a drift term. This drift has a direction associated with it, as t decreases, moving away from expiry, so the drift is towards smaller S. In a sense, this makes the forward price of the asset a better variable to use. Anyway, the numerical scheme can make use of this by using a one-sided difference. That is the situation for the Black–Scholes equation. More generally, the approximation that we use for delta in the equation could depend on the sign of b. For example, use the following

$$\text{if } b(S,t) \geq 0 \quad \text{then} \quad \frac{\partial V}{\partial S}(S,t) = \frac{V_{i+1}^k - V_i^k}{\delta S}$$

but if

$$b(S,t) < 0 \quad \text{then} \quad \frac{\partial V}{\partial S}(S,t) = \frac{V_i^k - V_{i-1}^k}{\delta S}.$$

This removes the limitation (77.4) on the asset step size, improving stability. However, since these one-sided differences are only accurate to $O(\delta S)$ the numerical method is less accurate.

The use of one-sided differences that depend on the sign of the coefficient b is called **upwind differencing**.[3] There is a small refinement of the technique in the choice of the value chosen for the function b:

$$\text{if } b(S,t) \geq 0 \quad \text{then} \quad b(S,t)\frac{\partial V}{\partial S}(S,t) = b_{i+\frac{1}{2}}^k \frac{V_{i+1}^k - V_i^k}{\delta S}$$

but if

$$b(S,t) < 0 \quad \text{then} \quad b(S,t)\frac{\partial V}{\partial S}(S,t) = b_{i-\frac{1}{2}}^k \frac{V_i^k - V_{i-1}^k}{\delta S}.$$

Notice how I have used the mid-point value for b.[4]

Below is a Visual Basic code fragment that uses a one-sided difference, depending on the sign of the drift term. This code fragment can be used for interest rate products, since it is solving

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma(r)^2 \frac{\partial^2 V}{\partial r^2} + (\mu(r) - \lambda(r)\sigma(r))\frac{\partial V}{\partial r} - rV = 0.$$

Note the arbitrary $\sigma(r) =$ `Volatility(IntRate(i))` and $\mu(r) - \lambda(r)\sigma(r) =$ `RiskNeutralDrift(IntRate(i))`.

This fragment of code is just the time stepping; above it would go declarations and setting up the payoff. Below it would go the outputting. It does not implement any boundary conditions at `i = 0` or at `i = NoIntRateSteps`; these would depend on the contract being valued.

```
For i = 1 To NoIntRateSteps - 1
    If RiskNeutralDrift(IntRate(i)) > 0 Then
         Delta(i) = (VOld(i + 1) - VOld(i)) / dr
         RNDrift = RiskNeutralDrift(IntRate(i) + 0.5 * dr)
    Else
         Delta(i) = (VOld(i) - VOld(i - 1)) / dr
         RNDrift = RiskNeutralDrift(IntRate(i) - 0.5 * dr)
    End If
    Gamma(i) = (VOld(i + 1) - 2 * VOld(i) + VOld(i - 1)) / (dr * dr)
    VNew(i) = VOld(i) + Timestep * (0.5 * Volatility(IntRate(i)) _
* Volatility(IntRate(i)) * Gamma(i) + RNDrift * Delta(i) - IntRate(i) * VOld(i))
Next i
```

[3] That's 'wind' as in breeze, not as in to wrap or coil.

[4] This choice won't make much difference to the result but it does help to make the numerical method 'conservative,' meaning that certain properties of the partial differential equation are retained by the solution of the difference equation. Having a conservative scheme is important in computational fluid dynamics applications, otherwise the scheme will exhibit mass 'leakage.'

To get back the $O(\delta S^2)$ accuracy of the central difference with a one-sided difference you can use the approximations described in Section 77.6.

We have seen that the explicit finite-difference method suffers from restrictions in the size of the grid steps. The explicit method is similar in principle to the binomial tree method, and the restrictions can be interpreted in terms of risk-neutral probabilities. The terms A, B and C are related to the risk-neutral probabilities of reaching the points $i-1$, i and $i+1$. Instability is equivalent to one of these being a negative quantity, and we can't allow negative probabilities. More sophisticated numerical methods exist that do not suffer from such restrictions, and I will describe these next.

The advantages of the explicit method

- It is very easy to program and hard to make mistakes
- When it does go unstable it is usually obvious
- It copes well with coefficients that are asset and/or time-dependent
- It is easy to incorporate accurate one-sided differences

The disadvantage of the explicit method

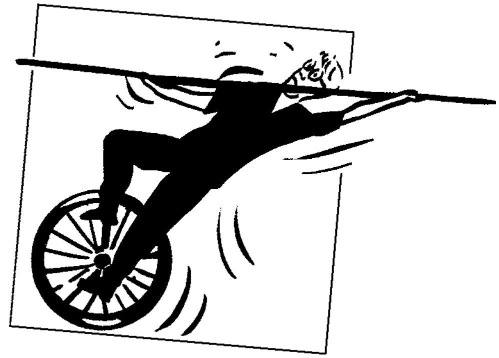
- There are restrictions on the time step so the method can be slower than other schemes

77.18 SUMMARY

The diffusion equation has been around for a *long* time. Numerical schemes for the solution of the diffusion equation have been around quite a while too, not as long as the equation itself but certainly a lot longer than the Black–Scholes equation and the binomial method. This means that there is a great deal of academic literature on the efficient solution of parabolic equations in general. I've introduced you to the subject with the explicit method, and in the next chapter I'll show you how the numerical methods can get more sophisticated.

FURTHER READING

- For general numerical methods see Johnson & Riess (1982) and Gerald & Wheatley (1992).
- The first use of finite-difference methods in finance was due to Brennan & Schwartz (1978). For its application in interest rate modeling see Hull & White (1990).
- An excellent, well written, book on numerical methods in finance is Ahmad (2006).



CHAPTER 78
further finite-difference methods for one-factor models

In this Chapter...

- implicit finite-difference methods including Crank–Nicolson
- Douglas schemes
- Richardson extrapolation
- American-style exercise
- exotic options

78.1 INTRODUCTION

We continue with one-factor numerical methods, discussing the more difficult to program implicit methods. The extra complexity of the methods is outweighed, though, by their superior stability properties. I also show how to extend the finite-difference method to cope with early exercise and path-dependent contracts.

78.2 IMPLICIT FINITE-DIFFERENCE METHODS

The **fully implicit method** uses the points as shown in Figure 78.1 to calculate the option value. The scheme is superficially just like the explicit method using finite-difference estimates of the option value, its delta and gamma but now at the time step $k+1$. The relationship between the option values on the mesh is simply

$$\frac{V_i^k - V_i^{k+1}}{\delta t}$$
$$+ a_i^{k+1}\left(\frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\delta S^2}\right)$$

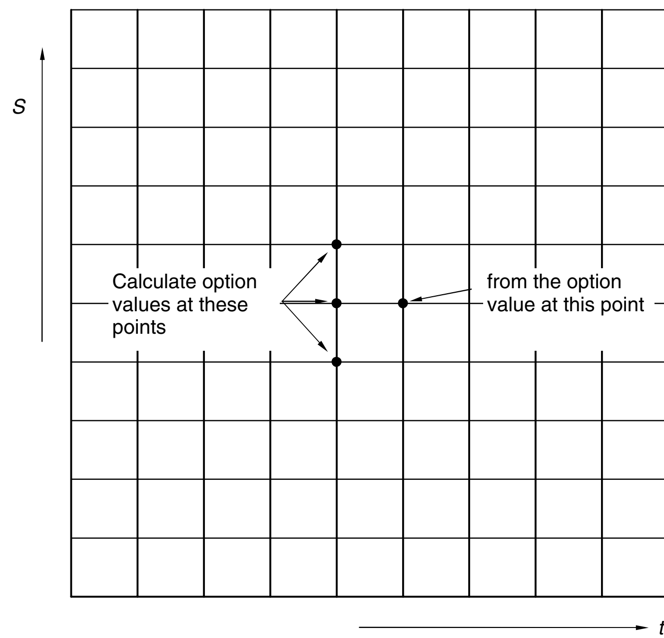


Figure 78.1 The relationship between option values in the fully implicit method.

$$+\, b_i^{k+1}\left(\frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\,\delta S}\right)$$
$$+\, c_i^{k+1} V_i^{k+1} = 0.$$

(It doesn't matter much whether the coefficients a, b and c are evaluated at the time step $k+1$ or k.) The method is still accurate to $O(\delta t, \delta S^2)$.

This can be written as

$$A_i^{k+1} V_{i-1}^{k+1} + (1 + B_i^{k+1}) V_i^{k+1} + C_i^{k+1} V_{i+1}^{k+1} = V_i^k \qquad (78.1)$$

where

$$A_i^{k+1} = -\nu_1 a_i^{k+1} - \tfrac{1}{2}\nu_2 b_i^{k+1},$$
$$B_i^{k+1} = 2\nu_1 a_i^{k+1} - \delta t\; c_i^{k+1}$$

and

$$C_i^{k+1} = -\nu_1 a_i^{k+1} + \tfrac{1}{2}\nu_2 b_i^{k+1}$$

where

$$\nu_1 = \frac{\delta t}{\delta S^2} \quad \text{and} \quad \nu_2 = \frac{\delta t}{\delta S}.$$

Again, Equation (78.1) does not hold for $i = 0$ or $i = I$, the boundary conditions supply the two remaining equations.

There is a world of difference between this scheme and the explicit finite-difference scheme. The two main differences concern the stability of the method and the solution procedure.

The method no longer suffers from the restriction on the time step. The asset step can be small and the time step large without the method running into stability problems.

The solution of the difference equation is no longer so straightforward. To get V_i^{k+1} from V_i^k means solving a set of linear equations; each V_i^{k+1} is directly linked to its two neighbors and thus indirectly linked to every option value at the same time step.

I am not going to pursue the fully implicit method any further since the method can be significantly improved upon with little extra computational effort.

78.3 THE CRANK–NICOLSON METHOD

The **Crank–Nicolson method** can be thought of as an average of the explicit method and the fully implicit method. It uses the six points shown in Figure 78.2.

The Crank–Nicolson scheme is

$$\begin{aligned}
&\frac{V_i^k - V_i^{k+1}}{\delta t} \\
&+ \frac{a_i^{k+1}}{2}\left(\frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\delta S^2}\right) + \frac{a_i^k}{2}\left(\frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}\right) \\
&+ \frac{b_i^{k+1}}{2}\left(\frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\,\delta S}\right) + \frac{b_i^k}{2}\left(\frac{V_{i+1}^k - V_{i-1}^k}{2\,\delta S}\right) \\
&+ \tfrac{1}{2}c_i^{k+1}V_i^{k+1} + \tfrac{1}{2}c_i^k V_i^k = O(\delta t^2, \delta S^2).
\end{aligned}$$

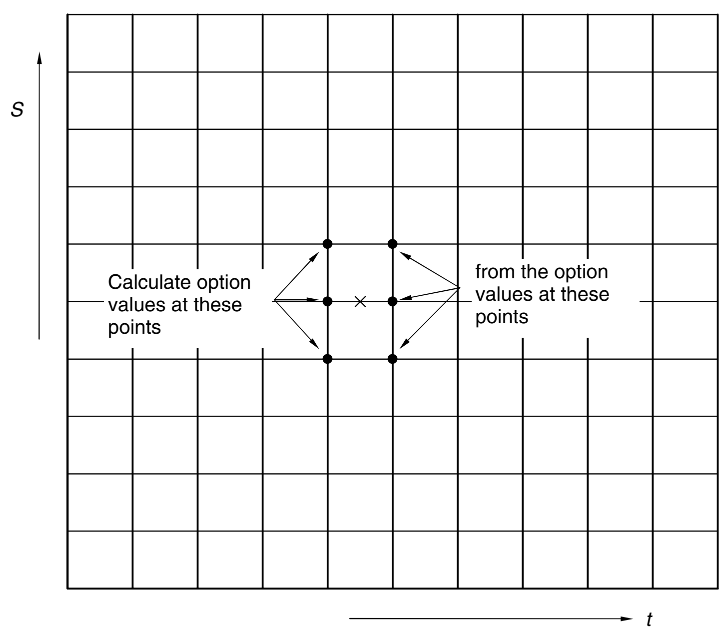


Figure 78.2 The relationship between option values in the Crank–Nicolson method.

This can be written as

$$-A_i^{k+1}V_{i-1}^{k+1} + (1 - B_i^{k+1})V_i^{k+1} - C_i^{k+1}V_{i+1}^{k+1} = A_i^k V_{i-1}^k + (1 + B_i^k)V_i^k + C_i^k V_{i+1}^k, \qquad (78.2)$$

where

$$A_i^k = \tfrac{1}{2}\nu_1 a_i^k - \tfrac{1}{4}\nu_2 b_i^k,$$
$$B_i^k = -\nu_1 a_i^k + \tfrac{1}{2}\delta t\, c_i^k,$$

and

$$C_i^k = \tfrac{1}{2}\nu_1 a_i^k + \tfrac{1}{4}\nu_2 b_i^k.$$

These equations only hold for $1 \le i \le I - 1$. The boundary conditions again supply the two missing equations. They are harder to handle than in the explicit method and I will discuss them in depth shortly.

Although this looks a bit messy, and it is harder to solve, the beauty of this method lies in its stability and accuracy. As with the fully implicit method there is no relevant limitation on the size of the time step for the method to converge. Better yet, the method is more accurate than the two considered so far. The error in the method is $O(\delta t^2, \delta S^2)$. So now we can use larger time steps, and still get an accurate solution. To see that the error due to the size of the time step is now $O(\delta t^2)$ expand Equation (78.2) about the point S, $t - \frac{1}{2}\delta t$ (this is the point 'x' in Figure 78.2).

The Crank–Nicolson method can be written in the matrix form

$$\begin{pmatrix} -A_1^{k+1} & 1-B_1^{k+1} & -C_1^{k+1} & 0 & \cdot & \cdot & \cdot \\ 0 & -A_2^{k+1} & 1-B_2^{k+1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} & 0 \\ \cdot & \cdot & \cdot & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} & -C_{I-1}^{k+1} \end{pmatrix} \begin{pmatrix} V_0^{k+1} \\ V_1^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^{k+1} \\ V_I^{k+1} \end{pmatrix}$$

$$= \begin{pmatrix} A_1^k & 1+B_1^k & C_1^k & 0 & \cdot & \cdot & \cdot \\ 0 & A_2^k & 1+B_2^k & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1+B_{I-2}^k & C_{I-2}^k & 0 \\ \cdot & \cdot & \cdot & 0 & A_{I-1}^k & 1+B_{I-1}^k & C_{I-1}^k \end{pmatrix} \begin{pmatrix} V_0^k \\ V_1^k \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^k \\ V_I^k \end{pmatrix} \qquad (78.3)$$

The two matrices have $I - 1$ rows and $I + 1$ columns. This is a representation of $I - 1$ equations in $I + 1$ unknowns.

The two equations that we are missing come from the boundary conditions. Using these conditions, I am going to convert this system of equations, (78.3), into a system of equations involving a square matrix. The aim is to write the system of equations in the form

$$\mathbf{M}_L^{k+1}\mathbf{v}^{k+1} + \mathbf{r}^k = \mathbf{M}_R^k\mathbf{v}^k,$$

for known *square* matrices $\mathbf{M}_L^{k+1}$ and $\mathbf{M}_R^k$, and a known vector $\mathbf{r}^k$, and where details of the boundary conditions have been fully incorporated.

I will consider three cases separately, depending on the form of the boundary condition that is to be implemented. I will also only deal with a boundary condition at $i = 0$; obviously boundary conditions at $i = I$ are treated similarly.

78.3.1 Boundary Condition Type I: V_0^{k+1} Given

Sometimes we know that our option has a particular value on the boundary $i = 0$, or on $i = I$. For example, if we have a European put we know that $V(0, t) = Ee^{-r(T-t)}$. This translates to knowing that $V_0^{k+1} = Ee^{-r(k+1)\delta t}$. We can write

$$\begin{pmatrix} -A_1^{k+1} & 1-B_1^{k+1} & -C_1^{k+1} & 0 & \cdot & \cdot & \cdot \\ 0 & -A_2^{k+1} & 1-B_2^{k+1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} & 0 \\ \cdot & \cdot & \cdot & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} & -C_{I-1}^{k+1} \end{pmatrix} \begin{pmatrix} V_0^{k+1} \\ V_1^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^{k+1} \\ V_I^{k+1} \end{pmatrix}$$

as

$$\begin{pmatrix} 1-B_1^{k+1} & -C_1^{k+1} & 0 & \cdot & \cdot & & \\ -A_2^{k+1} & 1-B_2^{k+1} & \cdot & \cdot & \cdot & & \\ & 0 & \cdot & \cdot & \cdot & 0 \\ & \cdot & \cdot & \cdot & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} \\ & \cdot & \cdot & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} \end{pmatrix} \begin{pmatrix} V_1^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ V_{I-1}^{k+1} \end{pmatrix} + \begin{pmatrix} -A_1^{k+1}V_0^{k+1} \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ \cdot \end{pmatrix} = \mathbf{M}_L^{k+1}\mathbf{v}^{k+1} + \mathbf{r}^k.$$

All I have done here is to multiply out the top and bottom rows of the matrix. The matrix $\mathbf{M}_L$ is square and of size $I - 1$. The vector $\mathbf{r}^k$, which is of length $I - 1$, only has non-zero elements at the top (and bottom) *and is completely known because it only depends on the function A and the prescribed value of V at the boundary.*

78.3.2 Boundary Condition Type II: Relationship Between V_0^{k+1} and V_1^{k+1}

If we have a barrier option for which a grid point does not coincide with the barrier we must use the approximation explained in Section 77.10:

$$V_0^{k+1} = \frac{1}{\alpha}\left(f - (1-\alpha)V_1^{k+1}\right).$$

This relationship between V_0^{k+1} and V_1^{k+1} for a 'down' barrier or between V_I^{k+1} and V_{I-1}^{k+1} for an 'up' barrier is seen in other contexts. Perhaps we know the slope of the option value for large or small S; this gives us a gradient boundary condition, which is also of this form. If we use a one-sided difference for the derivative then the boundary condition can also be written as a relationship between the last grid point and the last but one. More generally, suppose we have

$$V_0^{k+1} = a + bV_1^{k+1}.$$

This time the left-hand side of (78.3) can be written as

$$\begin{pmatrix} 1-B_1^{k+1} & -C_1^{k+1} & 0 & . & . & & \\ -A_2^{k+1} & 1-B_2^{k+1} & . & . & . & & \\ & 0 & . & . & . & 0 & \\ & . & . & . & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} & \\ & . & . & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} & \end{pmatrix} \begin{pmatrix} V_1^{k+1} \\ . \\ . \\ . \\ . \\ . \\ V_{I-1}^{k+1} \end{pmatrix} + \begin{pmatrix} -A_1^{k+1}(a+bV_1^{k+1}) \\ 0 \\ 0 \\ . \\ . \\ 0 \\ . \end{pmatrix}.$$

Again, this comes from multiplying out the top row of the matrix. By absorbing the V_1^{k+1} term into the matrix we can write this as

$$\begin{pmatrix} 1-B_1^{k+1}-bA_1^{k+1} & -C_1^{k+1} & 0 & . & . & & \\ -A_2^{k+1} & 1-B_2^{k+1} & . & . & . & & \\ & 0 & . & . & . & 0 & \\ & . & . & . & 1-B_{I-2}^{k+1} & -C_{I-2}^{k+1} & \\ & . & . & 0 & -A_{I-1}^{k+1} & 1-B_{I-1}^{k+1} & \end{pmatrix} \begin{pmatrix} V_1^{k+1} \\ . \\ . \\ . \\ . \\ . \\ V_{I-1}^{k+1} \end{pmatrix} + \begin{pmatrix} -aA_1^{k+1} \\ 0 \\ 0 \\ . \\ . \\ 0 \\ . \end{pmatrix}.$$

Again this is of the form

$$\mathbf{M}_L^{k+1}\mathbf{v}^{k+1} + \mathbf{r}^k$$

but for different $\mathbf{M}_L^{k+1}$ and $\mathbf{r}^k$ from before. This matrix and this vector are both known.

78.3.3 Boundary Condition Type III: $\partial^2 V/\partial S^2 = 0$

More complicated, but still perfectly manageable, is the boundary condition

$$\frac{\partial^2 V}{\partial S^2} = 0.$$

This condition is particularly useful since it is independent of the type of contract, as long as the contract has a payoff that is at most linear in the underlying. This condition is, in central difference form,

$$V_0^{k+1} = 2V_1^{k+1} - V_2^{k+1}.$$

Thus, this time we can write the left-hand side of (78.3) as

$$\begin{pmatrix} 1 - B_1^{k+1} - 2A_1^{k+1} & -C_1^{k+1} + A_1^{k+1} & 0 & . & . & & \\ -A_2^{k+1} & 1 - B_2^{k+1} & . & . & . & & \\ & 0 & . & . & . & 0 & \\ & . & . & . & 1 - B_{I-2}^{k+1} & -C_{I-2}^{k+1} \\ & . & . & 0 & -A_{I-1}^{k+1} & 1 - B_{I-1}^{k+1} \end{pmatrix} \begin{pmatrix} V_1^{k+1} \\ . \\ . \\ . \\ . \\ . \\ V_{I-1}^{k+1} \end{pmatrix}.$$

In this case there is no need to include an extra known vector $\mathbf{r}^k$. The left-hand side of our matrix equation is again

$$\mathbf{M}_L^{k+1}\mathbf{v}^{k+1} + \mathbf{r}^k$$

but again for different $\mathbf{M}_L^{k+1}$ and this time $\mathbf{r}^k$ is zero.

It will be instructive for the reader to absorb other boundary conditions into the matrix equation.

78.3.4 The Matrix Equation

Whichever of the boundary conditions we have, the Crank–Nicolson scheme, with boundary conditions incorporated, is

$$\mathbf{M}_L^{k+1}\mathbf{v}^{k+1} = \mathbf{M}_R^k\mathbf{v}^k - \mathbf{r}^k \tag{78.4}$$

where $\mathbf{v}^k$ is the column vector with $I - 1$ entries V_i^k ($i = 1$ at the top) and $\mathbf{M}_L^{k+1}$ and $\mathbf{M}_R^k$ are square matrices of size $I - 1$.

Remembering that we know $\mathbf{v}^k$, the matrix multiplication and vector addition on the right-hand side of (78.4) is simple enough to do. But how do we then find $\mathbf{v}^{k+1}$? This vector equation amounts to $I - 1$ equations in $I - 1$ unknowns, a set of linear simultaneous equations. In principle, the matrix $\mathbf{M}_L^{k+1}$ could be inverted to give

$$\mathbf{v}^{k+1} = \left(\mathbf{M}_L^{k+1}\right)^{-1}\left(\mathbf{M}_R^k\mathbf{v}^k - \mathbf{r}^k\right),$$

except that matrix inversion is very time consuming, and from a computational point of view extremely inefficient. There are two much better ways for solving (78.4) called LU decomposition and successive over-relaxation.

78.3.5 LU Decomposition

The matrix $\mathbf{M}_L^{k+1}$ is special in that it is **tridiagonal**, the only non-zero elements lie along the diagonal and the sub- and superdiagonals. This means that it is not too hard to decompose the matrix into the product of two other matrices, one having non-zero elements along the diagonal and the subdiagonal and the other having non-zero elements along the diagonal and the superdiagonal. I will call these two matrices **L** and **U** respectively. For the rest of this section I will drop all unimportant sub- and superscripts. Thus we can write

$$\mathbf{M} = \mathbf{L}\mathbf{U}. \tag{78.5}$$

I will first show how this **LU decomposition** is achieved and then show how it is used in the solution of (78.4). LU decomposition is an example of a 'direct method,' the aim being to find the exact solution of the equations in one go. The only error will be due to rounding.

Equation (78.5) is

$$\begin{pmatrix} 1-B_1 & -C_1 & 0 & . & . & . & 0 \\ -A_2 & 1-B_2 & -C_2 & 0 & . & . & . \\ 0 & -A_3 & 1-B_3 & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & 1-B_{I-3} & -C_{I-3} & 0 \\ . & . & . & 0 & -A_{I-2} & 1-B_{I-2} & -C_{I-2} \\ . & . & . & . & 0 & -A_{I-1} & 1-B_{I-1} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & . & . & . & 0 \\ l_2 & 1 & 0 & 0 & . & . & . \\ 0 & l_3 & 1 & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & 1 & 0 & 0 \\ . & . & . & 0 & l_{I-2} & 1 & 0 \\ . & . & . & . & 0 & l_{I-1} & 1 \end{pmatrix} \begin{pmatrix} d_1 & u_1 & 0 & . & . & . & 0 \\ 0 & d_2 & u_2 & 0 & . & . & . \\ 0 & 0 & d_3 & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & d_{I-3} & u_{I-3} & 0 \\ . & . & . & 0 & 0 & d_{I-2} & u_{I-2} \\ . & . & . & . & 0 & 0 & d_{I-1} \end{pmatrix}$$

where, without loss of generality, I have chosen the diagonal elements of **L** to be one. I will leave it to the reader to show that

$$d_1 = 1 - B_1$$

and then

$$l_i d_{i-1} = -A_i, \quad u_{i-1} = -C_{i-1} \quad \text{and} \quad d_i = 1 - B_i - l_i u_{i-1} \quad \text{for} \quad 2 \le i \le I-1. \tag{78.6}$$

Notice how we work from $i = 1$ to $i = I$.

Here is a Visual Basic code fragment that takes in the diagonal, the superdiagonal and the subdiagonal of the matrix **M** in the arrays `Diag()`, `SuperDiag()` and `SubDiag()` each of length `NoElements`, and calculates d, u and l.

```
Dim d(1 To NoElements) As Double
Dim u(1 To NoElements) As Double
Dim l(1 To NoElements) As Double
    d(1) = Diag(1)
    For i = 2 To NoElements
        u(i - 1) = SuperDiag(i - 1)
        l(i) = SubDiag(i) / d(i - 1)
        d(i) = Diag(i) - l(i) * SuperDiag(i - 1)
    Next i
```

So that's the first step, the decomposition of the matrix. Now we exploit the decomposition to solve (78.4).

Assuming that we have done the multiplication on the right-hand side of (78.4) to give

$$\mathbf{Mv} = \mathbf{q}$$

(where all sub- and superscripts have been dropped), we can write

$$\mathbf{LUv} = \mathbf{q}. \tag{78.7}$$

The vector $\mathbf{q}$ contains both the old option value array, at time step k, and details of the boundary conditions.

I am going to solve Equation (78.7) in two steps. First find $\mathbf{w}$ such that

$$\mathbf{Lw} = \mathbf{q}$$

and then $\mathbf{v}$ such that

$$\mathbf{Uv} = \mathbf{w}.$$

And then we are done.

The first step gives

$$w_1 = q_1$$

and

$$w_i = q_i - l_i w_{i-1} \quad \text{for} \quad 2 \leq i \leq I - 1.$$

Again we must work from $i = 2$ to $i = I - 1$ sequentially. The second step involves working *backwards* from $i = I - 2$ to $i = 1$:

$$v_{I-1} = \frac{w_{I-1}}{d_{I-1}}$$

and

$$v_i = \frac{w_i - u_i v_{i+1}}{d_i} \quad \text{for} \quad I - 2 \geq i \geq 1.$$

Here is a Visual Basic code fragment that takes the just-calculated `l` and `d`, and the right-hand side `q` and calculates the `v`.

```
Dim v(1 To NoElements) As Double
Dim w(1 To NoElements) As Double

    w(1) = q(1)
    For i = 2 To NoElements
        w(i) = q(i) - l(i) * w(i - 1)
    Next i

    v(NoElements) = w(NoElements) / d(NoElements)
    For i = NoElements - 1 To 1 Step -1
        v(i) = (w(i) - u(i) * v(i + 1)) / d(i)
    Next i
```

If our matrix equation is time homogeneous then the LU decomposition need only be done the once. That is, if a, b and c are functions of S only then do we do the decomposition once and store the necessary results. The method therefore works well for the classic Black–Scholes equation where the coefficients are only functions of S. These days it is common to have time dependency in the coefficients. The drift and volatility terms, whether in equity or fixed-income problems, usually have some term structure. This means that there is important time dependence in the coefficients. This time dependency means that the LU decomposition must be done afresh at each time step, significantly slowing down the computational time.

The advantages of the LU decomposition method

- It is quick
- The decomposition need only be done once if the matrix $\mathbf{M}$ is independent of time

The disadvantages of the LU decomposition method

- It is not immediately applicable to American options
- The decomposition must be done each time step if the matrix $\mathbf{M}$ is time-dependent

78.3.6 Successive Over-relaxation, SOR

We now come to an example of an 'indirect method.' With this we solve the equations iteratively. Although the resulting 'solution' will never be exact we can find the solution to whatever accuracy we want, and as long as this can be done fast then it will be good enough. Indirect methods are applicable to a wider range of problems—our left-hand matrix need not be tridiagonal for example—so I will describe the ideas in generality. This means I am going to drop all sub- and superscripts and change notation slightly.

Suppose that the matrix $\mathbf{M}$ in the matrix equation

$$\mathbf{M}\mathbf{v} = \mathbf{q}$$

has entries M_{ij} then the system of equations can be written as

$$M_{11}v_1 + M_{12}v_2 + \cdots + M_{1N}v_N = q_1$$
$$M_{21}v_1 + M_{22}v_2 + \cdots + M_{2N}v_N = q_2$$
$$\cdots$$
$$M_{N1}v_1 + M_{N2}v_2 + \cdots + M_{NN}v_N = q_N$$

where now N is the number of equations, the size of the matrix.

Rewrite this as

$$M_{11}v_1 = q_1 - (M_{12}v_2 + \cdots + M_{1N}v_N)$$
$$M_{22}v_2 = q_2 - (M_{21}v_1 + \cdots + M_{2N}v_N)$$
$$\cdots$$
$$M_{NN}v_N = q_N - (M_{N1}v_1 + \cdots)$$

This system is easily solved *iteratively* using

$$v_1^{n+1} = \frac{1}{M_{11}}\left(q_1 - (M_{12}v_2^n + \cdots + M_{1N}v_N^n)\right)$$
$$v_2^{n+1} = \frac{1}{M_{22}}\left(q_2 - (M_{21}v_1^n + \cdots + M_{2N}v_N^n)\right)$$
$$\cdots$$
$$v_N^{n+1} = \frac{1}{M_{NN}}\left(q_N - (M_{N1}v_1^n + \cdots)\right)$$

where the superscript n denotes the level of the iteration *and not the time step*. This iteration is started from some initial guess $\mathbf{v}^0$. (When solving our finite-difference problem it is usual to *start with the value of the option at the previous time step as the initial guess for the next time step*.) This iterative method is called the **Jacobi method**.

I can write the matrix $\mathbf{M}$ as the sum of a diagonal matrix $\mathbf{D}$, an upper triangular matrix $\mathbf{U}$ with zeros on the diagonal and a lower triangular matrix $\mathbf{L}$ also with zeros on the diagonal:

$$\mathbf{M} = \mathbf{D} + \mathbf{U} + \mathbf{L}.$$

I can use this representation to write the Jacobi and other methods quite elegantly. The Jacobi method is then

$$\mathbf{v}^{n+1} = \mathbf{D}^{-1}\left(\mathbf{q} - (\mathbf{U} + \mathbf{L})\mathbf{v}^n\right).$$

When we implement the Jacobi method in practice we find some of the values of v_i^{n+1} before others. In the **Gauss–Seidel** method we use the updated values as soon as they are calculated. This method can be written as

$$v_i^{n+1} = \frac{1}{M_{ii}}\left(q_i - \sum_{j=1}^{i-1} M_{ij}v_j^{n+1} - \sum_{j=i+1}^{N} M_{ij}v_j^n\right).$$

Observe that there are some terms on the right-hand side with the superscript $n+1$. These are values of v that were calculated earlier but at the same level of iteration. This method can be written more compactly as

$$\mathbf{v}^{n+1} = (\mathbf{D}+\mathbf{L})^{-1}\left(\mathbf{q}-\mathbf{U}\mathbf{v}^{n}\right).$$

When the matrix **M** has come from a finite-difference discretization of a parabolic equation (and that includes almost all finance problems) the above iterative methods usually converge to the correct solution *from one side*. This means that the corrections $v_i^{n+1} - v_i^n$ stay of the same sign as n increases, whether positive or negative depends on which side of the true solution was the initial guess. This is exploited in the **successive over-relaxation** or **SOR** method to speed up convergence.

The method can be written as

$$v_i^{n+1} = v_i^n + \frac{\omega}{M_{ii}}\left(q_i - \sum_{j=1}^{i-1} M_{ij}v_j^{n+1} - \sum_{j=i}^{N} M_{ij}v_j^n\right).$$

Again, the new values for v_i are used as soon as they are obtained. But now the factor ω, called the **acceleration** or **over-relaxation parameter** is included. This parameter, which must lie between 1 and 2, speeds up the convergence to the true solution.[1]

In matrix form we can write

$$\mathbf{v}^{n+1} = (\mathbf{I}+\omega\mathbf{D}^{-1}\mathbf{L})^{-1}((((1-\omega)\mathbf{I}-\omega\mathbf{D}^{-1}\mathbf{U})\mathbf{v}^{n})+\omega\mathbf{D}^{-1}\mathbf{q})$$

where **I** is the identity matrix.

Below is a Visual Basic code fragment that calculates `v` given the matrix **M** and the right-hand side `q`. The code assumes that our matrix **M** is tridiagonal with diagonal `MDiag(i)`, superdiagonal `MSuperDiag(i)` and subdiagonal `MSubDiag(i)`. The algorithm iterates until the mean square error `Error` is less than some pre-specified tolerance `tol`. Notice that I keep track of how many iteration were needed; this is `NoIts`.

```
Dim q(1 To N) As Double
Dim v(0 To N + 1) As Double
Dim temp(1 To N) As Double
Dim Diag(1 To N) As Double
Dim SuperDiag(1 To N + 1) As Double
Dim SubDiag(0 To N) As Double

SuperDiag(N + 1) = 0
SubDiag(0) = 0
NoIts = 0
While Err > tol
    Err = 0
    For i = 1 To N
        temp(i) = v(i) + omega * (q(i) - SuperDiag(i) * v(i + 1) _
                  - Diag(i) * v(i) - SubDiag(i) * v(i - 1)) / Diag(i)
        Err = Err + (temp(i) - v(i)) * (temp(i) - v(i))
        v(i) = temp(i) ' use as soon as calculated
```

[1] In iteration problems where successive results oscillate about the true solution the parameter ω would be chosen to be less than one; this would then be called **under-relaxation**.

```
    Next i
    NoIts = NoIts + 1
Wend
```

78.3.7 Optimal Choice of ω

The error $\mathbf{e}^n = \mathbf{v}^n - \mathbf{v}$, where $\mathbf{v}$ is the exact solution, satisfies

$$\mathbf{e}^{n+1} = (\mathbf{I} + \omega\mathbf{D}^{-1}\mathbf{L})^{-1}((1-\omega)\mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{U})\mathbf{e}^n.$$

The SOR method will converge provided that the largest of the moduli of the eigenvalues, a.k.a. the **spectral radius**, of the SOR matrix

$$(\mathbf{I} + \omega\mathbf{D}^{-1}\mathbf{L})^{-1}((1-\omega)\mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{U})$$

is less than one. There is a theoretical optimal value for the acceleration parameter ω, and that is when the spectral radius is a minimum. This value for ω maximizes the rate of convergence of the SOR method. How to find the optimal value of ω is discussed in depth in Smith (1985). In practice it is very simple to iterate on ω to find the optimal value. This is done as follows.

Start with $\omega = 1$. After the first time step is completed to the required accuracy, record the number of iterations taken to convergence. This is the quantity `NoIts` in the Visual Basic code. For the next time step increase ω by a small amount, say 0.05. Again record the number of iterations needed to get the required accuracy. If the number of iterations is fewer than for the first time step then, for the third time step, increase ω by another 0.05. Keep increasing ω at each time step until the number of iterations begins to increase again. Choose ω to be the value that minimizes the number of iterations. If the SOR matrix is time homogeneous then you won't need to modify the parameter again. If there is a very strong time dependence in the matrix then you may want to experiment with retesting for the optimal ω.

The advantages of the SOR method

- It is easier to program than LU decomposition
- It is easily applied to American options

The disadvantage of the SOR method

- It is slightly slower than LU decomposition for European options

78.4 COMPARISON OF FINITE-DIFFERENCE METHODS

Now that we have seen the explicit, the fully implicit and the Crank–Nicolson methods we can take a look at real, rather than theoretical, errors in the methods. In the next two figures are shown the errors for the three schemes in the solution of an at-the-money European call option, with strike 20, three months to expiry, a volatility of 20% and an interest rate of 5%. In Figure 78.3 I show the logarithm of the error against the logarithm of the asset step size. As expected the slope of all three lines is about two. The error in the three methods decreases like δS^2. In Figure 78.4 I show the logarithm of the error against the logarithm of the time

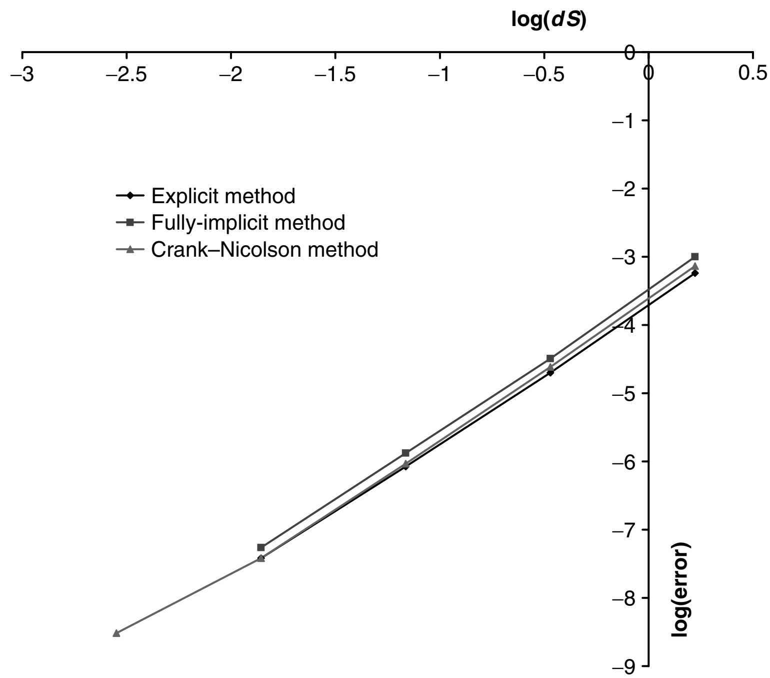


Figure 78.3 Log(Error) as a function of log(δS) for the three finite-difference schemes.

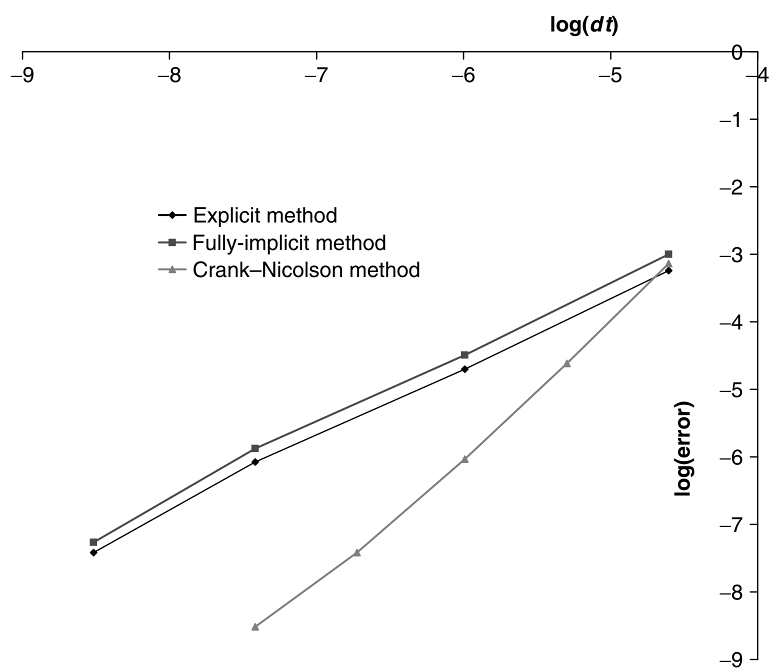


Figure 78.4 Log(Error) as a function of log(δt) for the three finite-difference schemes.

step. Now we can see the difference between the three methods. Both the explicit and the fully implicit have an error that decreases like δt but the Crank–Nicolson method is much better, having an error that decreases like δt^2.

78.5 OTHER METHODS

The methods described above are just three of many schemes that have been tried out over the years in computational fluid dynamics for example. Other schemes are possible using more than three asset points, more than two time points etc. The questions that arise in any method are the following:

- What is the error in the method in terms of δt and δS?
- What are the restrictions on the time step and/or asset step?
- Can I solve the resulting difference equations quickly?

We need answers to these three questions to decide which is the best method for our particular problem. Another question that arises, which may or may not be important, depending on how much effort you want to put into your programming:

- Is the method flexible enough to cope with changes in coefficients, boundary conditions etc.?

That is, do you have to begin again from scratch if the contract changes slightly, or can you simply change a subroutine to cope with the new contract? One important example of this is early exercise: Can the methods explained above handle American options? I will discuss this in a moment. But first, I want to describe a few more extensions of the finite difference ideas, ending up with a very, very simple technique that may speed up convergence. It doesn't always work, but it is definitely worth trying out because it involves no extra programming.

78.6 DOUGLAS SCHEMES

A method that has recently been rediscovered by financial researchers is the **Douglas scheme**. This is a method that manages to have a local truncation error of $O(\delta S^4, \delta t^2)$ for the same computational effort as the Crank–Nicolson scheme. It might be expected that to get a higher order of accuracy than Crank–Nicolson would require the use of more points in the S direction, but this is not so.

I will describe the method using the basic diffusion equation (see the reading list at the end of this chapter for details of how to extend the method to convection-diffusion equations and equations with non-constant coefficients).

The basic diffusion equation is

$$\frac{\partial V}{\partial t} + \frac{\partial^2 V}{\partial S^2} = 0.$$

This is not quite how it is usually written, but I want to keep the equation as close as possible to what we have got used to.

The explicit method applied to this equation is just

$$\frac{V_i^{k+1} - V_i^k}{\delta t} = \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}.$$

and the fully implicit is similarly

$$\frac{V_i^{k+1} - V_i^k}{\delta t} = \frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\delta S^2}.$$

The Crank–Nicolson scheme is just an average of these two methods. But is there any advantage in taking a *weighted* average? This leads to the idea of a θ method. Take a weighted average of the explicit and implicit methods to get

$$\frac{V_i^{k+1} - V_i^k}{\delta t} = \theta\left(\frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{\delta S^2}\right) + (1-\theta)\left(\frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}\right).$$

When $\theta = \frac{1}{2}$ we are back to the Crank–Nicolson method. For a general value of θ the local truncation error is

$$O\left(\tfrac{1}{2}\delta t - \tfrac{1}{12}\delta S^2 - \theta\delta t, \delta S^4, \delta t^2\right).$$

When $\theta = 0$, $\frac{1}{2}$ or 1 we get the results we have seen so far. But if

$$\theta = \tfrac{1}{2} - \frac{\delta S^2}{12\,\delta t}$$

then the local truncation error is improved. The implementation of the method is no harder than the Crank–Nicolson scheme.

78.7 THREE TIME-LEVEL METHODS

Numerical schemes are not restricted to the use of just two time levels. We can construct many algorithms using three or more time levels. Again, we would do this if it gave us a better local truncation error or had better convergence properties. For simplicity, I shall still concentrate on the basic diffusion equation.

The obvious first method to try uses a central difference for the time derivative in an explicit scheme; after all, this is more accurate than the time derivative used in our earlier explicit scheme. With this approximation to the time derivative we get

$$\frac{V_i^{k+1} - V_i^{k-1}}{2\,\delta t} = \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{\delta S^2}.$$

Although this uses a more accurate approximation it is unstable for any time step. This serves as a warning about the possible pitfalls in the search for good difference schemes.

An explicit scheme that is stable for all time steps is

$$\frac{V_i^{k+1} - V_i^{k-1}}{2\,\delta t} = \frac{V_{i+1}^k - V_i^{k+1} - V_i^{k-1} + V_{i-1}^k}{\delta S^2}$$

leading to

$$(1+2\nu_1)\,V_i^{k+1} = 2\nu_1\left(V_{i+1}^k + V_{i-1}^k\right) + (1-2\nu_1)V_i^{k-1}.$$

To get this, or any three time-level method started requires an initial condition and data at the first time level. The latter must be found by some two time-level method of the same accuracy as the three-level method that is to be used.

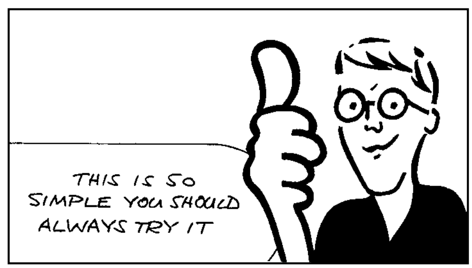
An implicit method with the same order of accuracy as the Crank–Nicolson method, but better for discontinuous initial data, is

$$-2\nu_1\left(V_{i+1}^{k+1} + V_{i-1}^{k+1}\right) + (3+2\nu_1)V_i^{k+1} = 4V_i^k - V_i^{k-1}.$$

And, of course, it is possible to improve the accuracy further in a three-level Douglas method:

$$\begin{aligned}(3-24\nu_1)V_{i+1}^{k+1} &+ (30+48\nu_1)V_i^{k+1} + (3-24\nu_1)V_{i-1}^{k+1}\\ &= 4V_{i-1}^k + 40V_i^k + 4V_{i+1}^k - V_{i+1}^{k-1} - 10V_i^{k-1} - V_{i-1}^{k-1}.\end{aligned}$$

78.8 RICHARDSON EXTRAPOLATION



In the explicit method the error is $O(\delta t, \delta S^2)$. If we assume that the approach to the correct solution as the time step and asset step tend to zero is in a sense 'regular' then we could postulate that[2]

$$\begin{aligned}\text{approximate solution} = \text{exact solution} &+ \epsilon_1\,\delta t + \epsilon_2\,\delta S^2\\ &+ \epsilon_3\,\delta t^2 + \cdots\end{aligned}$$

for some coefficients ϵ_i. Suppose that we have *two* approximate solution V_1 and V_2 using different grid sizes; we can write

$$\begin{aligned}V_1 &= \text{exact solution} + \epsilon_1\,\delta t_1 + \epsilon_2\,\delta S_1^2 + \epsilon_3\,\delta t_1^2 + \cdots\\ &= \text{exact solution} + \delta S_1^2\left(\epsilon_1\frac{\delta t_1}{\delta S_1^2} + \epsilon_2\right) + \cdots\end{aligned}$$

and

$$\begin{aligned}V_2 &= \text{exact solution} + \epsilon_1\,\delta t_2 + \epsilon_2\,\delta S_2^2 + \epsilon_3\,\delta t_2^2 + \cdots\\ &= \text{exact solution} + \delta S_2^2\left(\epsilon_1\frac{\delta t_2}{\delta S_2^2} + \epsilon_2\right) + \cdots.\end{aligned}$$

If we choose

$$\frac{\delta t_1}{\delta S_1^2} = \frac{\delta t_2}{\delta S_2^2},$$

[2] This is like a Taylor series. Think of $V(S, t; \delta S, \delta t)$ as the solution of the finite-difference equations, so that $V(S, t; 0, 0)$ is the exact solution. Now expand for small δt and δS.

i.e. ν_1 constant, where the subscripts denote the step sizes used in finding the solutions V_1 and V_2, then we can find a *better* solution than both V_1 and V_2 by eliminating the leading-order error terms in the above two equations. This better approximation is given by

$$\frac{\delta S_2^2\, V_1 - \delta S_1^2\, V_2}{\delta S_2^2 - \delta S_1^2}.$$

The accuracy of the method is now $O(\delta t^2, \delta S^3)$.[3]

Now refer back to Figures 78.3 and 78.4. In those figures you can see just how well the error is approximated by the time step and the square of the asset step size. Richardson extrapolation would work well in this case.

To give an example, use the Visual Basic function given in the previous chapter for the explicit scheme applied to the call option. For a strike of 100, maturity of one year, underlying asset at 100 with a volatility of 20% and an interest rate of 10%, the exact Black–Scholes value is 13.269. Using the program and 20 points gives a result of 13.067. Because the number of calculations is proportional to $\delta t^{-1}\delta S^{-1}$ the time taken would be 8,000 in some units. With 30 points the value improves becoming 13.183, but the time taken is 27,000. Using Richardson extrapolation on these two values gives value of 13.275. The total time taken is $8{,}000 + 27{,}000 = 35{,}000$. But the error is just 0.006. To get such an accuracy using the explicit method directly would require 100 asset steps, a time of 1,000,000, and hence 30 times as long.

Method	BS	Numerical	Error	Time taken
20 asset steps	13.269	13.067	0.202	8,000
30 asset steps	13.269	13.183	0.086	27,000
Richardson	13.269	13.275	0.006	35,000
100 asset steps	13.269	13.276	0.007	1,000,000

The same extrapolation principal can be applied to any of the methods described in this chapter. The method should be used with care. There is no guarantee that we can write the solution in a meaningful Taylor series expansion so the method could make matters worse. Also, the extrapolation method can increase rounding errors.

78.9 FREE BOUNDARY PROBLEMS AND AMERICAN OPTIONS

There is a great deal of theory behind what I am going to show you in the next couple of sections. This theory is obviously important, since it makes what I am going to do 'valid.' But the theory obscures the simplicity of the final result, so I am going to cut straight to the chase.

The value of American options must always be greater than the payoff, otherwise there will be an arbitrage opportunity

$$V(S, t) \geq \text{Payoff}(S).$$

[3] The accuracy with respect to the asset step may be even better than this because of symmetry in the central differences for the derivatives.

The payoff function may also be time-dependent. For example, if the option is Bermudan, i.e. exercise is only allowed on certain dates, then the payoff function is zero except on the special dates, when it is some prescribed function of the underlying. So I am going to write Payoff(S, t) and I need never mention Bermudan options again.

American options are examples of 'free boundary problems': We must solve a partial differential equation with an unknown boundary, the position of which is determined by having one more boundary condition than if the boundary were prescribed. In the American option problem we know that both the option value and its delta are continuous with the payoff function, the so called **smooth pasting condition**. Solving a problem where the boundary (or boundaries) is/are unknown is surely very complex. Wrong. If we approach the problem in the right way, the implementation could not be simpler.

The first point to note is that I am going to solve over a *fixed* range of S that I am sure contains the free boundary, and I am not going to try to find the boundary by any direct method.

78.9.1 Early Exercise and the Explicit Method

Let me tell you how to incorporate the early-exercise constraint into the explicit finite-difference scheme and then I will explain why it is clever.

Suppose that we have found V_i^k for all i at the time step k (e.g. at the expiry date); time step to find the option value at $k+1$ using the finite-difference scheme

$$V_i^{k+1} = A_i^k V_{i-1}^k + (1 + B_i^k) V_i^k + C_i^k V_{i+1}^k.$$

Don't worry about whether or not you have violated the American option constraint until you have found the option values V_i^{k+1} for all i. Now let's check whether the new option values are greater or less than the payoff. If they are less than the payoff then we have arbitrage. We can't allow that to happen so at every value of i for which the option value has allowed arbitrage, replace that value by the payoff at that asset value. That's all there is to it.

As far as the code is concerned, just put the following line of code instead of the line that updates the `VOld(i)`.

```
VOld(i) = max(VNew(i), Payoff(S(i), RealTime))
```

It is clear that this finite-difference solution will converge to something that has a value continuous with the payoff. What is not so clear is that the gradient constraint of continuous delta is also satisfied. But it is. The explicit method is just a fancy form of the binomial method and this simple replacement with the payoff is exactly what is done in the binomial method. The accuracy of the solution is still $O(\delta t, \delta S^2)$.

The early exercise boundary is somewhere between the asset values where to one side the option value is the payoff and to the other side the value is strictly greater than the payoff. It is perfectly possible for there to be more than one such boundary, and the method will find all of them with no extra effort.

78.9.2 Early Exercise and Crank–Nicolson

Implementing the American constraint in the Crank–Nicolson method is a bit harder than in the explicit method, but the rewards come in the accuracy. The method is still accurate to $O(\delta t^2, \delta S^2)$ with no limit on the time step.

The only complication arises because the Crank–Nicolson method is implicit, and every value of the option at the $k+1$ time step is linked to every other value at that time step. It is therefore not good enough to just replace the option value with the payoff after the values have all been calculated,[4] the replacement must be done at the same time as the values are found:

$$v_i^{n+1} = \max\left(v_i^n + \frac{\omega}{M_{ii}}\left(q_i - \sum_{j=1}^{i-1} M_{ij} v_j^{n+1} - \sum_{j=i}^{N} M_{ij} v_j^n\right), \text{Payoff}\right).$$

The payoff is evaluated at i and $k+1$ in the obvious manner. This method is called **projected SOR**.

Below is a Visual Basic code fragment that is identical to the SOR code above with the exception of one line of code that checks whether the option value is greater than the `Payoff`. If it is not, then the payoff is used as the value.

```
Dim q(1 To N) As Double
Dim v(1 To N) As Double
Dim temp(1 To N) As Double
Dim MDiag(1 To N) As Double
Dim MSuperDiag(1 To N + 1) As Double
Dim MSubDiag(0 To N) As Double

MSuperDiag(N + 1) = 0
MSubDiag(0) = 0
NoIts = 0
While Error < tol
    Error = 0
    For i = 1 To N
        temp(i) = v(i) + omega * (q(i) - MSuperdiag(i + 1) * v(i + 1) _
                  - MDiag(i) * v(i) - MSubDiag(i - 1) * v(i - 1)) / MDiag(i)
        temp(i) = max(temp(i), Payoff(S(i)))
        Error = Error + (temp(i) - v(i)) * (temp(i) - v(i))
        v(i) = temp(i) ' use as soon as calculated
    Next i
    NoIts = NoIts + 1
Wend
```

78.10 JUMP CONDITIONS

As well as having final conditions (the payoff), boundary conditions (at zero, infinity or at a barrier) and the partial differential equation to satisfy we often have jump conditions. These can be due to a jump in the underlying on a dividend date, the payment of some coupon, or because of the jump in a discretely sampled path-dependent quantity. One of these is very simple and the other two are basically the same as each other in their implementation.

78.10.1 A Discrete Cashflow

If our contract entitles us to a discretely paid sum on a specified date (a coupon on a bond for example), then the contract value must jump by the amount of the cashflow. In continuous time

[4] This can be done but the accuracy is then reduced to $O(\delta t)$.

we have

$$V(S, t_d^-) = V(S, t_d^+) + C. \tag{78.8}$$

The cashflow C may be a function of the underlying, or even of the option value, and as always you can read r for S if you are valuing an interest rate product.

From a numerical point of view we would ideally like the date t_d on which there is a cashflow to lie exactly on a grid point, i.e., we would like the date to be exactly an integer number of time steps before expiry. If that is the case then the implementation of (78.8) is straightforward. Simply calculate the option value up to and including the date t_d using whichever numerical method you prefer, then before moving on to the time step after[5] the payment just add C to every option value.

Complications arise if the date does not coincide with grid points. The simplest implementation is to just add C at the first time step after the date. This is accurate to $O(\delta t)$ and so is consistent for the explicit method. If you are using a more accurate numerical method such as Crank–Nicolson you should use a more accurate implementation of (78.8). If you are solving the basic Black–Scholes equation and the cashflow is a constant, independent of the underlying or option value, then it is simple to just present value C to make an equivalent cashflow that coincides with a time step. If the cashflow is proportional to the underlying then (in the absence of dividends) it doesn't matter if you apply the jump condition at the nearest time step since the underlying itself satisfies the Black–Scholes equation. There are many simple examples like these where there is an obvious, sometimes tedious, way of accurately incorporating the cashflow. But if the cashflow is more complicated or you are solving for an interest rate product then the best thing to do is to make the time step just after the cashflow (in real time) of the correct size to match up the grid with the cashflow date (see Figure 78.5). That way the cashflow is guaranteed to be incorporated at the same order of accuracy as the numerical method itself.

78.10.2 Discretely paid Dividend

More interesting jump conditions arise when there is a finite jump in one of the independent variables. If there is a jump in the underlying asset we must implement a more complicated condition. I am going to restrict my explanation of the implementation to the one case; all others are dealt with in the same fashion.

Suppose that a dividend of $D(S)$ is paid on date t_d. Across this dividend date the asset value drops by an amount $D(S)$ (this is no arbitrage) but the option value does not change. This is explained fully in Chapter 8. Mathematically we write

$$V(S, t_d^-) = V(S - D(S), t_d^+). \tag{78.9}$$

When we are solving numerically we will find the right-hand side of this equation having time stepped from expiry. Before continuing with the time stepping backwards in time we want to implement (78.9). I will assume that either t_d lies on a time step or a time step has been adjusted to ensure this. This means that in implementing (78.9) we are only concerned with accuracy due to the finiteness of the asset step size.

[5] In 'real' time *before*.

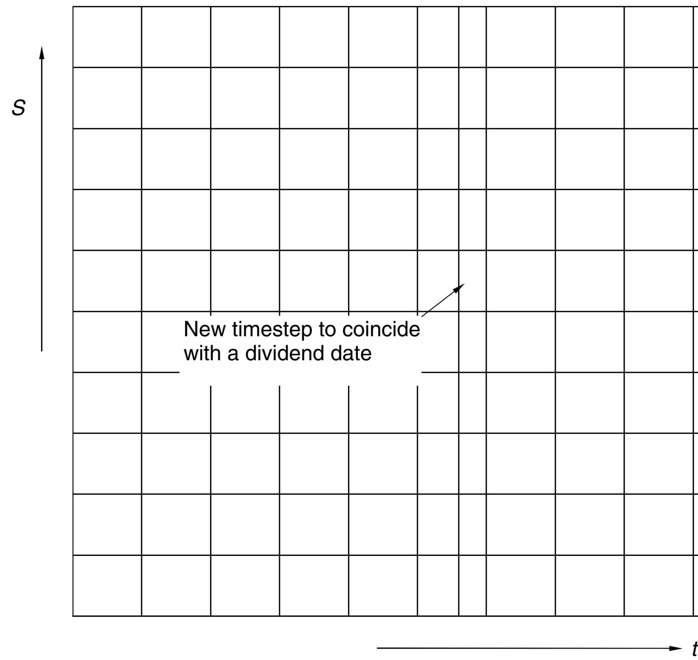


Figure 78.5 The grid near a discrete cashflow or a dividend date.

I will show how to use linear interpolation to implement the jump condition. This method is accurate to $O(\delta S^2)$ and is therefore of the same accuracy as the other methods described above.

From the finite-difference scheme we know V_i^{k+1} for all i, this is the option value just after the dividend date, that is, at time t_d^+. To implement the jump condition we must first find the grid points between which lies the point $S - D$. Then we interpolate between these two grid points to find an accurate value for the option value before the dividend is paid.

Here is Visual Basic code fragment that implements the jump condition. The dividend payment is `Div(S(i))`, a function of the underlying. The option values just after the dividend date are stored in `VOld(i)` and the values just before are stored in `VNew(i)`. `dummy` is the ratio of the distance between the point `S(i) - Div(S(i))` and the nearest grid point to the step size `AssetStep`.

```
For i = 0 To NoAssetSteps
    inew = Int((S(i) - Div(S(i))) / AssetStep)
    dummy = (S(i) - Div(S(i)) - inew * AssetStep) / AssetStep
    VNew(i) = (1 - dummy) * VOld(inew) + dummy * VOld(inew + 1)
Next i
```

There is a slight complication. I have assumed that $0 \le$ `inew` $\le$ `NoAssetSteps`. If this is not the case then `S(i) - Div(S(i))` lies outside the original grid. To deal with this it is usual to make some assumption about the behavior of V beyond the grid; usually that it is constant or linear in the underlying depending on the problem. The modification is simple to incorporate.

78.11 PATH-DEPENDENT OPTIONS

Many of the path-dependent contracts that I described in Part Two have simple partial differential equation representations. Often the problem is in three dimensions, the usual underlying (asset, interest rate etc.), time and the path-dependent quantity. These three-dimensional problems for exotic options come in two forms: In one there are no new terms in the equation because the path-dependent quantity is measured discretely, in the other there is an extra first derivative with respect to the new continuously sampled variable.[6]

Clearly the most important aspect of the path-dependent problem is the extra dimension. We must solve in three dimensions and must therefore introduce the option value as

$$V_{i,j}^{k}.$$

The superscript k still refers to time, the subscript i to the asset but now j refers to the new variable.

There is sufficient information above for the reader to be able to work out ways in which to solve these new problems. I will simply point the way and give a couple of concrete examples.

78.11.1 Discretely Sampled Quantities

When the path-dependent quantity is sampled discretely there are no new terms in the partial differential equation and thus no new terms in any difference equation approximation. This makes the finite-difference solution not too difficult, and any of the methods described above can be used. Remember, though, that you will have a three-dimensional grid and that you must solve for each value of j, the grid value for the path-dependent quantity. In this way you time step backwards until you reach one of the sampling dates.

A jump condition must be applied across the sampling date and the implementation is exactly as for the discretely paid dividend. Suppose that the jump condition is

$$V(S, I, t_d^-) = V(S, F(S, I), t_d^+).$$

This is again implemented by linear interpolation and is accurate to $O(\delta I^2)$ where δI is the step size for the path-dependent quantity I.

In the following Visual Basic code fragment the path-dependent quantity is called `P` and its step size is `PDQStep`. If `jnew` lies outside the range zero to `NoPDQSteps` then some extrapolation will be required. Note that I have interpolated the jump condition between grid points.

```
For j = 0 To NoPDQSteps
    For i = 0 To NoAssetSteps
        jnew = Int(F(S(i), P(j)) / PDQStep)
        dummy = (F(S(i), P(j)) - jnew * PDQStep) / PDQStep
        VNew(i, j) = (1 - dummy) * VOld(i, jnew) + dummy * VOld(i, jnew + 1)
    Next i
Next j
```

[6] Lookbacks are special in that even if the maximum is measured continuously there is no new term.

78.11.2 Continuously Sampled Quantities

When the path-dependent quantity is sampled continuously there is usually a new term in the partial differential equation and thus new terms in a difference equation approximation:

$$\frac{\partial V}{\partial t} + a(S, I, t)\frac{\partial^2 V}{\partial S^2} + b(S, I, t)\frac{\partial V}{\partial S} + f(S, I, t)\frac{\partial V}{\partial I} + c(S, I, t)V = 0.$$

(In practice the coefficients rarely depend on I.)

The implementation of anything but the explicit finite-difference method is more complicated so I shall concentrate on that here.

The updating of the scheme, the implementation of the boundary and final conditions are all very straightforward. The only point that I need comment on is the choice of the discrete version of the derivative

$$\frac{\partial V}{\partial I}.$$

Because there is no diffusion in the I direction (there is no second derivative of the option value with respect to I) the choice of difference is important. The derivative with respect to I represents convection in the I direction and the numerical scheme must be consistent with this. For this reason a one-sided difference must be used. If the coefficient $f(S, I, t)$ changes sign then the choice of difference must reflect this; upwind differencing *must* be used. Here is a possible choice for the difference:

$$\text{if } f(S, I, t) \geq 0 \quad \text{then} \quad f(S, I, t)\frac{\partial V}{\partial I}(S, I, t) = f^k_{i,j+\frac{1}{2}}\frac{V^k_{i,j+1} - V^k_{i,j}}{\delta I}$$

but if

$$\text{if } f(S, I, t) < 0 \quad \text{then} \quad f(S, I, t)\frac{\partial V}{\partial I}(S, I, t) = f^k_{i,j-\frac{1}{2}}\frac{V^k_{i,j} - V^k_{i,j-1}}{\delta I}.$$

Alternatively, one of the three-point one-sided differences could be used for better accuracy.

One final point about the numerical solution of three-dimensional exotic option problems. If there is a similarity reduction taking the problem down from three to two dimensions then it should be exploited. Not only will the solution be much faster, but the choice of scheme will be greater and the implementation easier.

78.12 SUMMARY

If you have worked your way successfully through the methods in this chapter then you have reached a level of sophistication that is far higher than the binomial method ... but there is still higher to go.

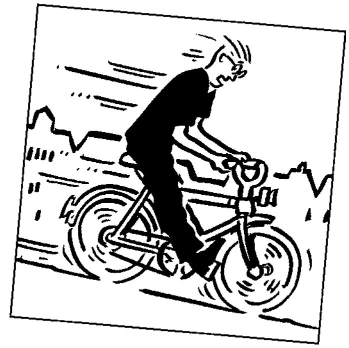
FURTHER READING

- Excellent books on the numerical solution of partial differential equations are Morton & Mayers (1994) and Smith (1985). Richtmyer & Morton (1976) and Mitchell & Griffiths

(1980) are very good on partial differential equations with non-constant coefficients, which is what we usually have in finance.

- Roache (1982) is a very interesting description of numerical methods as used in Computational Fluid Dynamics.
- See the original paper by Crank & Nicolson (1947).
- See the critique of the Crank–Nicolson scheme by Duffy (2004).
- See also Ahmad (2006) for further details of the above schemes.

CHAPTER 79
finite-difference methods for two-factor models



In this Chapter...

- the explicit finite-difference method for two-factor models
- the ADI and Hopscotch methods

79.1 INTRODUCTION

Many currently popular financial models have two random factors. Convertible bonds are usually priced with both random underlying and random risk of default or interest rate. Exotic equity derivatives are often priced with stochastic volatility, this is especially true of barrier options. Finite-difference methods are quite suited to such problems.

Once you get to three factors, or three sources of randomness, then finite-difference methods start to become slow and cumbersome. When there are four random factors then, if possible, you should use a Monte Carlo approach for pricing. These are discussed in Chapter 80. Monte Carlo works well in high dimensions.

Finite-difference methods work well in low dimensions, and handle early exercise very efficiently. Some path dependency is also easy to cope with.

In this chapter I will describe the simplest forms of finite-difference methods used with two or more factors.

79.2 TWO-FACTOR MODELS

I am going to refer throughout this chapter to the two-factor equation

$$\frac{\partial V}{\partial t} + a(S,r,t)\frac{\partial^2 V}{\partial S^2} + b(S,r,t)\frac{\partial V}{\partial S} + c(S,r,t)V$$
$$+ d(S,r,t)\frac{\partial^2 V}{\partial r^2} + e(S,r,t)\frac{\partial^2 V}{\partial S\,\partial r} + f(S,r,t)\frac{\partial V}{\partial r} = 0. \qquad (79.1)$$

As I wrote in the previous chapter, everything will be perfectly applicable to any problem, whatever the underlying variables. I write the problem to be solved as (79.1) assuming that my readers prefer the concrete 'asset' S and 'interest rate' r to the more abstract x and y. Actually, it will be quite helpful if we think of solving the two-factor convertible bond problem of Chapter 33. This is because that problem contains many important features such as choice of interest rate model and how it is discretized, and early exercise. For the general two-factor problem (79.1) to be parabolic we need

$$e(S, r, t)^2 < 4a(S, r, t)\, d(S, r, t).$$

As in the one-factor world, the variables must be discretized. That is, we solve on a three-dimensional grid with

$$S = i\ \delta S, \quad r = j\ \delta r, \quad \text{and} \quad t = T - k\ \delta t.$$

Expiry is $t = T$ or $k = 0$. The indices range from zero to I and J for i and j respectively. I have assumed that the interest rate model is only specified on $r \geq 0$. This may not be the case, as some simple interest rate models such as Vasicek are defined over negative r as well. If we have such a model we would redefine r. All the comments that I made in the last chapters about interest rate models and boundary conditions apply in the two-factor world.

The contract value is written as

$$V(S, r, t) = V_{ij}^k.$$

Whatever the problem to be solved, we must impose certain conditions on the solution. First of all, we must specify the final condition. This is the payoff function, telling us the value of the contract at the expiration of the contract.

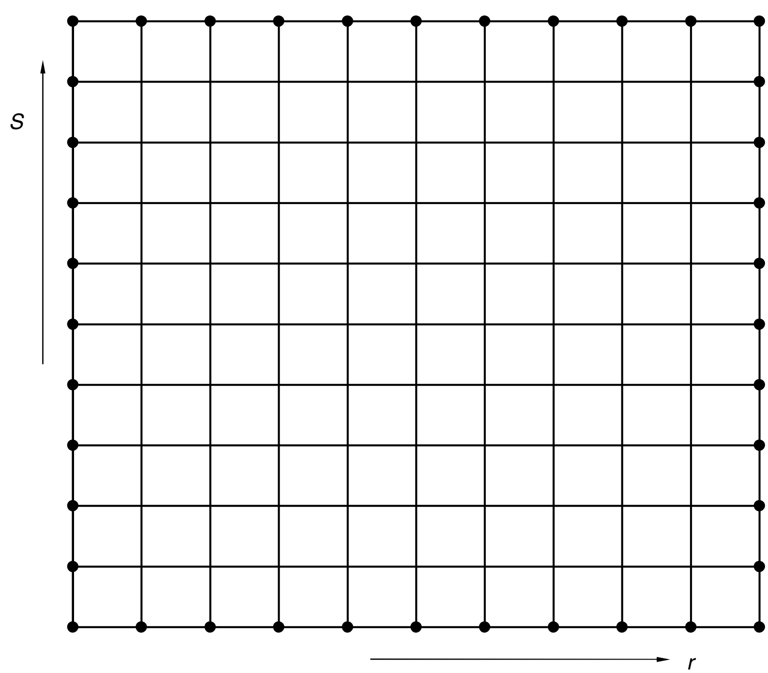


Figure 79.1 The S, r domain and boundary conditions.

Suppose that we are pricing a long-dated warrant with a call payoff. This is a good candidate for a two-factor model; we may want to incorporate a stochastic interest rate since we do not think that the assumption of constant interest rates is a good one over such a long period. The final condition for this problem is then

$$V(S, r, T) = V_{ij}^0 = \max(S - E, 0).$$

As well as a final condition, we must impose boundary conditions around our domain. The S, r domain is shown in Figure 79.1. Boundary conditions must be imposed at all the grid points marked with a dot. The boundary conditions will depend on the contract. Remember that there is also a time axis coming out of the page, and not drawn in this figure.

79.3 THE EXPLICIT METHOD

The one-factor explicit method can be extended to two-factors with very little effort. In fact, the ease of programming make it a very good method for those new to the subject. I am going to use symmetric central differences for all derivatives in (79.1). This is the best way to approximate the second derivatives but may not be the best for the first derivatives; I will discuss this later.

We have seen how to use central differences for all of the terms with the exception of the second derivative with respect to both S and r,

$$\frac{\partial^2 V}{\partial S\, \partial r}.$$

I can approximate this by

$$\frac{\partial \left(\dfrac{\partial V}{\partial r}\right)}{\partial S} \approx \frac{\dfrac{\partial V}{\partial r}(S + \delta S, r, t) - \dfrac{\partial V}{\partial r}(S - \delta S, r, t)}{2\,\delta S}.$$

But

$$\frac{\partial V}{\partial r}(S + \delta S, r, t) \approx \frac{V_{i+1,j+1}^k - V_{i+1,j-1}^k}{2\,\delta r}.$$

This suggests that a suitable discretization might be

$$\frac{\dfrac{V_{i+1,j+1}^k - V_{i+1,j-1}^k}{2\,\delta r} - \dfrac{V_{i-1,j+1}^k - V_{i-1,j-1}^k}{2\,\delta r}}{2\,\delta S}$$
$$= \frac{V_{i+1,j+1}^k - V_{i+1,j-1}^k - V_{i-1,j+1}^k + V_{i-1,j-1}^k}{4\,\delta S\,\delta r}.$$

This is particularly good since not only is the error of the same order as in the other derivative approximations but also it preserves the property that

$$\frac{\partial^2 V}{\partial S\, \partial r} = \frac{\partial^2 V}{\partial r\, \partial S}.$$

The resulting explicit difference scheme is

$$\frac{V_{ij}^k - V_{ij}^{k+1}}{\delta t} + a_{ij}^k \left(\frac{V_{i+1,j}^k - 2V_{ij}^k + V_{i-1,j}^k}{\delta S^2} \right) + b_{ij}^k \left(\frac{V_{i+1,j}^k - V_{i-1,j}^k}{2\,\delta S} \right) + c_{ij}^k V_{ij}^k$$

$$+ d_{ij}^k \left(\frac{V_{i,j+1}^k - 2V_{ij}^k + V_{i,j-1}^k}{\delta r^2} \right) + e_{ij}^k \left(\frac{V_{i+1,j+1}^k - V_{i+1,j-1}^k - V_{i-1,j+1}^k + V_{i-1,j-1}^k}{4\,\delta S\,\delta r} \right)$$

$$+ f_{ij}^k \left(\frac{V_{i,j+1}^k - V_{i,j-1}^k}{2\,\delta r} \right) = O(\delta t, \delta S^2, \delta r^2).$$

I could rewrite this in the form

$$V_{ij}^{k+1} = \cdots,$$

where the right-hand side is a linear function of the nine option values shown schematically in Figure 79.2. The coefficients of these nine values at time step k are related to a, b etc. It would not be very helpful to write the difference equation in this form, since the actual implementation is usually more transparent than this. Note that in general all nine points (ij), $(i \pm 1, j)$, $(i, j \pm 1)$, $(i \pm 1, j \pm 1)$ are used in the scheme. If there is no cross derivative term then only the five points (ij), $(i \pm 1, j \pm 1)$ are used. This simplifies some of the methods.

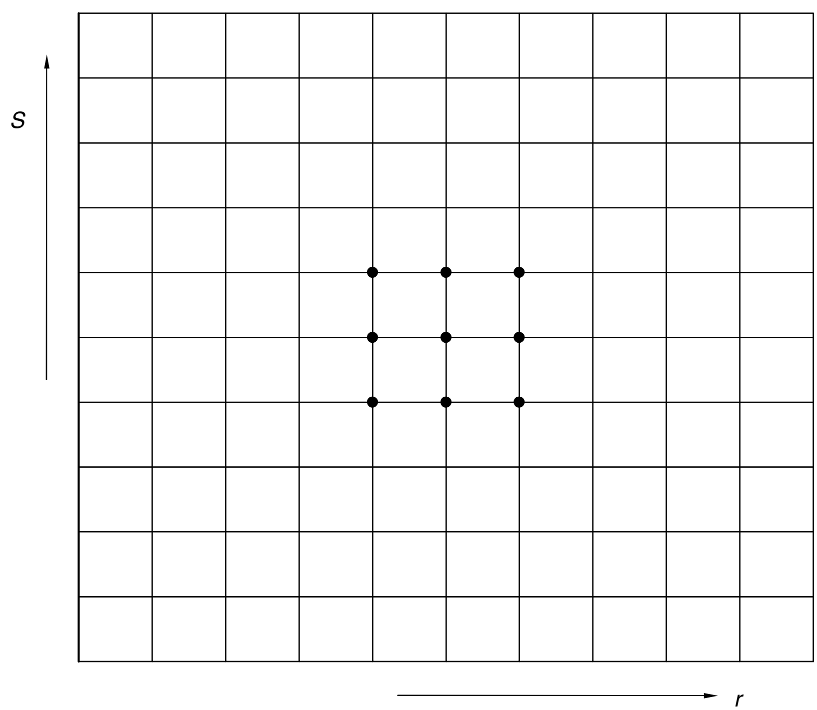


Figure 79.2 The nine option values used to time step in the explicit method.

I will give an example now. The following is a Visual Basic code fragment for the pricing of a convertible bond. The underlying asset is `S(i)` and the interest rate is `r(j)`.

```
For j = 1 To MR - 1
For i = 0 To MS - 1
    VS = (VOld(i + 1, j) - VOld(i - 1, j)) / (2 * dS)
    Vrp = (VOld(i, j + 1) - VOld(i, j)) / dr
    Vrm = (VOld(i, j) - VOld(i, j - 1)) / dr
    Vr = Vrm
    If RDdrift(r(j), RealTime) > 0 Then Vr = Vrp
    VSS = (VOld(i + 1, j) - 2 * VOld(i, j) + VOld(i - 1, j)) / (dS * dS)
    Vrr = (VOld(i, j + 1) - 2 * VOld(i, j) + VOld(i, j - 1)) / (dr * dr)
    VSr = (VOld(i + 1, j + 1) - VOld(i - 1, j + 1) - _
                VOld(i + 1, j - 1) + VOld(i - 1, j - 1)) / (4 * dS * dr)
    VNew(i, j) = VOld(i, j) + dt * _
                (0.5 * S(i) * S(i) * vol * vol * VSS _
                + 0.5 * ratevol * ratevol * Vrr _
                + rho * vol * ratevol * S(i) * VSr _
                    + r(j) * (S(i) * VS - VOld(i, j)) _
                    - div * S(i) * VS _
                    + RNDrift(r(j), RealTime) * Vr)
Next i
Next j
```

Notice how the derivatives of the option value `VOld(i, j)` are defined in turn. I have used a central difference for every derivative except the delta with respect to the interest rate. This is defined as a one-sided difference, using upwind differencing. I have used an inaccurate (error $O(\delta r)$) one-sided difference but this could be improved upon as suggested in Chapter 77. The reason for the use of upwind differencing for the interest rate first derivative is that the volatility of interest rates is usually quite small. The diffusion which makes the numerical solution easy is not very effective for the interest rate, making the problem almost hyperbolic rather than parabolic.

The volatility of the underlying is `vol` and the volatility of the interest rate is `ratevol`. The correlation is `rho`. The volatilities and correlation are constant.

The interest rate model that I have used is

$$dr = \texttt{RNDrift}(r, t)\, dt + \texttt{ratevol}\, dX.$$

I have allowed the risk-neutral drift `RNDrift` to be a function of both the interest rate and time, `RealTime`. the time dependence means that this code fragment could be used for a yield curve-fitted model.

The only other points I want to mention about this code fragment concern the boundary conditions. The interest rate index `j` ranges from `0` to `MR`. A boundary condition must be applied on `j = 0` and `j = MR` for all `i`. The same applies to the asset, with index `i` ranging from `0` to `MS`. In this code fragment you will notice that I have applied the time stepping algorithm at `i = 0`, i.e. at some boundary points. This is rather unusual. I can do this in some special cases because, as long as I allow `VOld(i, j)` to range from `i = -1`, the application of the equation at `i = 0` i.e. at $S = 0$ is still valid since the diffusion switches off at $S = 0$. This is another way of saying that the point $S = 0$ is singular.

After applying the relevant boundary conditions and before moving on to the next time step, the convertibility feature must be incorporated. This is easily done in the explicit method. Here is the code fragment that ensures that there are no arbitrage opportunities.

```
For i = 0 To MS
For j = 0 To MR
    VOld(i, j) = max(VNew(i, j), ConvertRate * S(i))
Next j
Next i
```

This code simultaneously updates the old values to the new values `VNew(i, j)` and makes the bond values greater than the payoff from conversion `convertrate * S(i)`. Often conversion is only allowed on or between certain dates. In that case, the constraint is only applied at the time steps on which conversion is permitted.

79.3.1 Stability of the Explicit Method

One of the advantages of the explicit method is again that it is easy to program. The main disadvantage comes in stability and speed. The method is only stable for sufficiently small time steps. Having an upper bound on the time step size seriously limits the speed of calculation.

We can again analyze the stability of the method by looking for solutions of the difference equation that are oscillatory in both the S and r directions. This means looking for a solution of the form

$$V_{ij}^k = \alpha^k e^{2\pi\sqrt{-1}(i/\lambda_S + j/\lambda_r)}$$

and assuming that all the coefficients A, b, c etc. are slowly varying over the δS, δr lengthscales. I will skip the details, and simply state the result in one special and important case. If we have a pure diffusion problem with no convection or decay terms, and no correlation between the variables, then only a and d are non-zero. In this case, the stability requirement becomes

$$a\frac{\delta t}{\delta S^2} + d\frac{\delta t}{\delta r^2} \le \frac{1}{2}.$$

This is an even more severe constraint than in the explicit method applied to one-factor problems.

To overcome this time step constraint we can try an implicit scheme, as we shall shortly see.

79.4 CALCULATION TIME

How long will it take to crunch a partial differential equation to price an option?

Let's look at the explicit finite-difference method applied to an option pricing problem. Since the explicit finite-difference method isn't just restricted to one or two dimensions let's look at solving in d dimensions. And suppose we want an accuracy of ϵ. How long will this take?

The error in the explicit method is $O(\delta t)$ and $O(\delta S^2)$. So we'll want to have both of these of order ϵ, our required accuracy.[1] So

$$\text{time step } = O(\epsilon)$$

and

$$\text{asset step } = O(\epsilon^{1/2}).$$

[1] Also must be consistent with stability.

The time taken will be proportional to the number of mesh points, that is order

$$\frac{1}{\delta t} \times \frac{1}{\delta S} \times \cdots \times \frac{1}{\delta S}$$

where there are d of the asset step terms. Thus the time taken will be

$$O\left(\epsilon^{-1}\,\epsilon^{-d/2}\right) = O\left(\epsilon^{-1-d/2}\right).$$

It won't take any longer than this to compute the greeks since we already calculate the option value at all mesh points.[2] But if we want to price many different options, M, say, it will take

$$O\left(M\epsilon^{-1-d/2}\right).$$

Early exercise won't add much to the computation time at all.

79.5 ALTERNATING DIRECTION IMPLICIT

There are many implicit methods used for one-factor problems. There are even more implicit methods for two-factor problems. Good numerical methods will be fast and require as little storage as possible. I am only going to describe two of the implicit finite-difference methods used for two factors; they are both intuitively appealing and use the same ideas that we saw in the previous chapter. The first method is called **Alternating Direction Implicit** or **ADI**.

We could try a two-factor extension of Crank–Nicolson to solve $(I-1)(J-1)$ equations in the same number of unknowns. This suffers from the problem that the resulting matrix does not have a nice form, and the solution is complicated and time consuming. If we want to keep the stability advantage of the implicit method and the ease of solution of the explicit method we could try to solve implicitly in one factor but explicitly in the other. This is the idea behind ADI. I will first explain the idea in words.

As well as V_{ij}^k, introduce an 'intermediate' value $V_{ij}^{k+(1/2)}$. Solve from time step k to the intermediate step $k+\frac{1}{2}$ using explicit differences in S and implicit differences in r. Since only one direction is implicit, the solution by LU decomposition or SOR is no harder than in one factor. Having found the intermediate value $V_{ij}^{k+(1/2)}$ step forward to time step $k+1$ using implicit differences in S and explicit differences in r. For this half time step I have changed around explicit and implicit from the previous half time step. Again the matrix equations are straightforward. The method is stable for all time steps and the error is $O(\delta t^2, \delta S^2, \delta r^2)$.

Now let us see how this is done in practice. I will demonstrate the idea with the simpler equation

$$\frac{\partial V}{\partial t} + a(S,r,t)\frac{\partial^2 V}{\partial S^2} + d(S,r,t)\frac{\partial^2 V}{\partial r^2} = 0. \tag{79.2}$$

It is no harder to use the ADI method on the full equation, with first derivatives etc., provided that the differences can be decomposed into either S or r derivatives. The cross derivative term causes some problems in the basic implementation of ADI. This can be got around but is tedious.

[2] Unlike Monte Carlo calculations, we get the basic greeks for free. The greeks involving sensitivity to parameters as opposed to variables, such as vega, *will* require further calculations to be done. Calculating the vega will double the computation time.

The explicit S, implicit r discretization looks like this

$$\frac{V_{ij}^{k} - V_{ij}^{k+(1/2)}}{\frac{1}{2}\delta t} + a_{ij}^{k}\left(\frac{V_{i+1,j}^{k} - 2V_{ij}^{k} + V_{i-1,j}^{k}}{\delta S^2}\right)$$
$$+ d_{ij}^{k+(1/2)}\left(\frac{V_{i,j+1}^{k+(1/2)} - 2V_{ij}^{k+(1/2)} + V_{i,j-1}^{k+(1/2)}}{\delta r^2}\right) = 0.$$

Putting all of the $k + \frac{1}{2}$ time step terms on the left and all k time step terms on the right I get

$$V_{ij}^{k+(1/2)} - \tfrac{1}{2}d_{ij}^{k+(1/2)}\left(\frac{V_{i,j+1}^{k+(1/2)} - 2V_{ij}^{k+(1/2)} + V_{i,j-1}^{k+(1/2)}}{\delta r^2}\right)\delta t$$
$$= V_{ij}^{k} + \tfrac{1}{2}a_{ij}^{k}\left(\frac{V_{i+1,j}^{k} - 2V_{ij}^{k} + V_{i-1,j}^{k}}{\delta S^2}\right)\delta t.$$

If we know all of the k time step terms we can find the $V_{ij}^{k+(1/2)}$ by solving a set of simultaneous equations. *This is the same as the fully implicit scheme mentioned in Chapter 78.* The only difference in this two-dimensional version is that the system must be solved for all i. Each set of equations, for each i, is of the same complexity as in one dimension, it is just that there are a lot more systems. The implicit scheme can be written as a matrix equation with a tridiagonal matrix on the left-hand side. The solution is found by LU decomposition or SOR.

Having found $V_{ij}^{k+(1/2)}$ we step forward to find V_{ij}^{k+1} by reversing the implicit and explicit roles:

$$\frac{V_{ij}^{k+1} - V_{ij}^{k+(1/2)}}{\frac{1}{2}\delta t} + a_{ij}^{k+1}\left(\frac{V_{i+1,j}^{k+1} - 2V_{ij}^{k+1} + V_{i-1,j}^{k+1}}{\delta S^2}\right)$$
$$+ d_{ij}^{k+(1/2)}\left(\frac{V_{i,j+1}^{k+(1/2)} - 2V_{ij}^{k+(1/2)} + V_{i,j-1}^{k+(1/2)}}{\delta r^2}\right) = 0.$$

This can be rewritten as

$$V_{ij}^{k+1} + \tfrac{1}{2}a_{ij}^{k+1}\left(\frac{V_{i+1,j}^{k+1} - 2V_{ij}^{k+1} + V_{i-1,j}^{k+1}}{\delta S^2}\right)\delta t$$
$$= V_{ij}^{k+(1/2)} - \tfrac{1}{2}d_{ij}^{k+(1/2)}\left(\frac{V_{i,j+1}^{k+(1/2)} - 2V_{ij}^{k+(1/2)} + V_{i,j-1}^{k+(1/2)}}{\delta r^2}\right)\delta t.$$

Again, this is a fully explicit scheme for finding V_{ij}^{k+1} from $V_{ij}^{k+(1/2)}$.

79.6 THE HOPSCOTCH METHOD

The final method I want to describe is the **Hopscotch method**. This is so called because of the way that grid points are used during the time stepping. Refer to Figure 79.3. I will again, for ease of exposition, assume that we are solving the equation (79.2).

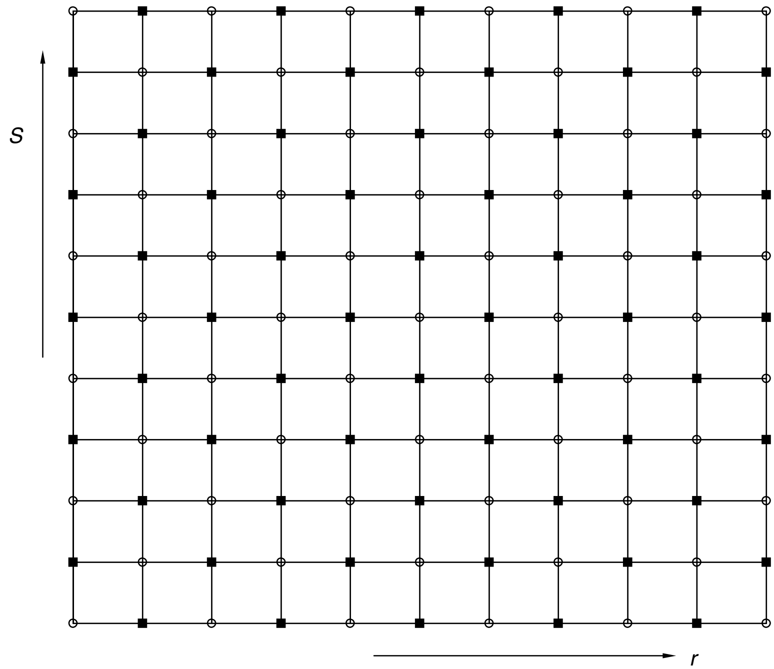


Figure 79.3 The use of grid points in the Hopscotch method.

At each time step we perform two 'sweeps.' For $k = 1$ and all future odd values, we find V_{ij}^{k+1} for all of the grid points marked with a circle in Figure 79.3. The points marked with circles are defined by $i + j$ being odd. This is done explicitly using

$$V_{ij}^{k+1} = V_{ij}^{k} + a_{ij}^{k}\left(\frac{V_{i+1,j}^{k} - 2V_{ij}^{k} + V_{i-1,j}^{k}}{\delta S^2}\right)\delta t + d_{ij}^{k}\left(\frac{V_{i,j+1}^{k} - 2V_{ij}^{k} + V_{i,j-1}^{k}}{\delta r^2}\right)\delta t.$$

Now we perform a second sweep, at the same time level, using the same scheme *but now at the grid points marked with squares, and using the just-calculated values* of V_{ij}^{k+1}:

$$V_{ij}^{k+1} = V_{ij}^{k+1} + a_{ij}^{k+1}\left(\frac{V_{i+1,j}^{k+1} - 2V_{ij}^{k+1} + V_{i-1,j}^{k+1}}{\delta S^2}\right)\delta t + d_{ij}^{k+1}\left(\frac{V_{i,j+1}^{k+1} - 2V_{ij}^{k+1} + V_{i,j-1}^{k+1}}{\delta r^2}\right)\delta t.$$

Although technically implicit, the scheme has not required the solution of any simultaneous equations. The error in the method is $O(\delta t, \delta S^2, \delta r^2)$.

When we come to the next time step, and all even numbered time steps, the roles of the grid points marked with circles and squares is reversed.

When there is a cross derivative term in the partial differential equation the explicitness of the scheme is lost.

79.7 SUMMARY

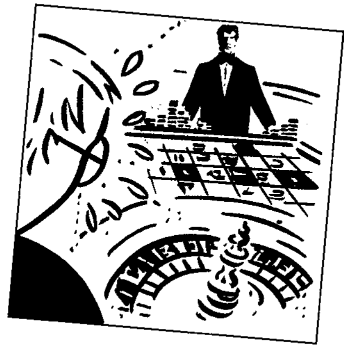
Finite-difference methods are suitable for solving financial problems with two or three random factors. Any more than that and Monte Carlo may be a better method. Finite-difference methods are particularly suited to contracts with early exercise features, or which are callable or puttable. Convertible bonds or American options with stochastic volatility (see Chapter 51) are the two obvious examples. Because simple explicit finite differences are similar to binomial trees, the methods in this chapter can do no worse than the binomial method. Some finite-difference methods can do a lot better though. The ADI and Hopscotch methods, although harder to code, are worth trying out as substitutes for the simpler but slower methods.

FURTHER READING

- See the books on the numerical solution of partial differential equations by Morton & Mayers (1994) and Mitchell & Griffiths (1980).
- Roache (1982) describes many more numerical methods with two or more factors.
- See McKee, Wall & Wilson (1996) for work on correlated ADI.
- See Ahmad (2006) for further details on methods for multiple factors.

CHAPTER 80

Monte Carlo simulation



In this Chapter...

- the relationship between option values and expectations for equities, currencies, commodities and indices
- the relationship between derivative products and expectations when interest rates are stochastic
- how to do Monte Carlo simulations to calculate derivative prices and to see the results of speculating with derivatives
- how to do simulations in many dimensions using Cholesky factorization

80.1 INTRODUCTION

The foundation of the theory of derivative pricing is the random walk of asset prices, interest rates etc. We have seen this foundation in Chapter 3 for equities, currencies and commodities, and the resulting option pricing theory from Chapter 5 onwards. This is the Black–Scholes theory leading to the Black–Scholes parabolic partial differential equation. We have also seen in Chapter 10 how the stochastic differential equation model for a random variable leads to a similar equation for the probability density function for the random variable. In that chapter I showed the relationship between option prices and the transition probability density. In this chapter we exploit this relationship, and see how derivative prices can be found from special simulations of the asset price, or interest rate, random walks. Briefly, the value of an option is the expected present value of the payoff. The catch in this is the precise definition of 'expected.'

I am going to distinguish between the valuation of options having equities, indices, currencies, or commodities as their underlying with interest rates assumed to be deterministic and those products for which it is assumed that interest rates are stochastic. First, I show the relationship between derivative values and expectations with deterministic interest rates.

80.2 RELATIONSHIP BETWEEN DERIVATIVE VALUES AND SIMULATIONS: EQUITIES, INDICES, CURRENCIES, COMMODITIES

Recall from Chapter 10 that the fair value of an option in the Black–Scholes world is the present value of the expected payoff at expiry under a *risk-neutral* random walk for the underlying.

The risk-neutral random walk for S is

$$dS = rS\,dt + \sigma S\,dX.$$

We can therefore write

$$\text{option value } = e^{-r(T-t)}E\left[\text{payoff}(S)\right]$$

provided that the expectation is with respect to the risk-neutral random walk, not the *real* one.
This result leads to an estimate for the value of an option by following these simple steps:

1. Simulate the risk-neutral random walk as discussed below, starting at today's value of the asset S_0, over the required time horizon. This time period starts today and continues until the expiry of the option. This gives one realization of the underlying price path.
2. For this realization calculate the option payoff.
3. Perform many more such realizations over the time horizon.
4. Calculate the average payoff over all realizations.
5. Take the present value of this average, this is the option value.

80.3 GENERATING PATHS

The initial part of this algorithm requires first of all the generation of random numbers from a standardized Normal distribution (or some suitable approximation). We discuss this issue below, but for the moment assume that we can generate such a series in sufficient quantities. Then one has to update the asset price at each time step using these random increments. Here we have a choice how to update S.

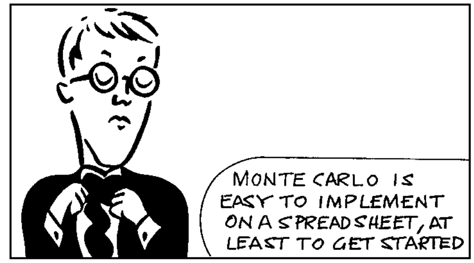
An obvious choice is to use

$$\delta S = rS\,\delta t + \sigma S\,\sqrt{\delta t}\,\phi,$$

where ϕ is drawn from a standardized Normal distribution. This discrete way of simulating the time series for S is called the **Euler method**. Simply put the latest value for S into the right-hand side to calculate δS and hence the next value for S. This method is easy to apply to any stochastic differential equation. This discretization method has an error of $O(\delta t)$.[1]

The above algorithm is illustrated in Figure 80.1. The stock begins at time $t = 0$ with a value of 100 and a volatility of 20%. The spreadsheet simultaneously calculates the values of a call and a put option. They both have an expiry of one year and a strike of 105. The interest rate is 5%. In this spreadsheet we see a small selection of a large number of Monte Carlo simulations of the random walk for S, *using a drift rate of 5%*. The time step was chosen to be 0.01. For each realization the final stock price is shown in row 102 (rows 13–93 have been hidden). The option payoffs are shown in rows 104 and 107. The mean of all these payoffs, over all the simulations, is shown in rows 105 and 108. In rows 106 and 109 we see the present values

[1] There are better approximations, for example the **Milstein method** which has an error of $O(\delta t^2)$.



	A	B	C	D	E	F	G	H	I
1	**Asset**	100		**Time**	**Sim 1**	**Sim 2**	**Sim 3**	**Sim 4**	**Sim 5**
2	**Drift**	5%		0	100.00	100.00	100.00	100.00	100.00
3	**Volatility**	20%		0.01	100.15	101.27	100.79	100.16	98.54
4	**Timestep**	0.01		0.02	99.80	100.84	102.72	102.31	101.66
5	**Int. rate**	5%		0.03	97.35	103.77	105.27	102.00	105.64
6				0.04	96.50	103.08	104.72	97.65	105.35
7				0.05	101.25	101.61	102.37	102.76	103.63
8				0.06	97.53	100.49	104.47	106.86	99.04
9				0.07	97.41	103.09	107.70	105.73	99.20
10				[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible].77
11				0.09	85.74	100.79	109.07	106.01	97.95
12				0.1	81.32	100.99	105.13	105.40	100.32
94				0.92	102.25	105.44	88.51	96.74	96.08
95				093	100.68	105.48	90.44	97.04	95.36
96				0.94	102.26	104.01	92.40	99.26	94.67
97				0.95	102.10	103.47	88.99	95.27	97.09
98				0.96	100.11	103.36	88.95	92.74	96.30
99				0.97	101.34	104.06	89.26	93.59	97.19
100				0[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]9
101				0.99	103.71	102.73	[illegible]	[illegible]	[illegible]8
102				1	104.94	104.47	91.86	95.05	98.79
103									
104	**Strike**	105	**CALL**	**Payoff**	0.00	0.00	0.00	0.00	0.00
105			**Mean**	8.43					
106			**PV**	8.02					
107			**PUTP**	**ayoff**	0.06	0.53	13.14	9.95	6.21
108			**Mean**	8.31					
109			**PV**	7.9					
110									
111									

= D3+$B$4

=E3*EXP(($B$5-0.5*$B$3*$B$3)*$B$4+$B$3*SQRT($B$4)*NORMSINV(RAND()))

= MAX($B$104-F102,0)

= MAX(G102-$B$104,0)

=AVERAGE(E104:IV104)

=D105*EXP(-$B$5*$D$102)

Figure 80.1 Spreadsheet showing a Monte Carlo simulation to value a call and a put option.

of the means; these are the option values. For serious option valuation you would not do such calculations on a spreadsheet. For the present example I took a relatively small number of sample paths.

The method is particularly suitable for path-dependent options. In the spreadsheet in Figure 80.2 I show how to value an Asian option. This contract pays an amount $\max(A - 105, 0)$ where A is the average of the asset price over the one-year life of the contract. The remaining details of the underlying are as in the previous example. How would the spreadsheet be modified if the average were only taken of the last six months of the contract's life?

	A	B	C	D	E	F	G	H	I
1	**Asset**	100		**Time**	**Sim 1**	**Sim 2**	**Sim 3**	**Sim 4**	**Sim 5**
2	**Drift**	5%		0	100.00	100.00	100.00	100.00	100.00
3	**Volatility**	20%		0.01	98.62	97.68	99.73	99.42	102.98
4	**Timestep**	0.01		0.02	100.69	96.45	101.13	101.28	101.36
5	**Int. rate**	5%		0.03	99.60	99.67	102.62	99.37	101.95
6				0.04	99.19	101.15	104.14	98.60	99.51
7				0.05	104.10	100.00	105.03	98.97	96.86
8				0.06	104.71	99.11	103.22	96.93	98.89
9				[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
10				[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
11				0.09	110.57	100.93	101.59	99.34	98.75
12				0.1	114.50	100.24	99.36	95.67	99.88
13				0.11	114.43	101.32	100.22	94.92	102.90
93				0.91	101.02	111.09	119.38	77.82	85.23
94				0.92	101.54	109.58	118.06	80.13	83.75
95				0.93	101.38	108.20	118.49	79.96	83.54
96				0.94	103.38	107.87	119.79	82.21	83.12
97				0.95	107.58	108.43	116.24	81.58	84.69
98				[illegible]	[illegible]	109.39	115.79	81.61	88.50
99				0.97	107.20	112.81	115.56	83.07	90.72
100				0.98	109.18	113.45	[illegible]	[illegible]	[illegible]
101				[illegible]	110.49	114.04	[illegible]	[illegible]	[illegible]
102				1	113.23	117.67	120.05	81.49	93.27
103									
104				**Average**	105.98	106.95	109.21	87.43	97.22
105									
106	**Strike**	105	**ASIAN**	**Payoff**	0.98	1.95	4.21	0.00	0.00
107			**Mean**	4.79					
108			**PV**	4.55					
109									
110									

= D3+$B$4

=E3*EXP(($B$5-0.5*$B$3*$B$3)*$B$4+$B$3*SQRT($B$4)*NORMSINV(RAND()))

=AVERAGE(E2:E102)

=AVERAGE(E106:IV106)

= MAX(G104-$B$106,0)

=D107*EXP(-$B$5*$D$102)

Figure 80.2 Spreadsheet showing a Monte Carlo simulation to value an Asian option.

80.4 LOGNORMAL UNDERLYING, NO PATH DEPENDENCY

For the lognormal random walk we are lucky that we can find a simple, and *exact*, time stepping algorithm. We can write the risk-neutral stochastic differential equation for S in the form

$$d(\log S) = \left(r - \tfrac{1}{2}\sigma^2\right) dt + \sigma\, dX.$$

This can be integrated exactly to give

$$S(t) = S(0) \exp\left(\left(r - \tfrac{1}{2}\sigma^2\right) t + \sigma \int_0^t dX\right).$$

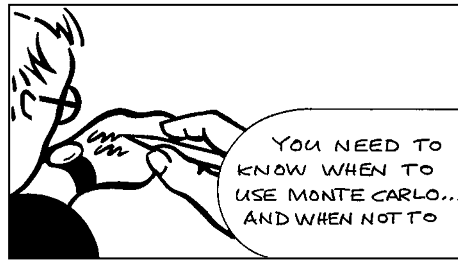
Or, over a time step δt,

$$S(t + \delta t) = S(t) + \delta S = S(t) \exp\left(\left(r - \tfrac{1}{2}\sigma^2\right) \delta t + \sigma \sqrt{\delta t}\, \phi\right). \tag{80.1}$$

Note that δt need not be small, since the expression is exact; and because it is exact and simple it is the best time-stepping algorithm to use. Also, because it is exact, if we have a payoff that only depends on the final asset value, i.e. is European and path independent, then we can simulate the final asset price in one giant leap, using a time step of T.

If the option is path-dependent then we have to go back to smaller time increments generally.

80.5 ADVANTAGES OF MONTE CARLO SIMULATION



Now that we have some idea of how Monte Carlo simulations are related to the pricing of options, I'll give you some of the benefits of using such simulations:

- The mathematics that you need to perform a Monte Carlo simulation can be very basic.
- Correlations can be easily modeled.
- There is plenty of software available, at the least there are spreadsheet functions that will suffice for most of the time.
- To get a better accuracy, just run more simulations.
- The effort in getting *some* answer is very low.
- The models can often be changed without much work.
- Complex path dependency can often be easily incorporated.
- People accept the technique, and will believe your answers.

80.6 USING RANDOM NUMBERS

The Black–Scholes theory as we have seen it has been built on the assumption of either a simple up-or-down move in the asset price, the binomial model, or a Normally distributed return. When it comes to simulating a random walk for the asset price it doesn't matter very much what distribution we use for the random increments as long as the time step is small and thus we have a large number of steps from the start to the finish of the asset price path. All we need are that the variance of the distribution must be finite and constant. (The constant must be such that the *annualized* volatility, i.e. the annualized standard deviation of returns, is the correct value. In particular, this means that it must scale with $\delta t^{1/2}$.) In the limit as the size of the time step goes to zero the simulations have the same probabilistic properties over a finite timescale regardless of the nature of the distribution over the infinitesimal timescale. This is a result of the central limit theorem.

Nevertheless, the most accurate model is the lognormal model with Normal returns. Since one has to worry about simulating sufficient paths to get an accurate option price one would ideally like not to have to worry about the size of the time step too much. As I said above, it is best to use the exact expression (80.1) and then the choice of time step does not affect the accuracy of the random walk. In some cases we can take a single time step since the time stepping algorithm is exact. If we do use such a large time step then we must generate Normally distributed random variables. I will discuss this below, where I describe the Box–Muller method.

If the size of the time step is δt then, for more complicated products, such as path-dependent ones, we may still introduce errors of $O(\delta t)$ by virtue of the discrete approximation to continuous events. An example would be of a continuous barrier. If we have a finite time step we miss the possibility of the barrier being triggered between steps. Generally speaking, the error due to the finiteness of the time step is $O(\delta t)$.

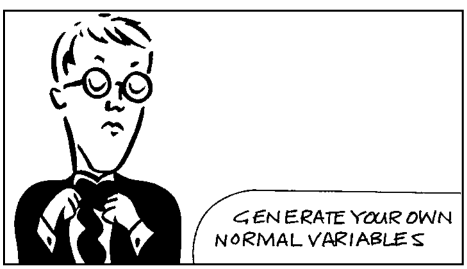
Because we are only simulating a finite number of an infinite number of possible paths, the error due to using N, say, realizations of the asset price paths is $O(N^{-1/2})$.

The total number of calculations required in the estimation of a derivative price is then $O(N/\delta t)$. This is also a measure of the time taken in the calculation of the price. The error in the price is

$$O\left(\max\left(\delta t, \frac{1}{\sqrt{N}}\right)\right),$$

i.e. the worst out of either the error due to the discreteness of the time step or the error in having only a finite number of realizations. To minimize this quantity, while keeping the total computing time fixed such that $O(N/\delta t) = K$, we must choose

$$N = O(K^{2/3}) \quad \text{and} \quad \delta t = O(K^{-1/3}).$$



80.7 GENERATING NORMAL VARIABLES

Some random number generators are good, others are bad, repeating themselves after a finite number of samples, or showing serial autocorrelation. Then they can be fast or slow. A particularly useful distribution that is easy to implement on a spreadsheet, and is fast, is the following approximation to the Normal distribution:

$$\left(\sum_{i=1}^{12} \psi_i\right) - 6,$$

where the ψ_i are independent random variables, drawn from a uniform distribution over zero to one. This distribution is close to Normal, having a mean of zero, a standard deviation of one, and a third moment of zero. It is in the fourth and higher moments that the distribution differs from Normal. I would use this in a spreadsheet when generating asset price paths with smallish time steps.

To get a better approximation you can just add up more uniform random variables. The general formula is

$$\sqrt{\frac{12}{N}}\left(\left(\sum_{i=1}^{N} \psi_i\right) - \frac{N}{2}\right).$$

This has a mean of zero and a standard deviation of one. As N gets larger so the approximation gets better. But if N gets too large we may as well use the following technique to produce genuinely Normal numbers.

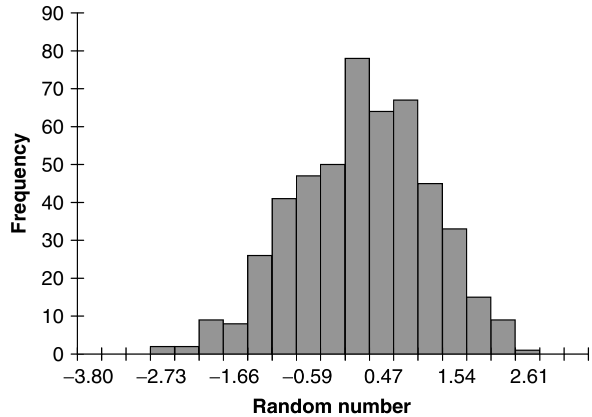


Figure 80.3 The approximation to the Normal distribution using 500 uniformly distributed points and the Box–Muller method.

80.7.1 Box–Muller

If you need to generate genuinely Normally-distributed random numbers then the simplest technique is the **Box–Muller method**. This method takes uniformly-distributed variables and turns them into Normal. The basic uniform numbers can be generated by any number of methods; see Press *et al*. (1992) for some algorithms. The Box–Muller method takes two uniform random numbers x_1 and x_2 between zero and one and combines them to give two numbers y_1 and y_2 that are both Normally distributed:

$$y_1 = \sqrt{-2\ln x_1}\ \cos(2\pi x_2) \quad \text{and} \quad y_2 = \sqrt{-2\ln x_1}\ \sin(2\pi x_2).$$

Here is a Visual Basic function that outputs a Normally-distributed variable.

```
Function BoxMuller()
Randomize
Do
        x = 2 * Rnd() - 1
        y = 2 * Rnd() - 1
        dist = x * x + y * y
Loop Until dist < 1
BoxMuller = x * Sqr(-2 * Log(dist) / dist)
End Function
```

In Figure 80.3 is the approximation to the Normal distribution using 500 points from the uniform distribution and the Box–Muller method.

For more efficient, more recent algorithms for generating Normal random variables see Jäckel (2002).

80.8 REAL VERSUS RISK NEUTRAL, SPECULATION VERSUS HEDGING

In Figure 80.4 are shown several realizations of a risk-neutral asset price random walk with 5% interest rate and 20% volatility. These are the thin lines. The bold lines in this figure

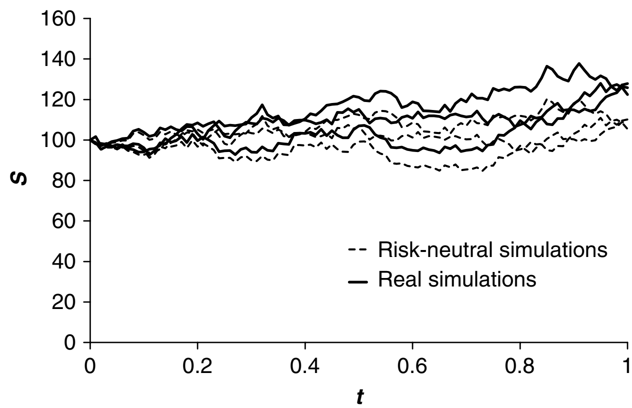


Figure 80.4 Several realizations of an asset price random walk.

are the corresponding *real* random walks using the same random numbers but here with a drift of 20% instead of the 5% interest rate drift. Although I am here emphasizing the use of Monte Carlo simulations in valuing options we can of course use them to estimate the payoff distribution from holding an *unhedged* option position. In this situation we are interested in the whole distribution of payoffs (and their present values) and not just the average or expected value. This is because in holding an unhedged position we cannot guarantee the return that we (theoretically) get from a hedged position. It is therefore valid and important to have the real drift as one of the parameters; it would be incorrect to estimate the probability density function for the return from an unhedged position using the risk-neutral drift.

In Figure 80.5 I show the estimated probability density function and cumulative distribution function for a call with expiry one year and strike 105 using Monte Carlo simulations with $\mu = 20\%$ and $\sigma = 20\%$. The probability density function curve does not show the zero payoffs. The probability of expiring out of the money and receiving no payoff is approximately 25%.

In Figure 80.6 I show the estimated probability density function and cumulative distribution function for a put with the same expiry and strike, again using Monte Carlo simulations with $\mu = 20\%$ and $\sigma = 20\%$. The probability density function curve does not show the zero payoffs. The probability of expiring out-of-the money and receiving no payoff is approximately 75%.

Valuing options for speculation is the subject of Chapter 59.

80.9 INTEREST RATE PRODUCTS

The relationship between expected payoffs and option values when the short-term interest rate is stochastic is slightly more complicated because there is the question of what rate to use for discounting.

The correct way to estimate option value with stochastic interest rates is as follows:

1. Simulate the random walk for the risk-adjusted spot interest rate r, as discussed below, starting at today's value of the spot rate, over the required time horizon. This time period starts today and continues until the expiry of the option. This gives one realization of the spot rate path.

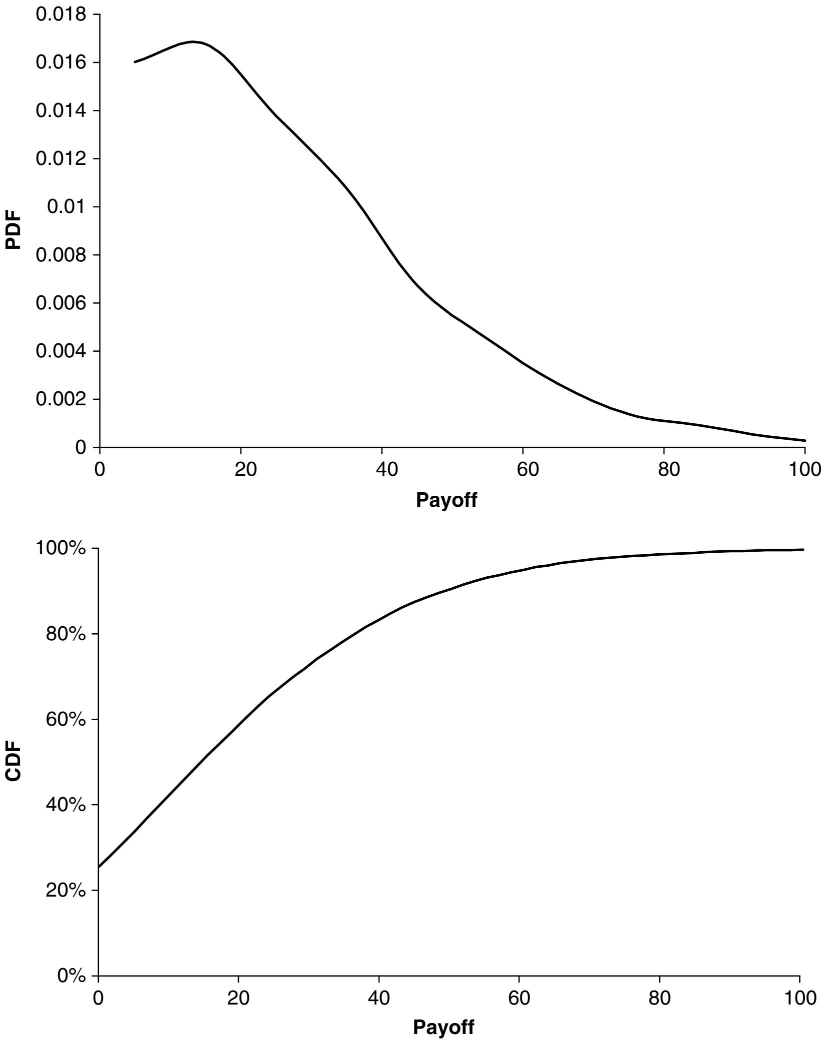
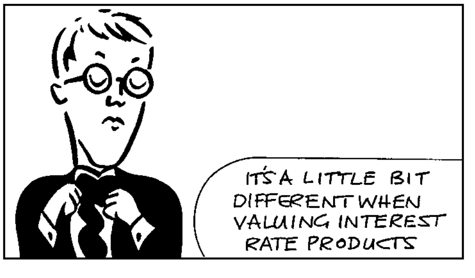


Figure 80.5 Real probability density function (top) and cumulative distribution function (bottom) for the payoff for a call.

2. For this realization calculate two quantities, the payoff and the *average* interest rate realized up until the payoff is received.
3. Perform many more such realizations.
4. For each realization of the r random walk calculate the present value of the payoff for this realization discounting at the average rate for this realization.
5. Calculate the average present value of the payoffs over all realizations; this is the option value.



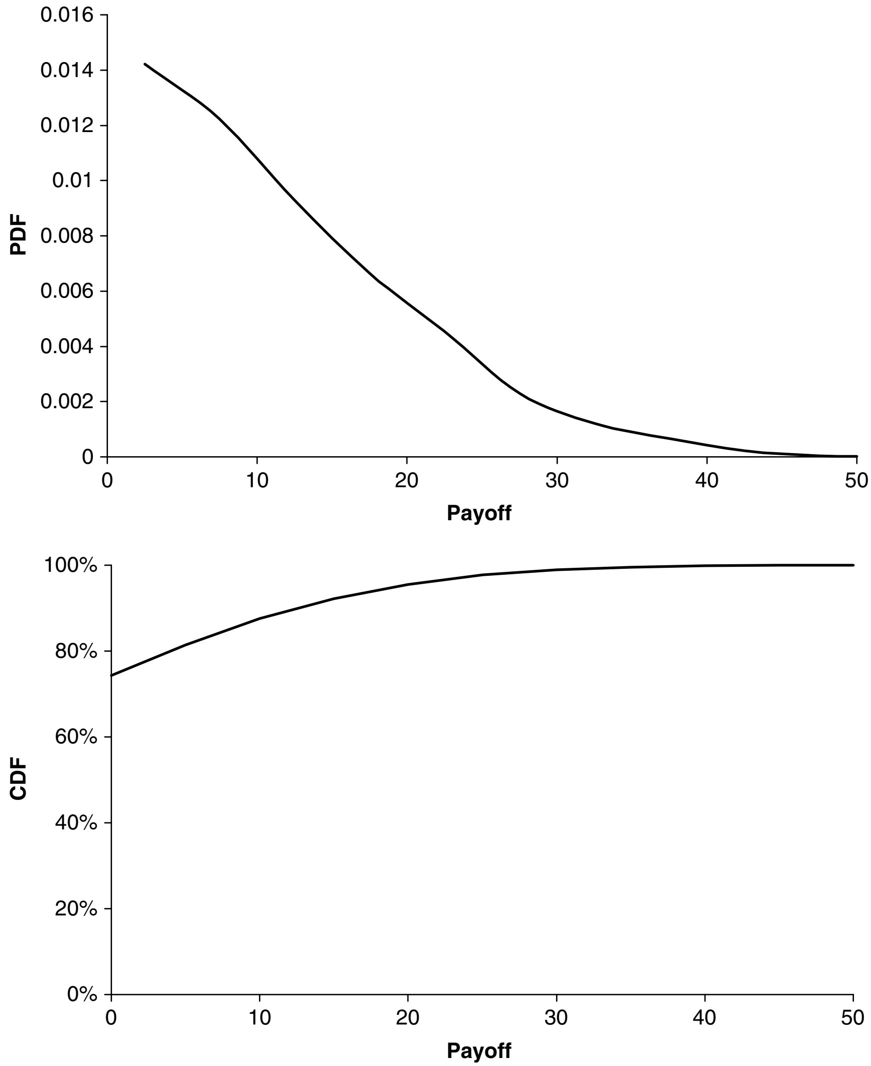


Figure 80.6 Real probability density function (top) and cumulative distribution function (bottom) for the payoff for a put.

In other words,

$$\text{option value} = E\left[e^{-\int_t^T r(\tau)\,d\tau}\ \text{payoff}(r)\right].$$

Why is this different from the deterministic interest rate case? Why discount at the average interest rate? We discount all cashflows at the average rate because this is the interest rate received by a money market account, and in the risk-neutral world all assets have the same risk-free growth rate. Recall that cash in the bank grows according to

$$\frac{dM}{dt} = r(t)M.$$

The solution of which is

$$M(t) = M(T)e^{-\int_t^T r(\tau)\,d\tau}.$$

This contains the same discount factor as in the option value.

The choice of discretization of spot rate models is usually limited to the Euler method

$$\delta r = (u(r,t) - \lambda(r,t)w(r,t))\delta t + w(r,t)\sqrt{\delta t}\,\phi.$$

Rarely can the spot rate equations be exactly integrated.

In the next spreadsheet (Figure 80.7) I demonstrate the Monte Carlo method for a contract with payoff $\max(r - 10\%, 0)$. Maturity is in one year. The model used to perform the simulations is Vasicek with constant parameters. The spot interest rate begins at 10%. The option value is the average present value in the last row.

	A	B	C	D	E	F	G	H
1	**Spot rate**	10%		**Time**	**Sim 1**	**Sim 2**	**Sim 3**	**Sim 4**
2	**Mean rate**	8%		0	10.00%	10.00%	10.00%	10.00%
3	**Reversion rate**	0.2		0.01	9.99%	10.04%	10.10%	9.88%
4	**Volatility**	0.007		0.02	9.95%	10.07%	10.14%	9.87%
5	**Timestep**	0.01		0.03	9.88%	10.08%	10.08%	9.81%
6				0.04	9.78%	10.12%	10.18%	9.84%
7			=D3+$B$5	0.05	9.87%	10.17%	10.15%	9.75%
8				0.06	9.87%	10.18%	10.10%	9.82%
9				0.07	9.73%	10.20%	10.11%	9.81%
10				0.08	9.72%	10.21%	10.27%	9.82%
11				0.09	9.60%	10.22%	10.25%	9.73%
12				[illegible]	[illegible]	[illegible]	[illegible]	9.79%
13				[illegible]	[illegible]	[illegible]	[illegible]	9.77%
14				[illegible]	[illegible]	[illegible]	[illegible]	9.75%
15				0.13	9.57%	10.15%	10.38%	9.73%
95				0.93	8.85%	10.44%	10.21%	9.27%
96	=AVERAGE(E2:E102)			0.94	8.81%	10.47%	10.23%	9.21%
97				0.95	8.83%	10.49%	10.17%	9.10%
98				0.96	8.81%	10.62%	10.19%	9.01%
99				0.97	8.73%	10.63%	10.29%	9.14%
100	=MAX(E102-$B$106,0)			0.98	8.68%	10.71%	10.28%	9.18%
101				0.99	8.56%	10.68%	10.15%	9.09%
102				1	8.42%	10.64%	10.33%	9.22%
103	=E106*EXP(-$D$102*E104)							
104				**Mean rate**	9.19%	10.32%	10.18%	9.56%
105								
106	**Strike**	10%		**Payoff**	0.0000	0.0064	0.0033	0.0000
107	=AVERAGE(E107:IV107)		**PV'd**		0.0000	0.0058	0.0030	0.0000
108			**Mean**	0.001352				
109								
110								
111								
112								
113								
114								

=F5+$B$3*($B$2-F5)*$B$5+$B$4*SQRT($B$5)*(RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()+RAND()-6)

Figure 80.7 Spreadsheet showing a Monte Carlo simulation to value a contract with a payoff $\max(r - 10\%, 0)$.

80.10 CALCULATING THE GREEKS

The simplest way to calculate the delta of an option using Monte Carlo simulation is to estimate the option's value twice. The delta of the option is the derivative of the option with respect to the underlying

$$\Delta = \lim_{h\to 0} \frac{V(S+h,t) - V(S-h,t)}{2h}.$$

This is a central difference, discussed in Chapter 77. This is an accurate estimate of the first derivative, with an error of $O(h^2)$. However, the error in the measurement of the two option values at $S+h$ and $S-h$ can be much larger than this for the Monte Carlo simulation. These Monte Carlo errors are then magnified when divided by h, resulting in an error of $O(1/hN^{1/2})$. To overcome this problem, estimate the value of the option at $S+h$ and $S-h$ *using the same values for the random numbers*. In this way the errors in the Monte Carlo simulation will cancel each other out. The same principal is used to calculate the gamma and the theta of the option.

If we are dealing with a lognormal random walk then things are much simpler than this in practice. Simulate many lognormal random paths as usual, starting with stock at S. Now imagine that the stock had started at $(1+\epsilon)S$. Do you have to resample the paths? No, because all you need do is multiply the final stock prices, at expiration, by the factor $1+\epsilon$. Of course, this only works for a lognormal random walk. Now use these new, scaled, prices to calculate payoffs and hence an option value. Do the same but with a factor of $1-\epsilon$. Then difference these two new option values and divide by $2\epsilon S$ and you have your delta, for almost no extra computational time.

Finally, we can also find the delta by observing what equation it satisfies.

The solution of the Black–Scholes partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{80.2}$$

has the interpretation of being the expected value of the present value of the payoff, under a risk-neutral random walk.

The first 'r' in this equation represents the 'r' in

$$dS = rS\,dt + \sigma S\,dX.$$

The second 'r' represents present valuing.

Differentiate Equation (80.2) with respect to S to get

$$\frac{\partial \Delta}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \Delta}{\partial S^2} + (r+\sigma^2)S\frac{\partial \Delta}{\partial S} = 0,$$

where

$$\Delta = \frac{\partial V}{\partial S}.$$

A comparison with the Black–Scholes equation shows that the solution of this equation represents the expected value of the final Δ (a step function) under the random walk

$$dS = (r+\sigma^2)S\,dt + \sigma S\,dX.$$

We can use the same paths for S as used for calculating the option value but shifted according to

$$Se^{\int_0^t \sigma^2(\tau)\,d\tau}.$$

Now calculate the expected value of the step function (delta at expiration) using this shifted random walk. Since there is no discounting term in the partial differential equation there is no need to take the present value.

80.11 HIGHER DIMENSIONS: CHOLESKY FACTORIZATION

Monte Carlo simulation is a natural method for the pricing of European-style contracts that depend on many underlying assets. Supposing that we have an option paying off some function of S_1, S_2,...,S_d then we could, in theory, write down a partial differential equation in $d+1$ variables. Such a problem would be horrendously time consuming to compute. The simulation methods discussed above can easily be extended to cover such a problem. All we need to do is to simulate

$$S_i(t+\delta t) = S_i(t)\exp\left(\left(r-\tfrac{1}{2}\sigma_i^2\right)\delta t + \sigma_i\sqrt{\delta t}\,\phi_i\right).$$

The catch is that the ϕ_i are correlated,

$$E[\phi_i\phi_j] = \rho_{ij}.$$

How can we generate *correlated* random variables? This is where **Cholesky factorization** comes in.

Let us suppose that we can generate d *uncorrelated* Normally distributed variables ϵ_1, $\epsilon_2, \ldots, \epsilon_d$. We can use these variables to get correlated variables with the transformation

$$\boldsymbol{\phi} = \mathbf{M}\boldsymbol{\epsilon} \tag{80.3}$$

where $\boldsymbol{\phi}$ and $\boldsymbol{\epsilon}$ are the column vectors with ϕ_i and ϵ_i in the ith rows. The matrix $\mathbf{M}$ is special and must satisfy

$$\mathbf{M}\mathbf{M}^T = \boldsymbol{\Sigma}$$

with $\boldsymbol{\Sigma}$ being the correlation matrix.

It is easy to show that this transformation will work. From (80.3) we have

$$\boldsymbol{\phi}\boldsymbol{\phi}^T = \mathbf{M}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\mathbf{M}^T. \tag{80.4}$$

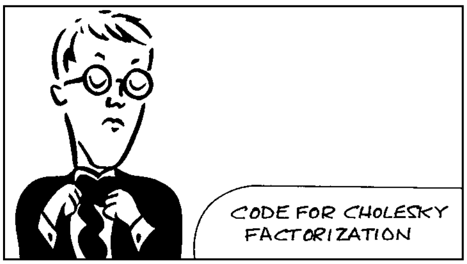
Taking expectations of each entry in this matrix equation gives

$$E\left[\boldsymbol{\phi}\boldsymbol{\phi}^T\right] = \mathbf{M}E\left[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\right]\mathbf{M}^T = \mathbf{M}\mathbf{M}^T = \boldsymbol{\Sigma}.$$

We can take expectations through the matrix multiplication in this because the right-hand side of (80.4) is linear in the terms $\epsilon_i\epsilon_j$.

This decomposition of the correlation matrix into the product of two matrices is not unique. The Cholesky factorization gives one way of choosing this decomposition. It results in a matrix **M** that is lower triangular. Here is an algorithm for the factorization.

The matrix `Sigma` contains the correlation matrix with dimension `n`. The output matrix is contained in `M`.




```
Function cholesky(Sigma As Object)
Dim n As Integer
Dim k As Integer
Dim i As Integer
Dim j As Integer
Dim x As Double
Dim a() As Double
Dim M() As Double
n = Sigma.Columns.Count
ReDim a(1 To n, 1 To n)
ReDim M(1 To n, 1 To n)
For i = 1 To n
    For j = 1 To n
        a(i, j) = Sigma.Cells(i, j).Value
        M(i, j) = 0
     Next j
Next i
For i = 1 To n
    For j = i To n
    x = a(i, j)
    For k = 1 To (i - 1)
        x = x - M(i, k) * M(j, k)
    Next k
    If j = i Then
        M(i, i) = Sqr(x)
    Else
        M(j, i) = x / M(i, i)
    End If
    Next j
Next i
cholesky = M
End Function
```

One health warning, Cholesky decomposition can become unstable if you have any perfect correlation. To get more insight into this and for other decomposition methods see Jäckel (2002).

80.12 CALCULATION TIME

If we want an accuracy of ϵ and we have d underlyings (dimensions), how long will it take for us to get the price of an options?

Since the accuracy is proportional to the time step, δt, and inversely proportional to the square root of the number of paths, N, then we must have

$$\delta t = O(\epsilon) \quad \text{and} \quad N = O\left(\epsilon^{-2}\right).$$

The total calculation time is therefore

$$O\left(\delta t^{-1} \times N \times d\right) = O\left(\epsilon^{-1} \times \epsilon^{-2} \times d\right) = O\left(d\epsilon^{-3}\right).$$

The factor d appears because we will have to simulate d correlated assets.

Note that we don't get the greeks from this calculation. If we want to calculate the greeks then we will have to repeat the calculation as many more times as we need greeks. But, on the other hand, we can price many different options at the same time with little extra time cost, since the time is taken up generating the paths and not in calculating payoffs.

80.13 SPEEDING UP CONVERGENCE

Monte Carlo simulation is inefficient, compared with finite-difference methods, in dimensions less than about three. It is natural, therefore, to ask how can one speed up the convergence. There are several methods in common use, two of which I now describe.

80.13.1 Antithetic Variables

In this technique one calculates two estimates for an option value using the one set of random numbers. We do this by using our Normal random numbers to generate one realization of the asset price path, an option payoff and its present value. Now take the same set of random numbers but change their signs, thus replace ϕ with $-\phi$. Again simulate a realization, and calculate the option payoff and its present value. Our estimate for the option value is the average of these two values. Perform this operation many times to get an accurate estimate for the option value.

This technique works because of the symmetry in the Normal distribution. This symmetry is guaranteed by the use of the antithetic variable.

80.13.2 Control Variate Technique

Suppose we have two similar derivatives, the former is the one we want to value by simulations and the second has a similar (but 'nicer') structure such that we have an explicit formula for its value. Use the *one* set of realizations to value *both* options. Call the values estimated by the Monte Carlo simulation V_1' and V_2'. If the accurate value of the second option is V_2 then a better estimate than V_1' for the value of the first option is

$$V_1' - V_2' + V_2.$$

The argument behind this method is that the error in V_1' will be the same as the error in V_2', and the latter is known.

A refinement of this technique is **martingale variance reduction**. In this method, one simulates one or more new dependent variables at the same time as the path of the underlying. This new stochastic variable is chosen so as to have an *expected value of zero* after each time step. This new variable, the 'variate,' is then added on to the value of the option. Since it has an expected value of zero it cannot make the estimate any worse, but if the variate is chosen carefully it can reduce the variance of the error significantly.

Let's see how this is done in practice using a single variate. Suppose we simulate

$$\delta S = \mu S\,\delta t + \sigma S\sqrt{\delta t}\,\phi$$

to price our contract. Now introduce the variate y, satisfying

$$\delta y = f(S,t)\,(\delta S - E[\delta S])\,,$$

with zero initial value. Note that this has zero expectation. The choice of $f(S, t)$ will be discussed in a moment. The new estimate for the option value is simply

$$\overline{V} - \alpha e^{-r(T-t)}\overline{y},$$

where $\overline{V}$ is our usual Monte Carlo estimate and $\overline{y}$ is the average over all the realizations of the new variate at expiry. The choice of α is simple; choose it to minimize the variance of the error i.e. to minimize

$$E\left[(V - \alpha e^{-r(T-t)}y)^2\right].$$

I leave the details to the reader.

And the function $f(S, t)$? The natural choice is the delta of an option that is closely related to the option in question, one for which there is a closed-form solution. Such a choice corresponds to an approximate form of delta hedging, and thus reduces the fluctuation in the contract value along each path.

80.14 PROS AND CONS OF MONTE CARLO SIMULATIONS

The Monte Carlo technique is clearly very powerful and general. The concept readily carries over to exotic and path-dependent contracts; just simulate the random walk and the corresponding cash flows, estimate the average payoff and take its present value.

The main disadvantages are twofold. First, the method is slow when compared with the finite-difference solution of a partial differential equation. Generally speaking this is true for problems up to three or four dimensions. When there are four or more stochastic or path-dependent variables the Monte Carlo method becomes relatively more efficient. Second, the application to American options is far from straightforward. The reason for the problem with American options is to do with the optimality of early exercise. To know when it is optimal to exercise the option one must calculate the option price *for all values of S and t up to expiry* in order to check that at no time is there any arbitrage opportunity. However, the Monte Carlo method in its basic form is only used to estimate the option price at one point in S, t-space, now and at today's value.

Because Monte Carlo simulation is based on the generation of a finite number of realizations using series of random numbers, the value of a option derived in this way will vary each time the simulations are run. Roughly speaking, the error between the Monte Carlo estimate and the correct option price is of the order of the inverse square root of the number of simulations. More precisely, if the standard deviation in the option value using a single simulation is ϵ then the standard deviation of the error after N simulations is $\epsilon/\sqrt{N}$. To improve our accuracy by a factor of 10 we must perform 100 times as many simulations.

80.15 AMERICAN OPTIONS

Applying Monte Carlo methods to the valuation of European contracts is simple, but applying them to American options is harder. The problem is to do with the time direction in which we are solving. We have seen how it is natural in the partial differential equation framework to work backwards from expiry to the present. If we do this numerically then we find the value of a contract at every mesh point between now and expiry. This means that along the way we can ensure that there is no arbitrage, and in particular ensure that the early-exercise constraint is satisfied.

When we use the Monte Carlo method in its basic form for valuing a European option we only ever find the option's value at the one point, the current asset level and the current time. We have no information about the option value at any other asset level or time. So if our contract is American we have no way of knowing whether or not we violated the early-exercise constraint somewhere in the future.

In principle, we could find the option value at each point in asset-time space using Monte Carlo. For every asset value and time that we require knowledge of the option value we start a new simulation. But when we have early exercise we have to do this at a large number of points in asset-time space, keeping track of whether the constraint is violated. If we find a value for the option that is below the payoff then we mark this point in asset-time space as one where we must exercise the option. And then for every other path that goes through this point we must exercise at this point, if not before. Such a procedure is possible, but the time taken grows exponentially with the number of points at which we value the option.

In Chapter 11 I said that there were no 'good' algorithms for Monte Carlo and early exercise. Well, that's not strictly true anymore, as we'll see in the next section.

80.16 LONGSTAFF & SCHWARTZ REGRESSION APPROACH FOR AMERICAN OPTIONS

There are several algorithms for pricing American options using Monte Carlo simulations, see Jäckel (2002) and Glasserman (2003) for details. But the one that is currently most popular with practitioners is that by Longstaff & Schwartz (2001). This method combines the forward simulation of stock price paths from start up to expiration, with the present valuing of cashflows along paths, but at each time step one looks at the benefit of exercising versus holding using a simple regression across stock prices. Let's walk through the algorithm in detail, with an example.

Consider an American put with strike price \$100, expiring in one year's time. The stock is currently \$100, it has a volatility of 20% and the risk-free interest rate is 5%.

Step one: Simulate many, N, realizations of the asset path from now to expiration using time steps of δt. If the random walk is lognormal then we will simulate

$$S_{j+1} = S_j \exp\left(\left(r - \tfrac{1}{2}\sigma^2\right)\delta t + \sigma\sqrt{\delta t}\,\phi\right)$$

where ϕ is a Normally distributed random variable, and S_j is the stock price after j time steps. We start each path at $j = 0$ with $S_0 = 100$. In Figure 80.8 are shown ten such paths, using five time steps from now to expiration. The numbers are shown in Figure 80.9.

Next step: Calculate the payoffs at expiration for each of these ten paths. These payoffs are shown in Figure 80.10. Notice that five paths end in the money and five out. Of course, this assumes that we haven't exercised the put option at any time before expiration. We'll get to that point in a moment. If we were to present value the average of all of these payoffs then we'd get the value of a European option.

The above tells us what we would receive if we exercised only at expiration, but we might have exercised earlier than that. Let's go back to the fourth time step, time 0.8, and look at the decision to exercise at that time for each of the paths. In what follows I use the notation of Longstaff & Schwartz.

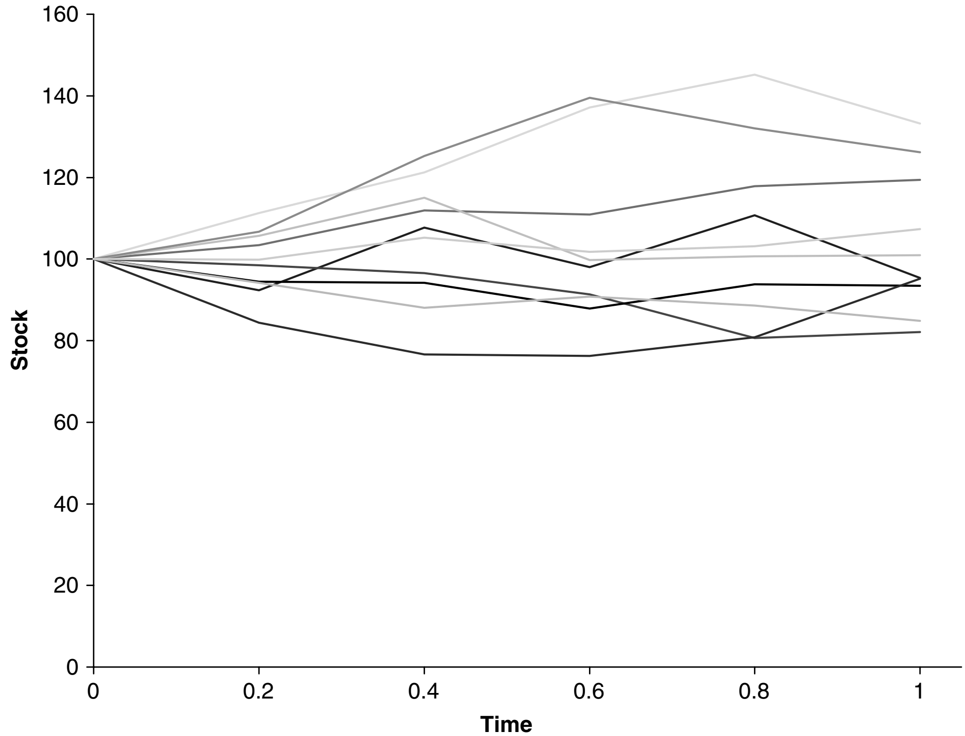


Figure 80.8 Ten stock price realizations.

Realization	0	0.2	0.4	0.6	0.8	1
1	100	92.30759	107.7357	98.04343	110.7416	95.34586
2	100	103.4446	111.9465	110.9322	117.8379	119.4419
3	100	111.2298	121.2417	137.1683	145.1687	133.1789
4	100	105.7152	115.0572	99.73054	100.6804	100.9471
5	100	98.47278	96.5825	91.32007	80.63689	82.1163
6	100	94.40168	94.16078	87.83702	93.84797	93.45847
7	100	106.7042	125.264	139.4822	132.0177	126.2041
8	100	84.37568	76.60055	76.21345	80.85454	95.19434
9	100	94.21698	88.00477	90.81541	88.63676	84.80556
10	100	99.81029	105.2631	101.747	103.1483	107.3703

Figure 80.9 The stock prices.

Figure 80.11 has an X and a Y column. The X column are the stock prices at time 0.8 *for those paths which are in-the-money at that time*. The Y column are the payoffs for these paths, discounted back from expiration to time 0.8. We want to calculate the cashflow from holding on to the option, conditional on the stock price at time 0.8. To do this we perform a regression of Y against a constant, X and X^2. In this example[2] we find the least-squares fit to be

$$Y = -0.1472\, X^2 + 25.347\, X - 1075.2. \tag{80.5}$$

[2] Other regressions are possible and to be preferred. We'll look at the subject of basis functions shortly.

Realization	0	0.2	0.4	0.6	0.8	1	Payoff
1	100	92.30759	107.7357	98.04343	110.7416	95.34586	4.654138
2	100	103.4446	111.9465	110.9322	117.8379	119.4419	0
3	100	111.2298	121.2417	137.1683	145.1687	133.1789	0
4	100	105.7152	115.0572	99.73054	100.6804	100.9471	0
5	100	98.47278	96.5825	91.32007	80.63689	82.1163	17.8837
6	100	94.40168	94.16078	87.83702	93.84797	93.45847	6.541526
7	100	106.7042	125.264	139.4822	132.0177	126.2041	0
8	100	84.37568	76.60055	76.21345	80.85454	95.19434	4.805663
9	100	94.21698	88.00477	90.81541	88.63676	84.80556	15.19444
10	100	99.81029	105.2631	101.747	103.1483	107.3703	0

Figure 80.10 The stock prices and payoffs.

Present value (at time 0.8) of the payoff if we hold on until expiration.

Stock prices at time 0.8, but only for those paths which are **in-the-money at this time.**

Realization	Y	X
1		
2		
3		
4		
5	0.99005 x 17.8837	80.63689
6	0.99005 x 6.541526	93.84797
7		
8	0.99005 x 4.805663	80.85454
9	0.99005 x 15.19444	88.63676
10		

Figure 80.11 Discounted payoffs and stock prices, for in-the-money at time 0.8 only. The discount factor is 0.99005.

Now, for these same in-the-money-at-time-0.8 paths we compare the value of exercising immediately, time 0.8, with that of holding on using (80.5) (see Figure 80.12). These numbers show that it is optimal to exercise now (time 0.8) for paths five and eight.

The next step is to build up a matrix of cashflows, as in Figure 80.13. This represents what is optimal *assuming that we don't exercise at any time step before 0.8*. Note that if there is a positive entry in the time 0.8 column it means that we should have exercised there (if not earlier) and so the cashflows for later times are set to zero.

Obviously we must continue to work backwards, next asking whether it would have been optimal to have exercised at time 0.6 for any of the paths. In Figure 80.14 are the cashflows, the Y column, present valued to time 0.6, and the stock prices, X, for those paths which are in-the-money at time 0.6. Note that there are now two discount factors because some cashflows has to be discounted one time step and others two time steps. The regression now gives

$$Y = -0.0361\ X^2 + 5.6613\ X - 203.95.$$

And off we go again; compare what you would get from exercising $\max(100 - X, 0)$ with the value from holding on $-0.0361\ X^2 + 5.6613\ X - 203.95$. After which we can draw up a new cashflow matrix.

Using the payoff for immediate exerice max(100–X,0)

Using the formula
$Y = -0.1472\,X^2 + 25.347\,X - 1075.2$
where X is the stock price at time 0.8

Realization	Exercise now	Hold on
1		
2		
3		
4		
5	19.36311255	11.5635
6	6.152033497	7.109119
7		
8	19.14546028	11.90641
9	11.36323851	15.0028
10		

Figure 80.12 What you get from exercising versus holding on.

Realization	0.2	0.4	0.6	0.8	1
1				0	4.654138
2				0	0
3				0	0
4				0	0
5				19.36311	0
6				0	6.541526
7				0	0
8				19.14546	0
8				0	15.19444
10				0	0

Figure 80.13 Cashflow matrix so far.

Present value (at time 0.6) of the cashflows if we hold on.

Stock prices at time 0.6, but only for those paths which are **in-the-money at this time**.

Realization	Y	X
1	0.980199 x 4.654138	98.04343
2		
3		
4	0	99.73054
5	0.99005 x 19.36311	91.32007
6	0.980199 x 6.541526	87.83702
7		
8	0.99005 x 19.14546	76.21345
9	0.980199 x 15.19444	90.81541
10		

Figure 80.14 Regression at time 0.6.

Realization	0.2	0.4	0.6	0.8	1
1	0	0	0	0	4.6541375
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	19.363113	0
6	0	0	0	0	6.541526
7	0	0	0	0	0
8	0	0	23.786554	0	0
9	0	11.995234	0	0	0
10	0	0	0	0	0

Figure 80.15 Final cashflow matrix.

Continue this process right back to time zero, and you should get the final cashflow matrix as shown in Figure 80.15. The last step in pricing the option is to present value all of these cashflows back to time zero and average. End result: 6.409. A finite-difference scheme gives the answer as 6.092.

VB code for this algorithm can be found in Chapter 83.

80.17 BASIS FUNCTIONS

I have used the simple quadratic basis function for the regression, since it is easier to explain, and does a decent job. There are, however, better ones you can use. The original paper by Longstaff and Schwartz discusses this matter. When it comes to using this method in higher dimensions (which is the point after all) the choice of basis functions can become important.

80.18 SUMMARY

Simulations are at the very heart of finance. With simulations you can explore the unknown future, and act accordingly. Simulations can also be used to price options; although the future is uncertain, the result of hedging an option is theoretically guaranteed.

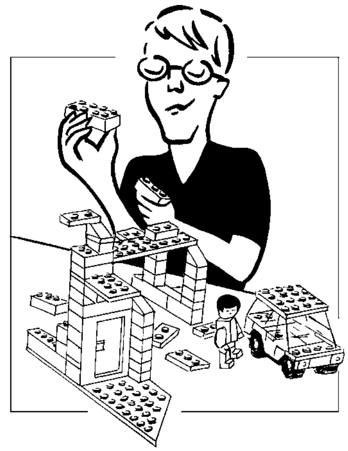
In this chapter I have tried to give a flavor of the potential of Monte Carlo and related methods. The reader should now be in a position to begin to use these methods in practice. The subject is a large, and growing one, and the reader is referred to the Further Reading section below for more information.

FURTHER READING

- See Boyle (1977) for the original application of Monte Carlo simulations to pricing derivatives.
- Duffie (1992) describes the important theory behind the validity of Monte Carlo simulations and also gives some clues about how to make the method efficient.
- The subject of Monte Carlo simulations is described straightforwardly and in detail by Vose (1997).

- For a review of Monte Carlo methods applied to American options see Boyle, Broadie & Glasserman (1995).
- For a very technical look at Monte Carlo see Jäckel (2002).
- For an American option algorithm see Longstaff & Schwartz (2001). This is a very good piece of work, nicely presented, understandable and with plenty of convincing examples. If only all papers were written like this.

CHAPTER 81
numerical integration



In this Chapter...

- how to do numerical integration in high dimensions to calculate the price of options on baskets

81.1 INTRODUCTION

Often the fair value of an option can be written down analytically as an integral. This is certainly the case for non-path-dependent European options contingent upon d lognormal underlyings, for which we have

$$V = e^{-r(T-t)} (2\pi(T-t))^{-d/2} (\mathrm{Det}\ \Sigma)^{-1/2} (\sigma_1 \cdots \sigma_d)^{-1}$$
$$\int_0^\infty \cdots \int_0^\infty \frac{\mathrm{Payoff}(S'_1 \cdots S'_d)}{S'_1 \cdots S'_d} \exp\left(-\frac{1}{2}\alpha^T \Sigma^{-1} \alpha\right) dS'_1 \cdots dS'_d$$

where

$$\alpha_i = \frac{1}{\sigma_i (T-t)^{1/2}} \left(\log\left(\frac{S_i}{S'_i}\right) + \left(r - D_i - \frac{\sigma_i^2}{2}\right)(T-t)\right)$$

and Σ is the correlation matrix for the d assets and Payoff$(\cdots)$ is the payoff function. Sometimes the value of path-dependent contracts can also be written as a multiple integral. American options, however, can rarely be expressed so simply.

If we do have such a representation of an option's value then all we need do to value it is to estimate the value of the multiple integral. Let us see how this can be done.

81.2 REGULAR GRID

We can do the multiple integration by evaluating the function on a uniform grid in the d-dimensional space of assets. There would thus be $N^{1/d}$ grid points in each direction where N is the total number of points used. Supposing we use the trapezium or mid-point rule, the error in the estimation of the integral will be $O(N^{-2/d})$ and the time taken approximately $O(N)$ since there are N function evaluations. As the dimension d increases, this method becomes prohibitively slow. Note that because the integrand is generally not smooth there is little point in using a higher order method than a mid-point rule unless one goes to the trouble of finding out the whereabouts of the discontinuities in the derivatives. To overcome this 'curse of dimensionality' we can use Monte Carlo integration or low-discrepancy sequences.

81.3 BASIC MONTE CARLO INTEGRATION

Suppose that we want to evaluate the integral

$$\int \ldots \int f(x_1, \ldots, x_d)\, dx_1 \ldots dx_d,$$

over some volume. We can very easily estimate the value of this by Monte Carlo simulation. The idea behind this is that the integral can be rewritten as

$$\int \ldots \int f(x_1, \ldots, x_d)\, dx_1 \ldots dx_d = \text{volume of region of integration} \times \text{average } f,$$

where the average of f is taken over the whole of the region of integration. To make life simple we can rescale the region of integration to make it the unit hypercube. Assuming that we have done this

$$\int_0^1 \ldots \int_0^1 f(x_1, \ldots, x_d)\, dx_1 \ldots dx_d = \text{average } f$$

because the volume is one. Such a scaling will obviously be necessary in our financial problems because the range of integration is typically from zero to infinity. I will return to this point later.

We can sample the average of f by Monte Carlo sampling using uniformly distributed random numbers in the d-dimensional space. After N samples we have

$$\text{average } f \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i) \tag{81.1}$$

where x_i is the vector of values of $x_1, \ldots, x_d$ at the ith sampling. As N increases, so the approximation improves. Expression (81.1) is only an approximation. The size of the error can be measured by the standard deviation of the correct average about the sampled average, this is

$$\sqrt{\frac{1}{N}\left(\overline{f^2} - \overline{f}^2\right)}$$

(which must be multiplied by the volume of the region), where

$$\overline{f} = \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

and

$$\overline{f^2} = \frac{1}{N} \sum_{i=1}^{N} f^2(x_i).$$

Thus the error in the estimation of the value of an integral using a Monte Carlo simulation is $O(N^{-1/2})$ where N is the number of points used, and is independent of the number of dimensions. Again there are N function evaluations and so the computational time is $O(N)$. The accuracy is much higher than that for a uniform grid if we have five or more dimensions.

I have explained Monte Carlo integration in terms of integrating over a d-dimensional unit hypercube. In financial problems we often have integrals over the range zero to infinity. The choice of transformation from zero-one to zero-infinity should be suggested by the problem under consideration. Let us suppose that we have d assets following correlated random walks. The risk-neutral value of these assets at a time t can be written as

$$S_i(T) = S_i(t)e^{\left(r-D_i-\frac{1}{2}\sigma_i^2\right)(T-t)+\sigma_i\phi_i\sqrt{T-t}},$$

in terms of their initial values at time t. The random variables ϕ_i are Normally distributed and correlated. We can now write the value of our European option as

$$e^{-r(T-t)}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\text{Payoff}(S_1(T),\ldots,S_d(T))\,p(\phi_1,\ldots,\phi_d)\,d\phi_1\ldots d\phi_d,$$

where $p(\phi_1,\ldots,\phi_d)$ is the probability density function for d correlated Normal variables with zero mean and unit standard deviation. I'm not going to write down p explicitly since we won't need to know its functional form *as long as we generate numbers from this distribution*. In effect, all that I have done here is to transform from lognormally distributed values of the assets to Normally distributed returns on the assets.

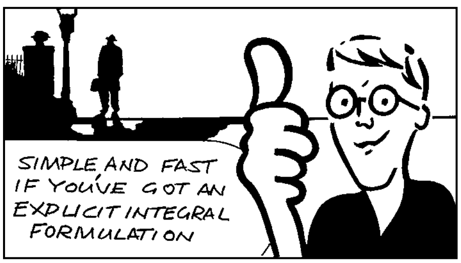
Now to value the option we must generate suitable Normal variables. The first step is to generate uncorrelated variables and then transform them into correlated variables. Both of these steps have been explained above; use Box–Muller and then Cholesky. The option value is then estimated by the average of the payoff over all the randomly generated numbers.

Here is a very simple code fragment for calculating the value of a European option in `NDim` assets using `NoPts` points. The interest rate is `IntRate`, the dividend yields are `Div(i)`, the volatilities are `Vol(i)`, time to expiry is `Expiry`. The initial values of the assets are `Asset(i)`. The Normally-distributed variables are the `x(i)` and the `S(i)` are the lognormally distributed future asset values.

```
a = Exp(-IntRate * Expiry) / NoPts
suma = 0
For k = 1 To NoPts
For i = 1 To NDim
If test = 0 Then
    Do
        y = 2 * Rnd() - 1
        z = 2 * Rnd() - 1
        dist = y * y + z * z
    Loop Until dist < 1
x(i) = y * Sqr(-2 * Log(dist) / dist)
test = 1
Else
x(i) = z * Sqr(-2 * Log(dist) / dist)
test = 0
End If
Next i
For i = 1 To NDim
```

```
    S(i) = Asset(i) * Exp((IntRate - Div(i) - _
                0.5 * Vol(i) * Vol(i)) * Expiry + _
                    Vol(i) * x(i) * Sqr(Expiry))
Next i
term = Payoff(S(1), S(2), S(3), S(4), S(5))
suma = suma + term
Next k
Value = suma * a
```

This code fragment is Monte Carlo in its most elementary form, and does not use any of the tricks described below. Some of these tricks are trivial to implement, especially those that are independent of the particular option being valued.



81.4 LOW-DISCREPANCY SEQUENCES

An obvious disadvantage of the basic Monte Carlo method for estimating integrals is that we cannot be certain that the generated points in the d-dimensional space are 'nicely' distributed. Indeed, there is inevitably a good deal of clumping. One way around this is to use a non-random series of points with better distributional properties.

Let us motivate the low-discrepancy sequence method by a Monte Carlo example. Suppose that we want to calculate the value of an integral in two dimensions and we use a Monte Carlo simulation to generate a large number of points in two dimensions at which to sample the integrand. The choice of points may look something like Figure 81.1. Notice how the points are not spread out evenly.

Now suppose we want to add a few hundred more points to improve the accuracy. Where should we put the new points? If we put the new points in the gaps between others then we increase the accuracy of the integral. If we put the points close to where there are already many points then we could make matters worse.

The above shows that we want a way of choosing points such that they are not too bunched, but nicely spread out. At the same time we want to be able to add more points later without spoiling our distribution. Clearly Monte Carlo is bad for evenness of distribution, but a uniform grid does not stay uniform if we add an arbitrary number of extra points. **Low discrepancy sequences** or **quasi-random sequences** have the properties we require.[1]

There are two types of low discrepancy sequences, open and closed. The open sequences are constructed on the assumption that we may want to add more points later. The closed sequences are optimized for a given size of sample, to give the best estimate of the integral for the number of points. The regular grid is an example of a closed low-discrepancy sequence. I will describe the open sequences here.

The first application of these techniques in finance was by Barrett, Moore & Wilmott (1992).[2] There are many such sequences with names such as **Sobol'**, **Faure**, **Haselgrove** and **Halton**. I shall describe the Halton sequence here, since it is by far the easiest to describe.

[1] There is actually nothing random about quasi-random sequences.

[2] Andy Morton says that this has been my best piece of work, knowing full well that the numerical analysts John Barrett and Gerald Moore should have all the credit.

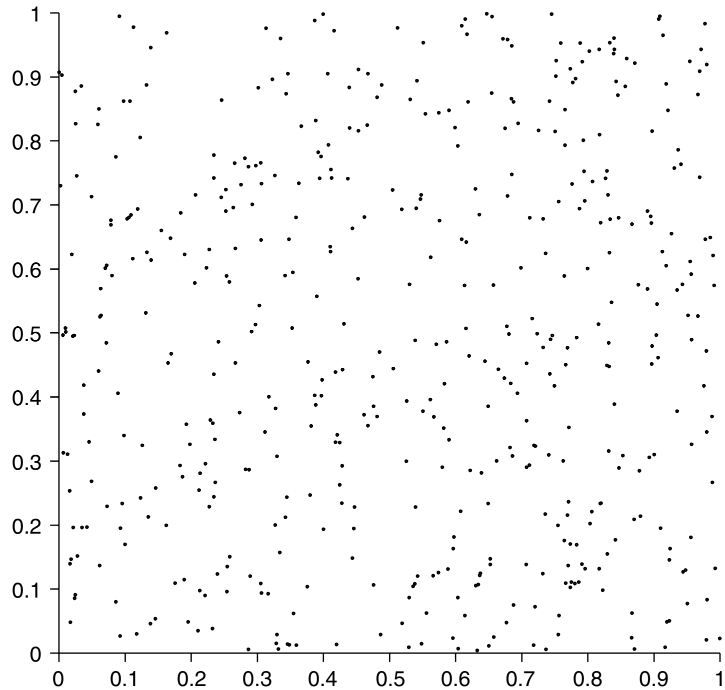


Figure 81.1 A Monte Carlo sample in two dimensions.

The Halton sequence is a sequence of numbers $h(i; b)$ for $i = 1, 2, \ldots$. The integer b is the base. The numbers all lie between zero and one.[3] The numbers are constructed as follows. First choose your base. Let us choose 2. Now write the positive integers in ascending order in base 2, i.e. 1, 10, 11, 100, 101, 110, 111 etc. The Halton sequence base 2 is the reflection of the positive integers in the decimal point i.e.

Integers base 10	Integers base 2	Halton sequence base 2	Halton number base 10
1	1	$1 \times \frac{1}{2}$	0.5
2	10	$0 \times \frac{1}{2} + 1 \times \frac{1}{4}$	0.25
3	11	$1 \times \frac{1}{2} + 1 \times \frac{1}{4}$	0.75
4	100	$0 \times \frac{1}{2} + 0 \times \frac{1}{4} + 1 \times \frac{1}{8}$	0.125
...	...	...	...

This has been called reflecting the numbers about the decimal point. If you plot the Halton points successively you will see that the next number in the sequence is always as far as possible

[3] So we must map our integrand onto the unit hypercube.

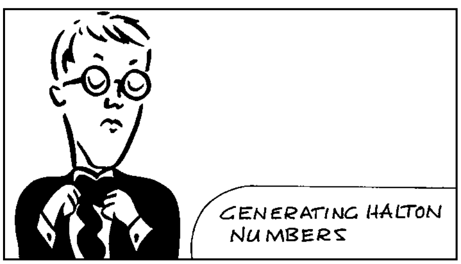
from the previous point. Generally, the integer n can be written as

$$i = \sum_{j=1}^{m} a_j b^j$$

in base b, where $0 \le a_j < b$. The Halton numbers are then given by

$$h(i; b) = \sum_{j=1}^{m} a_j b^{-j-1}.$$

Here is an algorithm for calculating Halton numbers of arbitrary base; the `n`th term in a Halton sequence of base `b` is given by `Halton(n, b)`.




```
Function Halton(n, b)
Dim n0, n1, r As Integer
Dim h As Double
Dim f As Double
    n0 = n
    h = 0
    f = 1 / b
While (n0 > 0)
    n1 = Int(n0 / b)
    r = n0 - n1 * b
    h = h + f * r
    f = f / b
    n0 = n1
Wend
Halton = h
End Function
```

The resulting sequence is nice because as we add more and more numbers, more and more 'dots,' we fill in the range zero to one at finer and finer levels.

In Figure 81.2 is the approximation to the Normal distribution using 500 points from a Halton sequence and the Box–Muller method. Compare this distribution with that in Figure 80.3. (You can get erroneous results if you combine the rejection-method Box–Muller with low-discrepancy numbers. See Jäckel (2002).)

When distributing numbers in two dimensions choose, for example, Halton sequence of bases 2 and 3 so that the integrand is calculated at the points $(h(i, 2), h(i, 3))$ for $i = 1, \ldots, N$. The bases in the two sequences should be prime numbers. The distribution of these points is shown in Figure 81.3; compare the distribution with that in the previous figure.

The estimate of the d-dimensional integral

$$\int_0^1 \cdots \int_0^1 f(x_1, \ldots, x_d)\, dx_1, \ldots, dx_d$$

is then

$$\frac{1}{N} \sum_{i=1}^{N} f(h(i, b_1), \ldots, h(i, b_n)),$$

where b_j are distinct prime numbers.

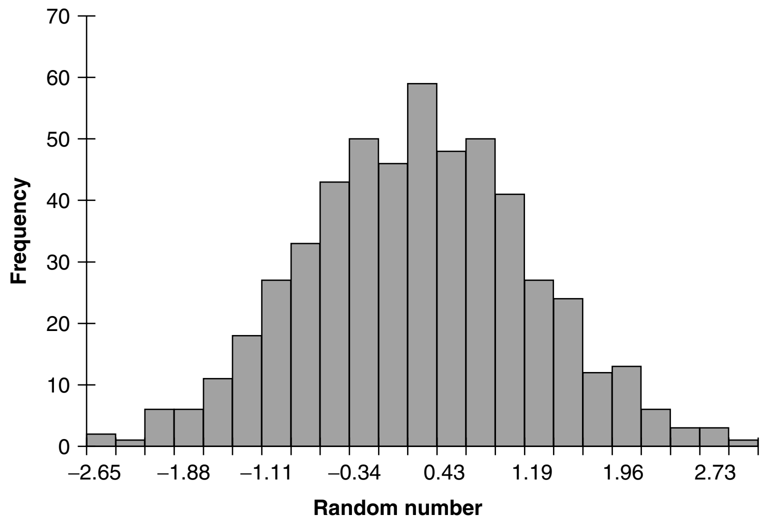


Figure 81.2 The approximation to the Normal distribution using 500 points from a Halton sequence and the Box–Muller method.

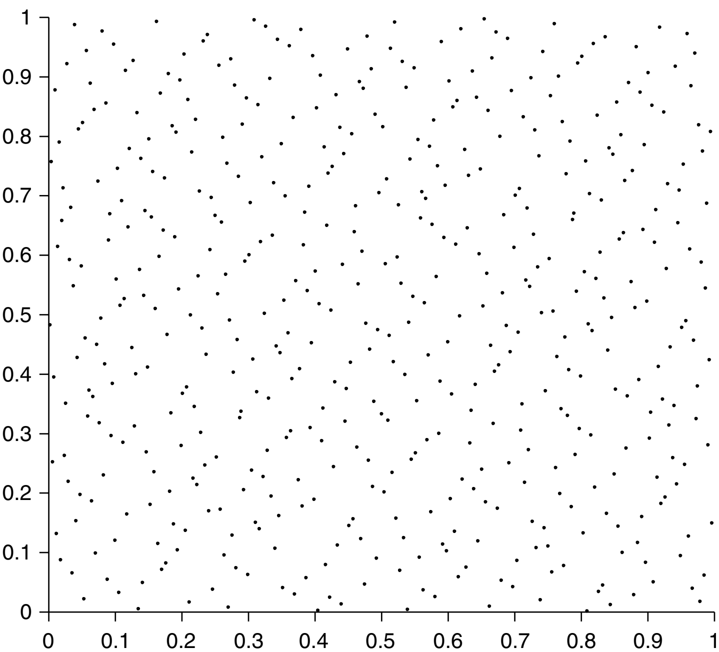


Figure 81.3 Halton points in two dimensions.

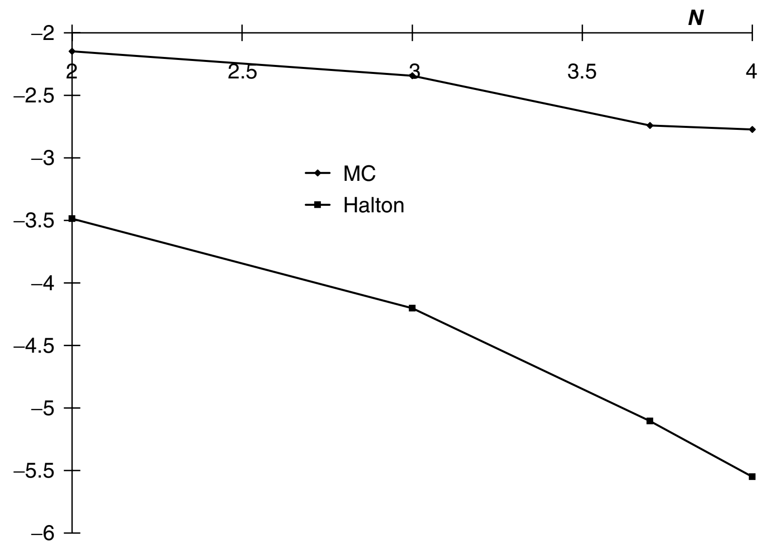


Figure 81.4 Estimate of the error in the value of a five-dimensional contract using basic Monte Carlo and a low-discrepancy sequence.

The error in these quasi-random methods is

$$O((\log N)^d N^{-1})$$

and is even better than Monte Carlo at all dimensions. The coefficient in the error depends on the particular low-discrepancy series being used. Sobol' is generally considered to be about the best sequence to use ... but it's much harder to explain. See Press *et al*. (1995) for code to generate Sobol' points. In three or more dimensions the method beats the uniform grid. The time taken is $O(N)$.

The error is clearly sensitive to the number of dimensions d. To fully appreciate the inverse relationship to N can require an awful lot of points. However, even with fewer points, in practice the method at its worst has the same error as Monte Carlo.

In Figure 81.4 is shown the relative error in the estimate of value of a five-dimensional contract as a function of the number of points used. The inverse relationship with the Halton sequence is obvious.

Another advantage of these low-discrepancy sequences is that if you collapse the points onto a lower dimension (for example, let all of the points in a two-dimensional plot fall down onto the horizontal axis) they will not be repeated, they will not fall on top of one another. This means that if there is any particularly strong dependence on one of the variables over the others then the method will still give an accurate answer because it will distribute points nicely over lower dimensions.

81.5 ADVANCED TECHNIQUES

There are several sophisticated techniques that can be used to improve convergence of Monte Carlo and related numerical integration methods. They can be generally classified as techniques for the **reduction of variance**, and hence for the increase in accuracy. None of these methods improve the, speed of convergence with respect to N (for example, the error remains $O(N^{-1/2})$ for Monte Carlo) but they can significantly reduce the coefficient in the error term.

The method of **antithetic variables** described above is *very* easily applied to numerical integration. It should always be used since it can do no harm and is completely independent of the product being valued.

Control variates can also be used in exactly the same way as described above. As with the pathwise simulation for pricing, the method depends on there being a good approximation to the product having an analytic formula.

The idea behind **importance sampling** is to change variables so that the resulting integrand is as close as possible to being constant. In the extreme case, when the integrand becomes exactly constant, the 'answer' is simply the volume of the region in the new variables. Usually, it is not possible to do so well. But the closer one gets to having a constant integrand, then the better the accuracy of the result. The method is rarely used in finance.

Stratified sampling involves dividing the region of integration into smaller subregions. The number of sampling points can then be allocated to the subregions in an optimal fashion, depending on the variance of the integral in each subregion. In more than one dimension it is not always obvious how to bisect the region, and can amount to laying down a grid, so defeating the purpose of Monte Carlo methods. The method can be improved upon by **recursive stratified sampling** in which a decision is made whether to bisect a region, based on the variance in the regions. Stratified sampling is rarely used in finance.

81.6 SUMMARY

Low-discrepancy sequences, where number theory and high finance meet.

FURTHER READING

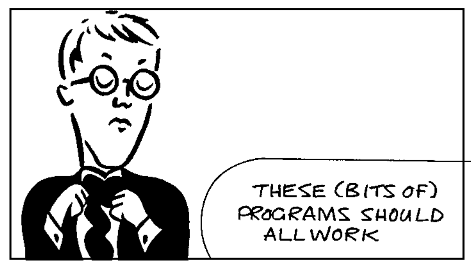
- See Sloan & Walsh (1990) and Stetson, Marshall & Loeball (1995) for details of how to optimize a grid.
- See Barrett, Moore & Wilmott (1992) for details of the Haselgrove method applied to options, and Haselgrove (1961) for more details of the method in abstract.
- For a practical example of pricing Mortgage Backed Securities see Ninomiya & Tezuka (1996).
- For more financial examples see Paskov & Traub (1995), Paskov (1996) and Traub & Wozniakowski (1994).
- See Niederreiter (1992) for an in-depth discussion of low-discrepancy sequences.
- See the amazing Press *et al*. (1992) for samples of code for random number generation and numerical integration. Make sure you use the latest edition; the random number generators in the first edition are not so good. They also describe more advanced integration techniques.

CHAPTER 82

finite-difference programs

In this Chapter...

- examples of code for many of the models and finite-difference methods described



82.1 INTRODUCTION

Finite-difference methods can be a bit daunting when you first start out, so here is a collection of programs that demonstrate some of the ideas in the book.

82.2 KOLMOGOROV EQUATION

This code calculates the probability that a stochastic volatility starting at `SigNow` leaves a given range, `SigMin` to `SigMax` before `TMax`. The model for the volatility is

$$d\sigma = \nu^2\sigma^{2\gamma-1}\left(\gamma - \tfrac{1}{2} - \frac{1}{2a^2}\log\left(\frac{\sigma}{\overline{\sigma}}\right)\right)dt + \nu\sigma^\gamma\,dX.$$

I've hard coded in the time step and the volatility step; you may want to change that.

```
Function Kolmogorov(SigNow, SigMax, SigMin, TMax, nu, SigBar, a, g)

Dim VOld(500) As Double
Dim VNew(500) As Double
Dim alph(500) As Double
Dim bet(500) As Double
Dim Sig(500) As Double
Dim dt As Double
Dim ds As Double
Dim NTS As Long
Dim NSS As Long
Dim SigNowInt As Long
Dim i As Long
```

```
Dim j As Long
Dim Delta As Double
Dim Gamma As Double
Dim frac As Double
Dim Answer As Double

dt = 0.001
ds = 0.01
NTS = Int(TMax / dt)
NSS = Int((SigMax - SigMin) / ds)
SigNowInt = Int((SigNow - SigMin) / ds)
frac = (SigNow - SigMin) / ds - SigNowInt

For i = 0 To NSS
VOld(i) = 0
Sig(i) = SigMin + i * ds
alph(i) = nu * nu * (Sig(i) ^ (2 * g - 1)) * (g - 0.5 - ((Log(Sig(i) / SigBar)) _
             / (a * a * 2)))
bet(i) = nu * (Sig(i) ^ g)
Next i
VOld(0) = 1
VOld(NSS) = 1

For j = 0 To NTS
For i = 1 To NSS - 1
Delta = (VOld(i + 1) - VOld(i - 1)) / 2 / ds
Gamma = (VOld(i + 1) - 2 * VOld(i) + VOld(i - 1)) / ds / ds
VNew(i) = VOld(i) + dt * (alph(i) * Delta + 0.5 * bet(i) * bet(i) * Gamma)
Next i

VNew(0) = 1
VNew(NSS) = 1

For i = 0 To NSS
VOld(i) = VNew(i)
Next i

Next j

If SigNow > SigMax Then
    Answer = 1
ElseIf SigNow < SigMin Then Answer = 1
Else: Answer = frac * VOld(SigNowInt + 1) + (1 - frac) * VOld(SigNowInt)
End If

Kolmogorov = Answer

End Function
```

82.3 EXPLICIT ONE-FACTOR MODEL FOR A CONVERTIBLE BOND

This Visual Basic code fragment shows an explicit finite-difference solution of Equation (33.1) for a CB with intermittent conversion. The code allows many conversion periods. The time steps on which conversion starts and ends for each period are stored in the arrays `ConvertDate-Start` and `ConvertDateEnd`, and between these time steps conversion is into `Convert-Rate` of the underlying. As the time step passes over the `ConvertDateStart` value the integer `itest` is increased by one so that we start to look at the next conversion period. In real time this period will be earlier since we are time stepping backwards.

The bond values are kept in `VOld` and updated in `VNew`. Note that we have used the boundary condition

$$\frac{\partial^2 V}{\partial S^2} = 0,$$

at the top of the asset range. This is represented by `VNew(M) = 2 * VNew(M - 1) - VNew(M - 2)`. At $S = 0$ we have

$$\frac{\partial V}{\partial t} = rV,$$

i.e. `VNew(0) = VOld(0) - dt * IntRate * VOld(0)`.

All you have to do to this to make it a fully-fledged function is to top and tail it with variable declarations etc. Make sure that you choose the time step `dt` sufficiently small to satisfy the stability constraint.

```
For j = 1 To N

        If (test = 0) Then
            itest = itest + 1
            test = 1
        End If

        For i = 1 To M - 1
            gamma = (VOld(i + 1) - 2 * VOld(i) + VOld(i - 1)) / dS / dS
            sdelta = S(i) * (VOld(i + 1) - VOld(i - 1)) / 2 / dS
            cash = sdelta - VOld(i)
            VNew(i) = VOld(i) + dt * _
                        (0.5 * S(i) * S(i) * sig * sig * gamma + IntRate * cash _
                         - div * sdelta)
        Next i
            VNew(0) = VOld(0) - dt * IntRate * VOld(0)
            VNew(M) = 2 * VNew(M - 1) - VNew(M - 2)

        For i1 = 0 To M
            VOld(i1) = VNew(i1)
            If j >= ConvertDateEnd(itest) And j <= ConvertDateStart(itest) Then
                VOld(i1) = max(VOld(i1), ConvertRate(itest) * S(i1))
                If j = ConvertDateStart(itest) Then test = 0
            End If
    Next i1
Next j
```

82.4 AMERICAN CALL, IMPLICIT

This code values an American call via an implicit scheme with successive over relaxation. The current asset value is `Asset`, volatility is `Sigma`, the spot interest rate is `IntRate`, dividend yield is Div. The option expiry is ExpTime, and the exercise price is Strike. The code also takes in the number of asset steps and the time step, NoAssetSteps and Timestep. Experiment with these. The solution range extends from $S = 0$ to $S = 3$ Asset; note the boundary conditions applied at the extremes of the range.

```
Function ImplicitUSCall(Asset, ExpTime, Sigma, IntRate, Div, Strike, _
  NoAssetSteps, Timestep)
  Dim VNew(0 To 1000), VOld(0 To 1000), g(0 To 1000), _
a(0 To 1000), b(0 To 1000), c(0 To 1000) As Double
  M = NoAssetSteps
  dS = Strike / M                  ' asset step
  frac = (Asset - M * dS) / dS
  M = 3 * M
  dt = Timestep
  kmax = Int(ExpTime / dt + 0.5)
  dt = ExpTime / kmax
  w = 1.5                             ' relaxation parameter
' set initial conditions for call payoff
  For i = 0 To M
    temp = i * dS
    VOld(i) = 0
    If (temp > Strike) Then VOld(i) = temp - Strike
    g(i) = VOld(i)
  Next i
' preset the tridiagonal elements for constant Sigma, IntRate, Div, dt.
  For i = 1 To M - 1
    a(i) = dt * 0.5 * i * (Sigma * Sigma * i + IntRate - Div)
    b(i) = 1 + dt * (Sigma * Sigma * i * i + IntRate)
    c(i) = dt * 0.5 * i * (Sigma * Sigma * i - IntRate + Div)
  Next i

' Start the time stepping.
  For j = 1 To kmax
    t = j * dt
    For i = 1 To M - 1  ' set initial guess for time m+1.
      VNew(i) = VOld(i)
    Next i
    k = 0               ' iteration count initialized
' set boundary conditions
    VNew(0) = 0
    VNew(M) = ds * M * Exp(-Div * t) - Strike * Exp(-IntRate * t)
10  norm = 0       ' set l2 norm to 0.
    For i = 1 To M - 1
        y = (VOld(i) + a(i) * VNew(i + 1) + c(i) * VNew(i - 1)) / b(i)
        temp = y - VNew(i)
        norm = norm + temp * temp
        VNew(i) = VNew(i) + w * temp
        If VNew(i) < g(i) Then VNew(i) = g(i)
      Next i
      k = k + 1
    If (norm > 0.0001) And (k < 50) Then GoTo 10
    For i = 1 To M - 1
      VOld(i) = VNew(i)
    Next i
  Next j
```

```
    ImplicitUSCall = frac * VNew(M / 3 + 1) + _
                     (1 - frac) * VNew(M / 3)
End Function
```

82.5 EXPLICIT PARISIAN OPTION

This code is the explicit method applied to a path-dependent contract, a Parisian option. Note that we need a two-dimensional array to keep track of the option value as a function of the underlying asset and the path-dependent quantity. All of the inputs to the function are obvious except for `BarTime`, which is the length of time the asset must be beyond the barrier for it to be triggered. What kind of Parisian option is this, up/down, in/out?

```
Function ParisianPut(Asset, ExpTime, Sigma, IntRate, Div, _
         BarTime, Barrier, Strike, NoAssetSteps)
  Dim VNew(0 To 300, 0 To 500), VOld(0 To 300, 0 To 500), _
         q(0 To 300, 0 To 500) As Double
  dS = Strike / NoAssetSteps                            ' asset step
  x = Int(Barrier / dS + 0.5)
  y = Int(Strike / dS + 0.5)
  z = Int(Asset / dS)
  l = 2 * y + 10                        ' number of asset points
  dt = 0.9 / (l * l * Sigma * Sigma)
  kmax = Int(ExpTime / dt + 1#)
  kmin = Int(BarTime / dt + 0.5)
  start1 = x
  end1 = l - 1
  start2 = 0
  end2 = x
' set initial conditions for put payoff
  For k = 0 To kmax
    For i = 0 To l
      temp = i * dS
      VOld(i, k) = 0#
      If (k <= kmin) And (temp < Strike) Then VOld(i, k) = Strike - temp
      q(i, k) = 0#
    Next i
  Next k
' reset boundary data at line tau=kmin
  For i = start2 To end2
    VOld(i, kmin) = 0#
  Next i
' Start the time stepping
  For j = 1 To kmax - 1
    t = j * dt
    ' solve BS equation in 1-d domain
    For i = start1 To end1
      VNew(i, 0) = VOld(i, 0) + dt * (0.5 * i * i * Sigma _
                   * Sigma * (VOld(i + 1, 0) - 2 _
                   * VOld(i, 0) + VOld(i - 1, 0)) _
                   + 0.5 * (IntRate - Div) * i _
                   * (VOld(i + 1, 0) - VOld(i - 1, 0)) _
                   - IntRate * VOld(i, 0))
    Next i
    ' solve modified BS equation in 2-d domain
```

```
    For k = 0 To kmin - 1
      For i = start2 + 1 To end2 - 1
        VNew(i, k) = VOld(i, k + 1) + dt * (0.5 * i * i * Sigma * Sigma _
                     * (VOld(i + 1, k + 1) - 2 * VOld(i, k + 1) _
                     + VOld(i - 1, k + 1)) + 0.5 * (IntRate - Div) * i _
                     * (VOld(i + 1, k + 1) - VOld(i - 1, k + 1)) _
                     - IntRate * VOld(i, k + 1))
      Next i
    Next k
    ' put data in q array on line k=kmin
    For i = 0 To l
      VNew(i, kmin) = q(i, j)
    Next i
    ' put boundary data in for tau>0, s=x.
    For k = 1 To kmin - 1
      VNew(x, k) = VNew(x, 0)
    Next k
    ' put boundary data in for tau>0, s=x.
    For k = 0 To kmin
      For i = 0 To l
        VOld(i, k) = VNew(i, k)
      Next i
      VOld(l, k) = 0#
      If (k <= kmin) Then VOld(0, k) = Strike * Exp(-IntRate * t)
    Next k
  Next j

  ParisianPut = VNew(z, 0)
End Function
```

82.6 PASSPORT OPTIONS

The vanilla Passport Option is an option on the balance of a trading account. The holder paper trades the underlying (up to the set `Limit`) until `Expiry`. At expiry he gets max(`Balance`, 0) for a call or max(−`Balance`, 0) for a put where `Balance` is the closing balance.

```
Function Passport(Asset, PType, Vol, Balance, N, Expiry, Limit)

Dim dt As Double
Dim dz As Double

Dim VOld(-500 To 500) As Double
Dim VNew(-500 To 500) As Double
Dim zminus, zplus, s As Integer
Dim i, Tim As Integer
Dim z(-500 To 500) As Double
Dim zmax As Double

Dim Arg As Integer
Dim frac As Double
Dim Answer As Double

dt = Expiry / N
zmax = 2
```

```
dz = 1.1 * Sqr(Vol * Vol * (1 + zmax) * (1 + zmax) * dt)
zplus = Int(zmax / dz)
zminus = -zplus

q = 1
If PType = "P" Then q = -1

For i = zminus To zplus
VOld(i) = Application.Max(q * i * dz, 0)
Next i

For Tim = 1 To N

For i = zminus + 1 To zplus - 1
      VNew(i) = VOld(i) + 0.5 * dt * Vol * Vol * _
            (1 + Abs(i * dz)) * (1 + Abs(i * dz)) * (VOld(i + 1) _
            - 2 * VOld(i) + VOld(i - 1)) / (dz * dz)
Next i
    VNew(zminus) = 2 * VNew(zminus + 1) - VNew(zminus + 2)
    VNew(zplus) = 2 * VNew(zplus - 1) - VNew(zplus - 2)

For i = zminus To zplus
      VOld(i) = VNew(i)
Next i

Next Tim

Arg = Int(Balance / Asset / dz)
frac = Balance / Asset / dz - Arg
Answer = frac * VOld(Arg + 1) + (1 - frac) * VOld(Arg)

Passport = Limit * Asset * Answer

End Function
```

82.7 CHOOSER PASSPORT OPTION

The Chooser Passport Option gives the holder the choice between the call and the put payoffs; the choice must be made at the choosing time `ChooseTime`, before `Expiry`.

```
Function ChooserPassport(Asset, Expiry, ChooseTime, Vol, Balance, Limit, NTS)

Dim Result As Double

Dim dt As Double
Dim dz As Double

Dim VOld(-500 To 500) As Double
Dim zminus, zplus, s As Integer
Dim i, Tim As Integer
Dim z(-500 To 500) As Double
Dim zmax As Double
Dim VNew(-500 To 500) As Double
Dim payoff(-500 To 500) As Double
```

```
Dim Arg As Integer
Dim frac As Double
Dim Answer As Double

Dim N1 As Integer
Dim N2 As Integer


dt = Expiry / NTS
N2 = Int(ChooseTime / dt)
N1 = NTS - N2
zmax = 0.5
dz = 1.2 * Vol * (1 + zmax) * Sqr(dt)
zplus = Int(zmax / dz)
zminus = -zplus

For i = zminus To 0
VOld(i) = 0
Next i
For i = 1 To zplus
VOld(i) = i * dz
Next i

For Tim = 1 To N1

For i = zminus + 1 To zplus - 1
      VNew(i) = VOld(i) + 0.5 * dt * Vol * Vol * (1 + Abs(i * dz)) _
                * (1 + Abs(i * dz)) * (VOld(i + 1) - 2 * VOld(i) _
                + VOld(i - 1)) / (dz * dz)
Next i
    VNew(zminus) = 2 * VNew(zminus + 1) - VNew(zminus + 2)
    VNew(zplus) = 2 * VNew(zplus - 1) - VNew(zplus - 2)

For i = zminus To zplus
      VOld(i) = VNew(i)
Next i

Next Tim

For i = zminus To 0
    VOld(i) = VOld(i) - i * dz
Next i

For Tim = N1 + 1 To NTS

For i = zminus + 1 To zplus - 1
      VNew(i) = VOld(i) + 0.5 * dt * Vol * Vol * (1 + Abs(i * dz)) _
                * (1 + Abs(i * dz)) * (VOld(i + 1) - 2 * VOld(i) _
                + VOld(i - 1)) / (dz * dz)
Next i
    VNew(zminus) = 2 * VNew(zminus + 1) - VNew(zminus + 2)
    VNew(zplus) = 2 * VNew(zplus - 1) - VNew(zplus - 2)

For i = zminus To zplus
      VOld(i) = VNew(i)
Next i
```

```
Next Tim

Arg = Int(Balance / Asset / dz)
frac = Balance / Asset / dz - Arg
Answer = frac * VOld(Arg + 1) + (1 - frac) * VOld(Arg)

ChooserPassport = Limit * Asset * Answer

End Function
```

82.8 EXPLICIT STOCHASTIC VOLATILITY

Here's a two-factor program for pricing a European call when volatility is stochastic. What is the risk-neutral volatility process I'm using?

```
Function stochvol(Expry, Strike, IntRate, SMax, SigMax, dt)

Dim ds, dsig, a, b, c, d, rho, DeltaS, DeltaSig, DeltaSig1, DeltaSig2, _
             GammaS, GammaSig, XGamma, Theta As Double
Dim N, i, j, k As Integer

ds = Strike / SMax * 2
dsig = 1 / SigMax

N = Int(Expry / dt) + 1
dt = Expry / N

a = 0.2
b = 1
c = 1
d = 1
rho = 0

ReDim S(0 To SMax) As Double
ReDim Sig(0 To SigMax) As Double
ReDim VOld(-1 To SMax, -1 To SigMax) As Double
ReDim VNew(-1 To SMax, -1 To SigMax) As Double

For i = 0 To SMax
S(i) = ds * i
Next i

For j = 0 To SigMax
Sig(j) = dsig * j
Next j

For i = 0 To SMax
For j = 0 To SigMax
VOld(i, j) = Application.Max(S(i) - Strike, 0)
Next j
Next i

For k = 1 To N
```

```
For i = 0 To SMax - 1
For j = 0 To SigMax - 1
DeltaS = (VOld(i + 1, j) - VOld(i - 1, j)) / 2 / ds
DeltaSig1 = (VOld(i, j + 1) - VOld(i, j)) / dsig
DeltaSig2 = (VOld(i, j) - VOld(i, j - 1)) / dsig
DeltaSig = DeltaSig1
If (a - b * Sig(j)) < 0 Then DeltaSig = DeltaSig2
GammaS = (VOld(i + 1, j) - 2 * VOld(i, j) + VOld(i - 1, j)) / ds / ds
GammaSig = (VOld(i, j + 1) - 2 * VOld(i, j) + VOld(i, j - 1)) / dsig / dsig
XGamma = (VOld(i + 1, j + 1) - VOld(i + 1, j - 1) - VOld(i - 1, j + 1) _
         + VOld(i - 1, j - 1)) / 4 / ds / dsig
Theta = 0.5 * Sig(j) * Sig(j) * S(i) * S(i) * GammaS + IntRate * S(i) * DeltaS _
          - IntRate * VOld(i, j) _
          + 0.5 * d * d * Sig(j) * GammaSig + c * (a - b * Sig(j)) * DeltaSig _
          + rho * Sig(j) * S(i) * d * Sqr(Sig(j)) * XGamma
VNew(i, j) = VOld(i, j) + dt * Theta
Next j
Next i

For j = 0 To SigMax - 1
VNew(SMax, j) = 2 * VNew(SMax - 1, j) - VNew(SMax - 2, j)
Next j

For i = 0 To SMax
VNew(i, SigMax) = 2 * VNew(i, SigMax - 1) - VNew(i, SigMax - 2)
Next i

For i = 0 To SMax
For j = 0 To SigMax
VOld(i, j) = VNew(i, j)
Next j
Next i

Next k

stochvol = VOld

End Function
```

82.9 UNCERTAIN VOLATILITY

This is a bit easy. Inside the engine, just after calculating `Gamma` and before calculating theta, you just need to add the lines

```
Vol = VolMax
If Gamma > 0 Then Vol = VolMin
```

and you are done.

82.10 CRASH MODELING

Below is a fragment of Visual Basic code that could be used for the crash model of Chapter 58. Here `S(i)` is the asset array, *lognormally* distributed about the current stock level i.e.

`S(i) = Exp(Log(Asset) + i * step)`. This is different from our usual, equally-spaced asset price grid. It's easier to use a logarithmically-spaced grid because we are assuming a percentage fall during the crash. `BSOld` is the array of Black–Scholes values calculated by another finite-difference scheme (or coming from a formula). `ik` is the number of grid points that the asset falls during a crash i.e. `ik = Int(-Log(1 - Crash) / Step)` where `Crash` is the percentage crash allowed. Then `kstar = 1 - Exp(-ik * Step)`. `testing` is used to see whether a crash is beneficial. `VOld` and `VNew` are the arrays holding the old and new values of the option. `Delta` is the optimal hedge ratio, either the 'Black–Scholes' delta or the crash delta.

```
For i = -N + ik To N

testing = VOld(i + 1) + (S(i) - S(i + 1) - kstar * S(i)) *  _
                (VOld(i + 1) - VOld(i - 1)) / (S(i + 1) - S(i - 1))

If BSOld(i - ik) > testing Then
' **************** delta hedge
        Delta(i) = (VOld(i + 1) - VOld(i - 1)) / (S(i + 1) - S(i - 1))
        VNew(i) = (VOld(i + 1) + (S(i) - S(i + 1) + IntRate * S(i) * dt) *  _
                (VOld(i + 1) - VOld(i - 1)) / (S(i + 1) - S(i - 1))) / _
                  (1 + IntRate * dt)
Else
' **************** crash hedge
        Delta(i) = (BSOld(i - ik) - VOld(i + 1)) / (S(i) - S(i + 1) - kstar _
                  * S(i))
        VNew(i) = (BSOld(i - ik) + _
                S(i) * (kstar + IntRate * dt) * (BSOld(i - ik) - VOld(i + 1)) / _
                (S(i) - S(i + 1) - kstar * S(i))) / (1 + IntRate * dt)
End If

Next i
```

This program needs to be topped and tailed.

How would you modify the code to allow for a range for the size of the crash?

82.11 EXPLICIT EPSTEIN–WILMOTT SOLUTION

This code solves the non-linear first-order Epstein–Wilmott model for uncertain interest rate. The coupon dates go in `input1` and the coupon amounts go in `input2`. The current spot interest rate is `rNow` with `rMin` and `rMax` being the minimum and maximum allowed rates. The rates are allowed to grow or decay at a rate of at most `growth`. How would you change this program to value the bond in the best case?

```
Function EpsteinWilmott(input1, input2, rNow, rMin, rMax, growth)
Dim VNew(500), VOld(500), CouponVal(20), CouponDate(20) As Double
numcoup = input1.Rows.Count
tMin = 0
tMax = 0
      For i = 0 To numcoup - 1
            CouponDate(i) = input1(numcoup - i)
            CouponVal(i) = input2(numcoup - i)
```

```
            tMax = Application.Max(CouponDate(i), tMax)
         Next i

    m = 1

' set up variable data
dr = 1 / 100
dt = dr / growth
rSteps = Int((rMax - rMin) / dr)
tSteps = Int((tMax - tMin) / dt)
rateinteger = (rNow - rMin) / dr
frac = (rNow - rateinteger * dr + rMin) / dr

' initial data for arrays
For i = 0 To rSteps
    VOld(i) = CouponVal(0)
Next i

' solve the equation

For j = 1 To tSteps

    ' boundary values
    If VOld(0) * (1 - rMin * dt) < VOld(1) * (1 - (rMin + 0.5 * dr) * dt) Then
        VNew(0) = VOld(0) * (1 - rMin * dt)
    Else: VNew(0) = VOld(1) * (1 - (rMin + 0.5 * dr) * dt)
    End If

    If VOld(rSteps) * (1 - rMax * dt) < VOld(rSteps - 1) _
      * (1 - (rMax - 0.5 * dr) * dt) Then
        VNew(rSteps) = VOld(rSteps) * (1 - rMax * dt)
    Else: VNew(rSteps) = VOld(rSteps - 1) * (1 - (rMax - 0.5 * dr) * dt)
    End If

    ' interior points
    Spacer = rMin + dr
    For i = 1 To rSteps - 1
        If VOld(i) * (1 - Spacer * dt) < VOld(i - 1) _
          * (1 - (Spacer - 0.5 * dr) * dt) _
            And VOld(i) * (1 - Spacer * dt) < VOld(i + 1) _
              * (1 - (Spacer + 0.5 * dr) * dt) Then
            VNew(i) = VOld(i) * (1 - Spacer * dt)
        ElseIf VOld(i - 1) * (1 - (Spacer - 0.5 * dr) * dt) < VOld(i) _
          * (1 - Spacer * dt) _
            And VOld(i - 1) * (1 - (Spacer - 0.5 * dr) * dt) < VOld(i + 1) _
              * (1 - (Spacer + 0.5 * dr) * dt) Then
            VNew(i) = VOld(i - 1) * (1 - (Spacer - 0.5 * dr) * dt)
        Else: VNew(i) = VOld(i + 1) * (1 - (Spacer + 0.5 * dr) * dt)
        End If

        Spacer = Spacer + dr
    Next i

' coupon dates
If j = Int((tMax - CouponDate(m)) / dt) Then
    For i = 0 To rSteps
```

```
        VNew(i) = VNew(i) + CouponVal(m)
    Next i
    m = m + 1
End If

' reset old values
For i = 0 To rSteps
        VOld(i) = VNew(i)
    Next i

Next j

EpsteinWilmott = (1 - frac) * VOld(rateinteger) + frac * VOld(rateinteger + 1)

End Function
```

82.12 RISKY-BOND CALCULATOR

This code values coupon-bearing bonds with stochastic risk of default (`hrate`, p). The hazard rate model is the mean reverting CIR model usually used for modeling the spot interest rate

$$dp = \texttt{rev}\,(\texttt{meanh} - p)\,dt + \texttt{vol}\;p^{1/2}\,dX$$

The principal is 1, paid at maturity `mat`, interest rate `intrate` is constant and continuously compounded. There can be a rolling guarantee on the next `nog` coupons, and/or a guaranteed principal. If the principal is guaranteed then one of the specified number of guaranteed coupons is at maturity, in other cases the guaranteed coupons are the earliest due. The bond can be callable.

```
Sub rbpricer(hrate, intrate, meanh, rev, vol, mat, coupon, cpa, nog, optpg, _
  optcall, callfor, nots, out)
Dim VOld(-1 To 1000) As Double
Dim VNew(0 To 1000) As Double
Dim p(0 To 1000) As Double
Dim dt As Double
Dim dp As Double
Dim pmax As Double
Dim nopsteps As Double
Dim inthrate As Integer
Dim frac As Double
Dim j As Integer
Dim realtime As Double
Dim delta As Double
Dim gamma As Double
Dim coupdate As Double
Dim Q As Double
Dim Q0 As Double
Dim Q2 As Double
Dim k As Integer
Dim lastone As Integer
Dim tc As Double

dt = mat / nots
```

```
pmax = 1
dp = 3 * 1.1 * Sqr(dt * pmax) * vol
nopsteps = Int(pmax / dp)

inthrate = Int(hrate / dp)
frac = (hrate - dp * inthrate) / dp

For i = 0 To nopsteps
    p(i) = i * dp
    VOld(i) = 1 + coupon / cpa
Next i

Q0 = 0
lastone = 0

If optpg = True Then
    Q0 = 1
    If nog > 0 Then
        Q0 = Q0 + coupon / cpa
        nog = nog - 1
    End If
Else
    lastone = 1
End If

coupdate = mat - 1 / cpa

For j = 1 To nots

realtime = mat - j * dt

Q = Q0 * Exp(-intrate * j * dt)
For k = 1 To min(nog, lastone + Int(j * dt * cpa))
tc = j * dt - Int(j * dt * cpa) / cpa + (k - 1) * 1 / cpa
Q = Q + coupon / cpa * Exp(-intrate * tc)
Next k

For i = 0 To nopsteps - 1


    If j * dt >= (nog - 1) / cpa Then

        delta = (VOld(i + 1) - VOld(i)) / dp
        If p(i) > meanh Then delta = (VOld(i) - VOld(i - 1)) / dp
        gamma = (VOld(i + 1) - 2 * VOld(i) + VOld(i - 1)) / dp / dp
        VNew(i) = VOld(i) + dt * (0.5 * vol * vol * p(i) * gamma + rev _
                    * (meanh - p(i)) * delta _
                    - (intrate + p(i)) * VOld(i) + p(i) * Q)

    Else

        If optpg Then
            VNew(i) = Q
        Else
            delta = (VOld(i + 1) - VOld(i)) / dp
            If p(i) > meanh Then delta = (VOld(i) - VOld(i - 1)) / dp
```

```
            gamma = (VOld(i + 1) - 2 * VOld(i) + VOld(i - 1)) / dp / dp
            VNew(i) = VOld(i) + dt * (0.5 * vol * vol * p(i) * gamma + rev _
                      * (meanh - p(i)) * delta _
                      - (intrate + p(i)) * VOld(i) + p(i) * (Q - Q0 _
                      * Exp(-intrate * j * dt)))
        End If

    End If

    Next i

    VNew(nopsteps) = VNew(nopsteps - 1) * VNew(nopsteps - 1) / VOld(nopsteps - 2)

    For i = 0 To nopsteps
        VOld(i) = VNew(i)
    Next i

    If realtime < coupdate And realtime + dt > coupdate Then
    coupdate = coupdate - 1 / cpa
    For i = 0 To nopsteps
        VOld(i) = VNew(i) + coupon / cpa
    Next i
    End If

    If optcall Then
    For i = 0 To nopsteps
        VOld(i) = min(VOld(i), callfor)
    Next i
    End If

Next j

out(1) = frac * VOld(inthrate) + (1 - frac) * VOld(inthrate + 1)
out(2) = frac * (VOld(inthrate + 1) - VOld(inthrate - 1)) / 2 / dp + _
        (1 - frac) * (VOld(inthrate + 2) - VOld(inthrate)) / 2 / dp
out(3) = Q

End Sub
```

CHAPTER 83
Monte Carlo programs

In this Chapter...

- a few illustrative simulation programs

83.1 INTRODUCTION

Here are a few Visual Basic codes to get you started with this important numerical method.

83.2 MONTE CARLO PRICING OF A BASKET

The following code values a European option on `NDim` lognormal assets. The payoff in this example is simply the maximum of all of the assets. Arrays must be input for `Asset`, the starting prices, `Vol` and `DivYld`. The correlation matrix `Correl` is input as a square array. This is decomposed by the subroutine `NewMat`.

```
Function Monte_Carlo_Basket(Asset, Correl, Vol, DivYld, IntRate, Expiry, _
                           NoEvals, NDim)

ReDim ux(1 To NDim) As Double
ReDim cx(1 To NDim) As Double
ReDim s(1 To NDim) As Double
ReDim CholM(1 To NDim, 1 To NDim) As Double

rootexpiry = Sqr(Expiry)
a = Exp(-IntRate * Expiry) / NoEvals
suma = 0

Call NewMat(Correl, CholM, NDim)

For k = 1 To NoEvals

'produce uncorrelated Normal variables
For i = 1 To NDim
If test = 0 Then
```

```
    Do
        y = 2 * Rnd() - 1
        z = 2 * Rnd() - 1
        dist = y * y + z * z
    Loop Until dist < 1
ux(i) = y * Sqr(-2 * Log(dist) / dist)
test = 1
Else
ux(i) = z * Sqr(-2 * Log(dist) / dist)
test = 0
End If
Next i

'turn them into correlated variables
For i = 1 To NDim
cx(i) = 0
    For j = 1 To i
    cx(i) = ux(j) * CholM(i, j) + cx(i)
    Next j
Next i

For i = 1 To NDim
    s(i) = Asset(i) * Exp((IntRate - DivYld(i) - _
                0.5 * Vol(i) * Vol(i)) * Expiry + _
                    Vol(i) * cx(i) * rootexpiry)
Next i

term = Application.Max(s)
suma = suma + term

Next k

Monte_Carlo_Basket = suma * a

End Function
```

Here is the Cholesky decomposition code.

```
Sub NewMat(Correl, CholM, NDim)
' Cholesky factorization
Dim x As Double
For i = 1 To NDim
    For j = 1 To NDim
        CholM(i, j) = 0
     Next j
Next i
For i = 1 To NDim
    For j = i To NDim
    x = Correl(i, j)
    For k = 1 To (i - 1)
        x = x - CholM(i, k) * CholM(j, k)
    Next k
    If j = i Then
        CholM(i, i) = Sqr(x)
    Else
        CholM(j, i) = x / CholM(i, i)
```

```
    End If
    Next j
Next i

End Sub
```

83.3 QUASI MONTE CARLO PRICING OF A BASKET

Same problem as above, but now using Halton numbers.

```
Function Quasi_Monte_Carlo_Basket(Asset, Correl, Vol, DivYld, IntRate, _
                           Expiry, NoEvals, NDim)
Dim prime(1 To 5, 1 To 2) As Integer
Dim en(1 To 5) As Long

ReDim ux(1 To NDim) As Double
ReDim cx(1 To NDim) As Double
ReDim s(1 To NDim) As Double
ReDim CholM(1 To NDim, 1 To NDim) As Double

prime(1, 1) = 2
prime(1, 2) = 13
prime(2, 1) = 3
prime(2, 2) = 17
prime(3, 1) = 5
prime(3, 2) = 19
prime(4, 1) = 7
prime(4, 2) = 23
prime(5, 1) = 11
prime(5, 2) = 29

rootexpiry = Sqr(Expiry)

a = Exp(-IntRate * Expiry) / NoEvals

suma = 0

Call NewMat(Correl, CholM, NDim)

For k = 1 To NoEvals

'************************
'produce uncorrelated Normal variables
For i = 1 To NDim
    Do
        y = 2 * Halton(en(i), prime(i, 1)) - 1
        z = 2 * Halton(en(i), prime(i, 2)) - 1
        dist = y * y + z * z
        en(i) = en(i) + 1
    Loop Until dist < 1
ux(i) = y * Sqr(-2 * Log(dist) / dist)
Next i
'*************************

'turn them into correlated variables
```

```
For i = 1 To NDim
cx(i) = 0
    For j = 1 To i
    cx(i) = ux(i) * CholM(i, j) + cx(i)
    Next j
Next i

For i = 1 To NDim
    s(i) = Asset(i) * Exp((IntRate - DivYld(i) - _
                0.5 * Vol(i) * Vol(i)) * Expiry + _
                    Vol(i) * cx(i) * rootexpiry)
Next i

term = Application.Max(s)
suma = suma + term

Next k

Quasi_Monte_Carlo_Basket = suma * a

End Function

Function Halton(n, b)
Dim n0, n1, r As Long
Dim h As Double
Dim f As Double
    n0 = n
    h = 0
    f = 1 / b
While (n0 > 0)
    n1 = Int(n0 / b)
    r = n0 - n1 * b
    h = h + f * r
    f = f / b
    n0 = n1
Wend
   Halton = h
End Function
```

83.4 MONTE CARLO FOR AMERICAN OPTIONS

This code implements the Longstaff & Schwartz least squares regression algorithm. Note that the code calls the Excel `LinEst` function to perform the required regression. The `Payoff` function is defined externally; in this case it is that for a put. You should also modify this code to use different basis functions.

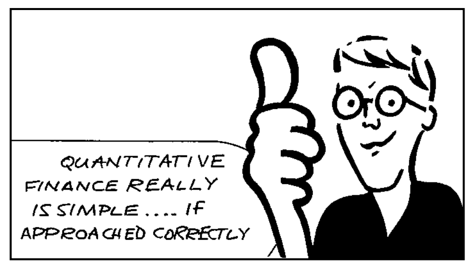
```
Function USOptionMC(SToday, Strike, Expn, Vol, IntRate, NTS, NPaths)

ReDim Stock(0 To NTS, 1 To NPaths)
ReDim Cashflow(1 To NTS, 1 To NPaths)
ReDim TempArrayX(1 To NPaths, 1 To 2)
ReDim TempArrayY(1 To NPaths)
Dim TempArrayXForRegression()
Dim TempArrayYForRegression()
ReDim Regress(1 To 3)
```

```
TStep = Expn / NTS
Drift = (IntRate - 0.5 * Vol * Vol) * TStep
SD = Vol * Sqr(TStep)
DF = Exp(-IntRate * TStep)

' Simulate stock
For i = 1 To NPaths
Stock(0, i) = SToday
For j = 1 To NTS
Stock(j, i) = Stock(j - 1, i) * Exp(Drift + SD * Norm)
Next j
Next i

For i = 1 To NPaths
Cashflow(NTS, i) = Payoff(Stock(NTS, i), Strike)
ValueEuro = ValueEuro + Cashflow(NTS, i) / NPaths
Next i

ValueEuro = DF ^ NTS * ValueEuro ' Just in case you want to output this as well

For j = NTS - 1 To 1 Step -1

Num = 0

For i = 1 To NPaths
If Payoff(Stock(j, i), Strike) > 0 Then Num = Num + 1
Next i

' Regression
ReDim TempArrayXForRegression(1 To 2, 1 To Num)
ReDim TempArrayYForRegression(1 To Num)
k = 1

For N = 1 To Num
TempArrayXForRegression(1, N) = 0
TempArrayXForRegression(2, N) = 0
TempArrayYForRegression(N) = 0
Next N

For i = 1 To NPaths

If Payoff(Stock(j, i), Strike) > 0 Then ' ITM
TempArrayXForRegression(1, k) = Stock(j, i)
TempArrayXForRegression(2, k) = Stock(j, i) * Stock(j, i)
For m = 1 To NTS - j
TempArrayYForRegression(k) = TempArrayYForRegression(k) _
                           + DF ^ m * Cashflow(j + m, i)
Next m
k = k + 1
End If

Next i

' Use Excel LinEst function to calculate regression coefficients
Regress = Application.LinEst(TempArrayYForRegression, TempArrayXForRegression)
```

```
For i = 1 To NPaths
If Payoff(Stock(j, i), Strike) > Regress(1) * Stock(j, i) * Stock(j, i) + _
            Regress(2) * Stock(j, i) + Regress(3) _
            And Payoff(Stock(j, i), Strike) > 0 Then
' Exercise
Cashflow(j, i) = Payoff(Stock(j, i), Strike)
For m = 1 To NTS - j
Cashflow(j + m, i) = 0
Next m
Else
Cashflow(j, i) = 0
End If
Next i

Next j

' Present Value all cashflows and average
For i = 1 To NPaths
For j = 1 To NTS
ValueUS = ValueUS + DF ^ j * Cashflow(j, i) / NPaths
Next j
Next i

USOptionMC = ValueUS

End Function



Function Payoff(S, E)
Payoff = 0
If S < E Then Payoff = E - S
End Function



Function Norm()
Dim fac As Double, R As Double, X As Double, Y As Double

1 X = 2 * Rnd - 1
Y = 2 * Rnd - 1
R = X * X + Y * Y
If (R >= 1) Then GoTo 1
Norm = Y * Sqr(-2 * Log(R) / R)

End Function
```

APPENDIX A

all the math you need. . . and no more (an executive summary)



In this Appendix. . .

- e
- log
- differentiation and Taylor series
- expectations and variances

A.1 INTRODUCTION

This book is for *everyone* interested in quantitative finance. This subject is becoming increasingly technical. Some people don't have a high-level math training, through no fault of their own, but may still be interested in the technical side of things. In this appendix we look at the mathematics that you need to cope with the vast majority of derivatives theory and practice. Although often couched in very high-level mathematics almost all finance theory can be interpreted using only the basics that I describe here.

The most useful math is the simplest math. This is particularly true in finance where the beauty of the mathematics can and does lead to people not seeing the wood for the trees. All basic finance theory and a great deal of the advanced research really require only elementary mathematics if approached in the right way. In this appendix I explain this elementary mathematics and show how to use it to understand the Black–Scholes derivatives theory. I am trying to do for Mathematical Finance what Seuss (1999) did for the English language.

A.2 THE DIFFERENT TYPES OF MATHEMATICS FOUND IN FINANCE

There are many different types of mathematics in use in quantitative finance. The most popular are differential equations, abstract probability theory (going by the name martingales) and simple adding, subtracting etc.

Obviously, if you can get away with just the adding and subtracting, and occasional other easy bits of arithmetic, well that's just fine and dandy. And that's what the simple option pricing model, the binomial model, is all about. However, we don't do an awful lot of that in this book, because it is rather limiting. You can't build much of a skyscraper with just a toy hammer and a bit of string. And I do want us to build skyscrapers in this book, as it were.

Then there's the abstract stuff; which is fantastic. The problem is that it is abstract *probability* theory. Sometimes we have problems that aren't probabilistic, then what? Probability theory isn't much use for that, is it? Sticking with the buildings analogy, I think of martingale theory not as a tool but as a material. Steel is a material; it's brilliant, you can build all sorts of things with it, ships, bridges, etc. But it's not that great for houses, and while it's useful as the skeleton of skyscrapers you will need other material to pad out that skeleton.

Differential calculus, now that's not a material, that's a box of tools. With the right tools and some imagination you can build anything. Calculus doesn't care whether a problem is deterministic or probabilistic or something completely different. Calculus is just about how things change or evolve, in time, space or with stock price. And that's mostly what we do in this book. Another advantage of focusing on the tools rather than the materials is that we don't have to limit ourselves in our modeling. Getting back to finance, many models at the cutting edge of finance research are non linear. Calculus has no problems with nonlinearity, whereas martingales do. If you are concentrating on the probabilistic models it seriously hampers your scope for creativity. After all, outside of finance most models are non linear.

The real-world subject of quantitative finance uses tools from many branches of mathematics, and financial modeling can be approached in a variety of different ways. For some strange reason the advocates of different branches of mathematics get quite emotional when discussing the merits and demerits of their methodologies and those of their 'opponents.' Is this a territorial thing; what are the pros and cons of martingales and differential equations; what is all this fuss about and will it end in tears before bedtime?

Here's a list of the various approaches to modeling and a selection of useful tools. The distinction between a 'modeling approach' and a 'tool' will start to become clear.

A.2.1 Modeling Approaches

Deterministic

The idea behind this approach is that our model will tell us everything about the future. Given enough data, and a big enough brain, we can write down some equations or an algorithm for predicting the future. Interestingly, the subject of dynamical systems and chaos fall into this category. And, as you know, chaotic systems show such sensitivity to initial conditions that predictability is in practice impossible. This is the 'butterfly effect,' that a butterfly flapping its wings in Brazil will 'cause' rainfall over Manchester. (Like what doesn't!) A topic popular in the early 1990s, this has not lived up to its promises in the financial world.

Probabilistic

One of the main assumptions about the financial markets, at least as far as quantitative finance goes, is that asset prices are random. We tend to think of describing financial variables as following some random path, with parameters describing the growth of the asset and its degree of randomness. We effectively model the asset path via a specified rate of growth, on average, and its deviation from that average. This approach to modeling has had the greatest impact over the last 30 years, leading to the explosive growth of the derivatives markets.

Discrete/Continuous

Whether probabilistic or deterministic the eventual model you write down can be discrete or continuous. Discrete means that asset prices and/or time can only be incremented in finite chunks, whether a dollar or a cent, a year or a day. Continuous means that no such lower increment exists. The mathematics of continuous processes is often easier than that of discrete ones. But then when it comes to number crunching you have to turn a continuous model into a discrete one anyway.

In discrete models we end up with difference equations. An example of this is the binomial model for asset pricing. Time progresses in finite amounts, the time step. In continuous models we end up with differential equations. The equivalent of the binomial model in discrete space is the Black–Scholes model, which has continuous asset price and continuous time. Whether binomial or Black–Scholes both of these models come from the probabilistic assumptions about the financial world.

A.2.2 The Tools

Now let's take a look at some of the tools available.

Simulations

If the financial world is random then we can experiment with the future by running simulations. For example, an asset price may be represented by its average growth and its risk, so let's simulate what could happen in the future to this random asset. If we were to take such an approach we would want to run many, many simulations. There'd be little point in running just the one, we'd like to see a range of possible future scenarios.

Simulations can also be used for non-probabilistic problems. Just because of the similarities between mathematical equations, a model derived in a deterministic framework may have a probabilistic interpretation.

Discretization methods

The complement to simulation methods, there are many types of these. The best known are the finite-difference methods which are discretizations of continuous models such as Black–Scholes. Depending on the problem you are solving, and unless it's very simple, you will probably go down the simulation or finite-difference routes for your number crunching.

Approximations

In modeling we aim to come up with a solution representing something meaningful and useful, such as an option price. Unless the model is really simple, we may not be able to solve it easily. This is where approximations come in. A complicated model may have approximate solutions; and these approximate solutions might be good enough for our purposes.

Asymptotic analysis

This is an incredibly useful technique, used in most branches of applicable mathematics, but almost unknown in finance. The idea is simple, find approximate solutions to a complicated problem by exploiting parameters or variables that are either large or small, or special in some way. For example, there are simple approximations for vanilla option values close to expiry.

Green's functions

This is a very special technique that only works in certain situations. The idea is that solutions to some difficult problems can be built up from solutions to special solutions of a similar problem.

A.3 e

The first bit of math you need to know about is e.

e is

- a number, 2.7183...
- a function when written e^x; this function is a.k.a. $\exp(x)$

The function e^x is just the number 2.7183 ... raised to the power x; e^2 is just $2.7183\ldots^2 = 7.3891\ldots$, e^1 is 2.7183 ... and $e^0 = 1$. What about non-integer powers?

The function e^x can be written as an infinite series

$$e^x = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \cdots = \sum_{i=0}^{\infty} \frac{x^i}{i!}.$$

This gets around the non-integer power problem.

A plot of e^x as a function of x is shown in Figure A.1.

The function e^x has the special property that the slope or gradient of the function is also e^x. Plot this slope as a function of x and for e^x you get the same curve again. It follows that the slope of the slope is also e^x, etc. etc.

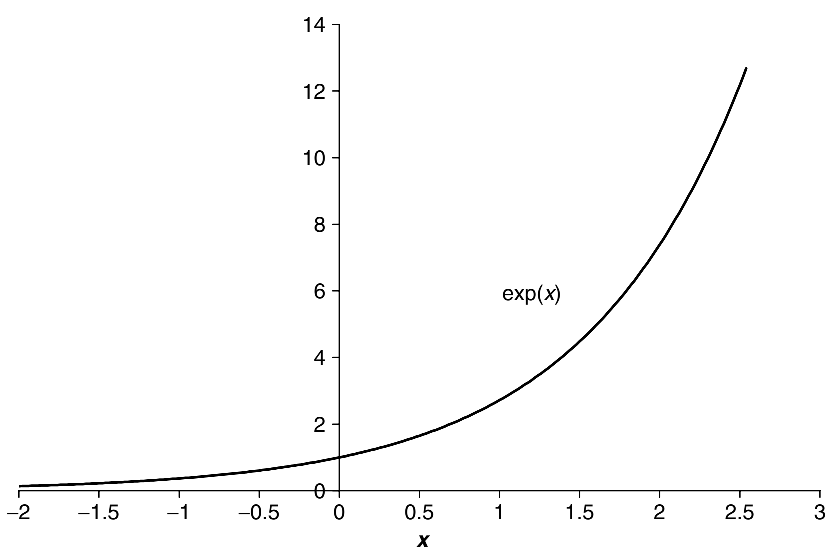


Figure A.1 The function e^x.

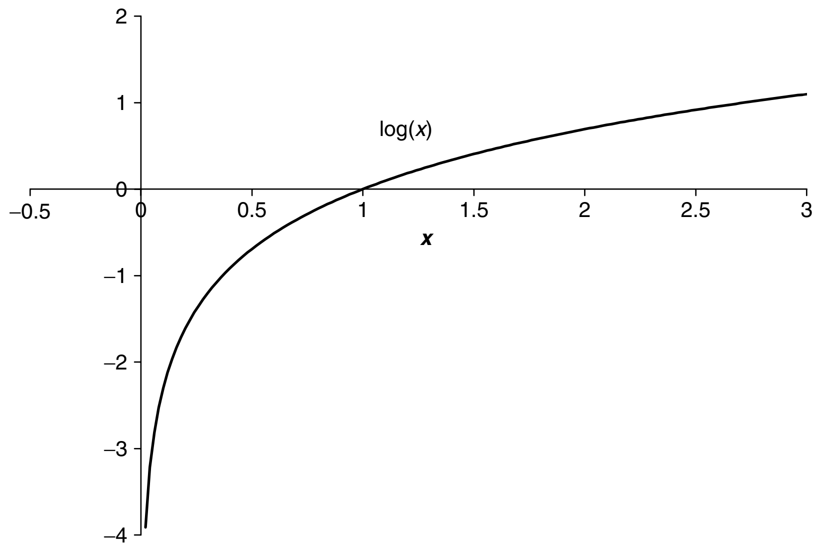


Figure A.2 The function $\log x$.

A.4 log

Take the plot of e^x in Figure A.1 and rotate it about a 45° line to get Figure A.2. This new function is $\ln x$, the Naperian logarithm of x. The relationship between ln and e is

$$e^{\ln x} = x \quad \text{or} \quad \ln(e^x) = x.$$

So, in a sense, they are inverses of each other.

The function $\ln x$ is also often denoted by $\log x$, as in this book. Sometimes $\log x$ refers to the function with the properties

$$10^{\log x} = x \quad \text{and} \quad \log(10^x) = x.$$

This function would be called 'logarithm base ten.' The most useful logarithm has base $e = 2.7183\ldots$ because of the properties of the gradient of e^x.

The slope of the $\log x$ function is x^{-1}.

From Figure A.2 you can see that there don't appear to be any values for $\log x$ for negative x. The function can be defined for these but you'd need to know about complex numbers, something we won't be requiring here.

A.5 DIFFERENTIATION AND TAYLOR SERIES

I've introduced the idea of a gradient or slope in the sections above. If we have a function denoted by $f(x)$, then we denote the gradient of this function at the point x by

$$\frac{df}{dx}.$$

Mathematically the slope is defined as

$$\frac{df}{dx} = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x}.$$

The action of finding the gradient is also called 'differentiating' and the slope can also be called the 'derivative' of the function. This use of 'derivative' shouldn't be confused with the use meaning an option contract.

The slope can also be differentiated, resulting in a second derivative of the function $f(x)$. This is denoted by

$$\frac{d^2 f}{dx^2}.$$

We can take this differentiation to higher and higher orders.

Take a look at Figure A.3. In particular, note the two dots marked on the bold curve. The bold curve is the function $f(x)$. The dot on the left is at the point x on the horizontal axis and the function value is $f(x)$, the distance up the vertical axis. The dot to the right of this is at $x + \delta x$ with function value $f(x + \delta x)$. What can we say about the vertical distance between the two dots in terms of the horizontal distance?

Start with a trivial example. If the distance δx is zero then the vertical distance is also zero. Now consider a very small but non-zero δx.

The straight line tangential to the bold curve $f(x)$ at the point x is shown in the figure. This line has slope df/dx evaluated at x. Notice that the right-hand hollow dot is almost on this bold line. This suggests that a good approximation to the value $f(x + \delta x)$ is

$$f(x + \delta x) \approx f(x) + \delta x \frac{df}{dx}(x).$$

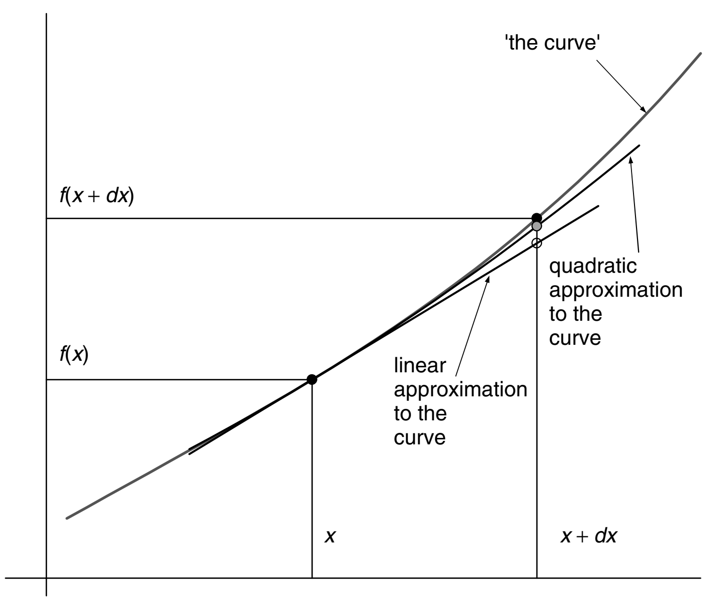


Figure A.3 A schematic diagram of Taylor series.

This is a linear relationship between $f(x+\delta x) - f(x)$ and δx. This makes sense since on rearranging we get

$$\frac{df}{dx} \approx \frac{f(x+\delta x) - f(x)}{\delta x}$$

which as δx goes to zero becomes our earlier definition of the gradient.

But the right-hand hollow dot is not *exactly* on the straight line. It is slightly above it. Perhaps a quadratic relationship between $f(x+\delta x) - f(x)$ and δx would be a more accurate approximation. This is indeed true (provided δx is small enough) and we can write

$$f(x+\delta x) \approx f(x) + \delta x \frac{df}{dx}(x) + \tfrac{1}{2}\delta x^2 \frac{d^2 f}{dx^2}(x).$$

This approximation, shown on the figure as the grey dot, is more accurate. One can take this approximation to cubic, quartic, ... The Taylor series representation of $f(x+\delta x)$ is the infinite sum

$$f(x+\delta x) = f(x) + \sum_{i=1}^{\infty} \frac{1}{i!}\delta x^i \frac{d^i f}{dx^i}(x).$$

Taylor series is incredibly useful in derivatives theory, where the function that we are interested in, instead of being f, is V, the value of an option. The independent variable is no longer x but is S, the price of the underlying asset. From day to day the asset price changes by a small, random amount. This asset price change is just δS (instead of δx). The first derivative of the option value with respect to the asset is known as the delta, and the second derivative is the gamma.

The value of an option is not only a function of the asset price S but also the time t: $V(S,t)$. This brings us into the world of partial differentiation.

Think of the function $V(S,t)$ as a surface with coordinates S and t on a horizontal plane. The *partial* derivative of $V(S,t)$ with respect to S is written

$$\frac{\partial V}{\partial S}$$

and is defined as

$$\frac{\partial V}{\partial S} = \lim_{\delta S \to 0} \frac{V(S+\delta S, t) - V(S,t)}{\delta S}.$$

Note that in this V is only ever evaluated at time t. This is like measuring the gradient of the function $V(S,t)$ in the S direction along a constant value of t.[1] The partial derivative of $V(S,t)$ with respect to t is similarly defined as

$$\frac{\partial V}{\partial t} = \lim_{\delta t \to 0} \frac{V(S, t+\delta t) - V(S,t)}{\delta t}.$$

Higher-order derivatives are defined in the obvious manner.

[1] Or think of it as the slope of a hill going North. The time derivative would be the slope going West.

The Taylor series expansion of the value of an option is then

$$V(S+\delta S, t+\delta t) \approx V(S,t) + \delta t \frac{\partial V}{\partial t} + \delta S \frac{\partial V}{\partial S} + \tfrac{1}{2}\delta S^2 \frac{\partial^2 V}{\partial S^2} + \cdots.$$

This series goes on for ever, but I've only written down the largest and most important terms, those which are required for the Black–Scholes analysis.

A.6 EXPECTATION AND VARIANCE

Much of the modeling in finance uses ideas from probability theory. Again you don't need to know that much to understand most of the theory.

The first important idea is that of expectation or mean. If you roll a die there is an equal, $\frac{1}{6}$, probability of each number coming up. What is the expected number or the average number if you roll the die many times. The answer is

$$1 \times \tfrac{1}{6} + 2 \times \tfrac{1}{6} + 3 \times \tfrac{1}{6} + 4 \times \tfrac{1}{6} + 5 \times \tfrac{1}{6} + 6 \times \tfrac{1}{6} = 3\tfrac{1}{2}.$$

Here we just multiply each of the possible numbers that could turn up by the probability of each, and sum. Although $3\frac{1}{2}$ is the expected value, it cannot, of course, be thrown since only integers are possible.

Generally, if we have a random variable X (the number thrown, say) which can take any of the values x_i (1, 2, 3, 4, 5, 6 in our example) for $i = 1, \ldots, N$ each of which has a probability $P(X = x_i)$ (in the example, $\frac{1}{6}$) then the expected value is

$$E[X] = \sum_{i=1}^{N} x_i P(X = x_i).$$

Expectations have the following properties:

$$E[cX] = cE[X]$$

and

$$E[X+Y] = E[X] + E[Y].$$

If the outcome of two random events X and Y have no impact on each other they are said to be independent. If X and Y are independent we have

$$E[XY] = E[X]E[Y].$$

Expectations are important in finance because we often want to know what we can expect to make from an investment on average.

The expectation or mean is also known as the first moment of the distribution of the random variable X. It can be thought of as being a typical value for X. The scatter of values around the mean can be measured by the second moment or the variance:

$$\mathrm{Var}(X) = E\left[(X - E[X])^2\right].$$

Variances have the following property:

$$\mathrm{Var}(cX) = c^2 \mathrm{Var}(X).$$

When X and Y are independent

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

The standard deviation is the square root of the variance and is perhaps more useful as a measure of dispersion since it has the same units as the variable X:

$$\text{Standard deviation}(X) = \sqrt{\mathrm{Var}(X)}.$$

If the standard deviation is small then values of X are concentrated around the mean, $E[X]$. If the standard deviation is large then values of X are more widely scattered.

Standard deviations are important in finance because they are often used as a measure of risk in an investment. The higher the standard deviation of investment returns the greater the dispersion of the returns and the greater the risk.

A.7 ANOTHER LOOK AT BLACK–SCHOLES

Now that we understand about differentiation we can take another look at the Black–Scholes equation:

$$\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

The option value $V(S, t)$ depends on (or 'is a function of') the asset price S and the time t.

The first derivative of the option value with respect to time is called the theta:

$$\Theta = \frac{\partial V}{\partial t}.$$

Notice that this is a partial derivative and so theta is the gradient of the option value in the direction of changing time, asset price fixed. It measures the rate of change of the option value with time if the asset price doesn't move, hence the other name 'time decay.'

The first derivative of the option value with respect to the asset price is called the delta:

$$\Delta = \frac{\partial V}{\partial S}.$$

This is the slope in the S direction with time fixed. Asset prices change very rapidly and so the dominant change in the option value from moment to moment is the delta multiplied by the change in the asset price. This is just a simple application of Taylor series; the difference between the option price at time t when the asset is at S and a later time $t + \delta t$ when the asset price is $S + \delta S$ is given by

$$V(S + \delta S, t + \delta t) - V(S, t) = \Delta\, \delta S + \cdots.$$

The $\cdots$ are terms which are, generally speaking, smaller than the leading term. They are still important, as we'll see in a moment.

Because the change in option value and the change in asset price are so closely linked we can see that holding a quantity Δ of the underlying asset short we can eliminate, to leading order, fluctuations in our net portfolio value. This is the basis of delta hedging.

The second derivative of the option value with respect to the asset price is called the gamma:

$$\Gamma = \frac{\partial^2 V}{\partial S^2}.$$

This is also just the S derivative of the delta. If the asset changes by an amount δS then the delta changes by an amount $\Gamma\ \delta S$. Thus the gamma is a measure of how much one might have to rehedge, and gives a measure of the amount of transaction costs from delta hedging.

Now we can interpret all the terms in the Black–Scholes equation, but what does the equation itself mean?

Written in terms of the greeks, the Black–Scholes equation is

$$\Theta + \tfrac{1}{2}\sigma^2 S^2 \Gamma + rS\Delta - rV = 0.$$

Reordering this we have

$$\Theta = rV - rS\Delta - \tfrac{1}{2}\sigma^2 S^2 \Gamma = r(V - S\Delta) - \tfrac{1}{2}\sigma^2 S^2 \Gamma.$$

When we have a delta hedged position we hold the option with value V and are short Δ of the underlying asset. Thus our portfolio value is at any time

$$V - S\Delta.$$

So we can write the Black–Scholes equation in words as

$$\text{Time decay} = (\text{interest received on cash equivalent of portfolio value}) - \tfrac{1}{2}\sigma^2 S^2 \text{ gamma}.$$

The option value grows by an equivalent of interest that would have been received by a riskless pure cash position. But the delta hedged option is not a cash position. That's where the final, gamma, term comes in.

Ignoring the interest on the cash equivalent, the theta and gamma terms add up to zero. Of course, you can't ignore this interest unless the portfolio has zero value or rates are zero.

The delta hedge is only accurate to leading order. If one is hedging with finite time intervals between rehedges then there is inevitably a little bit of randomness that we can't hedge away. We can see this if we go to higher order in the Taylor series expansion of $V(S + \delta S, t + \delta t)$:

$$V(S + \delta S, t + \delta t) - V(S, t) = \Delta\ \delta S + \Theta\ \delta t + \tfrac{1}{2}\Gamma\ \delta S^2 \ldots.$$

The Θ term is predictable if we know the time δt between hedges (and it has already appeared in the Black–Scholes equation). But the Γ term is multiplied by the random δS^2. We can't hedge this away perfectly. It is, in practice, the source of hedging errors. However, if we rehedge sufficiently frequently (i.e. δt is very small) then the combined effect of the gamma terms is via an *average* of the δS^2. And this average is $\sigma^2 S^2 \delta t$. Why is it the average that matters? It's like betting on the toss of a biased coin. If you have an advantage then you can exploit it by betting a small amount but very, very often. In the long run you will certainly win. (In terms of standard deviations, as the time between hedges decreases so does the standard deviation of the hedging error accumulated over the life of the option.)

We can now see that the gamma term in the Black–Scholes equation is to balance the higher-order fluctuations in the option value. Naturally, it therefore depends on the magnitude of these fluctuations, the volatility of the underlying asset.

A.8 SUMMARY

That wasn't hard, was it?

bibliography

Essential Books for Your Derivatives Library

Some of these books are for daily reference, others are insightful, others are simply 'classics.' Some are pure finance, some mathematics and some in between. They are my own favorites.

Abensur, N 1996 *The New Cranks Recipe Book*. Phoenix Illustrated

Ahmad, R 2006 *Numerical and Computational Methods for Derivative Pricing*. John Wiley & Sons

Alexander, CO 2001 *Market Models*. John Wiley & Sons

Cox, DR & Miller, HD 1965 *The Theory of Stochastic Processes*. Chapman & Hall

Cox, J & Rubinstein, M 1985 *Options Markets*. Prentice–Hall

Crank, JC 1989 *Mathematics of Diffusion*. Oxford

Derman, E 2004 *My Life As A Quant*. John Wiley & Sons

Dixit, AK & Pindyck, RS 1994 *Investment Under Uncertainty*. Princeton

Elton, EJ & Gruber, MJ 1995 *Modern Portfolio Theory and Investment Analysis*. John Wiley & Sons

Haug, EG 1997 *The Complete Guide to Option Pricing Formulas*. McGraw–Hill

Hull, J 1997 *Options, Futures and Other Derivative Securities*. Prentice–Hall

Ingersoll, JE Jr 1987 *Theory of Financial Decision Making*. Rowman & Littlefield

Jäckel, P 2002 *Monte Carlo Methods in Finance*. John Wiley & Sons

Jackson, M & Staunton, M 2001 *Advanced Modelling in Finance Using Excel and VBA*. John Wiley & Sons

Lewis, A 2000 *Option Valuation under Stochastic Volatility*. Finance Press

Malkiel, BG 1990 *A Random Walk Down Wall Street*. Norton

Mandelbrot, B & Hudson, R 2004 *The (Mis)Behaviour of Markets: A Fractal View of Risk, Ruin and Reward*. Profile Books

Markowitz, H 1959 *Portfolio Selection: Efficient Diversification of Investment*. John Wiley & Sons

Merton, RC 1992 *Continuous-time Finance*. Blackwell

Morton, KW & Mayers, DF 1994 *Numerical Solution of Partial Differential Equations*. Cambridge

Neftci, S 1996 *An Introduction to the Mathematics of Financial Derivatives*. Academic Press

Øksendal, B 1992 *Stochastic Differential Equations*. Springer–Verlag

Osband, K 2002 *Iceberg Risk: An Adventure in Portfolio Theory*. Texere

Oxford Financial Press www.oxfordfinancial.co.uk

Partnoy, F 1998 *F.I.A.S.C.O.* Profile Books

Poundstone, W 2005 *Fortune's Formula*. Hill & Wang

Press, WH, Teutolsky, SA, Vetterling, WT & Flannery, BP 1992 *Numerical Recipes in C*. Cambridge

Samuelson, P 1948 *Economics*. McGraw–Hill

Schönbucher, PJ 2003 *Credit Derivatives Pricing Models*. John Wiley & Sons

Schuss, Z 1980 *Theory and Applications of Stochastic Differential Equations*. John Wiley & Sons

Schwager, JD 1990 *Market Wizards*. HarperCollins

Schwager, JD 1992 *New Market Wizards*. HarperCollins

Seuss, Dr 1999 *The Cat in the Hat*. HarperCollins

Sharpe, WF 1985 *Investments*. Prentice–Hall

Smith, GD 1985 *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford

Taleb, NN 1997 *Dynamic Hedging*. John Wiley & Sons

Taleb, NN 2001 *Fooled by Randomness*. Texere

Thorp, EO 1962 *Beat the Dealer*. Vintage

Wilmott, P, Dewynne, J & Howison, SD 1993 *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press www.oxfordfinancial.co.uk

Wong, S 1981 *Professional Blackjack*. Pi Yee Press www.bj21.com

Other Books and Key Research Articles

Abboud, T & Zhang, Y 2004 Accurate early exercise free boundaries for American puts. *Wilmott* magazine January 2004 98–102

Ahn, H, Arkell, R, Choe, K, Holstad, E & Wilmott, P 1999 Optimal static vega hedge. MFG Working Paper, Oxford University

Ahn, H, Danilova, A & Swindle, G 2002 Storing arb. *Wilmott* magazine September 2002

Ahn, H, Dayal, M, Grannan, E & Swindle, G 1998 Option replication with transaction costs: general diffusion limits. *Annals of Applied Probability* **8** (3) 676–707

Ahn, H, Hua, P, Penaud, A & Wilmott, P 1999 Compensating traders and bonus maximization. Wilmott Associates Working Paper

Ahn, H, Muni, A, & Swindle, G 1997 Misspecified asset price models and robust hedging strategies. *Applied Mathematical Finance* **4** (1) 21–36

Ahn, H, Muni, A, & Swindle, G 1999 Optimal hedging strategies for misspecified asset price models. *Applied Mathematical Finance* **6** (3) 197–208

Ahn H, Khadem, V & Wilmott, P 1998 On the utility of risky bonds. MFG Working Paper, Oxford University

Ahn, H, Penaud, A & Wilmott, P 1998 Various passport options and their valuation. MFG Working Paper, Oxford University

Ahn, H & Wilmott, P 1998 On trading American options. MFG Working Paper, Oxford University

Ahn, H & Wilmott, P 2003a On exercising American options: the risk of making more money than you expected. *Wilmott* magazine March 2003 52–63

Ahn, H & Wilmott, P 2003b Stochastic volatility and mean-variance analysis. *Wilmott* magazine November 2003 84–90

Alexander, CO 1994 History debunked. *Risk* magazine **7** (12) 59–63

Alexander, CO 1995 Volatility and correlation forecasts. *Derivatives Week* (August)

Alexander, CO 1996a Evaluating the use of Riskmetrics as a risk measurement tool for your operation. *Derivatives: Use Trading and Regulation* **2** (3) 277–285

Alexander, CO 1996b Estimating and forecasting volatility and correlation: methods and applications. *Financial Derivatives and Risk Management* **7** (September) 64–72

Alexander, CO 1996c Volatility and correlation forecasting. In the *Handbook of Risk Management and Analysis* (C. Alexander, Ed) John Wiley & Sons 233–260

Alexander, CO 1997a Splicing methods for generating large covariance matrices. *Derivatives Week* (June)

Alexander, CO 1997b Estimating and forecasting volatility and correlation: methods and applications. In *Risk Management and Financial Derivatives: A Guide to the Mathematics* (S. Das, Ed) 337–354

Alexander, CO 1998 *The Handbook of Risk Management and Analysis*. John Wiley & Sons

Alexander, CO & Chibuma, A 1997 Orthogonal GARCH: An empirical validation in equities, foreign exchange and interest rates. Working Paper, Sussex University

Alexander, CO & Giblin, I 1997 Multivariate embedding methods: Forecasting high-frequency data in the first INFFC. Proceedings of the First International Nonlinear Financial Forecasting Competition, Finance and Technology Publishing

Alexander, CO & Johnson, A 1992 Are foreign exchange markets really efficient? *Economics Letters* **40** 449–453

Alexander, CO & Johnson, A 1994 Dynamic Links. *Risk* magazine **7** (2) 56–61

Alexander, CO & Leigh, C 1997 On the covariance matrices used in VAR models. *Journal of Derivatives* **4** (3) 50–62

Alexander, CO & Riyait, N 1992 The world according to GARCH. *Risk* magazine **5** (8) 120–125

Alexander, CO & Thillainathan, R 1996 The Asian Connections. *Emerging Markets Investor* **2** (6) 42–47

Alexander, CO & Williams, P 1997 Modelling the term structure of kurtosis: A comparison of neural network and GARCH methods. Working Paper, Sussex University

Amram, M & Kulatilaka, N 1999 *Real Options*. Harvard Business School Press

Andreasen, J, Jensen, B & Poulson, R 1998 Eight valuation methods in financial mathematics: the Black–Scholes formula as an example. *Mathematical Scientist* **23** 18–40

Angus, J 1999 A note on pricing Asian derivatives with continuous geometric averaging. *Journal of Future Markets* **19** (7) 845–858

Apabhai, MZ 1995 Term structure modelling and the valuation of yield curve derivative securities. D.Phil. thesis, Oxford University

Apabhai, MZ, Choe, K, Khennach, F & Wilmott, P 1995 Spot-on modelling. *Risk* magazine, December **8** (11) 59–63

Apabhai, MZ, Georgikopoulos, NI, Hasnip, D, Jamie, RKD, Kim, M & Wilmott, P 1998 A model for the value of a business, some optimisation problems in its operating procedures and the valuation of its debt. *IMA Journal of Applied Mathematics* **60** 1–13

Arditti, FD 1996 *Derivatives*. Harvard Business School Press

Artzner, P, Delbaen, F, Eber, J-M & Heath, D 1997 Thinking coherently. *Risk* magazine **10** (11) 68–72 (November)

Atkinson, C & Al-Ali, B 1997 On an investment-consumption model with transaction costs: an asymptotic analysis. *Applied Mathematical Finance* **4** 109–133

Atkinson, C & Wilmott, P 1993 Properties of moving averages of asset prices. *IMA Journal of Mathematics in Business and Industry* **4** 331–341

Atkinson, C & Wilmott, P 1995 Portfolio management with transaction costs: an asymptotic analysis. *Mathematical Finance* **5** 357–367

Atkinson, C, Pliska, S & Wilmott, P 1997 Portfolio management with transaction costs. *Proceedings of the Royal Society A*

Avellaneda, M & Buff, R 1997 Combinatorial implications of nonlinear uncertain volatility models: the case of barrier options. Courant Institute, NYU

Avellaneda, M, Friedman, C, Holmes, R & Samperi D, 1997 Calibrating volatility surfaces via relative-entropy minimization. *Applied Mathematical Finance* **4** 37–64

Avellaneda, M, Levy, A & Parás, A 1995 Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance* **2** 73–88

Avellaneda, M & Parás, A 1994 Dynamic hedging portfolios for derivative securities in the presence of large transaction costs. *Applied Mathematical Finance* **1** 165–194

Avellaneda, M & Parás, A 1996 Managing the volatility risk of derivative securities: the Lagrangian volatility model. *Applied Mathematical Finance* **3** 21–53

Ayache, E, Forsyth, PA & Vetzal, KR 2002 Next generation models for convertible bonds with credit risk. *Wilmott* magazine December 2002

Ayache, E, Henrotte, P, Nassar, S & Wang, X 2004 Can anyone solve the smile problem? *Wilmott* magazine January 2004 78–97

Babbs, S 1992 Binomial valuation of lookback options. Midland Montagu Working Paper

Baker, CTH 1977 *The Numerical Treatment of Integral Equations*. Oxford University Press

Bakstein, D & Wilmott, P 1999 Equity dividend models. Wilmott Associates Working Paper

Ball, C & Roma, A 1994 Stochastic volatility option pricing. *Journal of Financial and Quantitative Analysis* **29** 589–607

Banerjee, P 2003 Close-form pricing of plain and partial outside double barrier options. *Wilmott* magazine November 2003 46–49

Barles, G, Burdeau, J, Romano, M & Samsen, N 1995 Critical stock price near expiration. *Mathematical Finance* **5** 77–95

Barone-Adesi, G & Whaley, RE 1986 The valuation of American call options and the expected exdividend stock price decline. *Journal of Financial Economics* **17** 91–111

Barone-Adesi, G & Whaley, RE 1987 Efficient analytic approximation of American option values. *Journal of Finance* **41** 301–320

Barrett, JW, Moore, G & Wilmott, P 1992 Inelegant efficiency. *Risk* magazine **5** (9) 82–84

Beletski, T & Korn, R 2005 Optimal Investment with inflation-linked products. Working Paper, Universität Kairserslautern

Bergman, YZ, 1985 Pricing path contingent claims. *Research in Finance* **5** 229–241

Bergman, YZ, 1995 Option pricing with differential interest rates. *Review of Financial Studies* **8** 475–500

Bernstein PL 1998 *Against the Gods*. John Wiley & Sons

Berrahoui, M 2004 Pricing CMS spread options and digital CMS spread options with a smile. *Wilmott* magazine May 2004 63–69

Bhamra, HS 2000 Imitation in financial markets. *Int. J. Theor. Appl. Fin.* **3** 473–478

Black, F 1976 The pricing of commodity contracts. *Journal of Financial Economics* **3** 167–179

Black, F & Cox, J 1976 Valuing corporate securities: some effects of bond indenture provisions. *Journal of Finance* **31** 351–367

Black, F, Derman, E & Toy, W 1990 A one-factor model of interest rates and its application to Treasury bond options. *Financial Analysts Journal* **46** 33–39

Black, F & Scholes, M 1973 The pricing of options and corporate liabilities. *Journal of Political Economy* **81** 637–659

Blauer, I & Wilmott, P 1998 Risk of default in Latin American Brady bonds. *Net Exposure* **5** `www.netexposure.co.uk`

Bloch, D 1995 One-factor inflation rate modelling. M.Phil. dissertation, Oxford University

Bollerslev, T 1986 Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* **31** 307–327

Bouchaud, J-P, Potters, M & Cornalba, L 2002 Option pricing and hedging with temporal correlations. *Int. J. Theor. Appl. Fin.* **5** 1–14

Bowie, J & Carr, P 1994 Static Simplicity. *Risk* magazine **7** 45–49

Boyle, P 1977 Options: a Monte Carlo approach. *Journal of Financial Economics* **4** 323–338

Boyle, P 1991 Multi-asset path-dependent options. FORC conference, Warwick

Boyle, P, Broadie, M & Glasserman, P 1995 Monte Carlo methods for security pricing. Working Paper, University of Waterloo

Boyle, P & Emanuel, D 1980 Discretely adjusted option hedges. *Journal of Financial Economics* **8** 259–282

Boyle, P, Evnine, J & Gibbs, S 1989 Numerical evaluation of multivariate contingent claims. *Review of Financial Studies* **2** 241–250

Boyle, P & Tse, Y 1990 An algorithm for computing values of options on the maximum or minimum of several assets. *Journal of Financial and Quantitative Analysis* **25** 215–227

Boyle, P & Vorst, T 1992 Option replication in discrete time with transaction costs. *Journal of Finance* **47** 271

Brace, A, Gatarek, D & Musiela, M 1997 The market model of interest rate dynamics. *Mathematical Finance* **7** 127–154

Brandt, M & Jones, C 2002 Volatility Forecasting with Range-Based EGARCH Models. Working Paper, Fuqua

Brennan, M & Schwartz, E 1977 Convertible bonds: valuation and optimal strategies for call and conversion. *Journal of Finance* **32** 1699–1715

Brennan, M & Schwartz, E 1978 Finite-difference methods and jump processes arising in the pricing of contingent claims: a synthesis. *Journal of Financial and Quantitative Analysis* **13** 462–474

Brennan, M & Schwartz, E 1982 An equilibrium model of bond pricing and a test of market efficiency. *Journal of Financial and Quantitative Analysis* **17** 301–329

Brennan, M & Schwartz, E 1983 Alternative methods for valuing debt options. *Finance* **4** 119–138

Brenner, M & Subrahmanyam, MG 1994 A simple approach to option valuation and hedging in the Black–Scholes model. *Financial Analysts Journal* 25–28

Briys, EL, Mai, HM, Bellalah, MB & de Varenne, F 1998 *Options, Futures and Exotic Derivatives*. John Wiley & Sons

Brooks, M 1967 *The Producers*. MGM

Carr, P 1994 European put-call symmetry. Cornell University Working Paper

Carr, P 1995 Two extensions to barrier option pricing. *Applied Mathematical Finance* **2** 173–209

Carr, P 2005 FAQs in option pricing theory. *J. Derivatives*

Carr, P & Chou, A 1997 Breaking barriers. *Risk* magazine **10** 139–144

Carr, P, Ellis, K & Gupta, V 1998 Static hedging of exotic options. *Journal of Finance* 1165–1190

Carr, P & Madan, D 1999 Introducing the covariance swap. *Risk* magazine February 47–51

Carr, P & Verma, A 2005 Potential problems in market models of implied volatility. Working Paper

Carslaw, HS & Jaeger, JC 1989 *Conduction of Heat in Solids*. Oxford

Chan, K, Karolyi, A, Longstaff, F & Sanders, A 1992 An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* **47** 1209–1227

Chance, D 1990 Default risk and the duration of zero-coupon bonds. *Journal of Finance* **45** (1) 265–274

Cherubini, U, Luciano, E & Vecchiato, W 2004 *Copula Methods in Finance*. John Wiley & Sons

Chesney, M, Cornwall, J, Jeanblanc-Picqué, M, Kentwell, G & Yor, M 1997 Parisian pricing. *Risk* magazine **1** (1) 77–80

Chew, L 1996 *Managing Derivative Risk: The Use and Abuse of Leverage*. John Wiley & Sons

Clark, E 2002 Measuring country risk as implied volatility. *Wilmott* magazine September 2002

Connolly, KB 1997 *Buying and Selling Volatility*. John Wiley & Sons

Conze, A & Viswanathan 1991 Path-dependent options—the case of lookback options. *Journal of Finance* **46** 1893–1907

Cooper, I & Martin, M 1996 Default risk and derivative products. *Applied Mathematical Finance* **3** 53–74

Copeland, T, Koller, T & Murrin, J 1990 *Valuation: Measuring and Managing the Value of Companies*. John Wiley & Sons

Cox, J, Ingersoll, J & Ross, S 1980 An analysis of variable loan contracts. *Journal of Finance* **35** 389–403

Cox, J, Ingersoll, J & Ross, S 1981 The relationship between forward prices and futures prices. *Journal of Financial Economics* **9** 321–346

Cox, J, Ingersoll, J & Ross, S 1985 A theory of the term structure of interest rates. *Econometrica* **53** 385–467

Cox, JC, Ross, S & Rubinstein M 1979 Option pricing: a simplified approach. *Journal of Financial Economics* **7** 229–263

Crank, J & Nicolson, P 1947 A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type. *Proceedings of the Cambridge Philosophical Society* **43** 50–67

Curran, M 1992 Beyond average intelligence. *Risk* (October) 60

Das, S 1994 *Swaps and Financial Derivatives*. IFR

Das, S 1995 Credit risk derivatives. *Journal of Derivatives* **2** 7–23

Das, S & Tufano, P 1994 Pricing credit-sensitive debt when interest rates, credit ratings and credit spreads are stochastic. Working Paper, Harvard Business School Press

Davidson, AS & Herskovitz, MD 1996 *The Mortgage-backed Securities Handbook*. McGraw–Hill

Davis, MHA & Norman, AR 1990 Portfolio selection with transaction costs. *Mathematics of Operations Research* **15** 676–713

Davis, MHA, Panas, VG & Zariphopoulou, T 1993 European option pricing with transaction costs. *SIAM Journal of Control and Optimization* **31** 470–93

Demeterfi, K, Derman, E, Kamal, M & Zou, J 1999 Everything you ever wanted to know about volatility swaps. Goldman Sachs Quantitative Strategies Research Notes

Derman, E 1999 Regimes of volatility. Goldman Sachs Quantitative Strategies Technical Notes January 1999.

Derman, E & Kani, I 1994 Riding on a smile. *Risk* magazine **7** (2) 32–39

Derman, E, Ergener, D & Kani, I 1997 Static options replication. In *Frontiers in Derivatives.* (Konishi, A & Dattatreya, RE, Ed) Irwin

Derman, E & Kani, I 1997 Stochastic implied trees: arbitrage pricing with stochastic term and strike structure of volatility. Goldman Sachs Quantitative Strategies Technical Notes April 1997.

Derman, E & Zou, J 1997 Predicting the response of implied volatility to large index moves. Goldman Sachs Quantitative Strategies Technical Notes November 1997.

Dewynne, JN, Erhlichman, S & Wilmott, P 1998 A simple volatility surface parametrization. MFG Working Paper, Oxford University

Dewynne, JN, Whalley, AE & Wilmott, P 1994 Path-dependent options and transaction costs. *Philosophical Transactions of the Royal Society A*. **347** 517–529

Dewynne, JN, Whalley, AE & Wilmott, P 1995 Mathematical models and partial differential equations in finance. In *Quantitative Methods, Super Computers and AI in Finance* (Zenios, S. Ed) 95–124

Dewynne, JN & Wilmott, P 1993 Partial to the exotic. *Risk* magazine **6** (3) 38–46

Dewynne, JN & Wilmott, P 1994a Exotic financial options. *Proceedings of the 7th European Conference on Mathematics in Industry* 389–397

Dewynne, JN & Wilmott, P 1994b Modelling and numerical valuation of lookback options. MFG Working Paper, Oxford University

Dewynne, JN & Wilmott, P 1994c Untitled. MFG Working Paper, Oxford University

Dewynne, JN & Wilmott, P 1995a A note on American options with varying exercise price. *Journal of the Australian Mathematical Society* **37** 45–57

Dewynne, JN & Wilmott, P 1995b A note on average-rate options with discrete sampling. *SIAM Journal of Applied Mathematics* **55** 267–276

Dewynne, JN & Wilmott, P 1995c Asian options as linear complementarity problems: analysis and finite-difference solutions. *Advances in Futures and Options Research* **8** 145–177

Dewynne, JN & Wilmott, P 1996 Exotic options: mathematical models and computation. In *Frontiers in Derivatives* (Konishi, A and Dattatreya, R, Ed) 145–182

Dothan, MU 1978 On the term structure of interest rates. *Journal of Financial Economics* **6** 59–69

Duffee, G 1995 The variation of default risk with treasury yields. Working Paper, Federal Reserve Board, Washington

Duffie, D 1992 *Dynamic Asset Pricing Theory*. Princeton

Duffie, D & Harrison, JM 1992 Arbitrage pricing of Russian options and perpetual lookback options. *Annals of Applied Probability* **3** 641–651

Duffie, D, Ma, J & Yong, J 1994 Black's consol rate conjecture. Working Paper, Stanford

Duffie, D & Singleton, K 1994a Modeling term structures of defaultable bonds. Working Paper, Stanford

Duffie, D & Singleton, K 1994b An econometric model of the term structure of interest rate swap yields. Working Paper, Stanford

Duffy, D 2004 A critique of the Crank–Nicolson scheme. *Wilmott* magazine July 2004 68–77

Dumas, B, Fleming, J & Whaley, RE 1998 Implied volatility functions: empirical tests. *Journal of Finance* **53** (6) 2059–2106

Dunbar, N 1998 Meriwether's meltdown. *Risk* magazine **10** 32–36

Dupire, B 1993 Pricing and hedging with smiles. *Proc AFFI Conf*, La Baule June 1993

Dupire, B 1994 Pricing with a smile. *Risk* magazine **7** (1) 18–20

Eberlein, E & Keller, U 1995 Hyperbolic distributions in finance. *Bernoulli* **1** 281–299

El Karoui, Jeanblanc-Picqué, & Viswanathan 1991 Bounds for options. Lecture notes in *Control and Information Sciences* **117** 224–237, Springer–Verlag

Embrechts, P, Klüppelberg, C & Mikosch, T 1997 *Modelling Extremal Events*. Springer–Verlag

Engle, R 1982 Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007

Engle, R 1995 (ed.) *ARCH: Selected Readings*. Oxford

Engle, R & Bollerslev, T 1987 Modelling the persistence of conditional variances. *Econometric Reviews* **5** 1–50

Engle, R & Granger, C 1987 Cointegration and error correction: representation, estimation and testing. *Econometrica* **55** 251–276

Engle, R & Mezrich, J 1996 GARCH for groups. *Risk* magazine **9** (8) 36–40

Epstein, D, Mayor, N, Schönbucher, PJ, Whalley, AE & Wilmott, P 1997a The valuation of a firm advertising optimally. MFG Working Paper, Oxford University

Epstein, D, Mayor, N, Schönbucher, PJ, Whalley, AE & Wilmott, P 1997b The value of market research when a firm is learning: option pricing and optimal filtering. MFG Working Paper, Oxford University

Epstein, D & Wilmott, P 1997 Yield envelopes. *Net Exposure* **2** August `www.netexposure.co.uk`

Epstein, D & Wilmott, P 1998 A new model for interest rates. *International Journal Theoretical and Applied Finance* **1** 195–226

Epstein, D & Wilmott, P 1999 A nonlinear non-probabilistic spot interest rate model. *Philosophical Transactions of the Royal Society* A **357** 2109–2117

Fabozzi, FJ 1996 *Bond Markets, Analysis and Strategies*. Prentice–Hall

Fabozzi, FJ & Kipnis, GM (Eds) 1984 *Stock Index Futures*. Dow Jones–Irwin

Fama, E 1965 The behaviour of stock prices. *Journal of Business* **38** 34–105

Farmer, JD & Joshi, S 2000 The price dynamics of common trading strategies. SFI Working Paper 00-12-069

Farrell, JL Jr 1997 *Portfolio Management*. McGraw–Hill

Finger, C 2004 Issues in pricing of synthetic CDOs. *RiskMetrics* Working Paper

Fitt, AD, Dewynne, JN & Wilmott, P 1994 An integral equation for the value of a stop-loss option. *Proceedings of the 7th European Conference on Mathematics in Industry* 399–405

Fouque, JP, Papanicolaou, G & Sircar, KR 1998 Asymptotics of a two-scale stochastic volatility model. In Equations aux derivees partielles et applications 517–526

Fouque, JP, Papanicolaou, G & Sircar, KR 1999 Financial modeling in a fast mean-reverting stochastic volatility environment. *Asia Pacific Financial Models* **6** (1) 37–48

Fouque, JP, Papanicolaou, G & Sircar, KR 2000 Mean-reverting stochastic volatility, *International Journal of Theoretical and Applied Finance* **3** 101–142

Fouque, JP, Papanicolaou, G, Sircar, KR & Solna, K 2004 Timing the smile. *Wilmott* magazine March 2004 59–65

Fong, G & Vasicek, O 1991, Interest rate volatility as a stochastic factor. Working Paper

Frey, R & Stremme, A 1995 Market volatility and feedback effects from dynamic hedging. Working Paper, Bonn

Gabillon, J 1995 Analyzing the forward curve. In *Managing Energy Risk*. Financial Engineering, London

Garman, M 1976 A general theory of asset valuation under diffusion state processes, Technical Report 50, Center for Research in Management Science, Berkeley

Garman, M & Klass, M 1980 On the estimation of security price volatilities from historical data. *Journal of Business* **53** 67–78

Garman, MB & Kohlhagen, SW 1983 Foreign currency option values. *Journal of International Money and Finance* **2** 231–237

Geman, H & Yor, M 1993 Bessel processes, Asian options and perpetuities. *Mathematical Finance* **3** 349–375

Gemmill, G 1992 *Options Pricing*. McGraw–Hill

Gerald, CF & Wheatley, PO 1992 *Applied Numerical Analysis*. Addison Wesley

Geske, R 1977 The valuation of corporate liabilities as compound options. *Journal of Financial and Quantitative Analysis* **12** 541–552

Geske, R 1978 Pricing of options with stochastic dividend yield. *Journal of Finance* **33** 617–625

Geske, R 1979 The valuation of compound options. *Journal of Financial Economics* **7** 63–81

Glasserman, P 2003 *Monte Carlo Methods in Financial Engineering*. Springer–Verlag

Goldman, MB, Sosin, H & Gatto, M 1979 Path dependent options: buy at the low, sell at the high. *Journal of Finance* **34** 1111–1128

Gould, SJ 1985 The median is not the message. *Discover* (June) 40–42

Grindrod, P 1991 *Patterns and Waves: The Theory and Applications of Reaction-diffusion Equations*. Oxford

Haber, R, Schönbucher, PJ & Wilmott, P 1997 Parisian options. MFG Working Paper, Oxford University

Hagan, P 2002 Adjusters: Turning good prices into great prices. *Wilmott* magazine December 2002

Hagan, P 2003 Convexity conundrums: Pricing CMS swaps, caps and floors. *Wilmott* magazine March 2003

Hagan, P & Woodward, DE 1999 Equivalent Black Volatilities. *Applied Mathematical Finance* **6** 147–157

Hagan, P, Kumar, D, Lesniewski, A & Woodward, D 2002 Managing smile risk. *Wilmott* magazine September 2002

d'Halluin, Y, Pooley, D & Forsyth, P 2004 No fear of jumps. *Wilmott* magazine January 2004 62–70

Hamermesh, DS & Soss, NM 1974 An economic theory of suicide. *Journal of Political Economy* **82** 83–90

Hamilton, JD 1994 *Time Series Analysis*. Princeton

Harrison, JM & Kreps, D 1979 Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* **20** 381–408

Harrison, JM & Pliska, SR 1981 Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications* **11** 215–260

Haselgrove, CB 1961 A method for numerical integration. *Mathematics of Computation* **15** 323–337

Haug, E 2003 Know your weapon. *Wilmott* magazine May and July 2003

Haug, E & Haug, J 2002 Knock-in/out Margrabe. *Wilmott* magazine December 2002

Haug, E, Haug, J & Margrabe, W 2003 Asian pyramid power. *Wilmott* magazine March 2003

Heath, D, Jarrow, R & Morton, A 1992 Bond pricing and the term structure of interest rates: a new methodology. *Econometrica* **60** 77–105

Hendry, DF 1995 *Dynamic Econometrics*. Oxford

Henrard, M 2001 Parameter risk in the Black and Scholes model. Working paper

Henrotte, P 1993 Transaction costs and duplication strategies. Working Paper, Stanford University

Heston, S 1993 A closed-form solution for options with stochastic volatility with application to bond and currency options. *Review of Financial Studies* **6** 327–343

Heynen, RC & Kat, HM 1995 Lookback options with discrete and partial monitoring of the underlying price. *Applied Mathematical Finance* **2** 273–284

Hilliard, JE & Reis, J 1998 Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot. *Journal of Financial and Quantitative Analysis* **33** 61–86

Ho, T & Lee, S 1986 Term structure movements and pricing interest rate contingent claims. *Journal of Finance* **42** 1129–1142

Hodges, SD & Neuberger, A 1989 Optimal replication of contingent claims under transaction costs. *The Review of Futures Markets* **8** 222–239

Hogan, M 1993 Problems in certain two-factor term structure models. *Annals of Applied Probability* **3** 576

Hoggard, T, Whalley, AE & Wilmott, P 1994 Hedging option portfolios in the presence of transaction costs. *Advances in Futures and Options Research* **7** 21–35

Hua, P 1997 Modelling stock market crashes. Dissertation, Imperial College, London

Hua, P & Wilmott, P 1997 Crash courses. *Risk* magazine **10** (6) 64–67

Hull, JC & White, A 1987 The pricing of options on assets with stochastic volatilities. *Journal of Finance* **42** 281–300

Hull, JC & White, A 1988 An analysis of the bias in option pricing caused by a stochastic volatility, *Advances in Futures and Options Research*, **3**, 29–61,

Hull, JC & White, A 1990 Pricing interest rate derivative securities. *Review of Financial Studies* **3** 573–592

Hull, JC & White, A 1990 Valuing derivative securities using the finite difference method. *Journal of Financial and Quantitative Analysis* **25** 87–100

Hull, JC & White, A 1996 Finding the keys. In *Over the Rainbow* (Jarrow, R, Ed) *Risk* magazine

Hughston, LP 1998 Inflation Derivatives. Unpublished Merrill Lynch and King's College London Working Paper

Hyer, T, Lipton-Lifschitz, A & Pugachevsky, D 1997 Passport to success. *Risk* magazine **10** (9) 127–132

Jackwerth, JC & Rubinstein, M 1996 Recovering probability distributions from contemporaneous security prices. *Journal of Finance* **51** 1611–1631

Jäckel, P 2003 Mid the cap. *Wilmott* magazine September 2003

Jäckel, P & Kawai, A 2005 The future is convex. *Wilmott* magazine January 2005 44–55

Jamshidian, F 1989 An exact bond option formula. *Journal of Finance* **44** 205–209

Jamshidian, F 1990 Bond and option evaluation in the Gaussian interest rate model. Working Paper, Merrill Lynch Capital Markets

Jamshidian, F 1991 Forward induction and construction of yield curve diffusion models. *Journal of Fixed Income*, June 62–74

Jamshidian, F 1994 Hedging, quantos, differential swaps and ratios. *Applied Mathematical Finance* **1** 1–20

Jamshidian, F 1995 A simple class of square-root interest-rate models. *Applied Mathematical Finance* **2** 61–72

Jamshidian, F 1996a Sorting out swaptions. *Risk* magazine **9**

Jamshidian, F 1996b Bonds, futures and options evaluation in the quadratic interest rate model. *Applied Mathematical Finance* **3** 93–115

Jamshidian, F 1997 LIBOR and swap market models and measures. *Finance and Stochastics* **1** 293–330

Jarrow, R, Lando, D & Turnbull, S 1997 A Markov model for the term structure of credit spreads. *Review of Financial Studies* **10** 481–523

Jarrow, R & Turnbull, S 1990 Pricing options on financial securities subject to credit risk. Working paper, Cornell University, Ithaca, NY.

Jarrow, R & Turnbull, S 1995 pricing derivatives on securities subject to credit risk. *Journal of Finance* **50** 53–85

Javaheri, A 2005 *Inside Volatility Arbitrage*. John Wiley & Sons

Jefferies, S 1999 Hedging equity and interest rate derivatives in the presence of transaction costs. M.Sc. dissertation, Oxford

Johnson, HE 1983 An analytical approximation to the American put price. *Journal of Financial and Quantitative Analysis* **18** 141–148

Johnson, HE 1987 Options on the maximum or minimum of several assets. *Journal of Financial and Quantitative Analysis* **22** 277–283

Johnson, LW & Riess, RD 1982 *Numerical Analysis*. Addison–Wesley

Johnson, NF, Hui, PM & Lo, TS 1999 Self-organized segregation of traders within a market. *Phil. Trans*. A **357** 2013–2018

Jordan, DW & Smith, P 1977 *Nonlinear Ordinary Differential Equations*. Oxford

Jorion, P 1997 *Value at Risk*. Irwin

Jorion, P 1999 Risk management lessons from Long-Term Capital Management. Working Paper, University of California at Irvine

Joshi, M 2003 *The Concepts and Practice of Mathematical Finance*. CUP

Kaplanis, C 1986 Options, taxes and ex-dividend day behaviour. *Journal of Finance* **41** 411–424

Kelly, FP, Howison, SD & Wilmott, P (Eds) 1995 *Mathematical Models in Finance*. Chapman and Hall

Kelly, JL 1956 A new interpretation of information rate. *Bell Systems Tech. J.* **35** 917-926

Kemna, AGZ & Vorst, ACF 1990 A pricing method for options based upon average asset values. *Journal of Banking and Finance* (March) **14** 113–129

Khasminskii, RZ & Yin, G 1996 Asymptotic series for singularly perturbed Kolmogorov–Fokker–Planck equations, *SIAM J. Appl. Math.* **56**, 1766–1793

Kim, M 1995 Modelling company mergers and takeovers. MSc thesis, Imperial College, London

Klein, A 2002 Determining the future of mortality. *Contingencies* September/October 39–43

Klugman, R 1992 pricing interest rate derivative securities. M.Phil. thesis, Oxford University

Klugman, R & Wilmott, P 1994 A class of one-factor interest rate models. *Proceedings of the 7th European Conference on Mathematics in Industry* 419–426

Kolman, J 1999 LTCM speaks. *Derivatives Strategy* (April) 12–17

Korn, R 1997 *Optimal Portfolios*. World Scientific

Korn, R 2003 Worst-case investment with applications for banks and insurance companies, in: Conference Proceedings, ERC-Conference 2003, METU Ankara

Korn, R 2005 Worst-case scenario investment for insurers, *Insurance Mathematics and Economics* **36**, 1–11

Korn, R & Menkens, O 2004a On worst-case investment with applications for banks and insurance companies, in: *Interacting Stochastic Systems* (Deuschel, JD & Greven, A, Ed), Springer, 397–408

Korn, R & Menkens, O 2004b Worst-case scenario optimization: A new stochastic control approach, to appear in: *Mathematical Methods of Operations Research*

Korn, R & Wilmott, P 1996 Room for a view. MFG Working Paper, Oxford University

Korn, R & Wilmott, P 1998 A general framework for hedging and speculating with options. *Int. J. Theor. Appl. Fin.* **1** 507–522

Korn, R & Wilmott, P 2002 Optimal portfolios under the threat of a crash. *Int. J. Th. Appl. Fin.* **5** 171–188

Krekel, M 2003 The pricing of Asian options on average spot with average strike. *Wilmott* magazine July 2003

Krekel, M, de Kock, J, Korn, R & Man T-K, 2004 An analysis of pricing methods for basket options. *Wilmott* magazine 2004 82–89

Kuske, RA & Keller, JB 1998 Optimal exercise boundary for an American put option. To appear in *Applied Mathematical Finance* **5** 107–116

Kyprianou, A, Schoutens, W & Wilmott P (Eds) 2005 *Exotic Options and Advanced Levy Models*. John Wiley & Sons

Lacoste, V 1996 Wiener chaos: a new approach to option pricing. *Mathematical Finance* **6** 197–213

Lando, D 1994a Three essays on contingent claims pricing. Ph.D. thesis, Graduate School of Management, Cornell University

Lando, D 1994b On Cox processes and credit risky bonds. Working Paper, Institute of Mathematical Statistics, University of Copenhagen

Lawrence, D 1996 *Measuring and Managing Derivative Market Risk*. International Thompson Business Press

Leeson, N 1997 *Rogue Trader*. Warner

Leland, HE 1985 Option pricing and replication with transaction costs. *Journal of Finance* **40** 1283–1301

Levy, E 1990 Asian arithmetic. *Risk* magazine **3** (5) (May) 7–8

Levy, E 1992 Pricing European average rate currency options. *Journal of International Money and Finance* **14** 474–491

Lewicki, P & Avellaneda, M 1996 Pricing interest rate contingent claims in markets with uncertain volatilities. CIMS Preprint

Lewis, A 2002 Asian connections. *Wilmott* magazine September 2002

Lewis, A 2002 Fear of jumps. *Wilmott* magazine December 2002

Lewis, A 2003a Perpetual American options made easy. *Wilmott* magazine January 2003

Lewis, A 2003b American options under jump diffusions: An introduction. *Wilmott* magazine March 2003

Lewis, A 2003c Barrier options under jump diffusions: The analytic theory. *Wilmott* magazine May and July 2003

Lewis, A 2003d American-style options with one-sided jumps. *Wilmott* magazine November 2003

Lewis, A 2004a Volatility jump diffusions. *Wilmott* magazine January 2004 46–53

Lewis, A 2004b Double barrier options with jumps – a simple universal algorithm. *Wilmott* magazine March 2004 34–37

Lewis, A 2004c Diffusions, jumps and the distribution of the maximum. *Wilmott* magazine May 2004 48–53

Lewis, A 2004d Asian option pricing under jump diffusions. *Wilmott* magazine July 2004 52–55

Lewis, A, Haug, J & Haug, E 2003 Back to basics: A new approach to the discrete dividend problem. *Wilmott* magazine. September 2003

Lewis, M 1989 *Liar's Poker*. Penguin

Li, D 2000 On Default Correlation: A Copula Function Approach. *RiskMetrics* Working Paper

Lillo, F, Mantegna, R, Bouchaud, J-P & Potters, M 2002 Introducing variety in risk management. *Wilmott* magazine December 2002

Litterman, R & Iben, T 1991 Corporate bond valuation and the term structure of credit spreads. *Financial Analysts Journal* (Spring) 52–64

Longstaff, FA & Schwartz, ES 1992 A two-factor interest rate model and contingent claims valuation. *Journal of Fixed Income* **3** 16–23

Longstaff, FA & Schwartz, ES 1994 A simple approach to valuing risky fixed and floating rate debt. Working Paper, Anderson Graduate School of Management, University of California, Los Angeles

Longstaff, FA & Schwartz, ES 2001 Valuing American options by simulation: A simple least-squares approach. *Rev. Fin. Studies* **14** 113–147

Lucic, V 2003 Forward start options in stochastic volatility model. *Wilmott* magazine September 2003

Lyons, TJ 1995 Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance* **2** 117–133

Macaulay, F 1938 Some theoretical problems suggested by movement of interest rates, bond yields and stock prices since 1856. National Bureau of Economic Research, New York

Madan, D & Seneta, E 1990 The variance gamma model for share price returns. *J. Bus.* **63** 511–524

Madan, D & Unal, H 1994 Pricing the risks of default. Working Paper, College of Business and Management, University of Maryland

Madan, D, Carr, P & Change, E 1998 The variance gamma process and option pricing. *Euro. Fin. Rev.* **2** 79–105

Majd, S & Pindyck, RS 1987 Time to build, option value, and investment decisions. *Journal of Financial Economics* **18** 7–27

Mandelbrot, B 1963 The variation of certain speculative prices. *Journal of Business* **36** 394–419

Margrabe, W 1978 The value of an option to exchange one asset for another. *Journal of Finance* **33** 177–186

Matten, C 1996 *Managing Bank Capital*. John Wiley & Sons

McDonald, R & Siegel, D 1986 The value of waiting to invest. *Quarterly Journal of Economics* **101** 707–728

McKee, S, Wall, D & Wilson, S 1996 An alternating direction implicit scheme for parabolic equations with mixed derivative and convective terms. *J. Comp. Phys.* **126** 64–76

McMillan, LG 1996 *McMillan on Options*. John Wiley & Sons

McNeil, A 1998 On Extremes and Crashes. *Risk* magazine January 99

Mercurio, F 2004 Pricing Inflation-Indexed Derivatives. Working Paper

Mercurio, F & Vorst, TCF 1996 Option pricing with hedging at fixed trading dates. *Applied Mathematical Finance* **3** 135–158

Merton, RC 1973 Theory of rational option pricing. *Bell Journal of Economics and Management Science* **4** 141–183

Merton, RC 1974 On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance* **29** 449–470

Merton, RC 1976 Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* **3** 125–144

Merton, RC 1995 Influence of mathematical models in finance on practice: past, present and future. In *Mathematical Models in Finance* (Howison, SD, Kelly, FP & Wilmott, P, Ed), Chapman & Hall

Meyers, TA 1994 *The Technical Analysis Course*. Probus

Mikhailov, S & Nögel, U 2003 Heston's stochastic volatility model: Implementation, calibration and some extensions. *Wilmott* magazine July 2003

Miller, M 1997 *Merton Miller on Derivatives*. John Wiley & Sons

Miller, M & Modigliani, F 1961 Dividend policy, growth and the valuation of shares. *Journal of Business* **34** 411–433

Miron, P & Swannell, P 1991 *Pricing and Hedging Swaps*. Euromoney Publications

Mitchell, AR & Griffiths, DF 1980 *The Finite Difference Method in Partial Differential Equations*. John Wiley & Sons

Mohamed, B 1994 Simulation of transaction costs and optimal rehedging. *Applied Mathematical Finance* **1** 49–62

Morton, A & Pliska, SR 1995 Optimal portfolio management with fixed transaction costs. *Mathematical Finance* **5** 337–356

Murphy, G 1995 Generalised methods of moments estimation of the short process in the UK. Bank of England Working Paper

Murray, J 1989 *Mathematical Biology*. Springer–Verlag

Naik, V 1993 Option valuation and hedging strategies with jumps in the volatility of asset returns. *Journal of Finance* **48** 1969–1984

Natenberg, S 1994 *Option Volatility and Pricing*. McGraw–Hill

Nelken, I 1999 *Implementing Credit Derivatives*. McGraw–Hill

Nelsen, R 1999 *An Introduction to Copulas*. Springer–Verlag

Nelson, D 1990 Arch models as diffusion approximations. *Journal of Econometrics* **45** 7–38

Neuberger, A 1994 Option replication with transaction costs: an exact solution for the pure jump process. *Advances in Futures and Options Research* **7** 1–20

Neuberger, A 1994 The log contract: a new instrument to hedge volatility. *Journal of Portfolio Management* **20** 74–80

Niederreiter, H 1992 *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM

Nielson, JA & Sandmann, K 1996 The pricing of Asian options under stochastic interest rates. *Applied Mathematical Finance* **3** 209–236

Ninomiya, S & Tezuka, S 1996 Toward real-time pricing of complex financial derivatives. *Applied Mathematical Finance* **3** 1–20

Nyborg, KG 1996 The use and pricing of convertible bonds. *Applied Mathematical Finance* **3** 167–190

O'Hara, M 1995 *Market Microstructure Theory*. Blackwell

Options Institute 1995 *Options: Essential Concepts and Trading Strategies*. Irwin

Osband, K 2004 Support and resistance for support/resistance. *Wilmott* magazine November 2004 58–61

Ouachani, N & Zhang, Y 2004 Pricing cross-currency convertible bonds with pde. *Wilmott* magazine January 2004 54–61

Owen, G 1995 *Game Theory*. Academic Press

Oztukel, A 1996 Uncertain parameter models. M.Sc. dissertation, Oxford University

Oztukel, A & Wilmott, P 1998 Uncertain parameters, an empirical stochastic volatility model and confidence limits. *International Journal of Theoretical and Applied Finance* **1** 175–189

Parkinson M 1980 The extreme value method for estimating the variance of the rate of return. *Journal of Business* **53** (1) 61–65

Paskov 1996 Mew methodologies for valuing derivatives. In *Mathematics of Derivative Securities* (Pliska, SR and Dempster, M, Eds)

Paskov, SH & Traub, JF 1995 Faster valuation of financial derivatives. *Journal of Portfolio Management*. Fall 113–120

Pearson, N & Sun, T–S 1989 A test of the Cox, Ingersoll, Ross model of the term structure of interest rates using the method of moments. Sloan School of Management, MIT

Penaud, A 2004 Fast valuation of a portfolio of barrier options under Merton's jump diffusion hypothesis. *Wilmott* magazine September 2004 88–91

Penaud, A & Selfe, J 2003 First to default swaps. *Wilmott* magazine January 2003

Penaud, A, Wilmott, P & Ahn, H 1998 Exotic passport options. MFG Working Paper, Oxford University

Peters, EE 1991 *Chaos and Order in the Capital Markets*. John Wiley & Sons

Peters, EE 1994 *Fractal Market Analysis*. John Wiley & Sons

Pilopović, D 1998 *Energy Risk*. McGraw–Hill

Porter, DP & Smith, VL 1994 Stock market bubbles in the laboratory. *Applied Mathematical Finance* **1** 111–128

Prast, H 2000a Herding and financial panics: a role for cognitive psychology? De Nederlandische Bank report

Prast, H 2000b A cognitive dissonance models of financial market behaviour. De Nederlandische Bank report

Ramamurtie, S, Prezas, A & Ulman, S 1993 Consistency and identification problems in models of term structure of interest rates. Working Paper, Georgia State University

Rasmussen, H & Wilmott, P 2002 Asymptotic analysis of stochastic volatility models. In *New Directions in Mathematical Finance* (Wilmott, P & Rasmussen, H, Ed) John Wiley & Sons

Ravanelli, C 1999 MSc dissertation, Universita degli Studi di Milano

Rebonato, R 1996 *Interest-rate Option Models*. John Wiley & Sons

Rebonato, R 2004 *Volatility and Correlation*. John Wiley & Sons

Reiner, E & Rubinstein, M 1991 Breaking down the barriers. *Risk* magazine **4** 8

Reuters 1999 *Introduction to Technical Analysis* (Reuters Financial Training Series). John Wiley & Sons

Rich, D & Chance, D 1993 An alternative approach to the pricing of options on multiple assets. *Journal of Financial Engineering* **2** 271–285

Richtmyer, RD & Morton, KW 1976 *Difference Methods for Initial-value Problems*. John Wiley & Sons

Roache PJ 1982 *Computational Fluid Dynamics*. Hermosa, Albuquerque, NM

Rogers, LCG & Satchell, SE 1991 Estimating variance from high, low and closing prices. *Annals of Applied Probability* **1** 504–512

Roll, R 1977 An analytical formula for unprotected American call options on stocks with known dividends. *Journal of Financial Economics* **5** 251–258

Rubinstein, M 1991 Somewhere over the rainbow. *Risk* magazine **4** 10

Rubinstein, M 1994 Implied binomial trees. *Journal of Finance* **69** 771–818

Rupf, I, Dewynne, JN, Howison, SD & Wilmott, P 1993 Some mathematical results in the pricing of American options. *European Journal of Applied Mathematics* **4** 381–398

Sandmann, K & Sondermann, D 1994 On the stability of lognormal interest rate models. Working Paper, University of Bonn, Department of Economics

St Pierre, M 2004 Valuing and hedging synthetic CDO tranches using base correlation. Credit Derivatives, Bear Stearns, May

Schönbucher, PJ 1993 The feedback effect of hedging in illiquid markets. M.Sc. thesis, Oxford University

Schönbucher, PJ 1996 The term structure of defaultable bond prices. Discussion Paper B–384, University of Bonn

Schönbucher, PJ 1997a Modelling defaultable bond prices. Working Paper, London School of Economics, Financial Markets Group

Schönbucher, PJ 1997b Pricing credit risk derivatives. Working Paper, London School of Economics, Financial Markets Group

Schönbucher, PJ 1998 A review of credit risk and credit derivative modelling. To appear in *Applied Mathematical Finance* **5**

Schönbucher, PJ 1999 A market model for stochastic implied volatility. *Phil. Trans.* A **357** 2071–2092

Schönbucher, PJ & Schlögl, E 1996 Credit derivatives and competition in the loan market. Working Paper, University of Bonn, Department of Statistics

Schönbucher, PJ & Wilmott, P 1995a Hedging in illiquid markets: nonlinear effects. *Proceedings of the 8th European Conference on Mathematics in Industry*

Schönbucher, PJ & Wilmott, P 1995b The feedback effect of hedging in illiquid markets. MFG Working Paper, Oxford University

Schoutens, W, Simons, E & Tistaert, J 2004 A perfect calibration! Now what? *Wilmott* magazine March 2004 66–78

Sepp, A & Skachkov, I 2003 Option pricing with jumps. *Wilmott* magazine November 2003 50–58

Shore, S 1997 The modelling of credit risk and its applications to credit derivatives. M.Sc. dissertation, Oxford University

Shu, J & Zhang, J 2003 The relationship between implied and realized volatility of S&P500 index. *Wilmott* magazine January 2003

Sircar, KR & Papanicolaou, G 1996 General Black–Scholes models accounting for increased market volatility from hedging strategies. Stanford University, Working Paper

Sircar, KR & Papanicolaou, GC 1998 Stochastic volatility, smile and asymptotics, to appear in *Applied Mathematical Finance*

Sklar, A 1959 Fonctions de repartition á *n* dimensions et leur marges. *Publ. Int. Stat Univ.,* Paris, **8** 229–231

Skorokhod, AV 1987 *Asymptotic Methods in the Theory of Stochastic Differential Equations*, (Nauka, 1987; AMS, *Translations of Mathematical Monographs*, vol. 78)

Sloan, IH & Walsh, L 1990 A computer search of rank two lattice rules for multidimensional quadrature. *Mathematics of Computation* **54** 281–302

Sneddon, I 1957 *Elements of Partial Differential Equations*. McGraw–Hill

Soros, G 1987 *The Alchemy of Finance*. John Wiley & Sons

Stein, E & Stein, J 1991 Stock price distributions with stochastic volatility: an analytic approach, *Review of Financial Studies* **4**, 727–752

Stetson, C, Marshall, S & Loeball, D 1995 Laudable lattices. *Risk* magazine **8** (12) 60–63

Strang, G 1986 *Introduction to Applied Mathematics*. Wellesley–Cambridge

Stulz, RM 1982 Options on the minimum or maximum of two risky assets. *Journal of Financial Economics* **10** 161–185

Swift, J 1726 *Travels into Several Remote Nations of the World . . . by Lemuel Gulliver*. B. Motte, London

Swishchuk, A 2004 Modeling of variance and volatility swaps for financial markets with stochastic volatility. *Wilmott* magazine September 2004 64–77

Taylor, SJ & Xu, X 1994 The magnitude of implied volatility smiles: theory and empirical evidence for exchange rates, *The Review of Futures Markets*, **13**

Thompson, G 2000 Fast narrow bounds on the value of Asian options, Technical report, University of Cambridge

Thomson, R 1998 *Apocalypse Roulette: The Lethal World of Derivatives*. Pan

Thorp, E 2002 What I knew and when I knew it. *Wilmott* magazine September, December 2002, January 2003

Thorp, EO & Kassouf, S 1967 *Beat the Market*. Random House

Topper, J 2005 *Financial Engineering with Finite Elements*. John Wiley & Sons

Traub, JF & Wozniakowski, H 1994 Breaking intractability. *Scientific American* Jan 102–107

Trigeorgis, L 1998 *Real Options*. MIT Press

Turnbull, SM & Wakeman, LM 1991 A quick algorithm for pricing European average options, *Journal of Financial and Quantitative Finance* **26** 377–389

Ungar, E 1996 *Swap Literacy*. Bloomberg

Vasicek, OA 1977 An equilibrium characterization of the term structure. *Journal of Financial Economics* **5** 177–188

de la Vega, J 1688 *Confusión de Confusiones*. Republished by John Wiley & Sons

Vose, D 1997 *Quantitative Risk Analysis: A Guide to Monte Carlo Simulation Modelling*. John Wiley & Sons

Walsh, O 2003 The art and science of curve building. *Wilmott* magazine November 2003

Whaley, RE 1981 On the valuation of American call options on stocks with known dividends. *Journal of Financial Economics* **9** 207–211

Whaley, RE 1993 Derivatives on market volatility: hedging tools long overdue. *Journal of Derivatives* **1** 71–84

Whaley, RE 1997 Building on Black–Scholes. *Risk* magazine (December)

Whalley, AE & Wilmott, P 1993a Counting the costs. *Risk* magazine **6** (10) 59–66

Whalley, AE & Wilmott, P 1993b Option pricing with transaction costs. MFG Working Paper, Oxford

Whalley, AE & Wilmott, P 1994a Hedge with an edge. *Risk* magazine **7** (10) 82–85

Whalley, AE & Wilmott, P 1994b A comparison of hedging strategies. *Proceedings of the 7th European Conference on Mathematics in Industry* 427–434

Whalley, AE & Wilmott, P 1995 An asymptotic analysis of the Davis, Panas and Zariphopoulou model for option pricing with transaction costs. MFG Working Paper, Oxford University

Whalley, AE & Wilmott, P 1996 Key results in discrete hedging and transaction costs. In *Frontiers in Derivatives* (Konishi, A and Dattatreya, R, Ed) 183–196

Whalley, AE & Wilmott, P 1997 An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance* **7** 307–324

Wiggins, J 1987 Option values and stochastic volatility, *J. Financial Economics* **19**, 351–372

Wilmott, P 1994 Discrete charms. *Risk* magazine **7** (3) 48–51

Wilmott, P 1995 Volatility smiles revisited. *Derivatives Week* **4** (38) 8

Wilmott, P 1998 *Derivatives*. John Wiley & Sons

Wilmott, P 2002 Cliquet options and volatility models. *Wilmott* magazine December 2002

Wilmott, P & Wilmott, S 1990 Dispersion of pollutant along a river, an asymptotic analysis. OCIAM Working Paper

Zhang, PG 1997 *Exotic Options*. World Scientific

Zhu, Y & Avellaneda, M 1998 A risk-neutral stochastic volatility model, *International Journal of Theoretical and Applied Finance*, **1** (2) 289–310

index

Note: bold 1 refers to Volume 1; bold 2 refers to Volume 2 and bold 3 refers to Volume 3.

Index compiled by Annette Musker

WILEY COPYRIGHT INFORMATION AND TERMS OF USE

CD supplement to Paul Wilmott on Quantitative Finance, Second Edition

by Paul Wilmott

ISBN-13: 978-0-470-01870-5 (HB)
ISBN-10: 0-470-01870-4 (HB)

Published by John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England